



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 02:23 pm GMT

PDB ID : 3T60
Title : 5'-Diphenyl Nucleoside Inhibitors of Plasmodium falciparum dUTPase
Authors : Hampton, S.E.; Baragana, B.; Schipani, A.; Bosch-Navarrete, C.; Musso-Buendia, A.; Recio, E.; Kaiser, M.; Whittingham, J.L.; Roberts, S.M.; Shevtsov, M.; Brannigan, J.A.; Kahnberg, P.; Brun, R.; Wilson, K.S.; Gonzalez-Pacanowska, D.; Johansson, N.G.; Gilbert, I.H.
Deposited on : 2011-07-28
Resolution : 2.40 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

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	:	trunk28620
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)	:	recalc28949

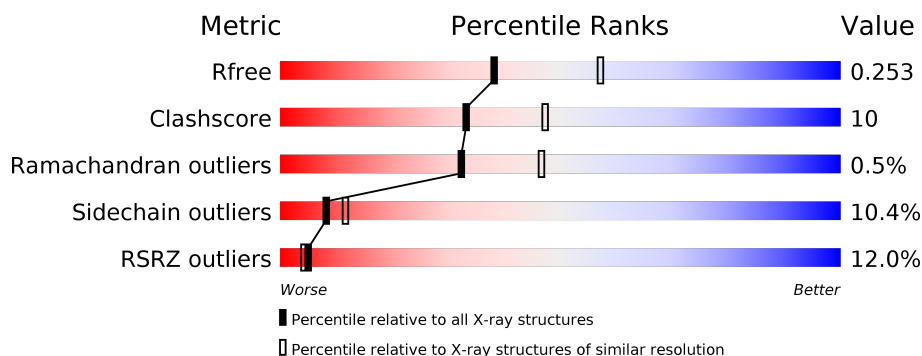
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

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 ≥ 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	181	<div> <div>6%</div> <div> <div>55%</div> <div>19%</div> <div>• •</div> <div>21%</div> </div> </div>
1	B	181	<div> <div>15%</div> <div> <div>47%</div> <div>20%</div> <div>• •</div> <div>28%</div> </div> </div>
1	C	181	<div> <div>7%</div> <div> <div>56%</div> <div>20%</div> <div>•</div> <div>20%</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	DUA	B	182	X	-	-	-
2	DUA	C	182	X	-	-	-
3	GOL	A	183	-	-	X	-
3	GOL	B	183	-	-	X	X
3	GOL	C	183	-	-	X	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3545 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 Deoxyuridine 5'-triphosphate nucleotidohydrolase, putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	143	Total	C	N	O	S	118	0	0
			1128	722	184	219	3			
1	B	131	Total	C	N	O	S	87	1	0
			1049	678	170	198	3			
1	C	145	Total	C	N	O	S	62	0	0
			1149	737	187	222	3			

There are 24 discrepancies between the modelled and reference sequences:

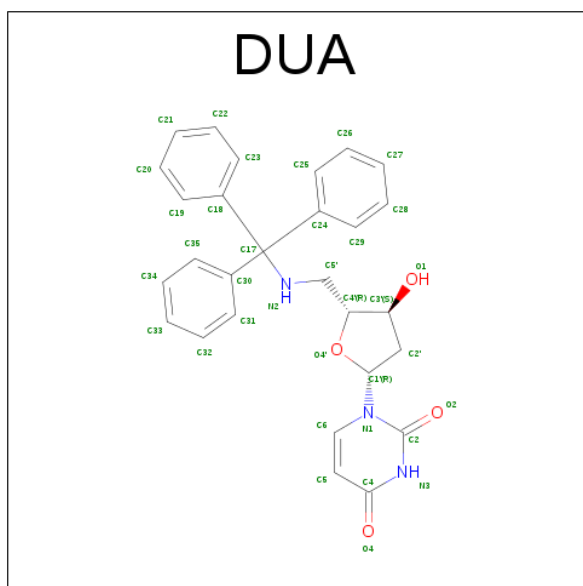
Chain	Residue	Modelled	Actual	Comment	Reference
A	174	LEU	-	EXPRESSION TAG	UNP Q8II92
A	175	GLU	-	EXPRESSION TAG	UNP Q8II92
A	176	HIS	-	EXPRESSION TAG	UNP Q8II92
A	177	HIS	-	EXPRESSION TAG	UNP Q8II92
A	178	HIS	-	EXPRESSION TAG	UNP Q8II92
A	179	HIS	-	EXPRESSION TAG	UNP Q8II92
A	180	HIS	-	EXPRESSION TAG	UNP Q8II92
A	181	HIS	-	EXPRESSION TAG	UNP Q8II92
B	174	LEU	-	EXPRESSION TAG	UNP Q8II92
B	175	GLU	-	EXPRESSION TAG	UNP Q8II92
B	176	HIS	-	EXPRESSION TAG	UNP Q8II92
B	177	HIS	-	EXPRESSION TAG	UNP Q8II92
B	178	HIS	-	EXPRESSION TAG	UNP Q8II92
B	179	HIS	-	EXPRESSION TAG	UNP Q8II92
B	180	HIS	-	EXPRESSION TAG	UNP Q8II92
B	181	HIS	-	EXPRESSION TAG	UNP Q8II92
C	174	LEU	-	EXPRESSION TAG	UNP Q8II92
C	175	GLU	-	EXPRESSION TAG	UNP Q8II92
C	176	HIS	-	EXPRESSION TAG	UNP Q8II92
C	177	HIS	-	EXPRESSION TAG	UNP Q8II92
C	178	HIS	-	EXPRESSION TAG	UNP Q8II92
C	179	HIS	-	EXPRESSION TAG	UNP Q8II92
C	180	HIS	-	EXPRESSION TAG	UNP Q8II92

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	181	HIS	-	EXPRESSION TAG	UNP Q8II92

- Molecule 2 is 2',5'-DIDEOXY-5'-(TRITYLAMINO)URIDINE (three-letter code: DUA) (formula: $C_{28}H_{27}N_3O_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			35	28	3	4		
2	B	1	Total	C	N	O	0	0
			35	28	3	4		
2	C	1	Total	C	N	O	0	0
			35	28	3	4		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	38	Total	O	0	0
			38	38		
4	B	23	Total	O	0	0
			23	23		
4	C	35	Total	O	0	0
			35	35		

4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α , β , γ	76.97Å 76.97Å 106.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.40 38.49 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.3 (40.00-2.40) 99.3 (38.49-2.40)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.10 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.3.0037	Depositor
R, R_{free}	0.196 , 0.254 0.194 , 0.253	Depositor DCC
R_{free} test set	1244 reflections (5.39%)	DCC
Wilson B-factor (Å ²)	42.5	Xtriage
Anisotropy	0.164	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 51.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.052 for h,-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	3545	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.61% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: GOL, DUA

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	1.49	13/1144 (1.1%)	1.49	21/1538 (1.4%)
1	B	1.58	12/1070 (1.1%)	1.15	9/1440 (0.6%)
1	C	1.67	9/1166 (0.8%)	1.45	19/1567 (1.2%)
All	All	1.58	34/3380 (1.0%)	1.38	49/4545 (1.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	1	3
1	C	0	1
All	All	1	4

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	166	GLY	CA-C	-27.09	1.08	1.51
1	B	127	GLU	CD-OE2	18.62	1.46	1.25
1	B	127	GLU	CD-OE1	-17.06	1.06	1.25
1	A	166	GLY	CA-C	15.62	1.76	1.51
1	C	80	ASN	C-N	-15.08	0.99	1.34

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	17	LYS	O-C-N	-18.35	93.34	122.70
1	A	63	TYR	N-CA-CB	16.29	139.93	110.60
1	A	80	ASN	C-N-CA	14.30	157.46	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	42	LYS	CD-CE-NZ	13.32	142.35	111.70
1	A	80	ASN	O-C-N	-11.66	104.04	122.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	133	ASN	CA

