wwPDB X-ray Structure Validation Summary Report (i)

Oct 31, 2023 – 11:08 AM JST

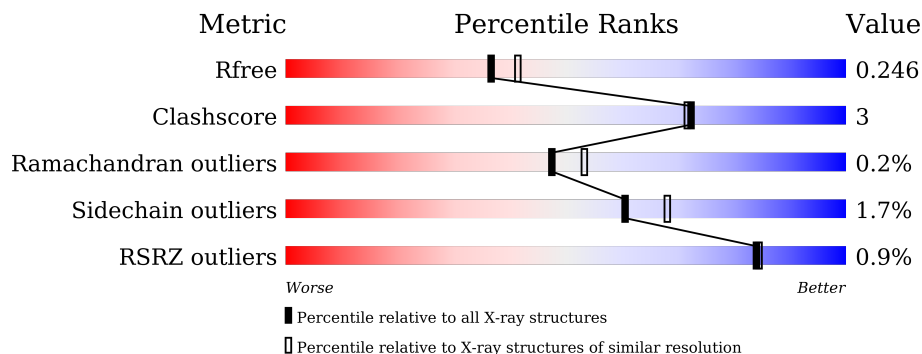|		
PDB ID	:	5CF1
Title	:	Crystal Structure of the M32V/M78V/I80V/L114F mutant of LEH
Authors	:	Wu, L.; Sun, Z.T.; Reetz, M.T.; Zhou, J.H.
Deposited on	:	2015-07-08
Resolution	:	2.24 Å(reported)

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 2.24 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	Percentile Ranks	Value
Rfree		0.246
Clashscore		3
Ramachandran outliers		0.2%
Sidechain outliers		1.7%
RSRZ outliers		0.9%

Worse — *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (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 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	155	81% 8% · 10%
1	B	155	85% 5% 10%
1	C	155	.% 83% 8% · 7%
1	D	155	85% 6% 9%
1	E	155	82% 7% · 10%
1	F	155	.% 76% 12% 12%

Continued on next page...

WORLDWIDE wwPDB PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	155	% 80% 7% · 12%
1	H	155	4% 78% 10% · 12%

2　Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 9182 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 Limonene-1,2-epoxide hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	139	Total 1094	C 699	N 174	O 219	S 2	0	2	0
1	B	139	Total 1091	C 697	N 174	O 218	S 2	0	1	0
1	E	139	Total 1092	C 698	N 174	O 218	S 2	0	1	0
1	F	137	Total 1076	C 687	N 172	O 215	S 2	0	1	0
1	C	144	Total 1133	C 724	N 182	O 225	S 2	0	1	0
1	D	141	Total 1102	C 704	N 176	O 220	S 2	0	0	0
1	G	137	Total 1068	C 684	N 172	O 210	S 2	0	0	0
1	H	137	Total 1074	C 686	N 172	O 214	S 2	0	0	0

