



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 06:18 PM UTC

PDB ID : 7VKW / pdb_00007vkw
Title : The apo structure of beta-1,2-glucosyltransferase from Ignavibacterium album
Authors : Kobayashi, K.; Shimizu, H.; Tanaka, N.; Kuramochi, K.; Nakai, H.; Nakajima, M.; Taguchi, H.
Deposited on : 2021-10-01
Resolution : 1.75 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

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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

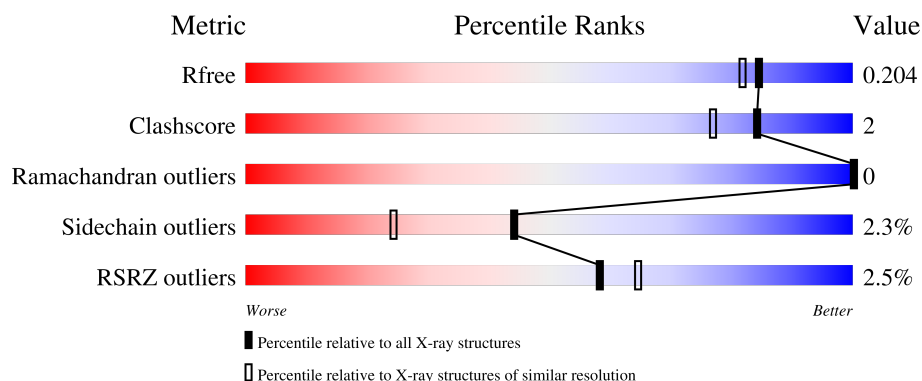
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.75 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}	180053	3183 (1.76-1.76)
Clashscore	190562	3299 (1.76-1.76)
Ramachandran outliers	187476	3274 (1.76-1.76)
Sidechain outliers	187428	3274 (1.76-1.76)
RSRZ outliers	180081	3183 (1.76-1.76)

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 ≥ 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	724	2% 91% 6% .
1	B	724	3% 90% 6% ..

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 12380 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 beta-1,2-glucosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	705	Total	C	N	O	S	0	0	0
			5830	3791	956	1067	16			
1	B	704	Total	C	N	O	S	0	0	0
			5821	3785	954	1066	16			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	717	LEU	-	expression tag	UNP I0AIT9
A	718	GLU	-	expression tag	UNP I0AIT9
A	719	HIS	-	expression tag	UNP I0AIT9
A	720	HIS	-	expression tag	UNP I0AIT9
A	721	HIS	-	expression tag	UNP I0AIT9
A	722	HIS	-	expression tag	UNP I0AIT9
A	723	HIS	-	expression tag	UNP I0AIT9
A	724	HIS	-	expression tag	UNP I0AIT9
B	717	LEU	-	expression tag	UNP I0AIT9
B	718	GLU	-	expression tag	UNP I0AIT9
B	719	HIS	-	expression tag	UNP I0AIT9
B	720	HIS	-	expression tag	UNP I0AIT9
B	721	HIS	-	expression tag	UNP I0AIT9
B	722	HIS	-	expression tag	UNP I0AIT9
B	723	HIS	-	expression tag	UNP I0AIT9
B	724	HIS	-	expression tag	UNP I0AIT9

