wwPDB X-ray Structure Validation Summary Report (i)

Apr 29, 2024 – 03:41 pm BST

PDB ID	:	4Z6K
Title	:	Alcohol dehydrogenase from the antarctic psychrophile Moraxella sp. TAE 123
Authors	:	Papanikolau, Y.; Bouriotis, V.; Petratos, K.
Deposited on	:	2015-04-05
Resolution	:	1.90 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:

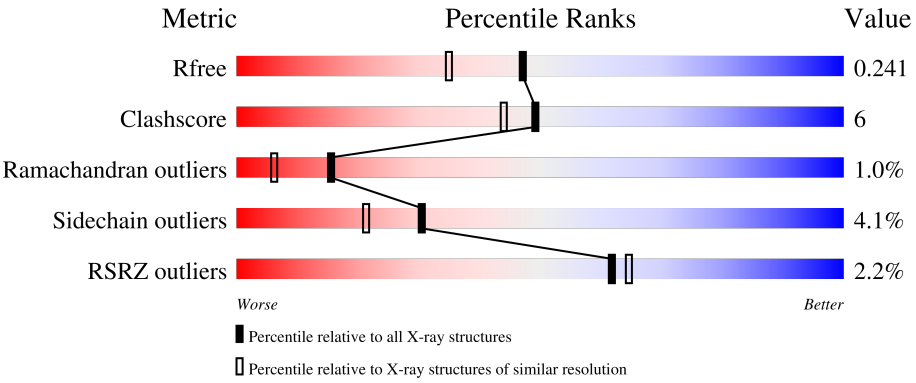
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	346	3% 87% 10% •
1	B	346	% 85% 10% • •
1	C	346	3% 89% 8% •
1	D	346	% 86% 9% • •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 11196 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 Alcohol dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	345	Total	C	N	O	S	0	2	0
			2577	1627	442	492	16			
1	B	340	Total	C	N	O	S	0	2	0
			2521	1593	427	485	16			
1	C	345	Total	C	N	O	S	0	1	0
			2567	1621	441	489	16			
1	D	340	Total	C	N	O	S	0	2	0
			2526	1596	428	486	16			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	339	LEU	-	expression tag	UNP Q8GIX7
A	340	GLU	-	expression tag	UNP Q8GIX7
A	341	HIS	-	expression tag	UNP Q8GIX7
A	342	HIS	-	expression tag	UNP Q8GIX7
A	343	HIS	-	expression tag	UNP Q8GIX7
A	344	HIS	-	expression tag	UNP Q8GIX7
A	345	HIS	-	expression tag	UNP Q8GIX7
A	346	HIS	-	expression tag	UNP Q8GIX7
B	339	LEU	-	expression tag	UNP Q8GIX7
B	340	GLU	-	expression tag	UNP Q8GIX7
B	341	HIS	-	expression tag	UNP Q8GIX7
B	342	HIS	-	expression tag	UNP Q8GIX7
B	343	HIS	-	expression tag	UNP Q8GIX7
B	344	HIS	-	expression tag	UNP Q8GIX7
B	345	HIS	-	expression tag	UNP Q8GIX7
B	346	HIS	-	expression tag	UNP Q8GIX7
C	339	LEU	-	expression tag	UNP Q8GIX7
C	340	GLU	-	expression tag	UNP Q8GIX7
C	341	HIS	-	expression tag	UNP Q8GIX7
C	342	HIS	-	expression tag	UNP Q8GIX7
C	343	HIS	-	expression tag	UNP Q8GIX7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	344	HIS	-	expression tag	UNP Q8GIX7
C	345	HIS	-	expression tag	UNP Q8GIX7
C	346	HIS	-	expression tag	UNP Q8GIX7
D	339	LEU	-	expression tag	UNP Q8GIX7
D	340	GLU	-	expression tag	UNP Q8GIX7
D	341	HIS	-	expression tag	UNP Q8GIX7
D	342	HIS	-	expression tag	UNP Q8GIX7
D	343	HIS	-	expression tag	UNP Q8GIX7
D	344	HIS	-	expression tag	UNP Q8GIX7
D	345	HIS	-	expression tag	UNP Q8GIX7
D	346	HIS	-	expression tag	UNP Q8GIX7

