wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 07:48 AM EDT

PDB ID	:	5EHK
Title	:	Crystal structure of tRNA dependent lantibiotic dehydratase MibB from Microbispora sp. 107891
Authors	:	Hao, Y.; Nair, S.K.
Deposited on	:	2015-10-28
Resolution	:	2.71 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 (i) 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 (1)) were used in the production of this report:

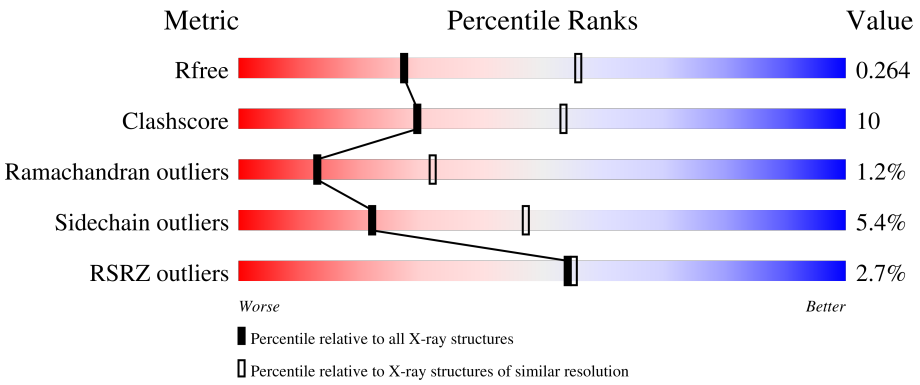
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.71 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 <=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	1115	2% 70% 20% • 7%
1	B	1115	3% 72% 19% • 7%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 16566 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 Lantibiotic dehydratase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1036	8044	5033	1533	1459	19	0	0	0
1	B	1036	8044	5033	1533	1459	19	0	0	0

- Molecule 2 is water.

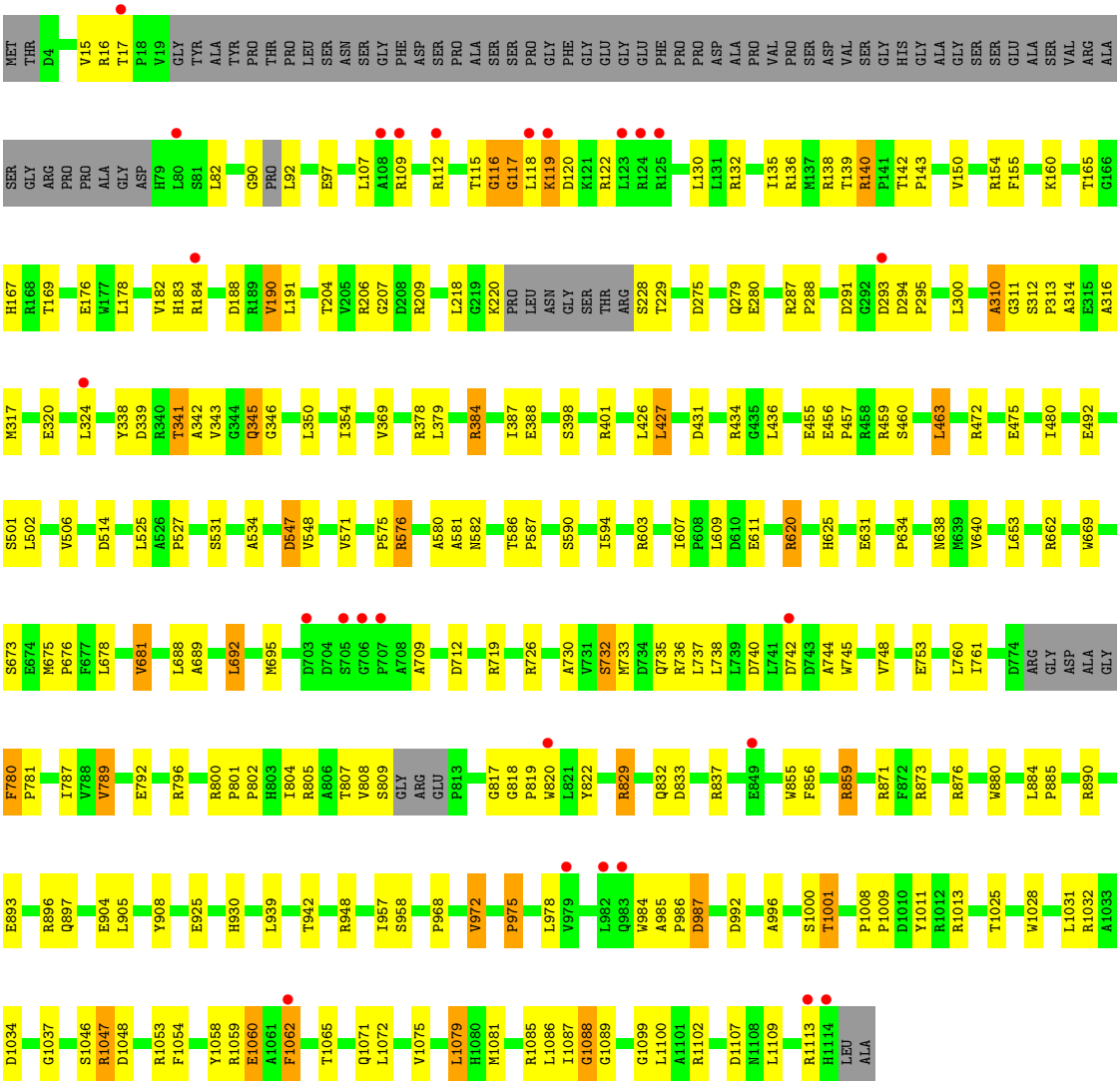
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	230	230	230	0	0
2	B	248	248	248	0	0

