



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 08:27 PM GMT

PDB ID : 4YKB  
Title : Structure of GUN4 from Chlamydomonas reinhardtii  
Authors : Tabriz, S.T.; Langley, D.B.; Willows, R.D.; Duff, A.P.; Harrop, S.J.  
Deposited on : 2015-03-04  
Resolution : 3.50 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 : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

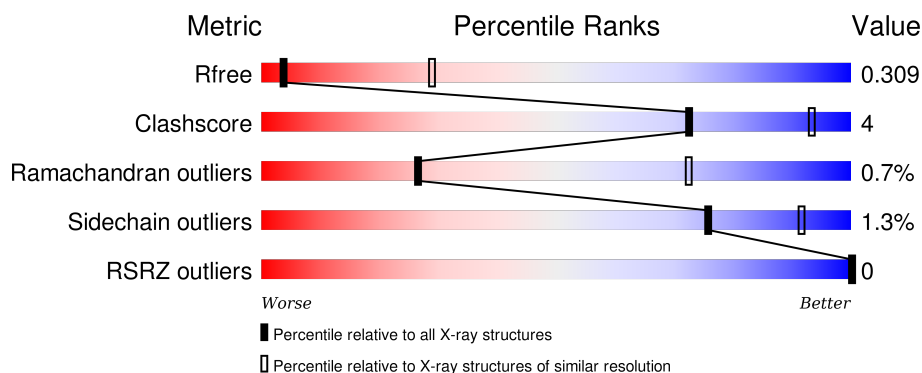
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)

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	229	
1	B	229	
1	C	229	
1	D	229	
1	E	229	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	229	 A horizontal bar chart showing the quality of chain F. The bar is divided into three segments: a green segment representing 56%, a yellow segment representing 5%, and a grey segment representing 38%. The percentages are labeled below the bar.

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 6241 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 Tetrapyrrole-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	151	Total	C	N	O	S	0	0	0
			1089	704	181	201	3			
1	B	147	Total	C	N	O	S	0	0	0
			1030	665	174	188	3			
1	C	146	Total	C	N	O	S	0	0	0
			1044	673	178	190	3			
1	D	148	Total	C	N	O	S	0	0	0
			1034	669	171	192	2			
1	E	145	Total	C	N	O	S	0	0	0
			1054	684	175	193	2			
1	F	141	Total	C	N	O	S	0	0	0
			990	642	163	183	2			