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	MET	-	expression tag	UNP Q9ZAG3
A	-4	HIS	-	expression tag	UNP Q9ZAG3
A	-3	HIS	-	expression tag	UNP Q9ZAG3
A	-2	HIS	-	expression tag	UNP Q9ZAG3
A	-1	HIS	-	expression tag	UNP Q9ZAG3
A	0	HIS	-	expression tag	UNP Q9ZAG3
A	1	HIS	-	expression tag	UNP Q9ZAG3
A	32	VAL	MET	engineered mutation	UNP Q9ZAG3
A	78	VAL	MET	engineered mutation	UNP Q9ZAG3
A	80	VAL	ILE	engineered mutation	UNP Q9ZAG3
A	114	PHE	LEU	engineered mutation	UNP Q9ZAG3
B	-5	MET	-	expression tag	UNP Q9ZAG3
B	-4	HIS	-	expression tag	UNP Q9ZAG3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	HIS	-	expression tag	UNP Q9ZAG3
B	-2	HIS	-	expression tag	UNP Q9ZAG3
B	-1	HIS	-	expression tag	UNP Q9ZAG3
B	0	HIS	-	expression tag	UNP Q9ZAG3
B	1	HIS	-	expression tag	UNP Q9ZAG3
B	32	VAL	MET	engineered mutation	UNP Q9ZAG3
B	78	VAL	MET	engineered mutation	UNP Q9ZAG3
B	80	VAL	ILE	engineered mutation	UNP Q9ZAG3
B	114	PHE	LEU	engineered mutation	UNP Q9ZAG3
E	-5	MET	-	expression tag	UNP Q9ZAG3
E	-4	HIS	-	expression tag	UNP Q9ZAG3
E	-3	HIS	-	expression tag	UNP Q9ZAG3
E	-2	HIS	-	expression tag	UNP Q9ZAG3
E	-1	HIS	-	expression tag	UNP Q9ZAG3
E	0	HIS	-	expression tag	UNP Q9ZAG3
E	1	HIS	-	expression tag	UNP Q9ZAG3
E	32	VAL	MET	engineered mutation	UNP Q9ZAG3
E	78	VAL	MET	engineered mutation	UNP Q9ZAG3
E	80	VAL	ILE	engineered mutation	UNP Q9ZAG3
E	114	PHE	LEU	engineered mutation	UNP Q9ZAG3
F	-5	MET	-	expression tag	UNP Q9ZAG3
F	-4	HIS	-	expression tag	UNP Q9ZAG3
F	-3	HIS	-	expression tag	UNP Q9ZAG3
F	-2	HIS	-	expression tag	UNP Q9ZAG3
F	-1	HIS	-	expression tag	UNP Q9ZAG3
F	0	HIS	-	expression tag	UNP Q9ZAG3
F	1	HIS	-	expression tag	UNP Q9ZAG3
F	32	VAL	MET	engineered mutation	UNP Q9ZAG3
F	78	VAL	MET	engineered mutation	UNP Q9ZAG3
F	80	VAL	ILE	engineered mutation	UNP Q9ZAG3
F	114	PHE	LEU	engineered mutation	UNP Q9ZAG3
C	-5	MET	-	expression tag	UNP Q9ZAG3
C	-4	HIS	-	expression tag	UNP Q9ZAG3
C	-3	HIS	-	expression tag	UNP Q9ZAG3
C	-2	HIS	-	expression tag	UNP Q9ZAG3
C	-1	HIS	-	expression tag	UNP Q9ZAG3
C	0	HIS	-	expression tag	UNP Q9ZAG3
C	1	HIS	-	expression tag	UNP Q9ZAG3
C	32	VAL	MET	engineered mutation	UNP Q9ZAG3
C	78	VAL	MET	engineered mutation	UNP Q9ZAG3
C	80	VAL	ILE	engineered mutation	UNP Q9ZAG3
C	114	PHE	LEU	engineered mutation	UNP Q9ZAG3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-5	MET	-	expression tag	UNP Q9ZAG3
D	-4	HIS	-	expression tag	UNP Q9ZAG3
D	-3	HIS	-	expression tag	UNP Q9ZAG3
D	-2	HIS	-	expression tag	UNP Q9ZAG3
D	-1	HIS	-	expression tag	UNP Q9ZAG3
D	0	HIS	-	expression tag	UNP Q9ZAG3
D	1	HIS	-	expression tag	UNP Q9ZAG3
D	32	VAL	MET	engineered mutation	UNP Q9ZAG3
D	78	VAL	MET	engineered mutation	UNP Q9ZAG3
D	80	VAL	ILE	engineered mutation	UNP Q9ZAG3
D	114	PHE	LEU	engineered mutation	UNP Q9ZAG3
G	-5	MET	-	expression tag	UNP Q9ZAG3
G	-4	HIS	-	expression tag	UNP Q9ZAG3
G	-3	HIS	-	expression tag	UNP Q9ZAG3
G	-2	HIS	-	expression tag	UNP Q9ZAG3
G	-1	HIS	-	expression tag	UNP Q9ZAG3
G	0	HIS	-	expression tag	UNP Q9ZAG3
G	1	HIS	-	expression tag	UNP Q9ZAG3
G	32	VAL	MET	engineered mutation	UNP Q9ZAG3
G	78	VAL	MET	engineered mutation	UNP Q9ZAG3
G	80	VAL	ILE	engineered mutation	UNP Q9ZAG3
G	114	PHE	LEU	engineered mutation	UNP Q9ZAG3
H	-5	MET	-	expression tag	UNP Q9ZAG3
H	-4	HIS	-	expression tag	UNP Q9ZAG3
H	-3	HIS	-	expression tag	UNP Q9ZAG3
H	-2	HIS	-	expression tag	UNP Q9ZAG3
H	-1	HIS	-	expression tag	UNP Q9ZAG3
H	0	HIS	-	expression tag	UNP Q9ZAG3
H	1	HIS	-	expression tag	UNP Q9ZAG3
H	32	VAL	MET	engineered mutation	UNP Q9ZAG3
H	78	VAL	MET	engineered mutation	UNP Q9ZAG3
H	80	VAL	ILE	engineered mutation	UNP Q9ZAG3
H	114	PHE	LEU	engineered mutation	UNP Q9ZAG3