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Zn 2 2	0	0
2	B	3	Total Zn 3 3	0	0
2	C	2	Total Zn 2 2	0	0
2	D	3	Total Zn 3 3	0	0

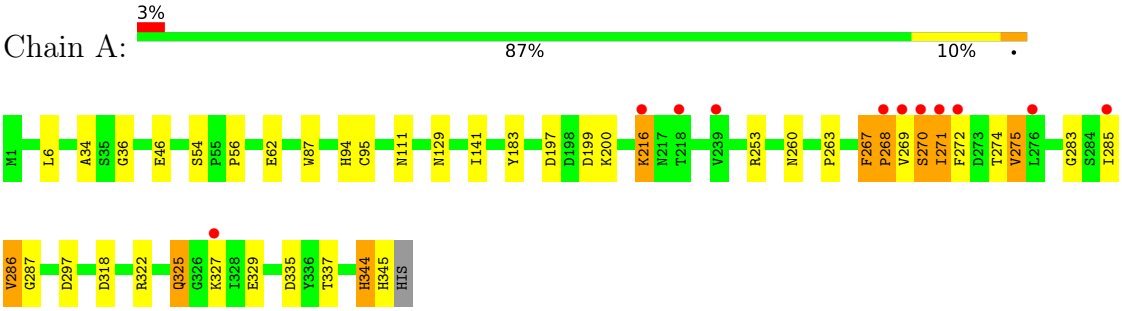
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	260	Total O 260 260	0	0
3	B	269	Total O 269 269	0	0
3	C	230	Total O 230 230	0	0
3	D	236	Total O 236 236	0	0

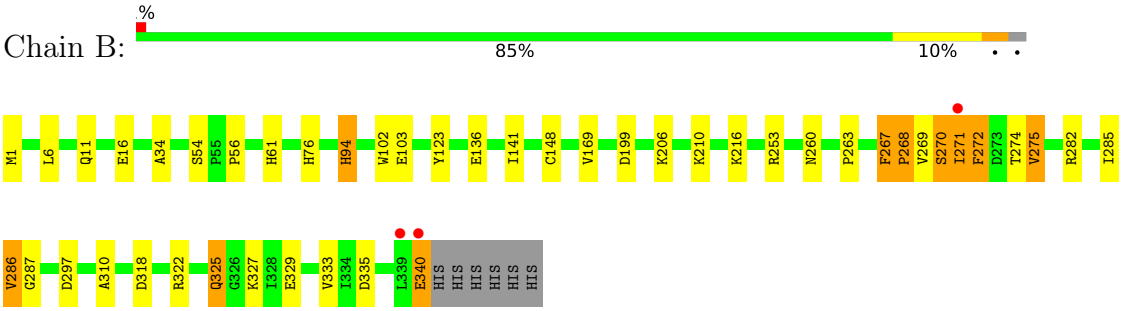
3 Residue-property plots (i)

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.