There are 126 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	32	MET	-	initiating methionine	UNP A8I5N5
A	33	GLY	-	expression tag	UNP A8I5N5
A	34	SER	-	expression tag	UNP A8I5N5
A	35	SER	-	expression tag	UNP A8I5N5
A	36	HIS	-	expression tag	UNP A8I5N5
A	37	HIS	-	expression tag	UNP A8I5N5
A	38	HIS	-	expression tag	UNP A8I5N5
A	39	HIS	-	expression tag	UNP A8I5N5
A	40	HIS	-	expression tag	UNP A8I5N5
A	41	HIS	-	expression tag	UNP A8I5N5
A	42	SER	-	expression tag	UNP A8I5N5
A	43	SER	-	expression tag	UNP A8I5N5
A	44	GLY	-	expression tag	UNP A8I5N5
A	45	LEU	-	expression tag	UNP A8I5N5
A	46	VAL	-	expression tag	UNP A8I5N5
A	47	PRO	-	expression tag	UNP A8I5N5
A	48	ARG	-	expression tag	UNP A8I5N5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	49	GLY	-	expression tag	UNP A8I5N5
A	50	SER	-	expression tag	UNP A8I5N5
A	51	HIS	-	expression tag	UNP A8I5N5
A	52	MET	-	expression tag	UNP A8I5N5
B	32	MET	-	initiating methionine	UNP A8I5N5
B	33	GLY	-	expression tag	UNP A8I5N5
B	34	SER	-	expression tag	UNP A8I5N5
B	35	SER	-	expression tag	UNP A8I5N5
B	36	HIS	-	expression tag	UNP A8I5N5
B	37	HIS	-	expression tag	UNP A8I5N5
B	38	HIS	-	expression tag	UNP A8I5N5
B	39	HIS	-	expression tag	UNP A8I5N5
B	40	HIS	-	expression tag	UNP A8I5N5
B	41	HIS	-	expression tag	UNP A8I5N5
B	42	SER	-	expression tag	UNP A8I5N5
B	43	SER	-	expression tag	UNP A8I5N5
B	44	GLY	-	expression tag	UNP A8I5N5
B	45	LEU	-	expression tag	UNP A8I5N5
B	46	VAL	-	expression tag	UNP A8I5N5
B	47	PRO	-	expression tag	UNP A8I5N5
B	48	ARG	-	expression tag	UNP A8I5N5
B	49	GLY	-	expression tag	UNP A8I5N5
B	50	SER	-	expression tag	UNP A8I5N5
B	51	HIS	-	expression tag	UNP A8I5N5
B	52	MET	-	expression tag	UNP A8I5N5
C	32	MET	-	initiating methionine	UNP A8I5N5
C	33	GLY	-	expression tag	UNP A8I5N5
C	34	SER	-	expression tag	UNP A8I5N5
C	35	SER	-	expression tag	UNP A8I5N5
C	36	HIS	-	expression tag	UNP A8I5N5
C	37	HIS	-	expression tag	UNP A8I5N5
C	38	HIS	-	expression tag	UNP A8I5N5
C	39	HIS	-	expression tag	UNP A8I5N5
C	40	HIS	-	expression tag	UNP A8I5N5
C	41	HIS	-	expression tag	UNP A8I5N5
C	42	SER	-	expression tag	UNP A8I5N5
C	43	SER	-	expression tag	UNP A8I5N5
C	44	GLY	-	expression tag	UNP A8I5N5
C	45	LEU	-	expression tag	UNP A8I5N5
C	46	VAL	-	expression tag	UNP A8I5N5
C	47	PRO	-	expression tag	UNP A8I5N5
C	48	ARG	-	expression tag	UNP A8I5N5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	49	GLY	-	expression tag	UNP A8I5N5
C	50	SER	-	expression tag	UNP A8I5N5
C	51	HIS	-	expression tag	UNP A8I5N5
C	52	MET	-	expression tag	UNP A8I5N5
D	32	MET	-	initiating methionine	UNP A8I5N5
D	33	GLY	-	expression tag	UNP A8I5N5
D	34	SER	-	expression tag	UNP A8I5N5
D	35	SER	-	expression tag	UNP A8I5N5
D	36	HIS	-	expression tag	UNP A8I5N5
D	37	HIS	-	expression tag	UNP A8I5N5
D	38	HIS	-	expression tag	UNP A8I5N5
D	39	HIS	-	expression tag	UNP A8I5N5
D	40	HIS	-	expression tag	UNP A8I5N5
D	41	HIS	-	expression tag	UNP A8I5N5
D	42	SER	-	expression tag	UNP A8I5N5
D	43	SER	-	expression tag	UNP A8I5N5
D	44	GLY	-	expression tag	UNP A8I5N5
D	45	LEU	-	expression tag	UNP A8I5N5
D	46	VAL	-	expression tag	UNP A8I5N5
D	47	PRO	-	expression tag	UNP A8I5N5
D	48	ARG	-	expression tag	UNP A8I5N5
D	49	GLY	-	expression tag	UNP A8I5N5
D	50	SER	-	expression tag	UNP A8I5N5
D	51	HIS	-	expression tag	UNP A8I5N5
D	52	MET	-	expression tag	UNP A8I5N5
E	32	MET	-	initiating methionine	UNP A8I5N5
E	33	GLY	-	expression tag	UNP A8I5N5
E	34	SER	-	expression tag	UNP A8I5N5
E	35	SER	-	expression tag	UNP A8I5N5
E	36	HIS	-	expression tag	UNP A8I5N5
E	37	HIS	-	expression tag	UNP A8I5N5
E	38	HIS	-	expression tag	UNP A8I5N5
E	39	HIS	-	expression tag	UNP A8I5N5
E	40	HIS	-	expression tag	UNP A8I5N5
E	41	HIS	-	expression tag	UNP A8I5N5
E	42	SER	-	expression tag	UNP A8I5N5
E	43	SER	-	expression tag	UNP A8I5N5
E	44	GLY	-	expression tag	UNP A8I5N5
E	45	LEU	-	expression tag	UNP A8I5N5
E	46	VAL	-	expression tag	UNP A8I5N5
E	47	PRO	-	expression tag	UNP A8I5N5
E	48	ARG	-	expression tag	UNP A8I5N5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	49	GLY	-	expression tag	UNP A8I5N5
E	50	SER	-	expression tag	UNP A8I5N5
E	51	HIS	-	expression tag	UNP A8I5N5
E	52	MET	-	expression tag	UNP A8I5N5
F	32	MET	-	initiating methionine	UNP A8I5N5
F	33	GLY	-	expression tag	UNP A8I5N5
F	34	SER	-	expression tag	UNP A8I5N5
F	35	SER	-	expression tag	UNP A8I5N5
F	36	HIS	-	expression tag	UNP A8I5N5
F	37	HIS	-	expression tag	UNP A8I5N5
F	38	HIS	-	expression tag	UNP A8I5N5
F	39	HIS	-	expression tag	UNP A8I5N5
F	40	HIS	-	expression tag	UNP A8I5N5
F	41	HIS	-	expression tag	UNP A8I5N5
F	42	SER	-	expression tag	UNP A8I5N5
F	43	SER	-	expression tag	UNP A8I5N5
F	44	GLY	-	expression tag	UNP A8I5N5
F	45	LEU	-	expression tag	UNP A8I5N5
F	46	VAL	-	expression tag	UNP A8I5N5
F	47	PRO	-	expression tag	UNP A8I5N5
F	48	ARG	-	expression tag	UNP A8I5N5
F	49	GLY	-	expression tag	UNP A8I5N5
F	50	SER	-	expression tag	UNP A8I5N5
F	51	HIS	-	expression tag	UNP A8I5N5
F	52	MET	-	expression tag	UNP A8I5N5