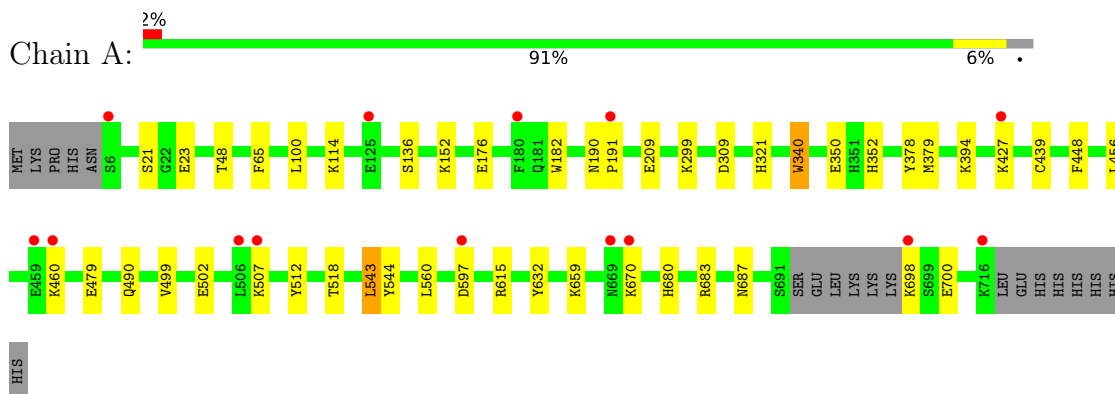
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	390	Total	O	0	0
			390	390		
2	B	339	Total	O	0	0
			339	339		

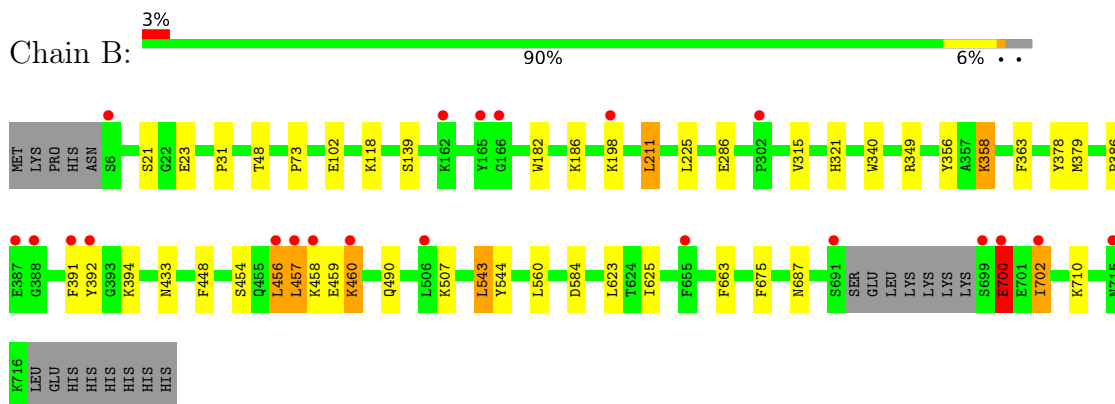
3 Residue-property plots [i](#)

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: beta-1,2-glucosyltransferase



- Molecule 1: beta-1,2-glucosyltransferase



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	164.65Å 71.68Å 129.98Å 90.00° 105.39° 90.00°	Depositor
Resolution (Å)	65.33 – 1.75 65.33 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.4 (65.33-1.75) 99.4 (65.33-1.75)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.44 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.162 , 0.197 0.174 , 0.204	Depositor DCC
R_{free} test set	7304 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å ²)	12.8	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 34.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	12380	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.58% 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	1.06	2/5980 (0.0%)	1.19	4/8074 (0.0%)
1	B	1.06	0/5971	1.21	10/8063 (0.1%)
All	All	1.06	2/11951 (0.0%)	1.20	14/16137 (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	B	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	340	TRP	C-O	5.24	1.30	1.23
1	A	209	GLU	CD-OE1	5.20	1.35	1.25

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	392	TYR	CB-CA-C	-6.74	98.27	109.99
1	B	391	PHE	CA-C-O	-5.96	113.36	120.10
1	B	700	GLU	CA-C-O	-5.58	114.85	121.32
1	B	392	TYR	N-CA-C	5.51	121.87	113.61
1	B	448	PHE	CA-CB-CG	5.43	119.23	113.80
1	B	459	GLU	CA-C-N	5.40	128.43	120.82
1	B	459	GLU	C-N-CA	5.40	128.43	120.82
1	A	502	GLU	CB-CG-CD	5.23	121.49	112.60
1	A	615	ARG	CG-CD-NE	5.20	123.43	112.00
1	A	448	PHE	CA-CB-CG	5.13	118.93	113.80
1	B	700	GLU	O-C-N	5.13	128.76	121.83
1	A	479	GLU	CB-CG-CD	5.06	121.20	112.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	700	GLU	CA-C-N	5.04	130.43	123.07
1	B	700	GLU	C-N-CA	5.04	130.43	123.07