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.

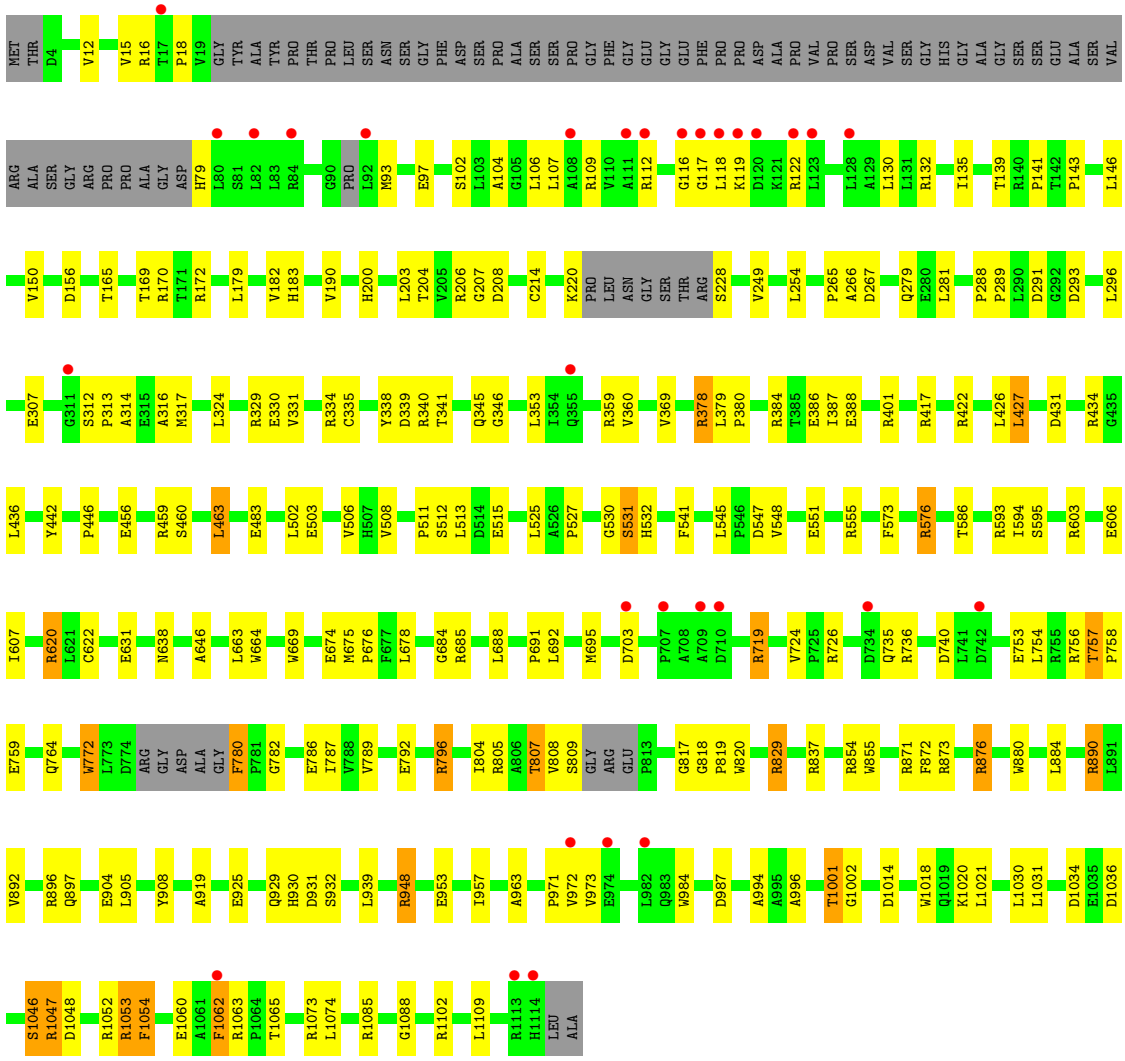
• Molecule 1: Lantibiotic dehydratase

Chain A: 2% 70% 20% • 7%



• Molecule 1: Lantibiotic dehydratase

Chain B: 3% 72% 19% • 7%


MET THR D4 V12 V15 R16 T17 P18 V19 GLY TYR ALA TYR PRO THR PRO LEU SER ASN SER GLY PHE ASP SER PRO ALA SER SER PRO GLY PHE GLY GLU GLY GLU PHE PRO PRO ASP ALA PRO VAL PRO SER ASP VAL SER GLY HIS GLY ALA GLY SER SER GLU ALA SER VAL
ARG ALA SER GLY ARG PRO PRO ALA GLY ASP H79 L80 S81 L82 L83 R84 G90 PRO L92 M93 E97 S102 L103 A104 G105 L106 L107 A108 R109 V110 A111 R112 G116 G117 L118 K119 D120 K121 R122 L123 L128 A129 L130 L131 R132 I135 T139 R140 P141 T142 P143 L146
V150 D156 T165 T169 R170 T171 R172 L179 V182 H183 V190 H200 L203 T204 V205 R206 G207 D208 C214 K220 PRO LEU ASN GLY SER THR ARG S228 V249 L254 P265 A266 D267 Q279 E280 L281 P288 P289 L290 D291 G292 D293 L296
E307 G311 S312 P313 A314 E315 A316 M317 L324 R329 E330 V331 R334 C335 Y338 D339 R340 T341 Q345 G346 L353 I354 Q355 R359 V360 V369 R378 L379 P380 R384 T385 E386 I387 E388 R401 R417 R422 L426 L427 D431 R434 G435
L436 Y442 P446 E456 R459 S460 L463 E483 L502 E503 V506 H507 V508 P511 S512 L513 D514 E515 L525 A526 P527 G530 S531 H532 F541 L545 P546 D547 V548 E551 R555 F573 R576 T586 R593 I594 S595 R603 E606
I607 R620 L621 C622 E631 N638 A646 L663 W664 W669 E674 M675 P676 F677 L678 G684 R685 L688 P691 L692 M695 D703 P707 A708 A709 D710 R719 V724 P725 R726 D734 Q735 R736 D740 L741 D742 E753 L754 R755 R756 T757 P758
