



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 27, 2017 – 05:37 PM EST

PDB ID : 6B57  
Title : tudor in complex with ligand  
Authors : Zhang, H.; Tempel, W.; Bountra, C.; Arrowsmith, C.H.; Edwards, A.M.; Min, J.; Structural Genomics Consortium (SGC)  
Deposited on : unknown  
Resolution : 1.93 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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  
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

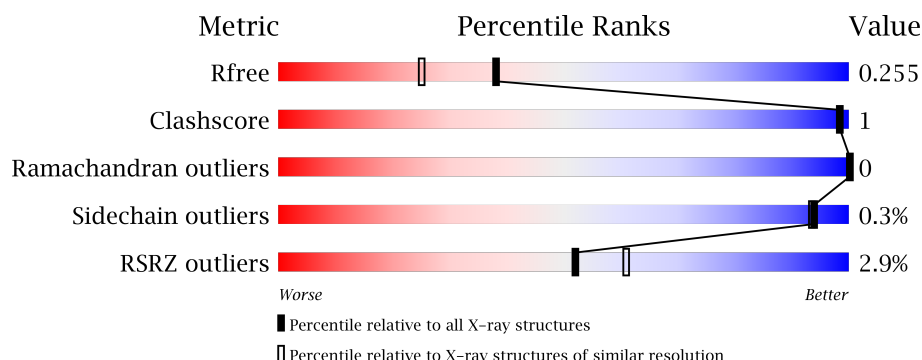
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.93 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	3233 (1.96-1.92)
Clashscore	112137	3430 (1.96-1.92)
Ramachandran outliers	110173	3395 (1.96-1.92)
Sidechain outliers	110143	3395 (1.96-1.92)
RSRZ outliers	101464	3250 (1.96-1.92)