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	82	Total 82	O 82	0	0
2	B	62	Total 62	O 62	0	0
2	E	49	Total 49	O 49	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	F	42	Total 42	O 42	0	0
2	C	80	Total 80	O 80	0	0
2	D	58	Total 58	O 58	0	0
2	G	37	Total 37	O 37	0	0
2	H	42	Total 42	O 42	0	0

3 Residue-property plots ⓘ

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: Limonene-1,2-epoxide hydrolase
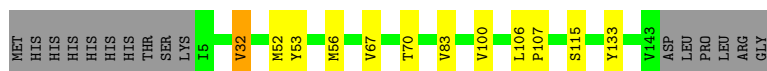
Chain A:

- Molecule 1: Limonene-1,2-epoxide hydrolase
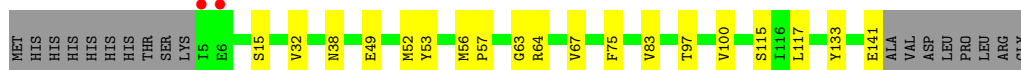
Chain B:

- Molecule 1: Limonene-1,2-epoxide hydrolase

Chain E:

- Molecule 1: Limonene-1,2-epoxide hydrolase

Chain F:

- Molecule 1: Limonene-1,2-epoxide hydrolase

Chain C:

- Molecule 1: Limonene-1,2-epoxide hydrolase

Chain D:

85%   6%   9%

MET HIS HIS HIS HIS HIS THR SER LYS I5 A34 K42 L43 I44 R64 E68 V78 V100 R137 L145 PRO LEU ARG GLY

• Molecule 1: Limonene-1,2-epoxide hydrolase

%

Chain G:

80%   7%   12%

MET HIS HIS HIS HIS HIS THR SER LYS I5 S15 P23 N38 E49 M52 Y53 M56 R64 Q69 F75 Y133 R137 E138 F139 GLU GLU A142 V143 ASP LEU PRO LEU ARG GLY

• Molecule 1: Limonene-1,2-epoxide hydrolase

4%

Chain H:

78%   10%   12%

MET HIS HIS HIS HIS HIS THR SER LYS I5 E6 A17 A34 K42 L43 I44 E45 M56 P59 P60 Q69 T76 I88 L106 P107 T108 G109 K110 S111 Y112 Y133 R137 E141 ALA VAL ASP LEU PRO LEU ARG GLY
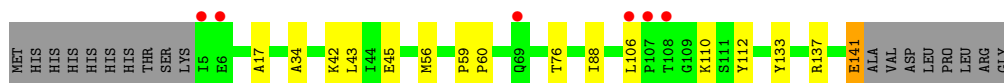
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	103.78Å 60.99Å 119.29Å 90.00° 90.70° 90.00°	Depositor
Resolution (Å)	23.91 – 2.24 23.91 – 2.24	Depositor EDS
% Data completeness (in resolution range)	97.1 (23.91-2.24) 97.2 (23.91-2.24)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.35 (at 2.24Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.188 , 0.245 0.192 , 0.246	Depositor DCC
R_{free} test set	3447 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	34.8	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 25.4	EDS
L-test for twinning[2]	$<|L|> = 0.50, <L^2> = 0.33$	Xtriage
Estimated twinning fraction	0.179 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	9182	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 59.93 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.6412e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 ⓘ