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	386	PRO	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	5830	0	5769	22	0
1	B	5821	0	5756	33	0
2	A	390	0	0	1	0
2	B	339	0	0	2	0
All	All	12380	0	11525	53	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 (53) 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:623:LEU:HD11	1:B:625:ILE:HD11	1.32	1.04
1:B:433:ASN:ND2	2:B:801:HOH:O	1.90	1.02
1:A:543:LEU:HD13	1:A:543:LEU:C	2.09	0.76
1:B:623:LEU:CD1	1:B:625:ILE:HD11	2.17	0.72
1:A:543:LEU:HD13	1:A:544:TYR:N	2.12	0.64
1:A:543:LEU:C	1:A:543:LEU:CD1	2.71	0.62
1:B:211:LEU:HD13	1:B:211:LEU:C	2.29	0.57
1:B:182:TRP:CE2	1:B:456:LEU:HD22	2.39	0.56
1:A:560:LEU:C	1:A:560:LEU:HD23	2.32	0.54
1:B:543:LEU:C	1:B:543:LEU:HD13	2.34	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:358:LYS:NZ	2:B:805:HOH:O	2.42	0.51
1:B:623:LEU:HD11	1:B:625:ILE:CD1	2.23	0.49
1:A:182:TRP:CE2	1:A:456:LEU:HD12	2.47	0.49
1:A:190:ASN:CG	1:A:191:PRO:HD2	2.37	0.49
1:B:31:PRO:HG2	1:B:73:PRO:HB2	1.95	0.49
1:B:457:LEU:HB3	1:B:460:LYS:NZ	2.27	0.49
1:B:560:LEU:C	1:B:560:LEU:HD23	2.38	0.48
1:B:543:LEU:HD13	1:B:544:TYR:N	2.29	0.48
1:A:439:CYS:SG	1:A:499:VAL:HG12	2.55	0.47
1:B:23:GLU:HG2	1:B:379:MET:SD	2.55	0.47
1:B:623:LEU:HG	1:B:625:ILE:HG12	1.97	0.47
1:A:597:ASP:OD1	1:A:597:ASP:C	2.56	0.46
1:A:23:GLU:HG2	1:A:379:MET:SD	2.55	0.46
1:A:394:LYS:HG2	1:B:584:ASP:OD2	2.16	0.46
1:B:663:PHE:HA	1:B:675:PHE:O	2.16	0.45
1:A:100:LEU:HD23	2:A:839:HOH:O	2.16	0.45
1:A:378:TYR:HA	1:A:379:MET:HA	1.82	0.45
1:B:457:LEU:HB3	1:B:460:LYS:HZ1	1.80	0.45
1:B:102:GLU:OE2	1:B:349:ARG:CG	2.65	0.45
1:A:321:HIS:HA	1:A:490:GLN:OE1	2.17	0.45
1:B:394:LYS:HE3	1:B:394:LYS:HB2	1.76	0.45
1:B:211:LEU:C	1:B:211:LEU:CD1	2.89	0.45
1:B:543:LEU:HB3	1:B:625:ILE:HD13	1.98	0.44
1:B:315:VAL:HG21	1:B:363:PHE:HB2	1.97	0.44
1:A:350:GLU:HB3	1:A:352:HIS:CE1	2.53	0.44
1:B:358:LYS:HA	1:B:358:LYS:HE2	1.97	0.44
1:A:687:ASN:HA	1:A:700:GLU:O	2.18	0.44
1:B:543:LEU:C	1:B:543:LEU:CD1	2.90	0.44
1:B:102:GLU:OE2	1:B:349:ARG:HG2	2.18	0.44
1:B:321:HIS:HA	1:B:490:GLN:OE1	2.18	0.43
1:A:176:GLU:OE2	1:A:309:ASP:OD2	2.37	0.43
1:B:378:TYR:HA	1:B:379:MET:HA	1.80	0.43
1:B:454:SER:OG	1:B:457:LEU:HB2	2.18	0.43
1:B:211:LEU:HD13	1:B:211:LEU:O	2.19	0.42
1:B:687:ASN:HA	1:B:700:GLU:O	2.19	0.42
1:A:65:PHE:CE1	1:A:152:LYS:HG3	2.55	0.42
1:B:21:SER:HA	1:B:48:THR:O	2.20	0.42
1:A:439:CYS:HB2	1:A:512:TYR:CG	2.55	0.41
1:A:632:TYR:C	1:A:632:TYR:CD1	2.97	0.41
1:B:702:ILE:HD12	1:B:710:LYS:HG2	2.03	0.41
1:A:518:THR:HA	1:A:544:TYR:O	2.21	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:21:SER:HA	1:A:48:THR:O	2.21	0.40
1:A:680:HIS:HB3	1:B:356:TYR:CE1	2.56	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	701/724 (97%)	683 (97%)	18 (3%)	0	100	100
1	B	700/724 (97%)	680 (97%)	20 (3%)	0	100	100
All	All	1401/1448 (97%)	1363 (97%)	38 (3%)	0	100	100