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	231	<div> <div>4%</div> <div> <div></div> <div>85%</div> <div>•</div> <div>13%</div> </div> </div>
1	B	231	<div> <div>%</div> <div> <div></div> <div>88%</div> <div>•</div> <div>10%</div> </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
2	UNX	B	601	-	-	-	X
2	UNX	B	602	-	-	-	X
2	UNX	B	606	-	-	-	X
2	UNX	B	607	-	-	-	X
2	UNX	B	608	-	-	-	X
2	UNX	B	609	-	-	-	X
2	UNX	B	610	-	-	-	X
2	UNX	B	612	-	-	-	X
2	UNX	B	613	-	-	-	X
2	UNX	B	614	-	-	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3287 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 Tudor and KH domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	202	Total	C	N	O	S	0	1	1
			1532	978	260	290	4			
1	B	207	Total	C	N	O	S	0	2	0
			1615	1026	284	301	4			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	291	MET	-	expression tag	UNP Q9Y2W6
A	292	HIS	-	expression tag	UNP Q9Y2W6
A	293	HIS	-	expression tag	UNP Q9Y2W6
A	294	HIS	-	expression tag	UNP Q9Y2W6
A	295	HIS	-	expression tag	UNP Q9Y2W6
A	296	HIS	-	expression tag	UNP Q9Y2W6
A	297	HIS	-	expression tag	UNP Q9Y2W6
A	298	SER	-	expression tag	UNP Q9Y2W6
A	299	SER	-	expression tag	UNP Q9Y2W6
A	300	GLY	-	expression tag	UNP Q9Y2W6
A	301	ARG	-	expression tag	UNP Q9Y2W6
A	302	GLU	-	expression tag	UNP Q9Y2W6
A	303	ASN	-	expression tag	UNP Q9Y2W6
A	304	LEU	-	expression tag	UNP Q9Y2W6
A	305	TYR	-	expression tag	UNP Q9Y2W6
A	306	PHE	-	expression tag	UNP Q9Y2W6
A	307	GLN	-	expression tag	UNP Q9Y2W6
A	308	GLY	-	expression tag	UNP Q9Y2W6
A	498	GLY	-	expression tag	UNP Q9Y2W6
A	499	SER	-	expression tag	UNP Q9Y2W6
A	500	ALA	-	expression tag	UNP Q9Y2W6
A	501	GLY	-	expression tag	UNP Q9Y2W6
A	502	SER	-	expression tag	UNP Q9Y2W6
A	503	ALA	-	expression tag	UNP Q9Y2W6
A	504	GLY	-	expression tag	UNP Q9Y2W6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	505	SER	-	expression tag	UNP Q9Y2W6
A	506	ALA	-	expression tag	UNP Q9Y2W6
A	507	GLY	-	expression tag	UNP Q9Y2W6
A	508	SER	-	expression tag	UNP Q9Y2W6
A	509	ALA	-	expression tag	UNP Q9Y2W6
A	510	THR	-	expression tag	UNP Q9Y2W6
A	511	GLY	-	expression tag	UNP Q9Y2W6
A	512	ARG	-	expression tag	UNP Q9Y2W6
A	513	ALA	-	expression tag	UNP Q9Y2W6
A	514	ARG	-	expression tag	UNP Q9Y2W6
A	515	ALA	-	expression tag	UNP Q9Y2W6
A	516	ARG	-	expression tag	UNP Q9Y2W6
A	517	ALA	-	expression tag	UNP Q9Y2W6
A	518	ARG	-	expression tag	UNP Q9Y2W6
A	519	GLY	-	expression tag	UNP Q9Y2W6
A	520	ARG	-	expression tag	UNP Q9Y2W6
A	521	ALA	-	expression tag	UNP Q9Y2W6
B	291	MET	-	expression tag	UNP Q9Y2W6
B	292	HIS	-	expression tag	UNP Q9Y2W6
B	293	HIS	-	expression tag	UNP Q9Y2W6
B	294	HIS	-	expression tag	UNP Q9Y2W6
B	295	HIS	-	expression tag	UNP Q9Y2W6
B	296	HIS	-	expression tag	UNP Q9Y2W6
B	297	HIS	-	expression tag	UNP Q9Y2W6
B	298	SER	-	expression tag	UNP Q9Y2W6
B	299	SER	-	expression tag	UNP Q9Y2W6
B	300	GLY	-	expression tag	UNP Q9Y2W6
B	301	ARG	-	expression tag	UNP Q9Y2W6
B	302	GLU	-	expression tag	UNP Q9Y2W6
B	303	ASN	-	expression tag	UNP Q9Y2W6
B	304	LEU	-	expression tag	UNP Q9Y2W6
B	305	TYR	-	expression tag	UNP Q9Y2W6
B	306	PHE	-	expression tag	UNP Q9Y2W6
B	307	GLN	-	expression tag	UNP Q9Y2W6
B	308	GLY	-	expression tag	UNP Q9Y2W6
B	498	GLY	-	expression tag	UNP Q9Y2W6
B	499	SER	-	expression tag	UNP Q9Y2W6
B	500	ALA	-	expression tag	UNP Q9Y2W6
B	501	GLY	-	expression tag	UNP Q9Y2W6
B	502	SER	-	expression tag	UNP Q9Y2W6
B	503	ALA	-	expression tag	UNP Q9Y2W6
B	504	GLY	-	expression tag	UNP Q9Y2W6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	505	SER	-	expression tag	UNP Q9Y2W6
B	506	ALA	-	expression tag	UNP Q9Y2W6
B	507	GLY	-	expression tag	UNP Q9Y2W6
B	508	SER	-	expression tag	UNP Q9Y2W6
B	509	ALA	-	expression tag	UNP Q9Y2W6
B	510	THR	-	expression tag	UNP Q9Y2W6
B	511	GLY	-	expression tag	UNP Q9Y2W6
B	512	ARG	-	expression tag	UNP Q9Y2W6
B	513	ALA	-	expression tag	UNP Q9Y2W6
B	514	ARG	-	expression tag	UNP Q9Y2W6
B	515	ALA	-	expression tag	UNP Q9Y2W6
B	516	ARG	-	expression tag	UNP Q9Y2W6
B	517	ALA	-	expression tag	UNP Q9Y2W6
B	518	ARG	-	expression tag	UNP Q9Y2W6
B	519	GLY	-	expression tag	UNP Q9Y2W6
B	520	ARG	-	expression tag	UNP Q9Y2W6
B	521	ALA	-	expression tag	UNP Q9Y2W6

