



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 07:33 am BST

PDB ID : 6K12  
Title : Babesia microti lactate dehydrogenase apo form (BmLDH)  
Authors : Long, Y.  
Deposited on : 2019-05-09  
Resolution : 2.79 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 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

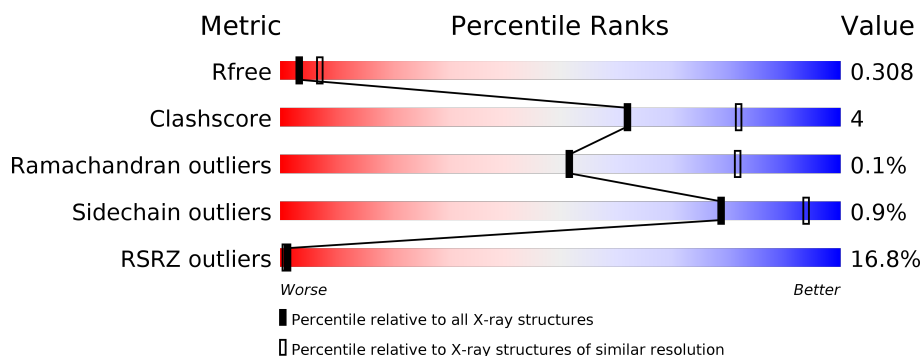
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.79 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	332	<div> <div>12%</div> <div>82%</div> <div>8%</div> <div>11%</div> </div>
1	B	332	<div> <div>10%</div> <div>79%</div> <div>5%</div> <div>17%</div> </div>
1	C	332	<div> <div>8%</div> <div>80%</div> <div>5%</div> <div>16%</div> </div>
1	D	332	<div> <div>18%</div> <div>71%</div> <div>9%</div> <div>20%</div> </div>
1	E	332	<div> <div>15%</div> <div>62%</div> <div>11%</div> <div>27%</div> </div>
1	F	332	<div> <div>19%</div> <div>69%</div> <div>10%</div> <div>20%</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 25316 atoms, of which 12881 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.

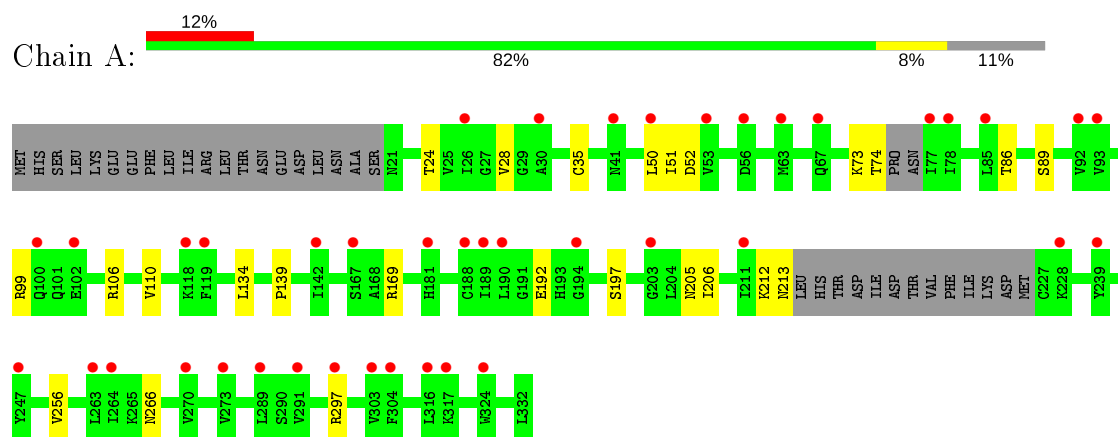
- Molecule 1 is a protein called L-lactate dehydrogenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	297	Total	C	H	N	O	S	0	0	0
			4646	1459	2364	385	426	12			
1	B	277	Total	C	H	N	O	S	0	0	0
			4316	1364	2195	352	394	11			
1	C	280	Total	C	H	N	O	S	0	0	0
			4362	1377	2221	354	398	12			
1	D	265	Total	C	H	N	O	S	0	0	0
			4131	1310	2096	334	380	11			
1	E	244	Total	C	H	N	O	S	0	0	0
			3754	1181	1908	306	348	11			
1	F	265	Total	C	H	N	O	S	0	0	0
			4107	1289	2097	333	377	11			

