



Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 04:23 PM GMT

PDB ID : 3GEK
Title : Crystal structure of putative thioesterase yhdA from *Lactococcus lactis*.
Northeast Structural Genomics Consortium Target KR113
Authors : Kuzin, A.P.; Su, M.; Seetharaman, R.; Lee, D.; Foote, E.L.; Ciccocanti, C.;
Janjua, H.; Xiao, R.; Nair, R.; Rost, B.; Acton, T.B.; Everett, J.K.; Monte-
lione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium
(NESG)
Deposited on : 2009-02-25
Resolution : 2.24 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

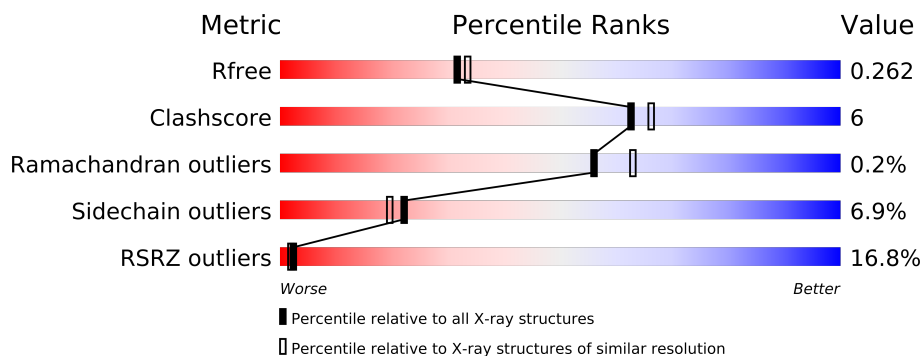
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.24 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}	66092	1112 (2.26-2.22)
Clashscore	79885	1317 (2.26-2.22)
Ramachandran outliers	78287	1282 (2.26-2.22)
Sidechain outliers	78261	1282 (2.26-2.22)
RSRZ outliers	66119	1112 (2.26-2.22)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	146	
1	B	146	
1	C	146	
1	D	146	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3986 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Putative thioesterase yhdA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	132	Total	C	N	O	S	Se	0	2	0
			1031	654	177	195	1	4			
1	B	125	Total	C	N	O	S	Se	0	0	0
			965	613	166	181	1	4			
1	C	127	Total	C	N	O	S	Se	0	0	0
			983	624	169	185	1	4			
1	D	123	Total	C	N	O	S	Se	0	1	0
			958	609	165	179	1	4			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	PHE	SER	ENGINEERED	UNP Q9CHK5
A	44	GLN	HIS	ENGINEERED	UNP Q9CHK5
A	70	GLY	SER	ENGINEERED	UNP Q9CHK5
A	75	PHE	LEU	ENGINEERED	UNP Q9CHK5
A	80	ASN	SER	ENGINEERED	UNP Q9CHK5
A	86	PRO	SER	ENGINEERED	UNP Q9CHK5
A	96	ARG	CYS	ENGINEERED	UNP Q9CHK5
A	139	LEU	-	EXPRESSION TAG	UNP Q9CHK5
A	140	GLU	-	EXPRESSION TAG	UNP Q9CHK5
A	141	HIS	-	EXPRESSION TAG	UNP Q9CHK5
A	142	HIS	-	EXPRESSION TAG	UNP Q9CHK5
A	143	HIS	-	EXPRESSION TAG	UNP Q9CHK5
A	144	HIS	-	EXPRESSION TAG	UNP Q9CHK5
A	145	HIS	-	EXPRESSION TAG	UNP Q9CHK5
A	146	HIS	-	EXPRESSION TAG	UNP Q9CHK5
B	23	PHE	SER	ENGINEERED	UNP Q9CHK5
B	44	GLN	HIS	ENGINEERED	UNP Q9CHK5
B	70	GLY	SER	ENGINEERED	UNP Q9CHK5
B	75	PHE	LEU	ENGINEERED	UNP Q9CHK5
B	80	ASN	SER	ENGINEERED	UNP Q9CHK5
B	86	PRO	SER	ENGINEERED	UNP Q9CHK5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	96	ARG	CYS	ENGINEERED	UNP Q9CHK5
B	139	LEU	-	EXPRESSION TAG	UNP Q9CHK5
B	140	GLU	-	EXPRESSION TAG	UNP Q9CHK5
B	141	HIS	-	EXPRESSION TAG	UNP Q9CHK5
B	142	HIS	-	EXPRESSION TAG	UNP Q9CHK5
B	143	HIS	-	EXPRESSION TAG	UNP Q9CHK5
B	144	HIS	-	EXPRESSION TAG	UNP Q9CHK5
B	145	HIS	-	EXPRESSION TAG	UNP Q9CHK5
B	146	HIS	-	EXPRESSION TAG	UNP Q9CHK5
C	23	PHE	SER	ENGINEERED	UNP Q9CHK5
C	44	GLN	HIS	ENGINEERED	UNP Q9CHK5
C	70	GLY	SER	ENGINEERED	UNP Q9CHK5
C	75	PHE	LEU	ENGINEERED	UNP Q9CHK5
C	80	ASN	SER	ENGINEERED	UNP Q9CHK5
C	86	PRO	SER	ENGINEERED	UNP Q9CHK5
C	96	ARG	CYS	ENGINEERED	UNP Q9CHK5
C	139	LEU	-	EXPRESSION TAG	UNP Q9CHK5
C	140	GLU	-	EXPRESSION TAG	UNP Q9CHK5
C	141	HIS	-	EXPRESSION TAG	UNP Q9CHK5
C	142	HIS	-	EXPRESSION TAG	UNP Q9CHK5
C	143	HIS	-	EXPRESSION TAG	UNP Q9CHK5
C	144	HIS	-	EXPRESSION TAG	UNP Q9CHK5
C	145	HIS	-	EXPRESSION TAG	UNP Q9CHK5
C	146	HIS	-	EXPRESSION TAG	UNP Q9CHK5
D	23	PHE	SER	ENGINEERED	UNP Q9CHK5
D	44	GLN	HIS	ENGINEERED	UNP Q9CHK5
D	70	GLY	SER	ENGINEERED	UNP Q9CHK5
D	75	PHE	LEU	ENGINEERED	UNP Q9CHK5
D	80	ASN	SER	ENGINEERED	UNP Q9CHK5
D	86	PRO	SER	ENGINEERED	UNP Q9CHK5
D	96	ARG	CYS	ENGINEERED	UNP Q9CHK5
D	139	LEU	-	EXPRESSION TAG	UNP Q9CHK5
D	140	GLU	-	EXPRESSION TAG	UNP Q9CHK5
D	141	HIS	-	EXPRESSION TAG	UNP Q9CHK5
D	142	HIS	-	EXPRESSION TAG	UNP Q9CHK5
D	143	HIS	-	EXPRESSION TAG	UNP Q9CHK5
D	144	HIS	-	EXPRESSION TAG	UNP Q9CHK5
D	145	HIS	-	EXPRESSION TAG	UNP Q9CHK5
D	146	HIS	-	EXPRESSION TAG	UNP Q9CHK5

