



wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 05:58 am BST

PDB ID : 2W45
Title : Epstein-Barr virus alkaline nuclease
Authors : Buisson, M.; Geoui, T.; Flot, D.; Tarbouriech, N.; Burmeister, W.P.
Deposited on : 2008-11-21
Resolution : 3.00 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 ⓘ 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

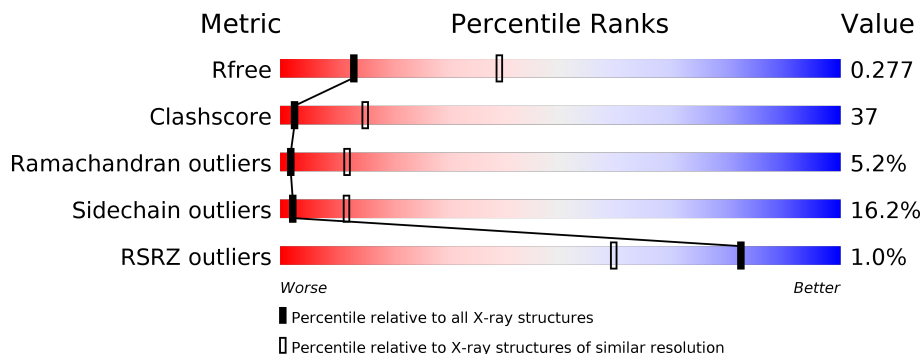
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 ≥ 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	470	
1	B	470	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7009 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 ALKALINE EXONUCLEASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	449	3539	2257	595	665	22	0	0	0
1	B	437	3446	2201	579	644	22	0	0	0

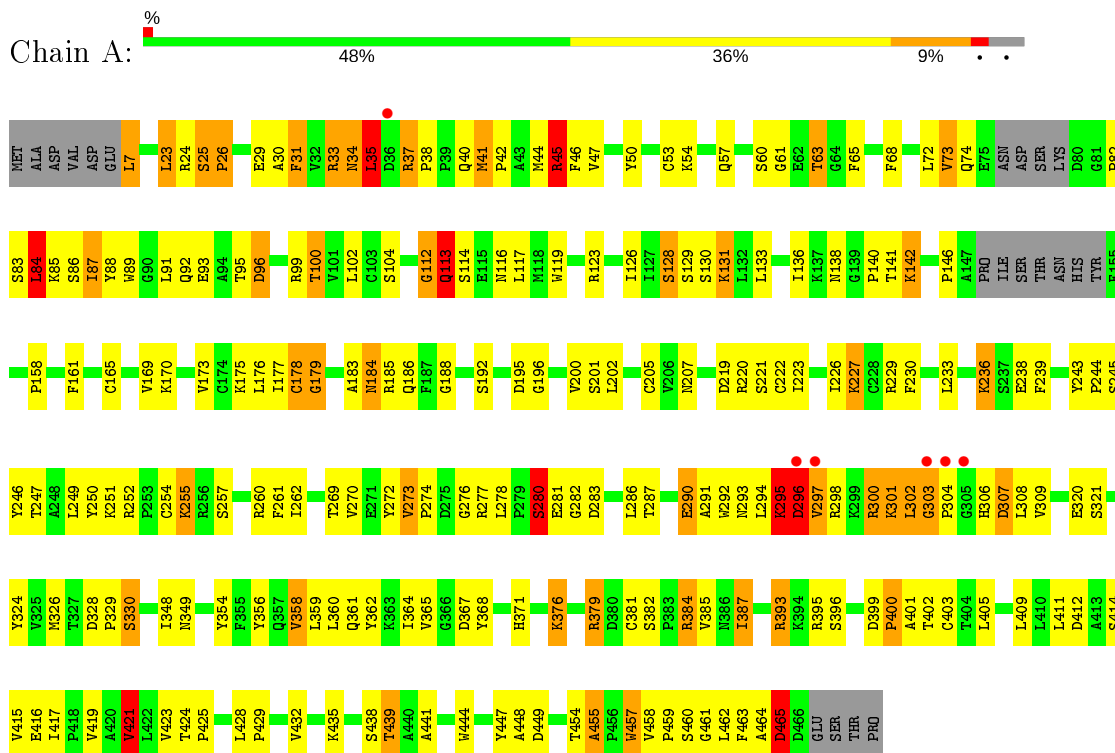
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	10	Total	O	0	0
			10	10		
2	B	14	Total	O	0	0
			14	14		

