



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 25, 2023 – 02:09 PM JST

PDB ID : 7YMF  
Title : Crystal Structure of DDX3X449\_450ET>DP  
Authors : Xiong, J.  
Deposited on : 2022-07-28  
Resolution : 2.30 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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

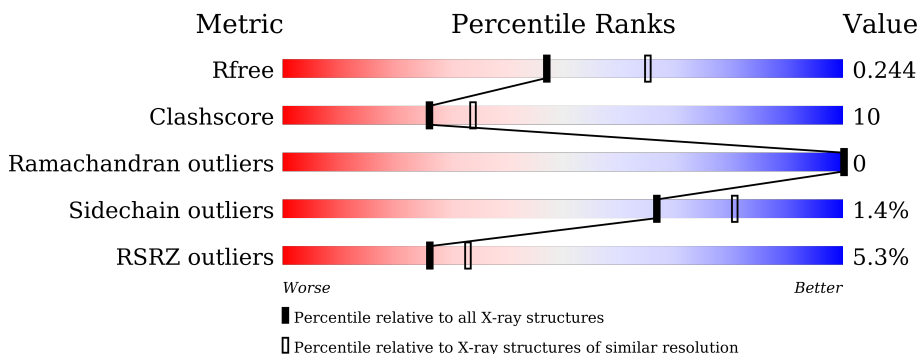
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

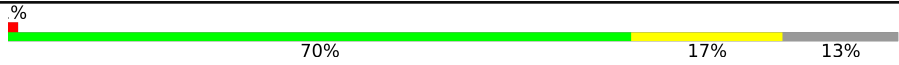

The reported resolution of this entry is 2.30 Å.

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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	303	 10% 70% 17% 13%
2	B	178	 10% 71% 15% 13%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3523 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 ATP-dependent RNA helicase DDX3X.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	263	2104	1333	363	392	16	0	3	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	112	MET	-	initiating methionine	UNP O00571
A	113	GLY	-	expression tag	UNP O00571
A	114	SER	-	expression tag	UNP O00571
A	115	SER	-	expression tag	UNP O00571
A	116	ASP	-	expression tag	UNP O00571
A	117	HIS	-	expression tag	UNP O00571
A	118	HIS	-	expression tag	UNP O00571
A	119	HIS	-	expression tag	UNP O00571
A	120	HIS	-	expression tag	UNP O00571
A	121	HIS	-	expression tag	UNP O00571
A	122	SER	-	expression tag	UNP O00571
A	123	SER	-	expression tag	UNP O00571
A	124	GLY	-	expression tag	UNP O00571
A	125	LEU	-	expression tag	UNP O00571
A	126	VAL	-	expression tag	UNP O00571
A	127	PRO	-	expression tag	UNP O00571
A	128	ARG	-	expression tag	UNP O00571
A	129	GLY	-	expression tag	UNP O00571
A	130	SER	-	expression tag	UNP O00571
A	131	HIS	-	expression tag	UNP O00571
A	132	MET	-	expression tag	UNP O00571

- Molecule 2 is a protein called ATP-dependent RNA helicase DDX3X.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	154	1230	778	211	239	2	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	449	ASP	GLU	engineered mutation	UNP O00571
B	450	PRO	THR	engineered mutation	UNP O00571
B	585	LEU	-	expression tag	UNP O00571
B	586	GLU	-	expression tag	UNP O00571
B	587	HIS	-	expression tag	UNP O00571
B	588	HIS	-	expression tag	UNP O00571
B	589	HIS	-	expression tag	UNP O00571
B	590	HIS	-	expression tag	UNP O00571
B	591	HIS	-	expression tag	UNP O00571
B	592	HIS	-	expression tag	UNP O00571