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	17	Total 17	O 17	0	0
2	B	18	Total 18	O 18	0	0
2	C	9	Total 9	O 9	0	0
2	D	5	Total 5	O 5	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	59.95Å 64.06Å 144.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.18 – 2.24 46.16 – 2.24	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.18-2.24) 99.6 (46.16-2.24)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.95 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.224 , 0.255 0.249 , 0.262	Depositor DCC
R_{free} test set	1385 reflections (5.03%)	DCC
Wilson B-factor (Å ²)	46.3	Xtriage
Anisotropy	0.043	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 63.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 27547 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3986	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	0.49	1/1052 (0.1%)	0.54	0/1413
1	B	0.40	0/978	0.56	0/1313
1	C	0.57	1/996 (0.1%)	0.54	0/1336
1	D	0.50	1/970 (0.1%)	0.53	0/1300
All	All	0.49	3/3996 (0.1%)	0.54	0/5362

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	69	SER	CB-OG	13.12	1.59	1.42
1	D	69	SER	CB-OG	-6.70	1.33	1.42
1	A	23	PHE	C-N	6.01	1.47	1.34

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1031	0	1031	21	0
1	B	965	0	963	12	0
1	C	983	0	982	9	0
1	D	958	0	960	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	17	0	0	0	0
2	B	18	0	0	0	0
2	C	9	0	0	0	0
2	D	5	0	0	0	0
All	All	3986	0	3936	44	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