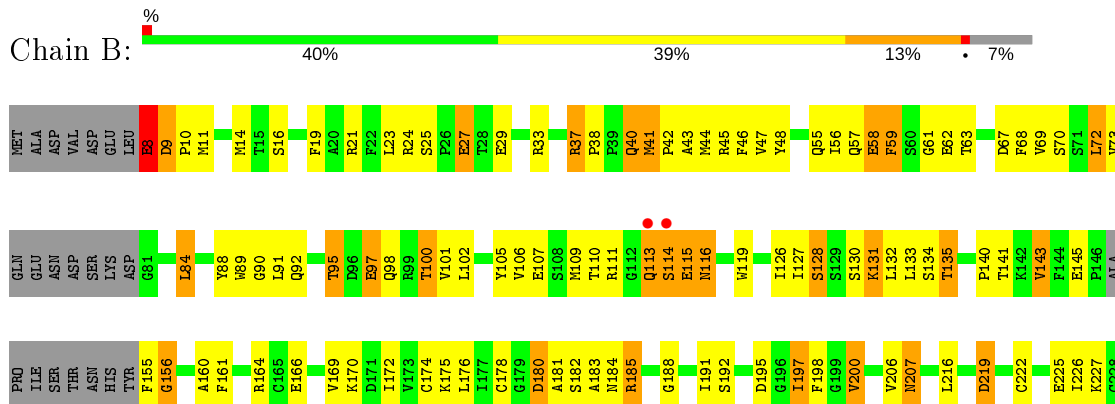
3 Residue-property plots

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: ALKALINE EXONUCLEASE



• Molecule 1: ALKALINE EXONUCLEASE



A446	G373	R229	R256	S414	S257	R267	R267	S415	R267	M292
Y447	G374	F250	S257	V415	Y297	R267	R267	V415	S257	M293
A448	G375	K231	F258	V416	R298	P268	P268	V416	S280	R294
D449	K376	K299	I259	E416	K299	T269	T269	E416	E281	L294
N450	PRO	R300	R260	I417	K301	R269	R269	I417	G282	L286
T451	GLY	K301	R261	P418	L302	V270	V270	P418	D283	T287
F452	ARG	L302	L262	V419	G303	E271	E271	V419	Q288	Q288
D453	D380	G303	M263	V423	P304	Y272	Y272	V423	M292	M292
T454	C381	G305	R267	T424	R304	V273	V273	T424	S280	S280
A455	S382	E238	R251	T425	P305	P274	P274	T425	E281	E281
P456	P383	H306	R252	P425	H306	D275	D275	P425	G276	G276
W457	R384	D240	R252	V426	V309	G276	G276	V426	R277	R277
V458	V385	P241	R252	V427	V309	R277	R277	V427	S280	S280
P459	M386	L242	R252	V427	A314	S280	S280	V427	E281	E281
S460	I387	Y243	R252	V427	A315	E281	E281	V427	G282	G282
L462	V388	P244	R252	V427	N316	L284	L284	V427	L284	L284
PHE	R393	Y246	R252	V427	R317	L284	L284	V427	L286	L286
ALA	K394	T247	R252	V427	S321	Q288	Q288	V427	T287	T287
ASP	R395	A248	R252	V427	M326	V358	V358	V427	Q288	Q288
ASP	D399	L249	R252	V427	T327	Q357	Q357	V427	Q288	Q288
GLU	P400	Y250	R252	V427	D328	V358	V358	V427	Q288	Q288
SER	A401	K251	R252	V427	P329	Q361	Q361	V427	Q288	Q288
THR	T402	R252	R252	V427	S330	Q361	Q361	V427	Q288	Q288
PRO	C403	R252	R252	V427	E331	Q361	Q361	V427	Q288	Q288
	T404	R252	R252	V427	M332	Q361	Q361	V427	Q288	Q288
	S407	R252	R252	V427	A333	Q361	Q361	V427	Q288	Q288
	D408	R252	R252	V427	G334	Q361	Q361	V427	Q288	Q288
	L409	R252	R252	V427	R335	Q361	Q361	V427	Q288	Q288
	L410	R252	R252	V427	T336	Q361	Q361	V427	Q288	Q288
	S414	R252	R252	V427	G337	Q361	Q361	V427	Q288	Q288
	V415	R252	R252	V427	I338	Q361	Q361	V427	Q288	Q288
	E416	R252	R252	V427	K339	Q361	Q361	V427	Q288	Q288
	I417	R252	R252	V427	D340	Q361	Q361	V427	Q288	Q288
	P418	R252	R252	V427	R341	Q361	Q361	V427	Q288	Q288
	V419	R252	R252	V427	V342	Q361	Q361	V427	Q288	Q288
	V423	R252	R252	V427	F343	Q361	Q361	V427	Q288	Q288
	T424	R252	R252	V427	V344	Q361	Q361	V427	Q288	Q288
	P425	R252	R252	V427	I348	Q361	Q361	V427	Q288	Q288
	V426	R252	R252	V427	N349	Q361	Q361	V427	Q288	Q288
	V427	R252	R252	V427	Y354	Q361	Q361	V427	Q288	Q288
	D430	R252	R252	V427	F355	Q361	Q361	V427	Q288	Q288
	S431	R252	R252	V427	Y356	Q361	Q361	V427	Q288	Q288
	K435	R252	R252	V427	Q357	Q361	Q361	V427	Q288	Q288
	T436	R252	R252	V427	V358	Q361	Q361	V427	Q288	Q288
	L437	R252	R252	V427	Q361	Q361	Q361	V427	Q288	Q288
	S438	R252	R252	V427	Y362	Q361	Q361	V427	Q288	Q288
	T439	R252	R252	V427	K363	Q361	Q361	V427	Q288	Q288
	G442	R252	R252	V427	I364	Q361	Q361	V427	Q288	Q288
	S443	R252	R252	V427	S438	Q361	Q361	V427	Q288	Q288
	W444	R252	R252	V427	T439	Q361	Q361	V427	Q288	Q288
	K445	R252	R252	V427	G442	Q361	Q361	V427	Q288	Q288
		R252	R252	V427	S443	Q361	Q361	V427	Q288	Q288
		R252	R252	V427	W444	Q361	Q361	V427	Q288	Q288
		R252	R252	V427	K445	Q361	Q361	V427	Q288	Q288