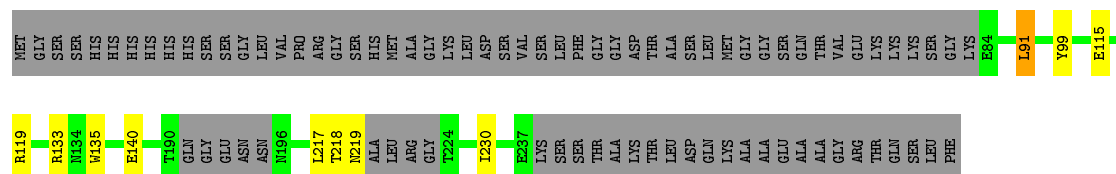
- Molecule 1: Tetrapyrrole-binding protein





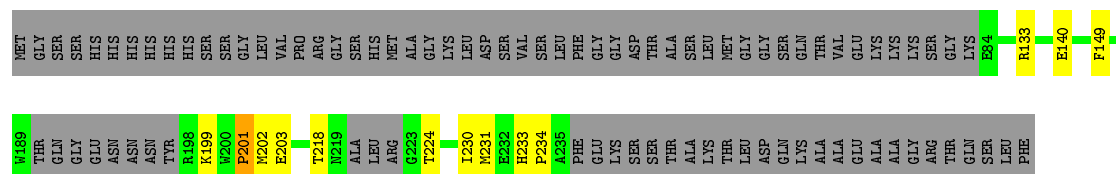
- Molecule 1: Tetrapyrrole-binding protein