All (44) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:48:ASN:HD21	1:D:58:ILE:HD11	1.51	0.76
1:C:54:ALA:O	1:C:58:ILE:HG12	1.86	0.75
1:A:105[B]:ARG:HE	1:A:105[B]:ARG:H	1.34	0.73
1:B:1:MSE:HE2	1:B:6:GLN:CG	2.24	0.68
1:A:9:ILE:HD12	1:A:55:LEU:HD13	1.75	0.68
1:A:53:LEU:HD21	1:D:79:ILE:HG21	1.76	0.67
1:A:94:ASN:HB2	1:A:114:THR:HG22	1.78	0.65
1:A:42:GLN:HE21	1:A:46:PHE:H	1.43	0.64
1:B:48:ASN:HD21	1:C:58:ILE:HD11	1.63	0.64
1:C:37:SER:H	1:C:40:HIS:CD2	2.16	0.63
1:C:105:ARG:HG3	1:C:106:ASN:HD22	1.66	0.61
1:A:103:GLY:HA3	1:A:105[B]:ARG:HH21	1.67	0.60
1:B:1:MSE:HE2	1:B:6:GLN:HG2	1.85	0.59
1:D:37:SER:H	1:D:40:HIS:HD2	1.51	0.58
1:C:37:SER:H	1:C:40:HIS:HD2	1.50	0.58
1:D:37:SER:H	1:D:40:HIS:CD2	2.22	0.57
1:A:37:SER:H	1:A:40:HIS:CD2	2.22	0.57
1:B:1:MSE:HE2	1:B:6:GLN:HG3	1.85	0.57
1:B:48:ASN:HD21	1:C:58:ILE:CD1	2.20	0.55
1:A:37:SER:H	1:A:40:HIS:HD2	1.57	0.53
1:A:7:LEU:HD13	1:A:39:PHE:HE2	1.75	0.52
1:B:7:LEU:HD13	1:B:39:PHE:HE2	1.75	0.51
1:D:42:GLN:HB2	1:D:43:PRO:CD	2.42	0.50
1:A:53:LEU:HD23	1:D:53:LEU:HB3	1.95	0.49
1:A:42:GLN:NE2	1:A:46:PHE:H	2.09	0.49
1:C:42:GLN:HB2	1:C:43:PRO:CD	2.43	0.49
1:A:58:ILE:O	1:A:62:MSE:HG3	2.13	0.48
1:B:7:LEU:O	1:B:40:HIS:HE1	1.97	0.48
1:D:54:ALA:O	1:D:58:ILE:HG12	2.14	0.47
1:B:60:ALA:HB2	1:B:111:ILE:HD11	1.97	0.47
1:B:123:GLN:C	1:B:124:ILE:HG13	2.34	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:42:GLN:HB2	1:A:43:PRO:HD2	1.97	0.46
1:B:42:GLN:HB2	1:B:43:PRO:CD	2.47	0.45
1:A:42:GLN:HB2	1:A:43:PRO:CD	2.47	0.45
1:A:53:LEU:HB3	1:D:53:LEU:HD23	1.97	0.45
1:A:100:LEU:HD11	1:A:110:GLU:HB2	2.00	0.44
1:A:79:ILE:HG21	1:D:53:LEU:HD21	2.00	0.44
1:B:1:MSE:CE	1:B:6:GLN:HG2	2.48	0.43
1:A:7:LEU:O	1:A:40:HIS:HE1	2.01	0.43
1:C:68:GLY:HA3	1:C:72:TYR:CZ	2.54	0.42
1:A:78:SER:HB2	1:A:127:VAL:HB	2.01	0.42
1:B:84:LEU:HG	1:D:101[A]:LYS:HB2	2.02	0.42
1:A:30:PHE:CE2	1:A:97:GLY:HA3	2.56	0.41
1:C:63:ALA:O	1:C:67:ILE:HG12	2.21	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	132/146 (90%)	128 (97%)	2 (2%)	2 (2%)	15	9
1	B	121/146 (83%)	119 (98%)	2 (2%)	0	100	100
1	C	123/146 (84%)	119 (97%)	4 (3%)	0	100	100
1	D	120/146 (82%)	116 (97%)	4 (3%)	0	100	100
All	All	496/584 (85%)	482 (97%)	12 (2%)	2 (0%)	56	46

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	69[A]	SER
1	A	69[B]	SER

5.3.2 Protein sidechains ⓘ

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	113/121 (93%)	102 (90%)	11 (10%)	12	8
1	B	104/121 (86%)	96 (92%)	8 (8%)	18	15
1	C	106/121 (88%)	101 (95%)	5 (5%)	36	40
1	D	103/121 (85%)	98 (95%)	5 (5%)	35	37
All	All	426/484 (88%)	397 (93%)	29 (7%)	22	20