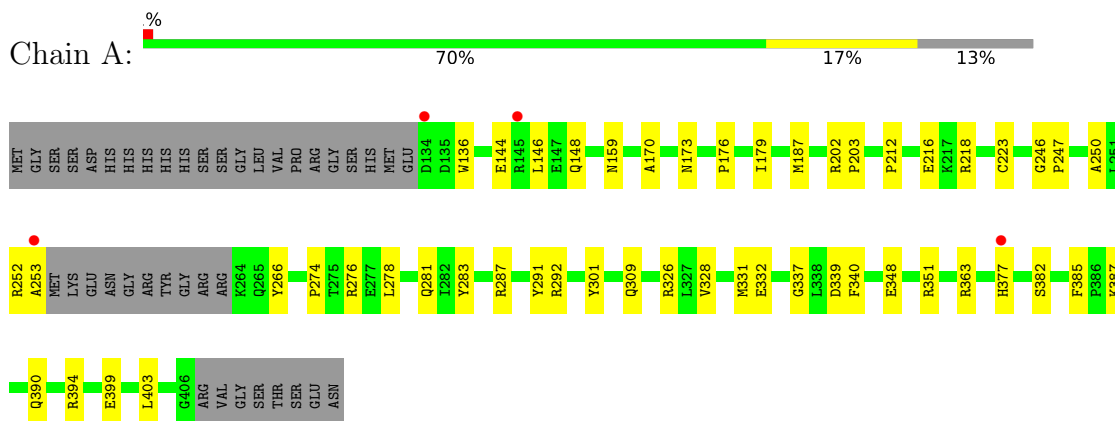
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	118	Total	O	0	0
			118	118		
3	B	71	Total	O	0	0
			71	71		

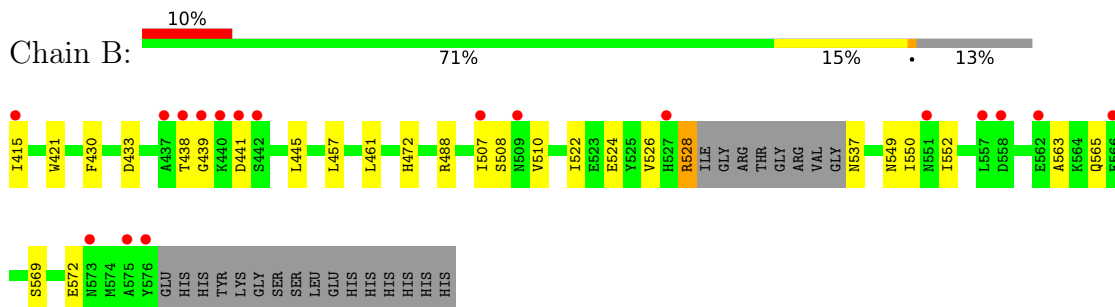
### 3 Residue-property plots [i](#)

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: ATP-dependent RNA helicase DDX3X



- Molecule 2: ATP-dependent RNA helicase DDX3X



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.56Å 106.56Å 116.71Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.32 – 2.30 49.32 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.32-2.30) 99.7 (49.32-2.30)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.15 (at 2.29Å)	Xtrriage
Refinement program	PHENIX 1.18.1_3865	Depositor
R, $R_{free}$	0.185 , 0.244 0.185 , 0.244	Depositor DCC
$R_{free}$ test set	1097 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.5	Xtrriage
Anisotropy	0.052	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 52.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.032 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3523	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.42% 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 i

### 5.1 Standard geometry i

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	0.59	0/2148	0.75	0/2902
2	B	0.61	0/1252	0.70	0/1694
All	All	0.59	0/3400	0.73	0/4596

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 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	2104	0	2097	42	1
2	B	1230	0	1209	29	1
3	A	118	0	0	6	0
3	B	71	0	0	3	0
All	All	3523	0	3306	69	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 10.