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	87.51Å 63.79Å 114.13Å 90.00° 93.59° 90.00°	Depositor
Resolution (Å)	71.43 – 3.00 71.51 – 3.00	Depositor EDS
% Data completeness (in resolution range)	94.7 (71.43-3.00) 94.2 (71.51-3.00)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.86 (at 3.01Å)	Xtrriage
Refinement program	REFMAC 5.5.0068	Depositor
R, R_{free}	0.192 , 0.282 0.189 , 0.277	Depositor DCC
R_{free} test set	1235 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å ²)	49.5	Xtrriage
Anisotropy	0.377	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 66.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	7009	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.67% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	3/3622 (0.1%)	0.99	8/4912 (0.2%)
1	B	0.92	3/3526 (0.1%)	1.00	4/4780 (0.1%)
All	All	0.92	6/7148 (0.1%)	1.00	12/9692 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	B	0	5
All	All	0	7

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	222	CYS	CB-SG	-6.57	1.71	1.82
1	A	165	CYS	CB-SG	-6.45	1.71	1.82
1	B	27	GLU	CB-CG	5.44	1.62	1.52
1	A	178	CYS	CB-SG	-5.33	1.73	1.81
1	B	457	TRP	CB-CG	-5.12	1.41	1.50

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	24	ARG	NE-CZ-NH1	-7.04	116.78	120.30
1	B	84	LEU	CB-CG-CD1	-6.92	99.24	111.00
1	B	128	SER	N-CA-CB	-6.48	100.78	110.50
1	A	7	LEU	CA-CB-CG	6.31	129.81	115.30
1	B	381	CYS	N-CA-C	6.19	127.72	111.00

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	303	GLY	Peptide
1	A	41	MET	Peptide
1	B	293	ASN	Peptide
1	B	41	MET	Peptide
1	B	8	GLU	Peptide

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	3539	0	3493	257	0
1	B	3446	0	3413	260	0
2	A	10	0	0	1	0
2	B	14	0	0	3	0
All	All	7009	0	6906	511	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 37.

The worst 5 of 511 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:301:LYS:HE2	1:A:301:LYS:C	1.37	1.40
1:A:294:LEU:O	1:A:295:LYS:CD	1.71	1.36
1:A:294:LEU:O	1:A:295:LYS:HD3	1.19	1.35
1:A:300:ARG:NH1	1:A:302:LEU:HD13	1.42	1.31
1:B:131:LYS:O	1:B:135:THR:HG22	1.32	1.26

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	443/470 (94%)	364 (82%)	57 (13%)	22 (5%)	2	12
1	B	429/470 (91%)	347 (81%)	59 (14%)	23 (5%)	2	11
All	All	872/940 (93%)	711 (82%)	116 (13%)	45 (5%)	2	12

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	42	PRO
1	A	295	LYS
1	A	296	ASP
1	A	465	ASP
1	B	42	PRO

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	394/414 (95%)	338 (86%)	56 (14%)	3	16
1	B	385/414 (93%)	315 (82%)	70 (18%)	1	9
All	All	779/828 (94%)	653 (84%)	126 (16%)	2	12

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	8	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	113	GLN
1	B	414	SER
1	B	24	ARG
1	B	70	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	345	ASN
1	B	55	GLN
1	B	211	GLN
1	A	293	ASN
1	B	116	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 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, 95th 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	449/470 (95%)	-0.05	6 (1%)	77 51	14, 31, 60, 87	0
1	B	437/470 (92%)	-0.08	3 (0%)	87 69	16, 34, 60, 77	0
All	All	886/940 (94%)	-0.06	9 (1%)	82 59	14, 33, 60, 87	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	303	GLY	4.2
1	A	305	GLY	3.5
1	A	296	ASP	3.4
1	A	297	VAL	2.9
1	B	114	SER	2.5

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.