



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 13, 2017 – 06:03 AM EST

PDB ID : 4R41  
Title : Complex Crystal structure of 4-nitro-2-phosphono-benzoic acid with sp-Aspartate-Semialdehyde Dehydrogenase and Nicotinamide-dinucleotide  
Authors : Pavlovsky, A.G.; Viola, R.E.  
Deposited on : unknown  
Resolution : 1.61 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : rb-20030345

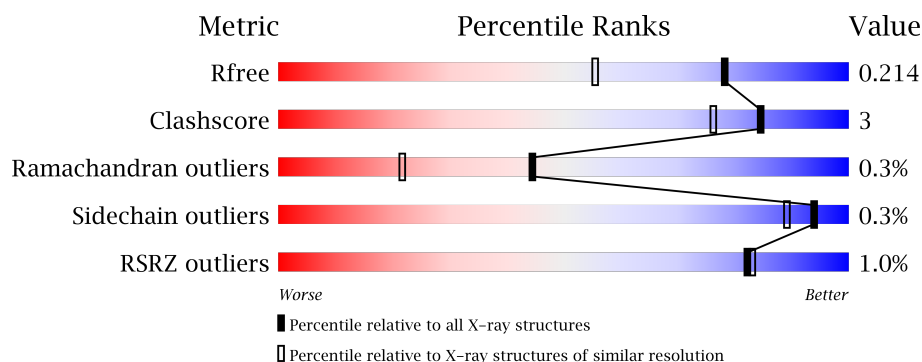
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*



The reported resolution of this entry is 1.61 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))
$R_{free}$	100719	3539 (1.64-1.60)
Clashscore	112137	3855 (1.64-1.60)
Ramachandran outliers	110173	3764 (1.64-1.60)
Sidechain outliers	110143	3763 (1.64-1.60)
RSRZ outliers	101464	3562 (1.64-1.60)

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	366	
1	B	366	

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
4	ACT	A	403	-	-	-	X
6	EDO	A	409	-	-	-	X
6	EDO	A	410	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 5909 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.

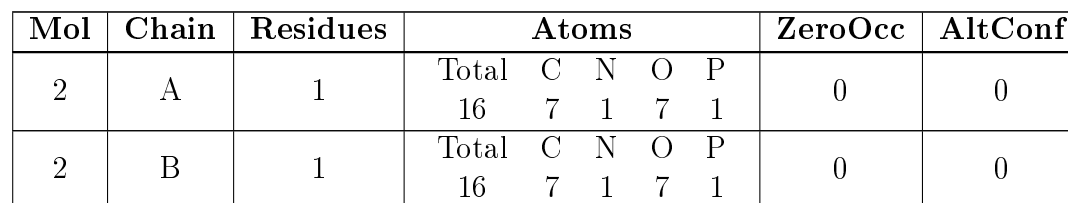
- Molecule 1 is a protein called Aspartate-semialdehyde dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	357	Total	C	N	O	S	0	6	0
			2758	1743	464	535	16			
1	B	347	Total	C	N	O	S	0	3	0
			2656	1676	450	517	13			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	359	LEU	-	EXPRESSION TAG	UNP A5MTN0
A	360	GLU	-	EXPRESSION TAG	UNP A5MTN0
A	361	HIS	-	EXPRESSION TAG	UNP A5MTN0
A	362	HIS	-	EXPRESSION TAG	UNP A5MTN0
A	363	HIS	-	EXPRESSION TAG	UNP A5MTN0
A	364	HIS	-	EXPRESSION TAG	UNP A5MTN0
A	365	HIS	-	EXPRESSION TAG	UNP A5MTN0
A	366	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	359	LEU	-	EXPRESSION TAG	UNP A5MTN0
B	360	GLU	-	EXPRESSION TAG	UNP A5MTN0
B	361	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	362	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	363	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	364	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	365	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	366	HIS	-	EXPRESSION TAG	UNP A5MTN0

- Molecule 2 is 4-nitro-2-phosphonobenzoic acid (three-letter code: 4NO) (formula:  $C_7H_6NO_7P$ ).