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	638/657 (97%)	626 (98%)	12 (2%)	50	31
1	B	637/657 (97%)	620 (97%)	17 (3%)	39	19
All	All	1275/1314 (97%)	1246 (98%)	29 (2%)	44	24

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	114	LYS
1	A	136	SER
1	A	299	LYS
1	A	340	TRP
1	A	427	LYS
1	A	460	LYS
1	A	507	LYS
1	A	543	LEU
1	A	659	LYS
1	A	670	LYS
1	A	683	ARG
1	A	698	LYS
1	B	118	LYS
1	B	139	SER
1	B	186	LYS
1	B	198	LYS
1	B	211	LEU
1	B	225	LEU
1	B	286	GLU
1	B	340	TRP
1	B	358	LYS
1	B	456	LEU
1	B	457	LEU
1	B	458	LYS
1	B	460	LYS
1	B	507	LYS
1	B	543	LEU
1	B	700	GLU
1	B	702	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	607	GLN
1	A	609	ASN
1	B	88	ASN
1	B	119	ASN
1	B	715	ASN

5.3.3 RNA ⓘ

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 oligosaccharides 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

6.1 Protein, DNA and RNA chains

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	705/724 (97%)	-0.25	14 (1%) 65 71	5, 12, 28, 48	0
1	B	704/724 (97%)	-0.07	21 (2%) 52 59	6, 14, 32, 70	0
All	All	1409/1448 (97%)	-0.16	35 (2%) 58 64	5, 13, 30, 70	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	165	TYR	6.4
1	B	699	SER	5.8
1	B	458	LYS	4.3
1	B	655	PHE	3.8
1	B	457	LEU	3.7
1	A	716	LYS	3.5
1	B	691	SER	3.5
1	B	391	PHE	3.5
1	B	388	GLY	3.4
1	B	6	SER	3.2
1	B	387	GLU	3.2
1	B	456	LEU	3.1
1	A	6	SER	2.9
1	B	460	LYS	2.9
1	A	125	GLU	2.8
1	B	702	ILE	2.8
1	B	302	PRO	2.6
1	A	507	LYS	2.6
1	A	191	PRO	2.6
1	B	700	GLU	2.5
1	B	506	LEU	2.4
1	A	669	ASN	2.4
1	A	180	PHE	2.3
1	A	698	LYS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	597	ASP	2.3
1	B	715	ASN	2.2
1	A	670	LYS	2.2
1	B	162	LYS	2.2
1	B	166	GLY	2.2
1	A	459	GLU	2.2
1	B	392	TYR	2.1
1	A	506	LEU	2.1
1	A	460	LYS	2.1
1	B	198	LYS	2.0
1	A	427	LYS	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 oligosaccharides 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.