



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 3, 2023 – 09:26 pm GMT

PDB ID : 2C9L  
Title : Structure of the Epstein-Barr virus ZEBRA protein  
Authors : Petosa, C.; Morand, P.; Baudin, F.; Moulin, M.; Artero, J.B.; Muller, C.W.  
Deposited on : 2005-12-13  
Resolution : 2.25 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](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  
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

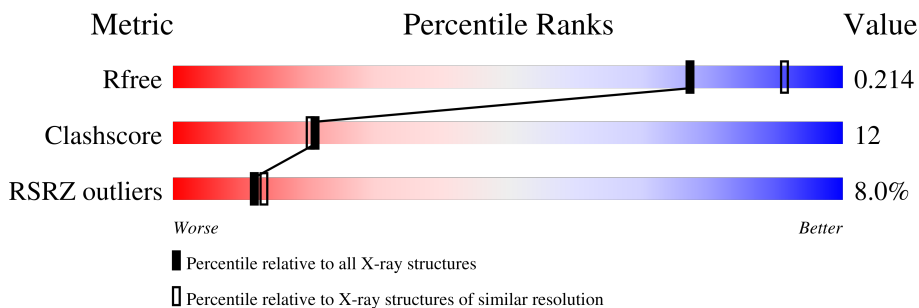
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.25 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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	19	
2	B	18	
3	Y	63	
3	Z	63	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 1878 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 DNA chain called 5'-D(\*AP\*AP\*GP\*CP\*AP\*CP\*TP\*GP\*AP\*CP \*TP\*C P\*AP\*TP\*GP\*AP\*AP\*GP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	19	388	186	75	109	18	0	0	0

- Molecule 2 is a DNA chain called 5'-D(\*AP\*CP\*TP\*TP\*CP\*AP\*CP\*TP\*GP\*AP \*GP\*T P\*CP\*AP\*GP\*TP\*GP\*CP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	18	368	176	64	110	18	0	0	0

- Molecule 3 is a protein called BZLF1 TRANS-ACTIVATOR PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	Y	63	529	332	102	92	3	0	1	0
3	Z	62	512	320	101	89	2	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Y	186	ALA	SER	engineered mutation	UNP P03206
Z	186	ALA	SER	engineered mutation	UNP P03206
Y	189	SER	CYS	engineered mutation	UNP P03206
Z	189	SER	CYS	engineered mutation	UNP P03206

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	17	Total	O	0	0
			17	17		

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	B	15	Total 15	O 15	0	0
4	Y	35	Total 35	O 35	0	0
4	Z	14	Total 14	O 14	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.17Å 26.52Å 98.09Å 90.00° 103.95° 90.00°	Depositor
Resolution (Å)	30.00 – 2.25 37.83 – 2.08	Depositor EDS
% Data completeness (in resolution range)	99.9 (30.00-2.25) 99.8 (37.83-2.08)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.35 (at 2.08Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.232 , 0.264 0.218 , 0.214	Depositor DCC
$R_{free}$ test set	735 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.9	Xtrriage
Anisotropy	0.273	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 62.7	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	1878	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.15% 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	0.81	0/436	1.14	0/671
2	B	0.81	0/411	1.13	2/632 (0.3%)
3	Y	0.32	0/536	0.49	0/714
3	Z	0.33	0/519	0.56	0/691
All	All	0.59	0/1902	0.87	2/2708 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	119	DT	O4'-C1'-N1	5.37	111.76	108.00
2	B	112	DT	O4'-C1'-N1	5.31	111.72	108.00

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	388	0	215	6	0
2	B	368	0	205	2	0
3	Y	529	0	560	16	0
3	Z	512	0	536	21	0
4	A	17	0	0	1	0
4	B	15	0	0	0	0
4	Y	35	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	Z	14	0	0	2	0
All	All	1878	0	1516	40	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 12.

The worst 5 of 40 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)
3:Z:236:ASP:O	4:Z:2014:HOH:O	1.72	1.04
3:Z:177:ILE:HG13	3:Z:179:ARG:HH21	1.31	0.91
3:Z:178:LYS:HD3	3:Z:178:LYS:H	1.38	0.86
3:Y:236:ASP:OXT	4:Y:2034:HOH:O	1.97	0.81
3:Z:177:ILE:HG13	3:Z:179:ARG:NH2	1.96	0.81

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

There are no protein backbone outliers to report in this entry.

### 5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

### 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](#)

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	19/19 (100%)	0.26	1 (5%) 26 29	26, 47, 76, 79	0
2	B	18/18 (100%)	0.22	1 (5%) 24 26	24, 51, 82, 91	0
3	Y	63/63 (100%)	0.43	2 (3%) 47 50	22, 46, 66, 86	0
3	Z	62/63 (98%)	0.89	9 (14%) 2 2	26, 52, 94, 106	0
All	All	162/163 (99%)	0.56	13 (8%) 12 13	22, 49, 89, 106	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Y	174	MET	7.9
3	Z	176	GLU	6.4
3	Z	236	ASP	5.4
3	Z	175	LEU	4.9
3	Z	230	ILE	3.7

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.