- # NAP
- 
- The chemical structure of Naproxen (NAP) is shown, featuring a naphthalene ring system. The structure includes a chiral center (C1) and a carboxylic acid group (COOH). The stereochemistry is indicated by wedged and dashed bonds. The structure is labeled with 'NAP' and 'NAP' at the top and bottom, and 'NAP' at the bottom right. The structure is also labeled with 'NAP' and 'NAP' at the top and bottom, and 'NAP' at the bottom right.

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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Na	0	0
			1	1		
5	A	2	Total	Na	0	0
			2	2		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is water.

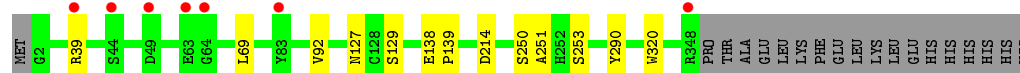
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	179	Total	O	0	0
			179	179		
7	B	149	Total	O	0	0
			149	149		



- Molecule 1: Aspartate-semialdehyde dehydrogenase



- Molecule 1: Aspartate-semialdehyde dehydrogenase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.59Å 97.73Å 63.94Å 90.00° 102.52° 90.00°	Depositor
Resolution (Å)	50.00 – 1.61 49.99 – 1.61	Depositor EDS
% Data completeness (in resolution range)	88.3 (50.00-1.61) 88.4 (49.99-1.61)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.37 (at 1.61Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.188 , 0.215 0.188 , 0.214	Depositor DCC
$R_{free}$ test set	4099 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	16.3	Xtriage
Anisotropy	0.474	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 52.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5909	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.14% 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, 0.333 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: NA, NAP, 4NO, EDO, ACT

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	0.52	0/2819	0.65	0/3829
1	B	0.49	0/2712	0.61	0/3684
All	All	0.50	0/5531	0.63	0/7513

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	2758	0	2772	20	0
1	B	2656	0	2665	8	0
2	A	16	0	5	0	0
2	B	16	0	5	0	0
3	A	48	0	25	2	0
3	B	48	0	25	4	0
4	A	4	0	3	0	0
5	A	2	0	0	0	0
5	B	1	0	0	0	0
6	A	20	0	30	2	0
6	B	12	0	18	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	179	0	0	1	1
7	B	149	0	0	1	1
All	All	5909	0	5548	30	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 3.

The worst 5 of 30 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:A:348[A]:ARG:NH2	7:A:549:HOH:O	1.95	0.97
1:A:86:LYS:HE2	1:A:353:LEU:HD22	1.58	0.85
1:B:69:LEU:HD23	1:B:92:VAL:HB	1.70	0.73
1:A:18[A]:MET:CE	1:A:94:ASN:HD21	2.03	0.72
1:B:39:ARG:HH11	3:B:402:NAP:H5N	1.56	0.70

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)
7:A:612:HOH:O	7:B:568:HOH:O[1_556]	1.52	0.68

## 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	361/366 (99%)	346 (96%)	14 (4%)	1 (0%)	44	22
1	B	348/366 (95%)	333 (96%)	14 (4%)	1 (0%)	44	22
All	All	709/732 (97%)	679 (96%)	28 (4%)	2 (0%)	44	22

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	251	ALA
1	B	251	ALA

### 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	299/302 (99%)	299 (100%)	0	100	100
1	B	287/302 (95%)	285 (99%)	2 (1%)	87	76
All	All	586/604 (97%)	584 (100%)	2 (0%)	94	89

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	250	SER
1	B	290	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	94	ASN
1	A	287	HIS
1	B	114	HIS

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 3 are monoatomic - leaving 13 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$
2	4NO	A	401	-	12,16,16	3.01	3 (25%)	14,24,24	1.98	3 (21%)
3	NAP	A	402	-	44,52,52	0.90	3 (6%)	51,80,80	1.37	3 (5%)
4	ACT	A	403	-	1,3,3	0.87	0	0,3,3	0.00	-
6	EDO	A	406	-	3,3,3	0.40	0	2,2,2	0.43	0
6	EDO	A	407	-	3,3,3	0.50	0	2,2,2	0.09	0
6	EDO	A	408	-	3,3,3	0.45	0	2,2,2	0.41	0
6	EDO	A	409	-	3,3,3	0.29	0	2,2,2	0.57	0
6	EDO	A	410	-	3,3,3	0.48	0	2,2,2	0.39	0
2	4NO	B	401	-	12,16,16	3.02	3 (25%)	14,24,24	1.54	3 (21%)
3	NAP	B	402	-	44,52,52	0.95	1 (2%)	51,80,80	1.37	4 (7%)
6	EDO	B	404	-	3,3,3	0.38	0	2,2,2	0.62	0
6	EDO	B	405	-	3,3,3	0.50	0	2,2,2	0.35	0
6	EDO	B	406	-	3,3,3	0.38	0	2,2,2	0.44	0