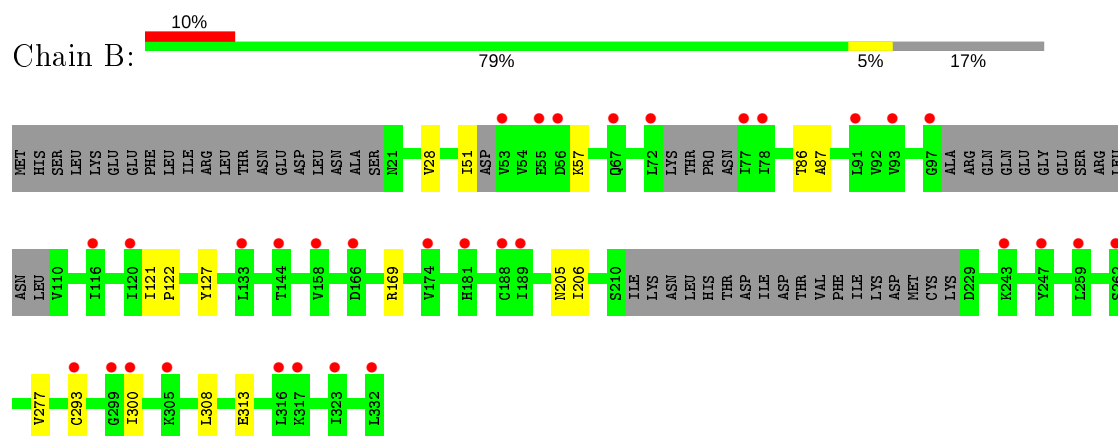
### 3 Residue-property plots [i](#)

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.