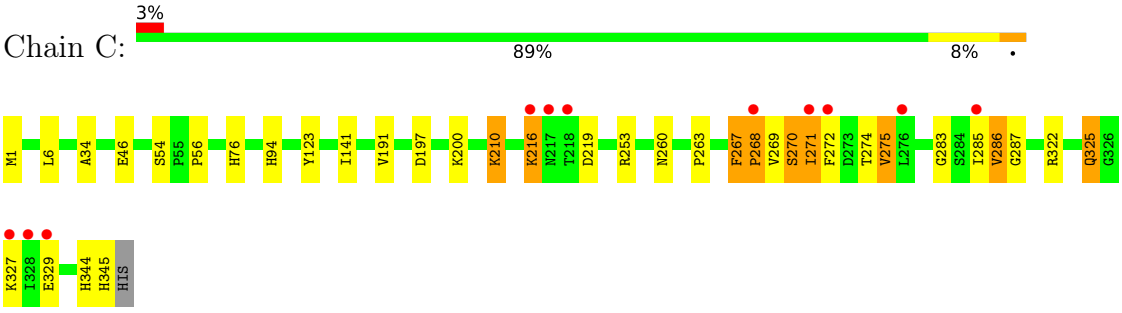
• Molecule 1: Alcohol dehydrogenase



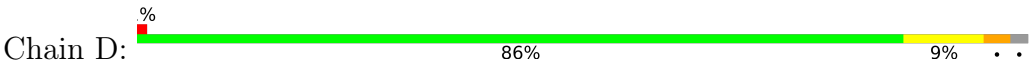
• Molecule 1: Alcohol dehydrogenase

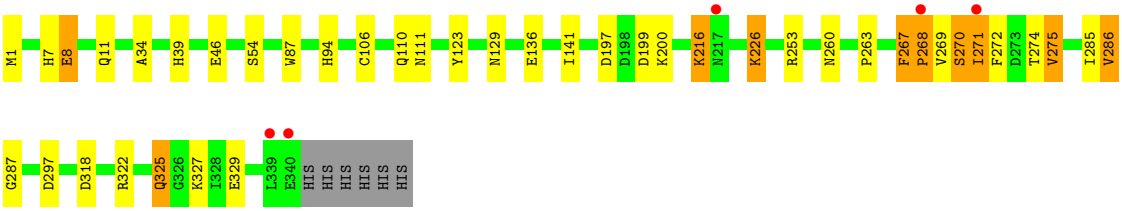


• Molecule 1: Alcohol dehydrogenase



• Molecule 1: Alcohol dehydrogenase




M1
H7
E8
Q11
A34
H39
E46
S54
W87
H94
C106
Q110
N111
Y123
N129
E136
I141
D197
D198
D199
K200
K216
N217
K226
R253
N260
P263
F267
P268
V269
S270
I271
F272
D273
T274
V275
I285
V286
G287
D297
D318
R322
Q325
G326
K327
I328
E329
L339
E340
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	136.57Å 136.57Å 210.83Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	118.34 – 1.90 78.70 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.6 (118.34-1.90) 96.6 (78.70-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	0.78 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.210 , 0.233 0.218 , 0.241	Depositor DCC
R_{free} test set	8613 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	37.0	Xtriage
Anisotropy	0.393	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 57.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.36$	Xtriage
Estimated twinning fraction	0.021 for -h,-k,l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	11196	wwPDB-VP
Average B, all atoms (Å^2)	60.0	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 42.57 % 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 2.0011e-04. 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

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	1.00	2/2631 (0.1%)	0.98	5/3576 (0.1%)
1	B	1.04	5/2579 (0.2%)	1.02	7/3505 (0.2%)
1	C	0.93	0/2625	0.97	2/3568 (0.1%)
1	D	0.99	1/2579 (0.0%)	0.99	4/3505 (0.1%)
All	All	0.99	8/10414 (0.1%)	0.99	18/14154 (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	A	0	1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	340	GLU	CG-CD	6.88	1.62	1.51
1	A	95	CYS	CB-SG	-6.60	1.71	1.82
1	B	103	GLU	CD-OE1	6.57	1.32	1.25
1	B	102	TRP	CB-CG	-6.06	1.39	1.50
1	D	106	CYS	CB-SG	-5.63	1.72	1.81

