



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:32 PM GMT

PDB ID : 1PIW  
Title : APO AND HOLO STRUCTURES OF AN NADP(H)-DEPENDENT CIN-  
NAMYL ALCOHOL DEHYDROGENASE FROM SACCHAROMYCES  
CEREVISIAE  
Authors : Valencia, E.; Larroy, C.; Ochoa, W.F.; Pares, X.; Fita, I.; Biosca, J.A.  
Deposited on : 2003-05-30  
Resolution : 3.00 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

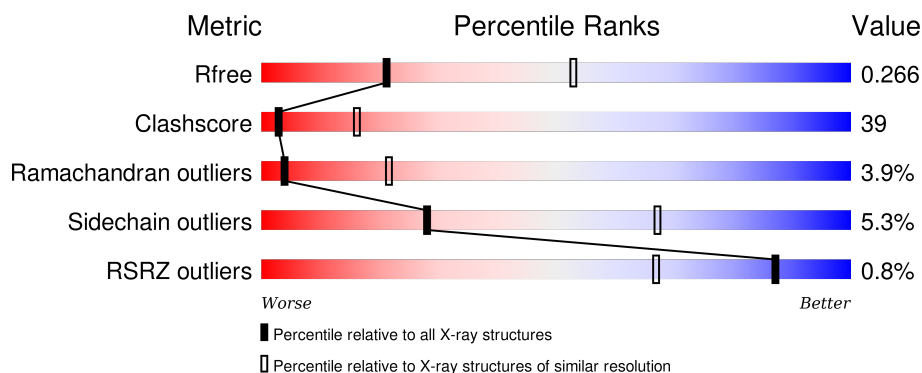
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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	360	<div> <div></div> <div>48%</div> <div>47%</div> <div>5%</div> </div>
1	B	360	<div> <div></div> <div>36%</div> <div>56%</div> <div>8%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NAP	B	4293	-	-	X	X

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5660 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.

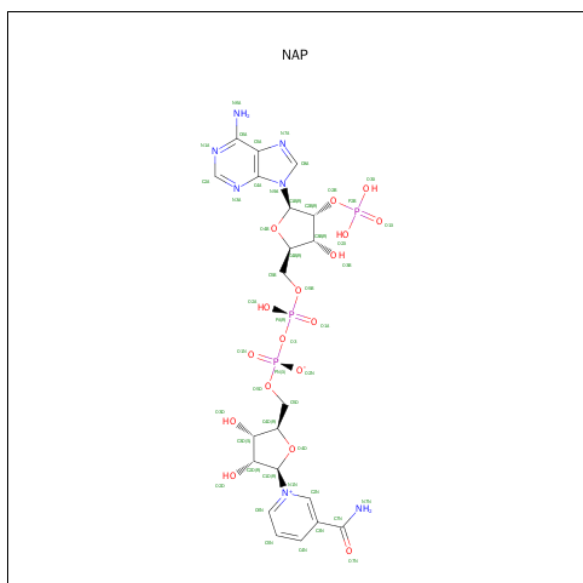
- Molecule 1 is a protein called Hypothetical zinc-type alcohol dehydrogenase-like protein in PRE5-FET4 intergenic region.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	360	Total	C	N	O	S	0	0	0
			2780	1770	466	523	21			
1	B	360	Total	C	N	O	S	0	0	0
			2780	1770	466	523	21			

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

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

- Molecule 3 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: C<sub>21</sub>H<sub>28</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>).