E759 Q764 W772 L773 D774 ARG GLY ASP ALA GLY F780 P781 G782 E786 T787 V788 V789 E792 R796 I804 R805 A806 T807 V808 S809 GLY ARG GLU P813 G817 G818 P819 W820 R829 R837 R854 W855 R871 F872 R873 R876 W880 L884 R890 L891
V892 R896 Q897 E904 L905 Y908 A919 E925 Q929 H930 D931 S932 L939 R948 E953 I957 A963 P971 V972 V973 E974 L982 Q983 W984 D987 A994 A995 A996 T1001 G1002 D1014 W1018 Q1019 K1020 L1021 L1030 L1031 D1034 E1035 D1036
S1046 R1047 D1048 R1052 R1053 F1054 E1060 A1061 F1062 R1063 P1064 T1065 R1073 L1074 R1085 G1088 R1102 L1109 R1113 H1114 LEU ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.14Å 151.11Å 205.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.24 – 2.71 45.24 – 2.71	Depositor EDS
% Data completeness (in resolution range)	96.1 (45.24-2.71) 96.1 (45.24-2.71)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.05 (at 2.73Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.204 , 0.264 0.204 , 0.264	Depositor DCC
R_{free} test set	3537 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	42.3	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 41.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	16566	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.30% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.46	0/8223	0.67	3/11182 (0.0%)
1	B	0.47	0/8223	0.67	0/11182
All	All	0.47	0/16446	0.67	3/22364 (0.0%)

