



wwPDB X-ray Structure Validation Summary Report ⓘ

May 27, 2020 – 01:03 am BST

PDB ID : 2W4B
Title : Epstein-Barr virus alkaline nuclease D203S mutant
Authors : Buisson, M.; Geoui, T.; Flot, D.; Tarbouriech, N.; Burmeister, W.P.
Deposited on : 2008-11-24
Resolution : 3.50 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

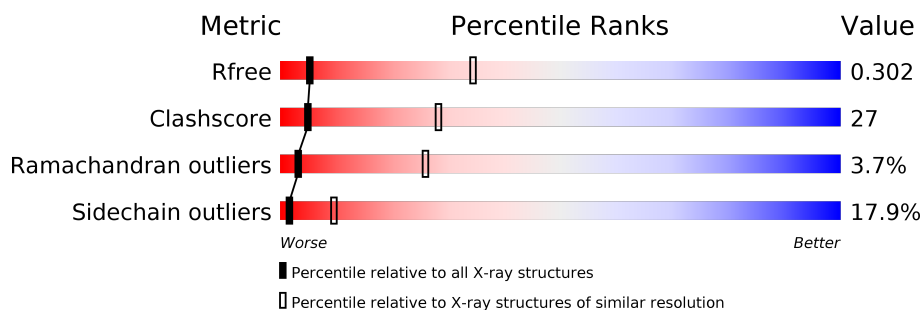
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)

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 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	470	
1	B	470	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 7041 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
1	A	456	Total	C	N	O	S	0	0	1
			3587	2288	605	672	22			
1	B	439	Total	C	N	O	S	0	0	1
			3454	2205	582	645	22			

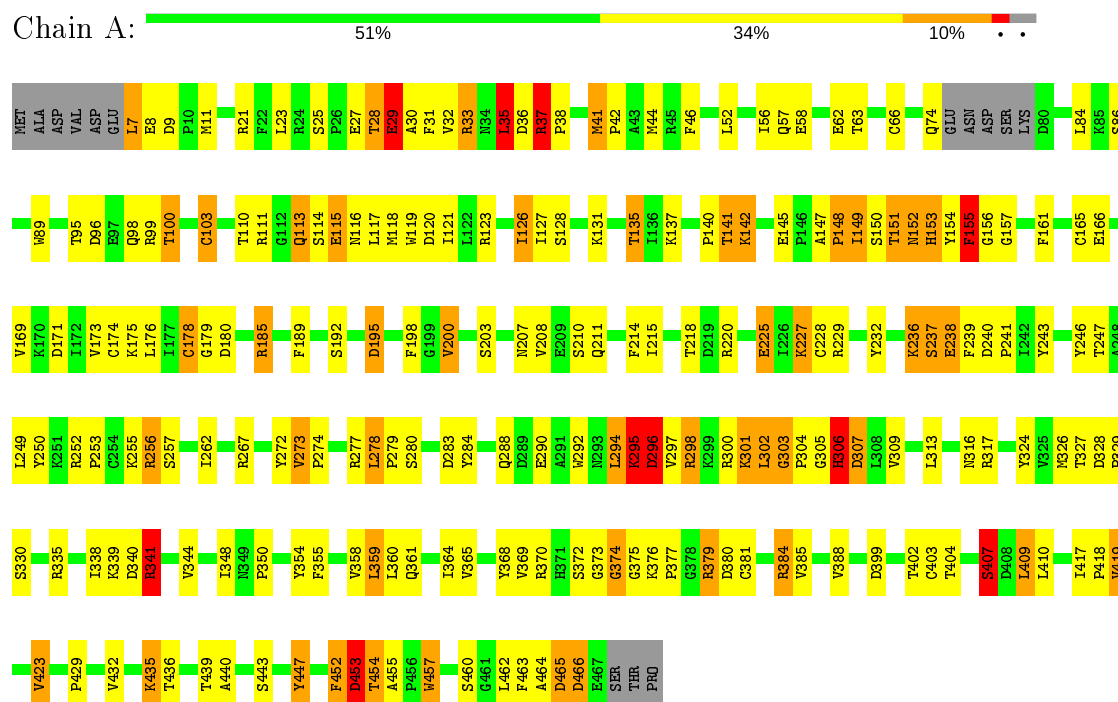
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	203	SER	ASP	engineered mutation	UNP P03217
B	203	SER	ASP	engineered mutation	UNP P03217

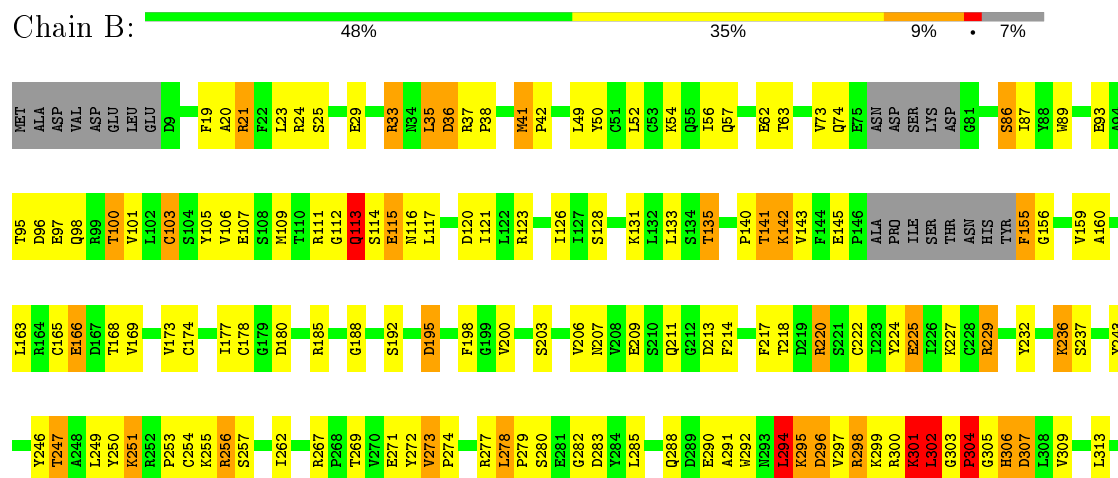
3 Residue-property plots

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. 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. 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



