



Full wwPDB X-ray Structure Validation Report

Apr 11, 2014 – 12:43 PM EDT

PDB ID : 4L1W
Title : Crystal Structuer of Human 3-alpha Hydroxysteroid Dehydrogenase Type 3 in Complex with NADP+ and Progesterone
Authors : Zhang, B.; Hu, X.-J.; Lin, S.-X.
Deposited on : 2013-06-03
Resolution : 2.20 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>

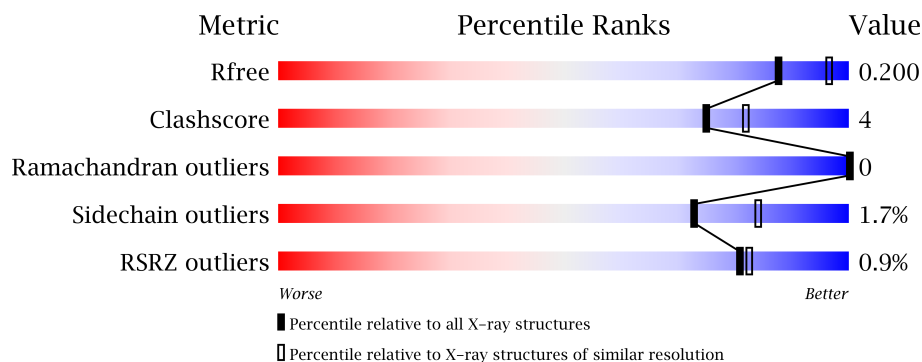
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : **FAILED**
Xtriage (Phenix) : dev-1439
EDS : stable22978
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) : stable22978

1 Overall quality at a glance

The reported resolution of this entry is 2.20 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	2938 (2.20-2.20)
Clashscore	79885	3751 (2.20-2.20)
Ramachandran outliers	78287	3681 (2.20-2.20)
Sidechain outliers	78261	3682 (2.20-2.20)
RSRZ outliers	66119	2939 (2.20-2.20)

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	325	
1	B	325	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	STR	A	402	-	X
3	STR	B	402	-	X
4	SO4	B	403	-	X

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 5741 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.

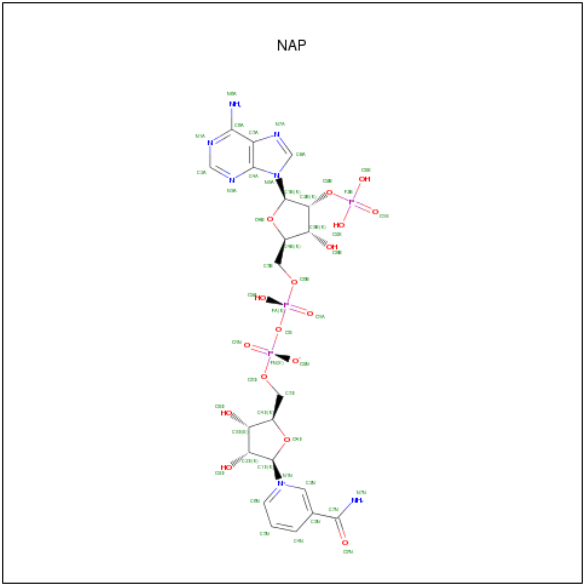
- Molecule 1 is a protein called Aldo-keto reductase family 1 member C2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	325	Total	C	N	O	S	0	1	0
			2613	1671	451	480	11			
1	B	325	Total	C	N	O	S	0	1	0
			2607	1670	447	479	11			

There are 6 discrepancies between the modelled and reference sequences:

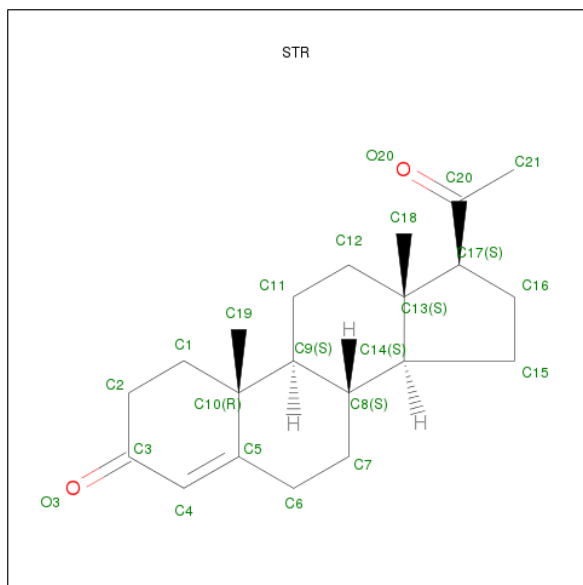
Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	EXPRESSION TAG	UNP P52895
A	0	VAL	-	EXPRESSION TAG	UNP P52895
A	1	ASP	-	EXPRESSION TAG	UNP P52895
B	-1	SER	-	EXPRESSION TAG	UNP P52895
B	0	VAL	-	EXPRESSION TAG	UNP P52895
B	1	ASP	-	EXPRESSION TAG	UNP P52895

- Molecule 2 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDEPHOSPHATE (three-letter code: NAP) (formula: C₂₁H₂₈N₇O₁₇P₃).



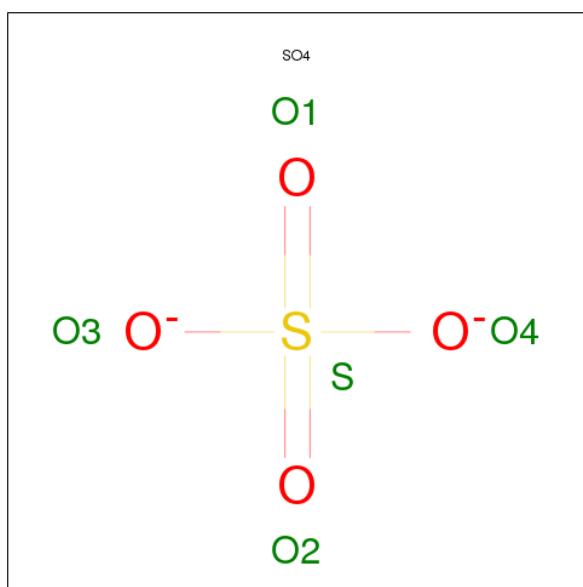
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	
			48	21	7	17	3	
2	B	1	Total	C	N	O	P	
			48	21	7	17	3	

- Molecule 3 is PROGESTERONE (three-letter code: STR) (formula: $C_{21}H_{30}O_2$).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	C O		
			23 21 2		0	0
3	B	1	Total	C O		
			23 21 2		0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is water.

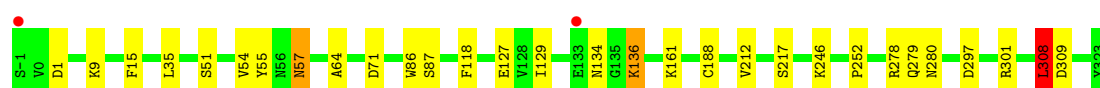
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	194	Total	O	0	0
			194	194		
5	B	175	Total	O	0	0
			175	175		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

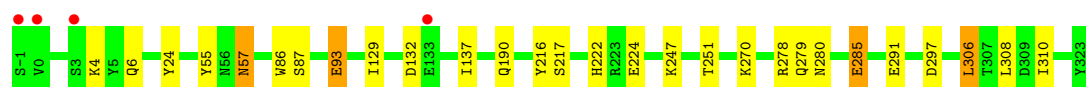
- Molecule 1: Aldo-keto reductase family 1 member C2

Chain A: 



- Molecule 1: Aldo-keto reductase family 1 member C2

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	52.80Å 86.82Å 76.49Å 90.00° 105.96° 90.00°	Depositor
Resolution (Å)	21.35 – 2.20 21.35 – 2.20	Depositor EDS
% Data completeness (in resolution range)	96.4 (21.35-2.20) 96.5 (21.35-2.20)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	17.02 (at 2.19Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.146 , 0.194 0.157 , 0.200	Depositor DCC
R_{free} test set	1640 reflections (5.30%)	DCC
Wilson B-factor (Å ²)	10.1	Xtriage
Anisotropy	0.189	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 26.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 32558 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5741	wwPDB-VP
Average B, all atoms (Å ²)	11.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.08% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NAP, SO4, STR