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	0/1124	0.59	0/1530
1	B	0.47	0/1118	0.56	0/1522
1	C	0.46	0/1161	0.60	0/1581
1	D	0.46	0/1126	0.58	0/1533
1	E	0.39	0/1119	0.50	0/1523
1	F	0.40	0/1103	0.51	0/1500
1	G	0.38	0/1091	0.52	0/1484
1	H	0.40	0/1098	0.51	0/1494
All	All	0.43	0/8940	0.55	0/12167

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1094	0	1055	9	0
1	B	1091	0	1050	4	0
1	C	1133	0	1096	8	0
1	D	1102	0	1061	6	0
1	E	1092	0	1052	8	0
1	F	1076	0	1033	10	0
1	G	1068	0	1033	5	0
1	H	1074	0	1032	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	82	0	0	2	0
2	B	62	0	0	1	0
2	C	80	0	0	0	0
2	D	58	0	0	1	0
2	E	49	0	0	1	0
2	F	42	0	0	1	0
2	G	37	0	0	0	0
2	H	42	0	0	0	0
All	All	9182	0	8412	58	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 3.

The worst 5 of 58 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:F:141:GLU:O	2:F:201:HOH:O	2.03	0.76
1:E:52:MET:SD	2:E:243:HOH:O	2.43	0.75
1:A:5:ILE:HG23	1:A:6:GLU:H	1.60	0.66
1:F:32:VAL:HG13	1:F:83:VAL:HG11	1.80	0.64
1:G:49:GLU:HA	1:G:64:ARG:HB2	1.83	0.60

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	139/155 (90%)	137 (99%)	2 (1%)	0	100	100
1	B	138/155 (89%)	136 (99%)	2 (1%)	0	100	100
1	C	143/155 (92%)	140 (98%)	1 (1%)	2 (1%)	11	6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	139/155 (90%)	137 (99%)	2 (1%)	0	100	100
1	E	138/155 (89%)	136 (99%)	2 (1%)	0	100	100
1	F	136/155 (88%)	135 (99%)	1 (1%)	0	100	100
1	G	133/155 (86%)	130 (98%)	3 (2%)	0	100	100
1	H	135/155 (87%)	133 (98%)	2 (2%)	0	100	100
All	All	1101/1240 (89%)	1084 (98%)	15 (1%)	2 (0%)	47	53

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	146	PRO
1	C	147	LEU

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	117/130 (90%)	114 (97%)	3 (3%)	46	52
1	B	116/130 (89%)	115 (99%)	1 (1%)	78	84
1	C	121/130 (93%)	119 (98%)	2 (2%)	60	68
1	D	117/130 (90%)	116 (99%)	1 (1%)	78	84
1	E	116/130 (89%)	115 (99%)	1 (1%)	78	84
1	F	115/130 (88%)	114 (99%)	1 (1%)	78	84
1	G	113/130 (87%)	107 (95%)	6 (5%)	22	21
1	H	114/130 (88%)	113 (99%)	1 (1%)	78	84
All	All	929/1040 (89%)	913 (98%)	16 (2%)	60	68

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	143	VAL

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type
1	G	69	GLN
1	D	78	VAL
1	G	52	MET
1	C	78	VAL

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

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 monosaccharides 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(Å²)	Q<0.9
1	A	139/155 (89%)	-0.46	0	100	100	22, 30, 48, 70	0
1	B	139/155 (89%)	-0.49	0	100	100	21, 31, 45, 65	0
1	C	144/155 (92%)	-0.45	1 (0%)	87	87	22, 32, 51, 74	0
1	D	141/155 (90%)	-0.47	0	100	100	23, 31, 48, 67	0
1	E	139/155 (89%)	-0.28	0	100	100	30, 44, 66, 74	0
1	F	137/155 (88%)	-0.26	2 (1%)	73	74	27, 47, 67, 86	0
1	G	137/155 (88%)	-0.32	1 (0%)	87	87	32, 46, 66, 86	0
1	H	137/155 (88%)	-0.13	6 (4%)	34	34	32, 49, 74, 84	0
All	All	1113/1240 (89%)	-0.36	10 (0%)	84	84	21, 37, 64, 86	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	107	PRO	4.2
1	F	5	ILE	3.9
1	H	6	GLU	2.5
1	C	6	GLU	2.2
1	H	106	LEU	2.1

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 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.