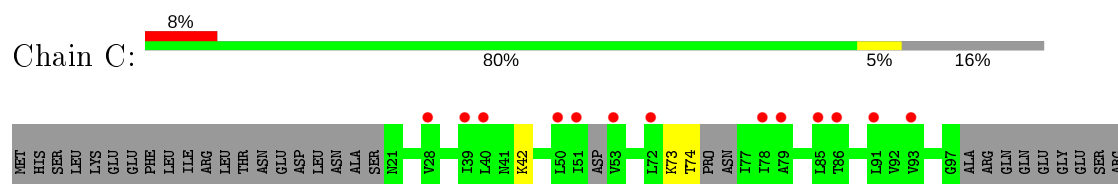
- Molecule 1: L-lactate dehydrogenase

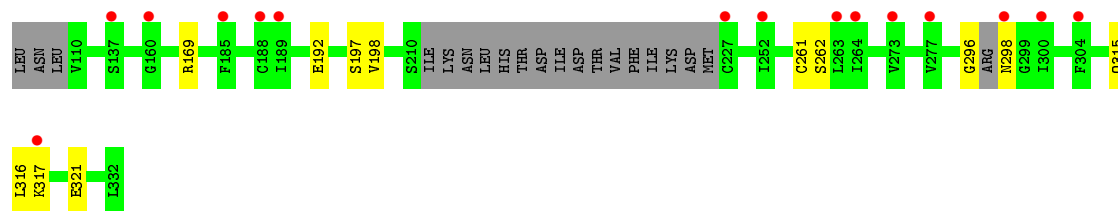


- Molecule 1: L-lactate dehydrogenase

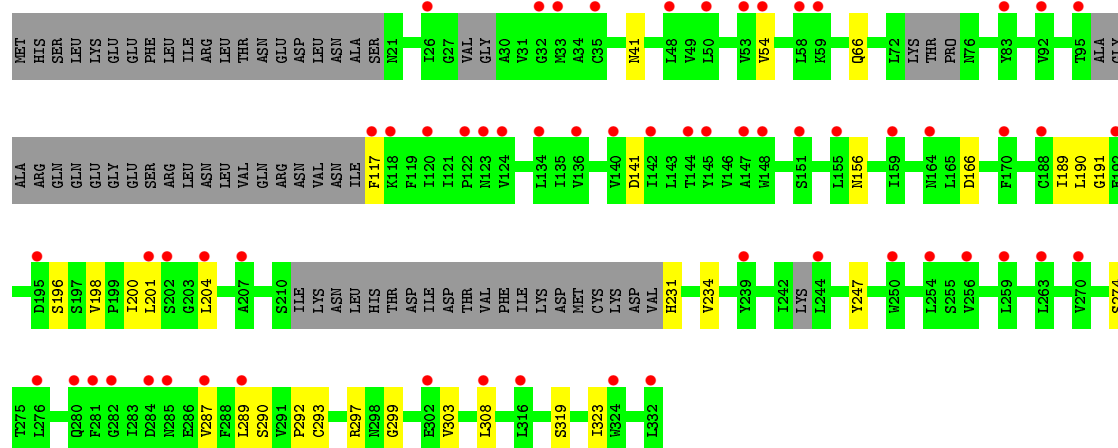


- Molecule 1: L-lactate dehydrogenase

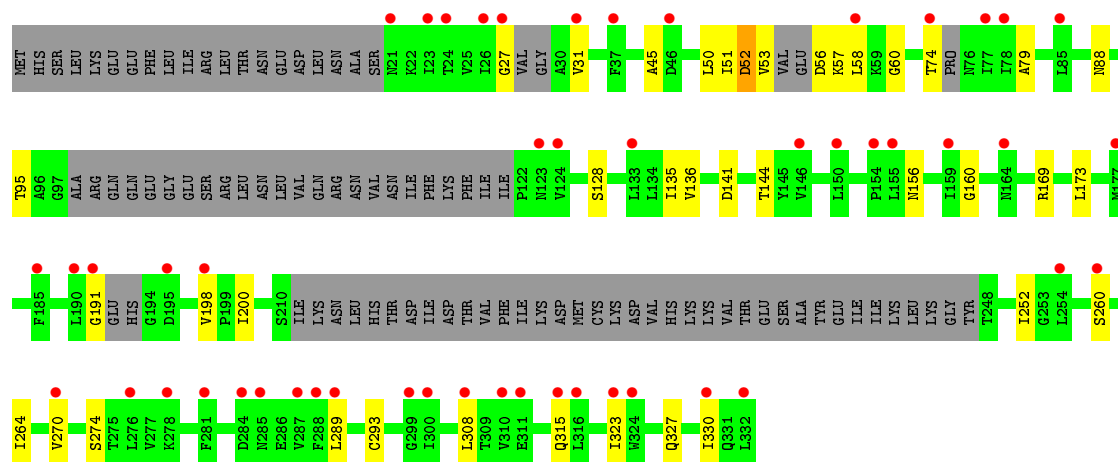




• Molecule 1: L-lactate dehydrogenase

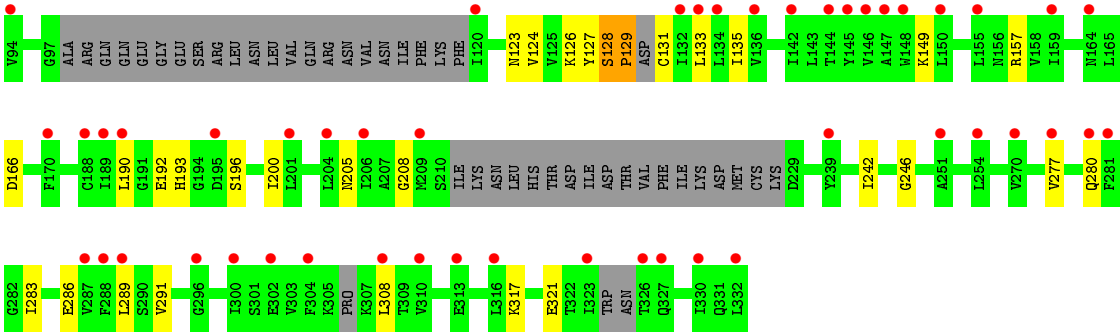


• Molecule 1: L-lactate dehydrogenase



• Molecule 1: L-lactate dehydrogenase





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	255.60Å 147.68Å 65.27Å 90.00° 90.01° 90.00°	Depositor
Resolution (Å)	45.66 – 2.79 45.66 – 2.79	Depositor EDS
% Data completeness (in resolution range)	99.2 (45.66-2.79) 99.1 (45.66-2.79)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.24 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.12 _2829	Depositor
R, $R_{free}$	0.280 , 0.308 0.280 , 0.308	Depositor DCC
$R_{free}$ test set	3009 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	64.0	Xtriage
Anisotropy	0.393	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 20.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.020 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.019 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.477 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.477 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.019 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	25316	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.52% 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	0.41	0/2315	0.47	0/3127
1	B	0.26	0/2152	0.44	0/2908
1	C	0.45	0/2171	0.46	0/2931
1	D	0.28	0/2065	0.44	0/2789
1	E	0.31	0/1869	0.48	1/2523 (0.0%)
1	F	0.42	0/2033	0.49	0/2741
All	All	0.36	0/12605	0.46	1/17019 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	53	VAL	N-CA-C	7.63	131.59	111.00

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	2282	2364	2363	17	0
1	B	2121	2195	2194	8	0
1	C	2141	2221	2218	7	0
1	D	2035	2096	2095	19	0
1	E	1846	1908	1904	33	0