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	0.96	1/2674 (0.0%)	0.92	6/3621 (0.2%)
1	B	0.92	3/2671 (0.1%)	0.90	10/3618 (0.3%)
All	All	0.94	4/5345 (0.1%)	0.91	16/7239 (0.2%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	285	GLU	CG-CD	-6.59	1.42	1.51
1	B	55	TYR	CZ-OH	-5.35	1.28	1.37
1	A	252	PRO	N-CD	5.11	1.55	1.47
1	B	291	GLU	CD-OE2	5.06	1.31	1.25

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	278	ARG	NE-CZ-NH1	10.15	125.38	120.30
1	B	285	GLU	OE1-CD-OE2	10.02	135.32	123.30
1	A	278	ARG	NE-CZ-NH2	-7.54	116.53	120.30
1	A	301	ARG	NE-CZ-NH2	-7.45	116.58	120.30
1	B	306[A]	LEU	CB-CG-CD1	6.48	122.02	111.00
1	B	306[B]	LEU	CB-CG-CD1	6.48	122.02	111.00
1	A	297	ASP	CB-CG-OD1	6.30	123.97	118.30
1	B	297	ASP	CB-CG-OD1	5.96	123.67	118.30
1	B	278	ARG	NE-CZ-NH1	5.69	123.14	120.30
1	A	71	ASP	CB-CG-OD1	5.69	123.42	118.30
1	A	308	LEU	CA-CB-CG	5.66	128.32	115.30
1	B	93	GLU	CA-CB-CG	5.29	125.03	113.40
1	B	297	ASP	CB-CG-OD2	-5.23	113.59	118.30
1	B	251	THR	C-N-CD	5.17	139.25	128.40
1	B	132	ASP	CB-CG-OD1	5.16	122.95	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	93	GLU	CB-CA-C	-5.12	100.15	110.40

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	2613	0	2620	18	0
1	B	2607	0	2619	18	0
2	A	48	0	25	5	0
2	B	48	0	25	11	0
3	A	23	0	30	4	0
3	B	23	0	30	6	0
4	A	5	0	0	0	0
4	B	5	0	0	0	0
5	A	194	0	0	3	0
5	B	175	0	0	3	0
All	All	5741	0	5349	43	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 4.