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

Mol	Chain	Res	Type
1	A	7	LEU
1	A	21	ASP
1	A	24	VAL
1	A	31	SER
1	A	53	LEU
1	A	78	SER
1	A	79	ILE
1	A	84	LEU
1	A	104	LYS
1	A	114	THR
1	A	130	LEU
1	B	7	LEU
1	B	38	ASP
1	B	42	GLN
1	B	84	LEU
1	B	99	LEU
1	B	105	ARG
1	B	130	LEU
1	B	131	VAL
1	C	7	LEU
1	C	35	ILE
1	C	42	GLN
1	C	84	LEU
1	C	99	LEU
1	D	7	LEU
1	D	8	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	42	GLN
1	D	53	LEU
1	D	84	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	6	GLN
1	A	40	HIS
1	A	42	GLN
1	A	85	ASN
1	A	106	ASN
1	B	40	HIS
1	C	40	HIS
1	C	106	ASN
1	D	6	GLN
1	D	8	ASN
1	D	40	HIS
1	D	77	GLN
1	D	94	ASN
1	D	106	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

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	132/146 (90%)	0.60	12 (9%) 9 11	39, 49, 63, 65	0
1	B	125/146 (85%)	0.88	23 (18%) 2 1	44, 50, 56, 62	0
1	C	127/146 (86%)	1.01	28 (22%) 1 1	45, 51, 60, 66	0
1	D	123/146 (84%)	0.94	22 (17%) 2 1	44, 54, 62, 67	0
All	All	507/584 (86%)	0.85	85 (16%) 2 2	39, 51, 61, 67	0

All (85) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	73	PHE	6.5
1	D	73	PHE	5.7
1	C	35	ILE	5.6
1	A	22	LYS	5.2
1	C	73	PHE	4.7
1	A	21	ASP	4.5
1	B	17	ASP	4.3
1	C	18	GLU	4.2
1	A	75	PHE	4.2
1	C	81	ALA	4.2
1	A	71	GLN	4.0
1	D	15	PHE	4.0
1	C	71	GLN	3.9
1	C	75	PHE	3.9
1	A	70	GLY	3.9
1	D	71	GLN	3.6
1	C	68	GLY	3.5
1	B	81	ALA	3.4
1	C	28	TYR	3.3
1	D	92	PHE	3.3
1	B	132	PRO	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	15	PHE	3.2
1	D	35	ILE	3.1
1	D	1	MSE	3.1
1	D	81	ALA	3.1
1	C	134	LYS	3.1
1	C	67	ILE	3.1
1	D	70	GLY	3.1
1	C	124	ILE	3.0
1	B	80	ASN	3.0
1	D	17	ASP	3.0
1	D	80	ASN	3.0
1	C	104	LYS	2.9
1	B	78	SER	2.9
1	B	79	ILE	2.9
1	C	105	ARG	2.9
1	B	76	GLY	2.9
1	D	105	ARG	2.9
1	A	19	ASN	2.8
1	C	125	THR	2.8
1	B	127	VAL	2.8
1	C	26	LYS	2.8
1	D	72	TYR	2.8
1	B	126	VAL	2.8
1	A	132	PRO	2.8
1	A	18	GLU	2.8
1	C	80	ASN	2.7
1	B	133	GLN	2.7
1	C	79	ILE	2.6
1	D	104	LYS	2.6
1	D	13	GLN	2.6
1	B	16	THR	2.6
1	C	133	GLN	2.6
1	B	131	VAL	2.6
1	B	43	PRO	2.6
1	D	33	LYS	2.5
1	D	28	TYR	2.5
1	B	38	ASP	2.5
1	C	27	ILE	2.5
1	C	131	VAL	2.5
1	C	10	THR	2.5
1	D	16	THR	2.5
1	B	15	PHE	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	59	THR	2.4
1	A	81	ALA	2.4
1	D	75	PHE	2.4
1	C	50	GLY	2.3
1	D	11	ASP	2.3
1	D	124	ILE	2.3
1	C	78	SER	2.3
1	B	108	VAL	2.2
1	C	82	ASN	2.2
1	B	124	ILE	2.2
1	B	53	LEU	2.2
1	C	49	GLY	2.1
1	B	125	THR	2.1
1	B	55	LEU	2.1
1	C	17	ASP	2.1
1	B	105	ARG	2.1
1	C	83	HIS	2.1
1	B	82	ASN	2.1
1	D	38	ASP	2.1
1	A	72	TYR	2.0
1	A	124	ILE	2.0
1	D	78	SER	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.