1	F	2010	2097	2091	28	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	12435	12881	12865	111	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 4.

The worst 5 of 111 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:E:289:LEU:HD21	1:E:323:ILE:CD1	1.42	1.48
1:E:289:LEU:CD2	1:E:323:ILE:HD11	1.67	1.25
1:F:128:SER:O	1:F:131:CYS:SG	1.93	1.25
1:E:289:LEU:CD2	1:E:323:ILE:CD1	2.16	1.23
1:F:124:VAL:O	1:F:128:SER:OG	1.64	1.14

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	291/332 (88%)	274 (94%)	16 (6%)	1 (0%)	41	72
1	B	267/332 (80%)	258 (97%)	9 (3%)	0	100	100
1	C	268/332 (81%)	259 (97%)	9 (3%)	0	100	100
1	D	253/332 (76%)	237 (94%)	16 (6%)	0	100	100
1	E	230/332 (69%)	220 (96%)	10 (4%)	0	100	100
1	F	249/332 (75%)	223 (90%)	25 (10%)	1 (0%)	34	66
All	All	1558/1992 (78%)	1471 (94%)	85 (6%)	2 (0%)	51	81

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	28	VAL
1	A	52	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	259/293 (88%)	258 (100%)	1 (0%)	91	97
1	B	241/293 (82%)	240 (100%)	1 (0%)	91	97
1	C	244/293 (83%)	242 (99%)	2 (1%)	81	94
1	D	232/293 (79%)	228 (98%)	4 (2%)	60	87
1	E	211/293 (72%)	208 (99%)	3 (1%)	67	90
1	F	230/293 (78%)	228 (99%)	2 (1%)	78	94
All	All	1417/1758 (81%)	1404 (99%)	13 (1%)	78	94

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	117	PHE
1	D	247	TYR
1	E	156	ASN
1	D	41	ASN
1	E	56	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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	297/332 (89%)	1.07	41 (13%) 2 1	45, 56, 68, 95	0
1	B	277/332 (83%)	1.08	32 (11%) 4 2	43, 54, 63, 70	0
1	C	280/332 (84%)	1.05	28 (10%) 7 4	45, 56, 67, 73	0
1	D	265/332 (79%)	1.41	60 (22%) 0 0	54, 75, 103, 111	0
1	E	244/332 (73%)	1.38	50 (20%) 1 0	52, 71, 100, 108	0
1	F	265/332 (79%)	1.41	62 (23%) 0 0	53, 73, 96, 109	0
All	All	1628/1992 (81%)	1.23	273 (16%) 1 1	43, 62, 95, 111	0

The worst 5 of 273 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	308	LEU	8.3
1	F	59	LYS	7.3
1	E	285	ASN	7.2
1	D	332	LEU	6.0
1	D	144	THR	5.9

### 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.