



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2017 – 05:39 PM EDT

PDB ID : 1R6Z  
Title : The Crystal Structure of the Argonaute2 PAZ domain (as a MBP fusion)  
Authors : Song, J.J.; Liu, J.; Tolia, N.H.; Schneiderman, J.; Smith, S.K.; Martienssen, R.A.; Hannon, G.J.; Joshua-Tor, L.  
Deposited on : unknown  
Resolution : 2.80 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-20029824  
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-20029824

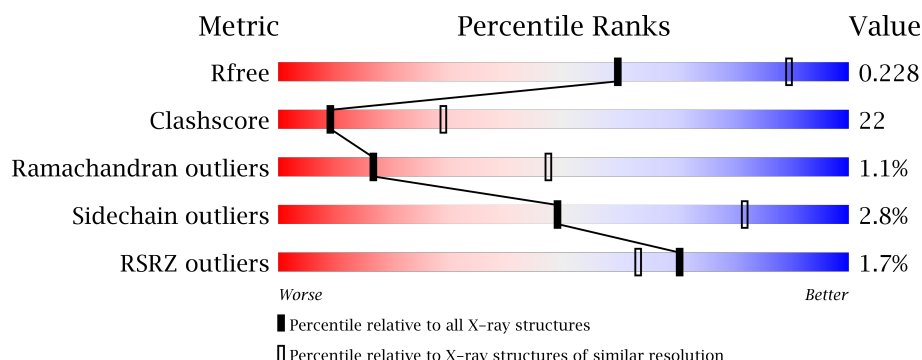
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	2583 (2.80-2.80)
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)
RSRZ outliers	101464	2610 (2.80-2.80)

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	509	<div> <div>%</div> <div> <div></div> <div>59%</div> <div>37%</div> <div>..</div> </div> </div>
1	P	509	<div> <div>3%</div> <div> <div></div> <div>59%</div> <div>36%</div> <div>..</div> </div> </div>
1	Z	509	<div> <div>%</div> <div> <div></div> <div>61%</div> <div>35%</div> <div>..</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11834 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.

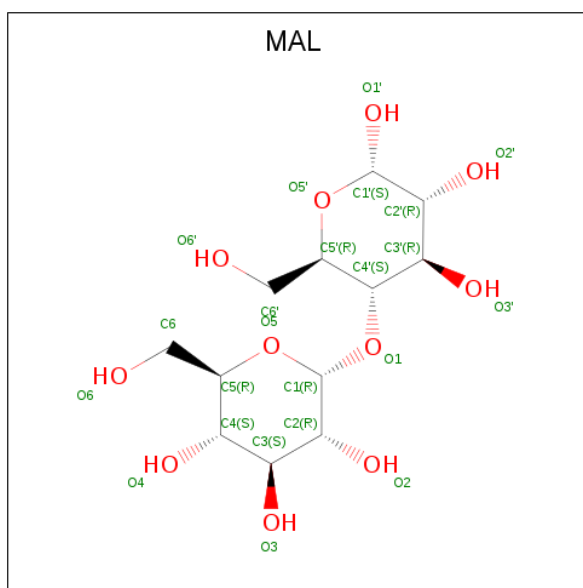
- Molecule 1 is a protein called Chimera of Maltose-binding periplasmic protein and Argonaute 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	P	498	Total	C	N	O	S	0	0	0
			3901	2511	642	738	10			
1	A	497	Total	C	N	O	S	0	0	0
			3892	2506	641	735	10			
1	Z	499	Total	C	N	O	S	0	0	0
			3906	2512	644	740	10			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	-1	MET	-	CLONING ARTIFACT	UNP P02928
P	366	ASN	-	LINKER	UNP P02928
P	367	ALA	-	LINKER	UNP P02928
P	368	ALA	-	LINKER	UNP P02928
P	369	ALA	-	LINKER	UNP P02928
P	370	GLU	-	LINKER	UNP P02928
P	371	PHE	-	LINKER	UNP P02928
A	-1	MET	-	CLONING ARTIFACT	UNP P02928
A	366	ASN	-	LINKER	UNP P02928
A	367	ALA	-	LINKER	UNP P02928
A	368	ALA	-	LINKER	UNP P02928
A	369	ALA	-	LINKER	UNP P02928
A	370	GLU	-	LINKER	UNP P02928
A	371	PHE	-	LINKER	UNP P02928
Z	-1	MET	-	CLONING ARTIFACT	UNP P02928
Z	366	ASN	-	LINKER	UNP P02928
Z	367	ALA	-	LINKER	UNP P02928
Z	368	ALA	-	LINKER	UNP P02928
Z	369	ALA	-	LINKER	UNP P02928
Z	370	GLU	-	LINKER	UNP P02928
Z	371	PHE	-	LINKER	UNP P02928