Chain E:  59% 1% 37%



- Molecule 1: Tetrapyrrole-binding protein

Chain F:  56% 5% 38%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.00Å 115.00Å 141.19Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.50 – 3.50 42.55 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.0 (44.50-3.50) 99.1 (42.55-3.50)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.78 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, $R_{free}$	0.256 , 0.313 0.255 , 0.309	Depositor DCC
$R_{free}$ test set	752 reflections (5.71%)	DCC
Wilson B-factor (Å <sup>2</sup> )	79.9	Xtriage
Anisotropy	0.003	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 89.5	EDS
Estimated twinning fraction	0.049 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	1 of 13941 reflections (0.007%)	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	6241	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.14% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.56	0/1118	0.76	2/1534 (0.1%)
1	B	0.54	0/1055	0.79	3/1453 (0.2%)
1	C	0.54	0/1069	0.77	1/1468 (0.1%)
1	D	0.52	0/1062	0.70	1/1467 (0.1%)
1	E	0.56	0/1084	0.77	2/1492 (0.1%)
1	F	0.56	0/1018	0.76	2/1405 (0.1%)
All	All	0.54	0/6406	0.76	11/8819 (0.1%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	187	ILE	CG1-CB-CG2	-9.54	90.42	111.40
1	E	91	LEU	CA-CB-CG	6.87	131.10	115.30
1	A	119	ARG	NE-CZ-NH1	6.74	123.67	120.30
1	C	208	MET	CG-SD-CE	6.66	110.86	100.20
1	E	119	ARG	NE-CZ-NH1	6.17	123.38	120.30

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	1089	0	923	15	0
1	B	1030	0	875	9	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1044	0	900	7	0
1	D	1034	0	855	8	0
1	E	1054	0	887	12	0
1	F	990	0	808	6	0
All	All	6241	0	5248	48	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 4.

The worst 5 of 48 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:C:115:GLU:OE2	1:C:119:ARG:NH1	1.79	1.14
1:D:189:TRP:CZ2	1:D:217:LEU:HD12	2.35	0.61
1:B:115:GLU:OE2	1:B:217:LEU:HG	2.02	0.59
1:A:137:TYR:CZ	1:E:135:TRP:CH2	2.93	0.57
1:B:218:THR:HG21	1:B:230:ILE:HD12	1.86	0.57

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	145/229 (63%)	132 (91%)	13 (9%)	0	100	100
1	B	141/229 (62%)	133 (94%)	8 (6%)	0	100	100
1	C	140/229 (61%)	130 (93%)	8 (6%)	2 (1%)	14	58
1	D	144/229 (63%)	132 (92%)	10 (7%)	2 (1%)	14	58
1	E	139/229 (61%)	130 (94%)	9 (6%)	0	100	100
1	F	135/229 (59%)	124 (92%)	9 (7%)	2 (2%)	13	56

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	844/1374 (61%)	781 (92%)	57 (7%)	6 (1%)	26 72

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	198	ARG
1	D	219	ASN
1	D	224	THR
1	C	218	THR
1	F	224	THR

### 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	93/193 (48%)	92 (99%)	1 (1%)	80 92
1	B	86/193 (45%)	85 (99%)	1 (1%)	78 92
1	C	90/193 (47%)	88 (98%)	2 (2%)	60 85
1	D	85/193 (44%)	83 (98%)	2 (2%)	57 85
1	E	90/193 (47%)	90 (100%)	0	100 100
1	F	80/193 (42%)	79 (99%)	1 (1%)	76 91
All	All	524/1158 (45%)	517 (99%)	7 (1%)	76 91

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	219	ASN
1	F	149	PHE
1	D	149	PHE
1	B	149	PHE
1	D	217	LEU

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](#)

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	151/229 (65%)	-0.47	0 100 100	40, 61, 85, 97	0
1	B	147/229 (64%)	-0.50	0 100 100	44, 64, 82, 93	0
1	C	146/229 (63%)	-0.55	0 100 100	35, 61, 78, 94	0
1	D	148/229 (64%)	-0.41	0 100 100	36, 61, 81, 102	0
1	E	145/229 (63%)	-0.51	0 100 100	39, 62, 79, 108	0
1	F	141/229 (61%)	-0.27	0 100 100	40, 66, 94, 104	0
All	All	878/1374 (63%)	-0.45	0 100 100	35, 63, 84, 108	0

There are no RSRZ outliers to report.

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.