N316	N317	G318	V319	Y324	T327	D328	P329	N332	R335	K339	D340	R341	V344	N345	I346	P350	Y364	F365	Y368	V369	R370	H371	S372	G373	G374	G375	K376	PRO	GLY	ARG	D380	C381	R384	V385	N386	I387	F391	F392	R393								
K394	R395	T402	C403	T404	S407	D408	L409	L410	L411	D412	A413	S414	V415	E416	I417	P418	V419	A420	V423	T424	P429	V432	I433	R434	K435	T439	A440	S443	W444	K445	A446	Y447	T454	A455	P456	W457	V458	P459	L462	F463	ALA	ASP	ASP	GLU	SER	THR	PRO

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)	31.19 – 3.50 31.19 – 3.00	Depositor EDS
% Data completeness (in resolution range)	84.8 (31.19-3.50) 99.8 (31.19-3.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.24 (at 3.00Å)	Xtriage
Refinement program	REFMAC 5.5.0038	Depositor
R, R_{free}	0.190 , 0.256 0.262 , 0.302	Depositor DCC
R_{free} test set	1295 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	52.9	Xtriage
Anisotropy	0.346	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 46.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	7041	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.35% 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

5.1 Standard geometry

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	1.18	8/3674 (0.2%)	1.07	14/4987 (0.3%)
1	B	0.99	6/3534 (0.2%)	1.02	19/4791 (0.4%)
All	All	1.09	14/7208 (0.2%)	1.05	33/9778 (0.3%)

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	8
1	B	0	4
All	All	0	12

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	452	PHE	C-N	-40.57	0.40	1.34
1	A	453	ASP	C-N	-23.24	0.80	1.34
1	B	35	LEU	C-N	-21.17	0.85	1.34
1	B	36	ASP	C-N	-11.74	1.07	1.34
1	B	381	CYS	CB-SG	-9.48	1.66	1.82

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	452	PHE	O-C-N	-19.02	92.27	122.70
1	A	35	LEU	C-N-CA	18.66	168.35	121.70
1	A	452	PHE	C-N-CA	14.54	158.05	121.70
1	A	35	LEU	O-C-N	-13.65	100.86	122.70
1	B	36	ASP	O-C-N	-11.61	104.12	122.70

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	149	ILE	Peptide
1	A	150	SER	Peptide
1	A	35	LEU	Peptide
1	A	407	SER	Peptide
1	A	41	MET	Peptide

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	3587	0	3537	204	5
1	B	3454	0	3417	188	1
All	All	7041	0	6954	384	5

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 27.

The worst 5 of 384 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:271:GLU:OE1	1:B:298:ARG:NH1	1.60	1.32
1:A:453:ASP:C	1:A:454:THR:CA	2.00	1.30
1:A:116:ASN:OD1	1:A:117:LEU:N	1.65	1.28
1:B:302:LEU:O	1:B:304:PRO:HD3	1.34	1.26
1:A:95:THR:HG21	1:B:178:CYS:O	1.38	1.23

All (5) 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)
1:A:341:ARG:NH1	1:B:36:ASP:O[2_645]	1.69	0.51
1:A:58:GLU:OE2	1:A:156:GLY:N[2_656]	1.72	0.48
1:A:58:GLU:OE2	1:A:156:GLY:CA[2_656]	1.81	0.39
1:A:58:GLU:OE2	1:A:156:GLY:C[2_656]	1.94	0.26

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:58:GLU:OE2	1:A:157:GLY:N[2_656]	2.17	0.03

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	452/470 (96%)	370 (82%)	65 (14%)	17 (4%)	3	25
1	B	431/470 (92%)	369 (86%)	46 (11%)	16 (4%)	3	26
All	All	883/940 (94%)	739 (84%)	111 (13%)	33 (4%)	3	26

5 of 33 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	9	ASP
1	A	42	PRO
1	A	148	PRO
1	A	295	LYS
1	A	296	ASP

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	400/414 (97%)	325 (81%)	75 (19%)	1	8
1	B	386/414 (93%)	320 (83%)	66 (17%)	2	12

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	786/828 (95%)	645 (82%)	141 (18%)	2 9

5 of 141 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	402	THR
1	B	33	ARG
1	B	402	THR
1	A	409	LEU
1	A	447	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	40	GLN
1	B	55	GLN
1	B	332	ASN
1	A	357	GLN
1	B	316	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	3
1	B	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	36:ASP	C	37:ARG	N	1.13
1	B	36:ASP	C	37:ARG	N	1.07
1	B	35:LEU	C	36:ASP	N	0.85
1	A	453:ASP	C	454:THR	N	0.80
1	A	452:PHE	C	453:ASP	N	0.40

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.