- Molecule 2 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	19	Total X 19 19	0	0
2	A	15	Total X 15 15	0	0

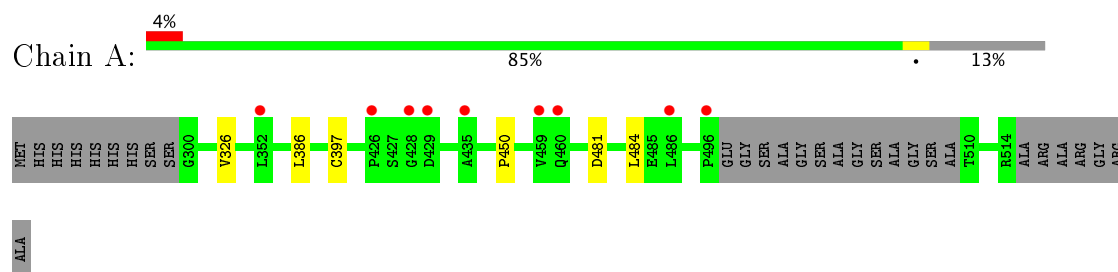
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	54	Total O 54 54	0	0
3	B	52	Total O 52 52	0	0

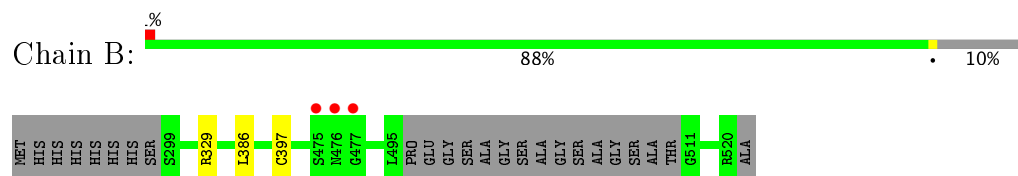
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Tudor and KH domain-containing protein



- Molecule 1: Tudor and KH domain-containing protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.07Å 72.66Å 130.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	65.35 – 1.93 65.35 – 1.93	Depositor EDS
% Data completeness (in resolution range)	99.0 (65.35-1.93) 99.4 (65.35-1.93)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.85 (at 1.94Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.206 , 0.234 0.220 , 0.255	Depositor DCC
$R_{free}$ test set	1683 reflections (5.17%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.6	Xtriage
Anisotropy	0.313	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 40.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3287	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.64% 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: UNX

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	0.51	0/1572	0.67	0/2146
1	B	0.52	0/1662	0.69	0/2262
All	All	0.51	0/3234	0.68	0/4408

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	1532	0	1407	3	0
1	B	1615	0	1519	1	0
2	A	15	0	0	0	0
2	B	19	0	0	0	0
3	A	54	0	0	0	0
3	B	52	0	0	0	0
All	All	3287	0	2926	4	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 1.

All (4) 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:481:ASP:HB3	1:A:484:LEU:HB2	1.91	0.51
1:B:386:LEU:HD11	1:B:397:CYS:HB2	1.93	0.49
1:A:326:VAL:HG11	1:A:450:PRO:HG3	1.98	0.46
1:A:386:LEU:HD11	1:A:397:CYS:HB2	2.00	0.43

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	199/231 (86%)	193 (97%)	6 (3%)	0	100	100
1	B	205/231 (89%)	202 (98%)	3 (2%)	0	100	100
All	All	404/462 (87%)	395 (98%)	9 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	150/191 (78%)	150 (100%)	0	100	100
1	B	163/191 (85%)	162 (99%)	1 (1%)	89	87
All	All	313/382 (82%)	312 (100%)	1 (0%)	94	93