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
2	4NO	A	401	-	-	0/8/14/14	0/1/1/1
3	NAP	A	402	-	-	0/27/67/67	0/5/5/5
4	ACT	A	403	-	-	0/0/0/0	0/0/0/0
6	EDO	A	406	-	-	0/1/1/1	0/0/0/0
6	EDO	A	407	-	-	0/1/1/1	0/0/0/0
6	EDO	A	408	-	-	0/1/1/1	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	EDO	A	409	-	-	0/1/1/1	0/0/0/0
6	EDO	A	410	-	-	0/1/1/1	0/0/0/0
2	4NO	B	401	-	-	0/8/14/14	0/1/1/1
3	NAP	B	402	-	-	0/27/67/67	0/5/5/5
6	EDO	B	404	-	-	0/1/1/1	0/0/0/0
6	EDO	B	405	-	-	0/1/1/1	0/0/0/0
6	EDO	B	406	-	-	0/1/1/1	0/0/0/0

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	4NO	P13-C4	-3.25	1.75	1.80
2	A	401	4NO	C2-N1	-3.04	1.39	1.45
2	B	401	4NO	C2-N1	-2.79	1.40	1.45
2	B	401	4NO	P13-C4	-2.20	1.77	1.80
3	A	402	NAP	O4D-C1D	2.29	1.44	1.41

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	402	NAP	N3A-C2A-N1A	-6.43	123.25	128.86
3	B	402	NAP	N3A-C2A-N1A	-6.22	123.44	128.86
2	A	401	4NO	C6-C5-C10	-3.90	114.22	120.23
2	A	401	4NO	O14-P13-C4	-3.83	104.25	110.84
2	B	401	4NO	O14-P13-C4	-3.64	104.57	110.84

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	402	NAP	2	0
6	A	406	EDO	1	0
6	A	407	EDO	1	0
3	B	402	NAP	4	0

## 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 [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	357/366 (97%)	-0.45	0 100 100	9, 16, 29, 44	0
1	B	347/366 (94%)	-0.30	7 (2%) 65 65	10, 20, 41, 59	0
All	All	704/732 (96%)	-0.37	7 (0%) 82 83	9, 18, 35, 59	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	39	ARG	3.3
1	B	63	GLU	2.6
1	B	49	ASP	2.5
1	B	348	ARG	2.3
1	B	64	GLY	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
6	EDO	A	409	4/4	0.78	0.23	23.89	21,23,25,34	0
4	ACT	A	403	4/4	0.86	0.41	18.80	23,24,25,25	4
6	EDO	A	410	4/4	0.85	0.16	3.23	34,37,37,38	0
6	EDO	B	404	4/4	0.95	0.10	1.90	29,30,31,31	0
3	NAP	A	402	48/48	0.95	0.13	1.76	19,27,93,95	0
2	4NO	B	401	16/16	0.93	0.13	1.64	34,36,38,39	0
6	EDO	B	405	4/4	0.96	0.10	1.63	25,25,26,28	0
6	EDO	A	406	4/4	0.97	0.07	1.18	23,23,23,24	0
6	EDO	A	407	4/4	0.88	0.14	1.16	26,27,27,28	0
3	NAP	B	402	48/48	0.92	0.15	0.96	20,25,60,61	48
6	EDO	A	408	4/4	0.96	0.07	-0.22	31,32,34,34	0
2	4NO	A	401	16/16	0.98	0.07	-0.23	21,22,25,25	0
5	NA	B	403	1/1	0.95	0.08	-0.26	23,23,23,23	0
5	NA	A	405	1/1	0.99	0.05	-0.83	15,15,15,15	0
5	NA	A	404	1/1	0.99	0.05	-1.53	14,14,14,14	0
6	EDO	B	406	4/4	0.85	0.12	-	42,43,44,45	0

## 6.5 Other polymers

There are no such residues in this entry.