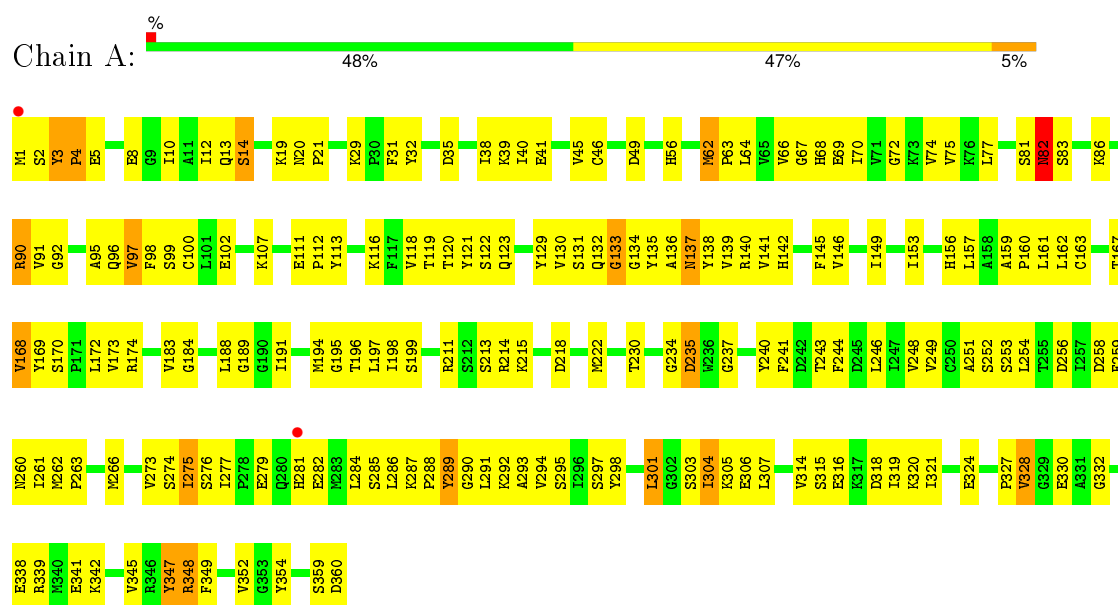


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

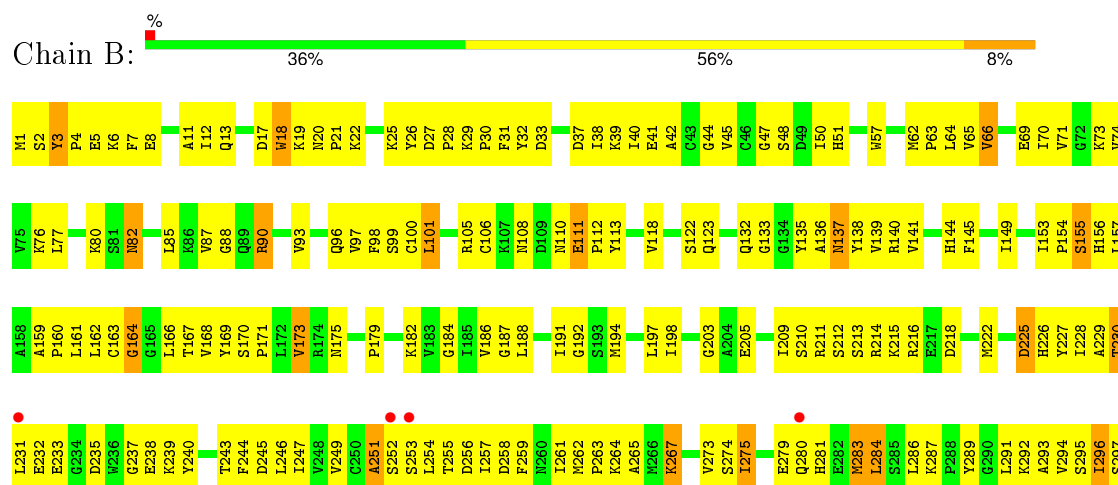
### 3 Residue-property plots

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: Hypothetical zinc-type alcohol dehydrogenase-like protein in PRE5-FET4 intergenic region



- Molecule 1: Hypothetical zinc-type alcohol dehydrogenase-like protein in PRE5-FET4 intergenic region



Y298	Y299	A300	L301	G302	S303	L304	A305	E306	L307		L310	L311	K312	L313	V314	S315	E316	K317	D318	L319	K320	L321	K322	V323	E324	T325	L326	P327	V328	G329	E330	A331	G332	V333	H334	E335	A336	F337	E338	K339	M340	E341	K342	G343	D344	V345	R346	Y347	R348	F349	T350	L351	Y352	G353	Y354		E357	D360
------	------	------	------	------	------	------	------	------	------	--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--	------	------

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	234.98Å 100.35Å 52.49Å 90.00° 92.18° 90.00°	Depositor
Resolution (Å)	30.00 – 3.00 29.35 – 3.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-3.00) 98.8 (29.35-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.35 (at 3.00Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.218 , 0.271 0.219 , 0.266	Depositor DCC
$R_{free}$ test set	2393 reflections (9.88%)	DCC
Wilson B-factor (Å <sup>2</sup> )	60.7	Xtriage
Anisotropy	0.762	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 53.8	EDS
Estimated twinning fraction	0.021 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 24236 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5660	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.69% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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, NAP

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	0.42	0/2844	0.71	0/3839
1	B	0.39	0/2844	0.63	0/3839
All	All	0.41	0/5688	0.67	0/7678

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	2780	0	2758	169	0
1	B	2780	0	2758	282	1
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	48	0	25	11	0
3	B	48	0	25	26	0
All	All	5660	0	5566	440	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 39.