All (69) 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)
2:B:507:ILE:CG2	2:B:510:VAL:HG11	1.84	1.08

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:507:ILE:CG2	2:B:510:VAL:CG1	2.39	1.01
2:B:507:ILE:HG22	2:B:510:VAL:CG1	2.02	0.90
2:B:507:ILE:HG22	2:B:510:VAL:HG12	1.53	0.89
1:A:144:GLU:O	1:A:148:GLN:HG2	1.79	0.82
2:B:507:ILE:HG22	2:B:507:ILE:O	1.83	0.77
2:B:507:ILE:HG23	2:B:510:VAL:CG1	2.13	0.77
2:B:507:ILE:O	2:B:510:VAL:HG12	1.86	0.76
1:A:281:GLN:HG2	3:A:557:HOH:O	1.89	0.72
2:B:507:ILE:HG23	2:B:510:VAL:HG11	1.67	0.72
1:A:247:PRO:HG2	1:A:252:ARG:HG3	1.73	0.69
1:A:394:ARG:NH1	3:A:503:HOH:O	2.21	0.68
2:B:433:ASP:OD2	3:B:601:HOH:O	2.12	0.67
1:A:223:CYS:HA	1:A:382:SER:O	1.95	0.66
2:B:549:ASN:O	2:B:552:ILE:HG13	1.95	0.66
1:A:276:ARG:NH1	3:A:505:HOH:O	2.28	0.66
1:A:136:TRP:CE2	1:A:247:PRO:HD3	2.33	0.63
2:B:438:THR:CG2	2:B:441:ASP:HA	2.29	0.61
1:A:179:ILE:HG23	1:A:212:PRO:HG3	1.84	0.60
1:A:176:PRO:HG3	1:A:216:GLU:HG3	1.85	0.59
1:A:387:LYS:HA	1:A:390:GLN:OE1	2.04	0.57
2:B:507:ILE:CG2	2:B:510:VAL:HG12	2.19	0.57
1:A:173:ASN:ND2	1:A:399:GLU:OE2	2.32	0.56
1:A:176:PRO:CG	1:A:216:GLU:HG3	2.35	0.56
2:B:438:THR:HG22	2:B:441:ASP:HA	1.88	0.56
2:B:507:ILE:CG2	2:B:507:ILE:O	2.51	0.56
1:A:266:TYR:OH	2:B:569:SER:HB3	2.07	0.55
2:B:488:ARG:NH1	3:B:603:HOH:O	2.30	0.54
1:A:328:VAL:HA	1:A:331:MET:CE	2.39	0.53
1:A:390:GLN:O	1:A:394:ARG:HG2	2.11	0.51
1:A:328:VAL:HA	1:A:331:MET:HE2	1.93	0.51
1:A:332:GLU:OE2	1:A:363:ARG:NH1	2.45	0.50
1:A:291:TYR:CD2	1:A:292:ARG:HG3	2.46	0.50
2:B:507:ILE:HG21	2:B:510:VAL:HG11	1.86	0.50
2:B:439:GLY:C	2:B:441:ASP:N	2.65	0.49
1:A:266:TYR:HB3	1:A:340:PHE:HD2	1.77	0.49
1:A:348:GLU:OE1	1:A:351:ARG:HD3	2.13	0.48
1:A:283:TYR:CZ	1:A:287:ARG:HD2	2.49	0.48
2:B:445:LEU:CD2	2:B:507:ILE:HD13	2.44	0.48
1:A:385:PHE:CE2	1:A:390:GLN:HG2	2.49	0.47
1:A:385:PHE:CZ	1:A:390:GLN:HG2	2.48	0.47
2:B:522:ILE:O	2:B:526:VAL:HG23	2.14	0.47