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	199	ASP	CB-CG-OD2	8.28	125.75	118.30
1	C	253	ARG	NE-CZ-NH2	6.80	123.70	120.30
1	B	318	ASP	CB-CG-OD2	-6.72	112.25	118.30
1	B	199	ASP	CB-CG-OD2	6.62	124.26	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	297	ASP	CB-CG-OD1	6.57	124.22	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	344	HIS	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	2577	0	2548	36	0
1	B	2521	0	2510	32	0
1	C	2567	0	2541	34	0
1	D	2526	0	2513	39	0
2	A	2	0	0	0	0
2	B	3	0	0	0	0
2	C	2	0	0	0	0
2	D	3	0	0	0	0
3	A	260	0	0	3	0
3	B	269	0	0	9	0
3	C	230	0	0	4	0
3	D	236	0	0	4	0
All	All	11196	0	10112	119	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 6.

The worst 5 of 119 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:94[B]:HIS:HD2	3:C:677:HOH:O	1.34	1.08
1:C:219:ASP:HA	3:C:633:HOH:O	1.55	1.05
3:B:687:HOH:O	1:C:94[B]:HIS:HD2	1.57	0.87

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:687:HOH:O	1:C:94[B]:HIS:CD2	2.34	0.79
1:B:271:ILE:HG21	1:D:263:PRO:HG3	1.66	0.78

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	345/346 (100%)	327 (95%)	15 (4%)	3 (1%)	17	7
1	B	340/346 (98%)	324 (95%)	12 (4%)	4 (1%)	13	4
1	C	344/346 (99%)	328 (95%)	13 (4%)	3 (1%)	17	7
1	D	340/346 (98%)	326 (96%)	11 (3%)	3 (1%)	17	7
All	All	1369/1384 (99%)	1305 (95%)	51 (4%)	13 (1%)	15	7

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	270	SER
1	A	271	ILE
1	B	270	SER
1	B	271	ILE
1	C	270	SER

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	273/272 (100%)	263 (96%)	10 (4%)	34	25
1	B	268/272 (98%)	256 (96%)	12 (4%)	27	18
1	C	272/272 (100%)	261 (96%)	11 (4%)	31	22
1	D	268/272 (98%)	256 (96%)	12 (4%)	27	18
All	All	1081/1088 (99%)	1036 (96%)	45 (4%)	30	20

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	268	PRO
1	D	46	GLU
1	C	275	VAL
1	C	345	HIS
1	D	216	LYS

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

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)

Of 10 ligands modelled in this entry, 10 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 95^{th} 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	345/346 (99%)	0.13	11 (3%) 47 50	40, 54, 90, 114	0
1	B	340/346 (98%)	0.14	3 (0%) 84 85	40, 53, 89, 116	0
1	C	345/346 (99%)	0.19	11 (3%) 47 50	42, 58, 93, 119	0
1	D	340/346 (98%)	0.10	5 (1%) 73 76	40, 57, 89, 117	0
All	All	1370/1384 (98%)	0.14	30 (2%) 62 64	40, 56, 91, 119	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	268	PRO	4.9
1	A	268	PRO	4.6
1	B	339	LEU	4.4
1	D	339	LEU	4.1
1	C	218	THR	3.9

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

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. The B-factors column lists the minimum, median, 95^{th} 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	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	ZN	B	401	1/1	0.97	0.13	54,54,54,54	0
2	ZN	B	402	1/1	0.97	0.17	41,41,41,41	0
2	ZN	C	401	1/1	0.97	0.12	56,56,56,56	0
2	ZN	D	401	1/1	0.97	0.12	57,57,57,57	0
2	ZN	B	403	1/1	0.98	0.16	50,50,50,50	0
2	ZN	A	401	1/1	0.98	0.12	53,53,53,53	0
2	ZN	C	402	1/1	0.98	0.15	44,44,44,44	0
2	ZN	A	402	1/1	0.98	0.15	42,42,42,42	0
2	ZN	D	402	1/1	0.98	0.16	40,40,40,40	0
2	ZN	D	403	1/1	0.98	0.14	59,59,59,59	0

6.5 Other polymers (i)

There are no such residues in this entry.