- Molecule 2 is MALTOSE (three-letter code: MAL) (formula:  $C_{12}H_{22}O_{11}$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	P	1	Total	C	O	0	0
			23	12	11		
2	A	1	Total	C	O	0	0
			23	12	11		
2	Z	1	Total	C	O	0	0
			23	12	11		

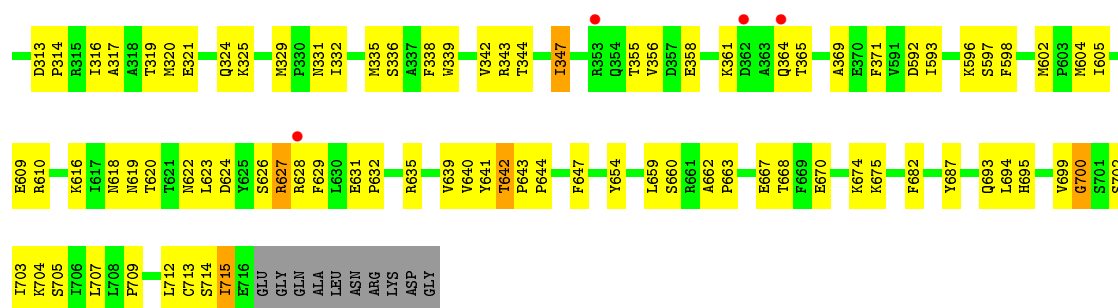
- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	P	3	Total	Ni	0	0
			3	3		
3	Z	3	Total	Ni	0	0
			3	3		
3	A	3	Total	Ni	0	0
			3	3		

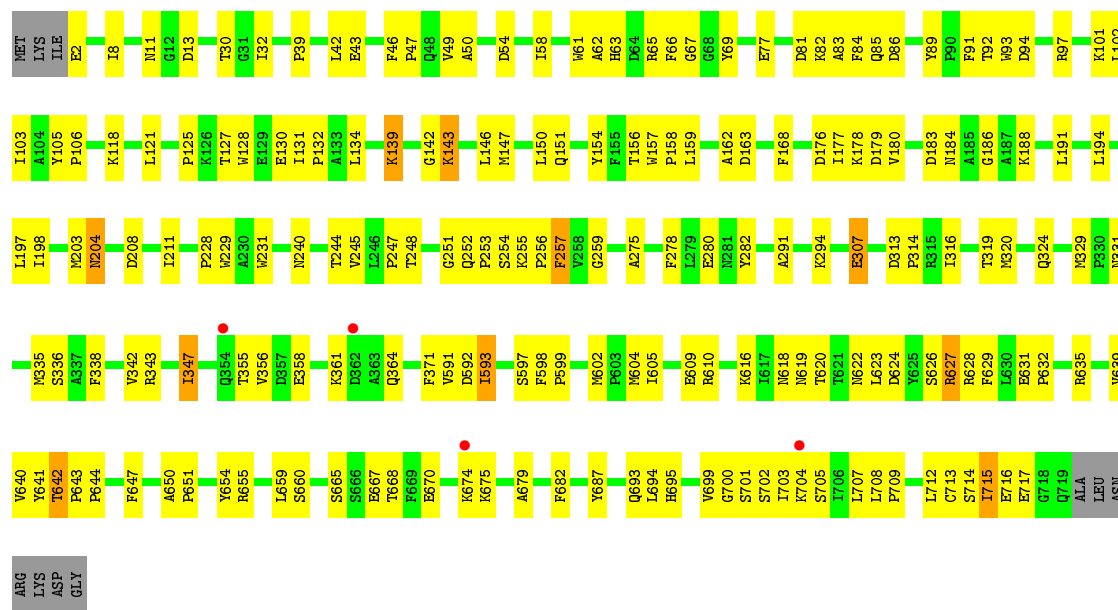
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	P	17	Total	O	0	0
			17	17		
4	A	22	Total	O	0	0
			22	22		
4	Z	18	Total	O	0	0
			18	18		





• Molecule 1: Chimera of Maltose-binding periplasmic protein and Argonaute 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	89.81Å 89.81Å 380.33Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.91 – 2.80 44.91 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (44.91-2.80) 97.3 (44.91-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.96 (at 2.81Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.227 , 0.281 0.234 , 0.228	Depositor DCC
$R_{free}$ test set	2088 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.9	Xtriage
Anisotropy	0.093	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 57.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.065 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	11834	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.19% 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: NI, MAL

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	0/3990	0.62	0/5417
1	P	0.40	0/3999	0.64	1/5429 (0.0%)
1	Z	0.43	0/4004	0.63	0/5435
All	All	0.42	0/11993	0.63	1/16281 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	P	595	HIS	N-CA-C	-6.06	94.65	111.00

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	3892	0	3862	169	0
1	P	3901	0	3868	183	0
1	Z	3906	0	3865	171	0
2	A	23	0	21	2	0
2	P	23	0	21	0	0
2	Z	23	0	21	2	0
3	A	3	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	P	3	0	0	0	0
3	Z	3	0	0	0	0
4	A	22	0	0	3	0
4	P	17	0	0	3	0
4	Z	18	0	0	0	0
All	All	11834	0	11658	519	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 22.