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	329	ARG

Some 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 34 ligands modelled in this entry, 34 are unknown - 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 [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	202/231 (87%)	0.34	9 (4%) 34 42	24, 45, 85, 96	0
1	B	207/231 (89%)	0.14	3 (1%) 75 82	23, 41, 75, 87	0
All	All	409/462 (88%)	0.24	12 (2%) 52 61	23, 43, 80, 96	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	428	GLY	5.2
1	A	496	PRO	4.7
1	A	435	ALA	3.8
1	A	429	ASP	3.8
1	A	426	PRO	3.3
1	A	459	VAL	3.1
1	B	475	SER	3.1
1	A	486	LEU	2.9
1	B	476	ASN	2.9
1	B	477	GLY	2.8
1	A	352	LEU	2.4
1	A	460	GLN	2.2

### 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 ⓘ

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(Å <sup>2</sup> )	Q<0.9
2	UNX	B	612	1/1	0.94	0.24	14.01	28,28,28,28	0
2	UNX	B	607	1/1	0.94	0.21	11.31	33,33,33,33	0
2	UNX	B	601	1/1	0.82	0.19	9.80	31,31,31,31	0
2	UNX	B	614	1/1	0.92	0.17	5.90	30,30,30,30	0
2	UNX	B	613	1/1	0.97	0.17	5.55	30,30,30,30	0
2	UNX	B	608	1/1	0.95	0.19	4.84	33,33,33,33	0
2	UNX	B	609	1/1	0.96	0.16	4.47	30,30,30,30	0
2	UNX	B	606	1/1	0.92	0.17	3.77	39,39,39,39	0
2	UNX	B	610	1/1	0.98	0.15	3.51	33,33,33,33	0
2	UNX	B	602	1/1	0.89	0.17	2.48	30,30,30,30	0
2	UNX	A	606	1/1	0.82	0.20	1.48	39,39,39,39	0
2	UNX	B	619	1/1	0.91	0.14	1.38	33,33,33,33	0
2	UNX	A	601	1/1	0.96	0.13	1.23	26,26,26,26	0
2	UNX	B	615	1/1	0.90	0.15	1.00	33,33,33,33	0
2	UNX	A	608	1/1	0.82	0.20	0.80	48,48,48,48	0
2	UNX	B	618	1/1	0.91	0.35	-	52,52,52,52	0
2	UNX	A	611	1/1	0.93	0.26	-	30,30,30,30	0
2	UNX	A	609	1/1	0.97	0.16	-	33,33,33,33	0
2	UNX	A	602	1/1	0.95	0.16	-	41,41,41,41	0
2	UNX	A	612	1/1	0.93	0.21	-	29,29,29,29	0
2	UNX	A	610	1/1	0.88	0.19	-	38,38,38,38	0
2	UNX	A	614	1/1	0.87	0.36	-	50,50,50,50	0
2	UNX	A	613	1/1	0.85	0.18	-	40,40,40,40	0
2	UNX	A	604	1/1	0.80	0.46	-	46,46,46,46	0
2	UNX	B	603	1/1	0.88	0.36	-	40,40,40,40	0
2	UNX	B	617	1/1	0.90	0.40	-	45,45,45,45	0
2	UNX	A	605	1/1	0.90	0.17	-	38,38,38,38	0
2	UNX	B	605	1/1	0.91	0.20	-	34,34,34,34	0
2	UNX	A	607	1/1	0.86	0.25	-	50,50,50,50	0
2	UNX	A	603	1/1	0.92	0.24	-	46,46,46,46	0
2	UNX	B	616	1/1	0.97	0.21	-	31,31,31,31	0
2	UNX	B	611	1/1	0.94	0.16	-	29,29,29,29	0
2	UNX	B	604	1/1	0.87	0.44	-	58,58,58,58	0
2	UNX	A	615	1/1	0.87	0.42	-	60,60,60,60	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.