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:472:HIS:HA	3:B:634:HOH:O	2.15	0.46
1:A:291:TYR:CE2	1:A:292:ARG:HG3	2.49	0.46
1:A:250:ALA:HB2	2:B:421:TRP:CE2	2.51	0.46
2:B:524:GLU:OE2	2:B:528:ARG:NH2	2.47	0.46
1:A:387:LYS:HD2	1:A:390:GLN:OE1	2.16	0.45
1:A:246:GLY:HA3	1:A:252:ARG:NH2	2.32	0.45
2:B:550:ILE:HD12	2:B:550:ILE:HA	1.85	0.45
1:A:252:ARG:O	1:A:253:ALA:HB2	2.16	0.45
1:A:274:PRO:HD2	1:A:278:LEU:HD23	1.98	0.44
2:B:457:LEU:HD12	2:B:457:LEU:HA	1.77	0.44
1:A:187:MET:HB2	1:A:187:MET:HE2	1.78	0.43
1:A:281:GLN:OE1	3:A:502:HOH:O	2.20	0.43
2:B:439:GLY:O	2:B:441:ASP:N	2.52	0.43
1:A:309:GLN:OE1	1:A:326:ARG:NH1	2.48	0.43
1:A:170:ALA:HB2	1:A:403[B]:LEU:CD2	2.49	0.42
1:A:218:ARG:O	1:A:377[B]:HIS:CE1	2.74	0.41
1:A:276:ARG:HG3	1:A:301:TYR:CD1	2.54	0.41
2:B:461:LEU:HD23	2:B:461:LEU:HA	1.84	0.41
1:A:266:TYR:CE1	1:A:337:GLY:HA3	2.54	0.41
1:A:144:GLU:HG3	1:A:148:GLN:NE2	2.35	0.41
2:B:508:SER:O	2:B:537:ASN:HB3	2.21	0.41
1:A:159:ASN:ND2	3:A:506:HOH:O	2.36	0.41
1:A:202:ARG:HA	1:A:203:PRO:HD3	1.97	0.41
1:A:218:ARG:O	1:A:377[B]:HIS:HE1	2.03	0.41
1:A:276:ARG:HD3	3:A:505:HOH:O	2.22	0.40
1:A:399:GLU:HG3	1:A:399:GLU:O	2.22	0.40
2:B:563:ALA:HB3	2:B:565:GLN:HG2	2.03	0.40

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:A:387:LYS:NZ	2:B:415:ILE:N[3_555]	2.07	0.13

## 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	262/303 (86%)	259 (99%)	3 (1%)	0	100	100
2	B	150/178 (84%)	147 (98%)	3 (2%)	0	100	100
All	All	412/481 (86%)	406 (98%)	6 (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	229/260 (88%)	227 (99%)	2 (1%)	78	89
2	B	136/156 (87%)	133 (98%)	3 (2%)	52	69
All	All	365/416 (88%)	360 (99%)	5 (1%)	67	81

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	146	LEU
1	A	339	ASP
2	B	430	PHE
2	B	528	ARG
2	B	572	GLU

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

Mol	Chain	Res	Type
1	A	241	GLN
1	A	281	GLN
2	B	573	ASN

### 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](#)

There are no ligands in this entry.

### 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	263/303 (86%)	-0.12	4 (1%) 73 79	13, 29, 50, 74	0
2	B	154/178 (86%)	0.29	18 (11%) 4 6	17, 33, 66, 90	0
All	All	417/481 (86%)	0.03	22 (5%) 26 33	13, 30, 60, 90	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	415	ILE	6.3
2	B	576	TYR	6.3
2	B	575	ALA	5.9
2	B	440	LYS	5.7
2	B	441	ASP	4.8
2	B	439	GLY	4.3
2	B	557	LEU	3.6
2	B	507	ILE	3.0
2	B	566	GLU	2.8
2	B	573	ASN	2.7
1	A	145	ARG	2.6
2	B	562	GLU	2.5
2	B	509	ASN	2.5
2	B	551	ASN	2.5
2	B	527	HIS	2.4
1	A	134	ASP	2.3
1	A	253	ALA	2.3
2	B	438	THR	2.2
1	A	377[A]	HIS	2.2
2	B	442	SER	2.2
2	B	558	ASP	2.2
2	B	437	ALA	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 monosaccharides in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.