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	1
All	All	0	3

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	859	ARG	NE-CZ-NH2	-5.80	117.40	120.30
1	A	859	ARG	NE-CZ-NH1	5.74	123.17	120.30
1	A	1089	GLY	N-CA-C	-5.29	99.88	113.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1088	GLY	Peptide
1	A	780	PHE	Peptide
1	B	780	PHE	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	8044	0	8012	156	1
1	B	8044	0	8012	158	1
2	A	230	0	0	26	0
2	B	248	0	0	31	0
All	All	16566	0	16024	306	1

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 10.

The worst 5 of 306 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:541:PHE:O	2:B:1201:HOH:O	1.80	0.98
1:B:808:VAL:O	2:B:1202:HOH:O	1.84	0.93
1:A:688:LEU:O	2:A:1201:HOH:O	1.86	0.93
1:B:220:LYS:O	2:B:1203:HOH:O	1.85	0.93
1:B:786:GLU:OE2	2:B:1204:HOH:O	1.85	0.93

All (1) 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:1060:GLU:OE2	1:B:1053:ARG:NH2[4_445]	2.18	0.02

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	1024/1115 (92%)	963 (94%)	50 (5%)	11 (1%)	14	34
1	B	1024/1115 (92%)	963 (94%)	48 (5%)	13 (1%)	12	30
All	All	2048/2230 (92%)	1926 (94%)	98 (5%)	24 (1%)	13	32

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	116	GLY
1	A	427	LEU
1	A	1113	ARG
1	B	427	LEU
1	B	684	GLY

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	812/868 (94%)	760 (94%)	52 (6%)	17	39
1	B	812/868 (94%)	777 (96%)	35 (4%)	29	57
All	All	1624/1736 (94%)	1537 (95%)	87 (5%)	22	47

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	378	ARG
1	B	829	ARG
1	B	463	LEU
1	B	620	ARG
1	B	904	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	183	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	832	GLN
1	A	1019	GLN
1	A	1080	HIS

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, 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	1036/1115 (92%)	-0.09	26 (2%)	57	59	6, 32, 68, 108	0
1	B	1036/1115 (92%)	-0.10	30 (2%)	51	52	6, 31, 74, 106	0
All	All	2072/2230 (92%)	-0.09	56 (2%)	54	55	6, 32, 71, 108	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1114	HIS	8.9
1	B	118	LEU	7.9
1	A	707	PRO	5.5
1	B	742	ASP	5.4
1	A	1113	ARG	5.0

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 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.