The worst 5 of 519 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:Z:593:ILE:H	1:Z:593:ILE:HD12	1.21	1.06
1:P:313:ASP:HB3	1:P:316:ILE:HD13	1.38	1.03
1:A:596:LYS:O	1:A:596:LYS:HD3	1.57	1.03
1:P:715:ILE:HG22	1:P:716:GLU:H	1.23	1.02
1:Z:313:ASP:HB3	1:Z:316:ILE:HD13	1.39	1.02

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	495/509 (97%)	456 (92%)	34 (7%)	5 (1%)	18	50
1	P	496/509 (97%)	453 (91%)	35 (7%)	8 (2%)	11	36
1	Z	497/509 (98%)	461 (93%)	33 (7%)	3 (1%)	28	62
All	All	1488/1527 (97%)	1370 (92%)	102 (7%)	16 (1%)	17	47

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	P	704	LYS
1	Z	703	ILE
1	P	619	ASN
1	A	619	ASN
1	Z	619	ASN

### 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	412/421 (98%)	401 (97%)	11 (3%)	50	83
1	P	413/421 (98%)	401 (97%)	12 (3%)	48	81
1	Z	413/421 (98%)	401 (97%)	12 (3%)	48	81
All	All	1238/1263 (98%)	1203 (97%)	35 (3%)	49	82

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	257	PHE
1	A	628	ARG
1	Z	642	THR
1	A	307	GLU
1	A	371	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	240	ASN
1	Z	595	HIS
1	Z	172	ASN
1	A	99	ASN
1	Z	240	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 9 are monoatomic - leaving 3 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	MAL	A	727	-	24,24,24	1.01	1 (4%)	35,35,35	1.71	8 (22%)
2	MAL	P	727	-	24,24,24	1.08	1 (4%)	35,35,35	1.62	7 (20%)
2	MAL	Z	727	-	24,24,24	1.04	1 (4%)	35,35,35	1.62	7 (20%)

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	MAL	A	727	-	-	0/8/48/48	0/2/2/2
2	MAL	P	727	-	-	0/8/48/48	0/2/2/2
2	MAL	Z	727	-	-	0/8/48/48	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	727	MAL	O5'-C1'	-2.54	1.38	1.43
2	Z	727	MAL	O5'-C1'	-2.22	1.39	1.43
2	A	727	MAL	C4'-C5'	2.23	1.58	1.52

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	Z	727	MAL	O1'-C1'-O5'	-3.03	101.25	110.20
2	A	727	MAL	O1'-C1'-O5'	-3.00	101.32	110.20
2	P	727	MAL	O1'-C1'-O5'	-2.99	101.36	110.20
2	Z	727	MAL	O6'-C6'-C5'	-2.52	102.86	111.34
2	P	727	MAL	O6'-C6'-C5'	-2.31	103.57	111.34

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	727	MAL	2	0
2	Z	727	MAL	2	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	497/509 (97%)	-0.07	7 (1%) 75 69	15, 43, 98, 176	0
1	P	498/509 (97%)	0.08	15 (3%) 51 39	9, 46, 100, 192	0
1	Z	499/509 (98%)	-0.07	4 (0%) 86 81	12, 38, 91, 152	0
All	All	1494/1527 (97%)	-0.02	26 (1%) 70 63	9, 42, 94, 192	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	P	625	TYR	4.1
1	A	172	ASN	3.6
1	P	30	THR	3.5
1	Z	674	LYS	3.2
1	P	631	GLU	3.0

### 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	NI	A	728	1/1	0.98	0.29	0.67	61,61,61,61	0
2	MAL	Z	727	23/23	0.95	0.19	0.15	30,30,30,30	0
2	MAL	P	727	23/23	0.95	0.17	-0.10	29,29,29,45	0
3	NI	Z	728	1/1	0.96	0.13	-0.65	67,67,67,67	0
2	MAL	A	727	23/23	0.94	0.15	-0.86	23,23,23,23	0
3	NI	A	730	1/1	0.85	0.17	-	106,106,106,106	0
3	NI	Z	730	1/1	0.49	0.18	-	123,123,123,123	0
3	NI	P	729	1/1	0.83	0.20	-	100,100,100,100	0
3	NI	A	729	1/1	0.94	0.27	-	77,77,77,77	0
3	NI	Z	729	1/1	0.97	0.15	-	58,58,58,58	0
3	NI	P	730	1/1	0.83	0.12	-	91,91,91,91	0
3	NI	P	728	1/1	0.99	0.14	-	39,39,39,39	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.