All (43) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:280:ASN:HD21	2:B:401:NAP:H61A	1.35	0.73
1:B:129:ILE:HD12	3:B:402:STR:H12	1.73	0.71
1:A:309:ASP:OD2	5:A:660:HOH:O	2.10	0.69
2:B:401:NAP:C4N	5:B:572:HOH:O	2.42	0.67
1:A:129:ILE:HD12	3:A:402:STR:H61	1.78	0.66
1:B:190:GLN:OE1	2:B:401:NAP:H2N	2.03	0.58
2:B:401:NAP:C3N	5:B:572:HOH:O	2.51	0.58
1:B:137:ILE:HD11	1:B:310:ILE:HD12	1.88	0.56
2:A:401:NAP:H3D	2:A:401:NAP:H6N	1.87	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:401:NAP:H3D	2:B:401:NAP:H6N	1.90	0.54
1:A:308:LEU:HD11	3:A:402:STR:C16	2.39	0.53
1:B:270:LYS:O	2:B:401:NAP:H8A	2.08	0.52
1:B:280:ASN:ND2	2:B:401:NAP:H61A	2.05	0.52
1:A:279:GLN:NE2	2:A:401:NAP:H62A	2.07	0.51
1:A:1:ASP:OD2	5:A:551:HOH:O	2.19	0.51
1:B:222:HIS:HD2	1:B:224:GLU:H	1.58	0.51
1:A:57:ASN:C	1:A:57:ASN:HD22	2.15	0.51
1:B:217:SER:HA	2:B:401:NAP:O1A	2.12	0.49
1:B:129:ILE:HD12	3:B:402:STR:C1	2.42	0.49
1:A:118:PHE:CE1	1:A:308:LEU:HD23	2.49	0.48
1:B:129:ILE:CD1	3:B:402:STR:H12	2.42	0.48
1:A:134:ASN:HB2	1:A:136:LYS:HD3	1.96	0.48
1:A:9:LYS:H	1:B:6:GLN:NE2	2.13	0.47
3:B:402:STR:H111	3:B:402:STR:H182	1.58	0.46
3:A:402:STR:H182	3:A:402:STR:H111	1.55	0.45
2:B:401:NAP:C3D	2:B:401:NAP:H6N	2.47	0.44
2:A:401:NAP:C3D	2:A:401:NAP:H6N	2.46	0.44
1:B:279:GLN:NE2	2:B:401:NAP:H62A	2.15	0.44
1:A:15:PHE:HB2	1:B:4:LYS:HD2	1.99	0.43
1:A:280:ASN:HD21	2:A:401:NAP:H61A	1.65	0.43
1:B:24:TYR:OH	3:B:402:STR:H72	2.19	0.43
1:A:217:SER:HA	2:A:401:NAP:O1A	2.18	0.43
1:B:57:ASN:HD22	1:B:57:ASN:C	2.20	0.43
1:B:86:TRP:CG	1:B:87:SER:N	2.86	0.42
1:A:86:TRP:CG	1:A:87:SER:N	2.88	0.42
1:A:161:LYS:HD3	5:A:609:HOH:O	2.19	0.42
1:B:216:TYR:CE1	1:B:306[B]:LEU:HD12	2.55	0.41
1:A:35:LEU:HD12	1:A:64:ALA:HB2	2.03	0.41
1:A:129:ILE:CD1	3:A:402:STR:H61	2.50	0.41
2:B:401:NAP:H51N	5:B:515:HOH:O	2.19	0.41
1:A:51:SER:HG	1:A:55:TYR:HD2	1.70	0.40
1:B:308:LEU:HD11	3:B:402:STR:H213	2.03	0.40
1:A:188:CYS:HA	1:A:212:VAL:O	2.22	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	324/325 (100%)	316 (98%)	8 (2%)	0	100	100
1	B	324/325 (100%)	315 (97%)	9 (3%)	0	100	100
All	All	648/650 (100%)	631 (97%)	17 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	288/287 (100%)	282 (98%)	6 (2%)	66	78
1	B	288/287 (100%)	284 (99%)	4 (1%)	78	88
All	All	576/574 (100%)	566 (98%)	10 (2%)	73	84

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	54	VAL
1	A	57	ASN
1	A	127	GLU
1	A	136	LYS
1	A	246	LYS
1	A	308	LEU
1	B	57	ASN
1	B	93	GLU
1	B	247	LYS
1	B	285	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	57	ASN
1	A	105	ASN
1	A	279	GLN
1	A	280	ASN
1	B	57	ASN
1	B	105	ASN
1	B	222	HIS
1	B	279	GLN
1	B	280	ASN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

Mogul failed to run properly - this section will therefore be empty.

5.5 Carbohydrates ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.6 Ligand geometry ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.7 Other polymers ⓘ

Mogul failed to run properly - this section will therefore be empty.

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	325/325 (100%)	-0.60	2 (0%) 86 88	3, 8, 26, 43	0
1	B	325/325 (100%)	-0.44	4 (1%) 75 76	5, 11, 25, 59	0
All	All	650/650 (100%)	-0.52	6 (0%) 81 82	3, 9, 26, 59	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	-1	SER	5.8
1	B	0	VAL	4.8
1	A	-1	SER	2.9
1	B	133	GLU	2.8
1	B	3	SER	2.2
1	A	133	GLU	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 ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	STR	B	402	23/23	0.25	8.58	28,33,38,40	23
3	STR	A	402	23/23	0.16	5.23	30,38,41,47	0
4	SO4	B	403	5/5	0.17	4.19	18,21,27,30	0
2	NAP	A	401	48/48	0.11	1.03	6,9,14,16	0
2	NAP	B	401	48/48	0.10	0.04	8,10,15,17	0
4	SO4	A	403	5/5	0.10	-0.46	11,12,14,15	0

6.5 Other polymers ⓘ

There are no such residues in this entry.