The worst 5 of 440 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:252:SER:HB2	3:B:4293:NAP:C6A	1.26	1.56
1:B:252:SER:CB	3:B:4293:NAP:C6A	2.23	1.17
1:B:300:ALA:HA	3:B:4293:NAP:H72N	1.15	1.08
1:B:252:SER:HB2	3:B:4293:NAP:N6A	1.69	1.06
1:B:301:LEU:HD23	3:B:4293:NAP:O7N	1.53	1.06

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:B:338:GLU:OE2	1:B:338:GLU:OE2 2_756	1.91	0.29

## 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	358/360 (99%)	302 (84%)	45 (13%)	11 (3%)	5	28
1	B	358/360 (99%)	290 (81%)	51 (14%)	17 (5%)	3	17
All	All	716/720 (99%)	592 (83%)	96 (13%)	28 (4%)	4	21

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	TYR
1	A	83	SER
1	A	235	ASP
1	B	251	ALA
1	B	283	MET

### 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	303/303 (100%)	286 (94%)	17 (6%)	26	65
1	B	303/303 (100%)	288 (95%)	15 (5%)	30	70
All	All	606/606 (100%)	574 (95%)	32 (5%)	28	67

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	315	SER
1	B	82	ASN
1	B	301	LEU
1	A	348	ARG
1	B	90	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	13	GLN
1	B	20	ASN
1	B	137	ASN
1	A	309	GLN
1	B	123	GLN

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	NAP	A	4292	-	42,52,52	1.80	9 (21%)	54,80,80	1.81	12 (22%)
3	NAP	B	4293	1	42,52,52	1.80	8 (19%)	54,80,80	1.81	12 (22%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAP	A	4292	-	-	0/27/67/67	0/5/5/5
3	NAP	B	4293	1	-	0/27/67/67	0/5/5/5

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	4292	NAP	PA-O1A	-2.41	1.42	1.51
3	B	4293	NAP	PA-O1A	-2.41	1.42	1.51
3	A	4292	NAP	C3N-C7N	2.00	1.53	1.50
3	B	4293	NAP	O4B-C4B	2.45	1.50	1.45
3	A	4292	NAP	O4B-C4B	2.46	1.50	1.45

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	4292	NAP	N3A-C2A-N1A	-7.29	123.31	128.89
3	B	4293	NAP	N3A-C2A-N1A	-7.25	123.34	128.89

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	4293	NAP	C1B-N9A-C4A	-3.51	121.65	126.94
3	A	4292	NAP	C1B-N9A-C4A	-3.49	121.68	126.94
3	A	4292	NAP	O7N-C7N-C3N	-2.90	116.42	119.59

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 37 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	4292	NAP	11	0
3	B	4293	NAP	26	0

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	360/360 (100%)	-0.46	2 (0%) 90 73	23, 45, 71, 93	0
1	B	360/360 (100%)	-0.27	4 (1%) 82 58	36, 67, 89, 100	0
All	All	720/720 (100%)	-0.37	6 (0%) 87 67	23, 54, 82, 100	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	252	SER	3.1
1	A	1	MET	2.9
1	B	253	SER	2.4
1	B	280	GLN	2.3
1	A	281	HIS	2.1

### 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 carbohydrates in this entry.

### 6.4 Ligands [i](#)

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, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAP	B	4293	48/48	0.67	0.50	3.44	50,50,50,50	48
3	NAP	A	4292	48/48	0.91	0.21	0.61	30,58,74,75	9
2	ZN	B	1502	1/1	0.98	0.17	-0.82	65,65,65,65	0
2	ZN	A	1502	1/1	0.99	0.20	-1.14	32,32,32,32	0
2	ZN	B	1501	1/1	0.99	0.07	-1.22	52,52,52,52	0
2	ZN	A	1501	1/1	1.00	0.06	-2.05	40,40,40,40	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.