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	127	GLU	Sidechain
1	B	26	ASP	Mainchain,Peptide
1	C	17	LYS	Mainchain

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	1128	0	1144	20	0
1	B	1049	0	1082	24	0
1	C	1149	0	1166	16	2
2	A	35	0	26	3	0
2	B	35	0	26	0	0
2	C	35	0	26	2	0
3	A	6	0	8	6	0
3	B	6	0	8	4	0
3	C	6	0	8	5	0
4	A	38	0	0	0	0
4	B	23	0	0	0	0
4	C	35	0	0	2	2
All	All	3545	0	3494	62	2

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.

The worst 5 of 62 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:92:SER:HB2	3:C:183:GOL:H32	1.51	0.92
1:B:88:LEU:HD11	1:B:137:VAL:HG13	1.57	0.86
1:B:31:LEU:HD21	1:B:139:LEU:HD22	1.65	0.78
1:B:8:LEU:HD21	1:B:52:LYS:HE3	1.67	0.77
1:B:143:THR:OG1	1:B:145:GLU:HG2	1.89	0.73

All (2) 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:C:131:LYS:NZ	4:C:218:HOH:O[4_564]	1.02	1.18
1:C:131:LYS:CE	4:C:218:HOH:O[4_564]	1.97	0.23

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	137/181 (76%)	132 (96%)	4 (3%)	1 (1%)	25	37
1	B	126/181 (70%)	123 (98%)	2 (2%)	1 (1%)	22	33
1	C	139/181 (77%)	134 (96%)	5 (4%)	0	100	100
All	All	402/543 (74%)	389 (97%)	11 (3%)	2 (0%)	32	46

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	104	SER
1	B	133	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	127/164 (77%)	115 (91%)	12 (9%)	10	15
1	B	120/164 (73%)	103 (86%)	17 (14%)	4	4
1	C	129/164 (79%)	119 (92%)	10 (8%)	15	23
All	All	376/492 (76%)	337 (90%)	39 (10%)	8	12

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	80	ASN
1	B	124	SER
1	C	139	LEU
1	B	81	ILE
1	B	84	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	138	GLN
1	B	57	GLN
1	B	126	GLN
1	C	138	GLN

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

6 ligands are modelled in this entry.

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	DUA	A	182	-	34,39,39	1.35	2 (5%)	44,55,55	2.93	19 (43%)
3	GOL	A	183	-	5,5,5	0.82	0	5,5,5	1.72	1 (20%)
2	DUA	B	182	-	34,39,39	1.30	5 (14%)	44,55,55	3.21	11 (25%)
3	GOL	B	183	-	5,5,5	0.91	0	5,5,5	0.38	0
2	DUA	C	182	-	34,39,39	1.53	7 (20%)	44,55,55	3.44	18 (40%)
3	GOL	C	183	-	5,5,5	0.87	0	5,5,5	1.08	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	DUA	A	182	-	-	0/24/40/40	0/5/5/5
3	GOL	A	183	-	-	0/4/4/4	0/0/0/0
2	DUA	B	182	-	1/1/4/5	0/24/40/40	0/5/5/5
3	GOL	B	183	-	-	0/4/4/4	0/0/0/0
2	DUA	C	182	-	1/1/4/5	0/24/40/40	0/5/5/5
3	GOL	C	183	-	-	0/4/4/4	0/0/0/0

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	182	DUA	C17-C24	-3.78	1.48	1.54
2	C	182	DUA	C2-N3	-3.17	1.31	1.38
2	C	182	DUA	C17-C24	-2.85	1.49	1.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	182	DUA	C17-C30	-2.59	1.50	1.54
2	C	182	DUA	C35-C30	2.02	1.42	1.39

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	182	DUA	C18-C17-N2	-7.51	97.03	108.72
2	C	182	DUA	C18-C17-N2	-7.24	97.46	108.72
2	B	182	DUA	C3'-C2'-C1'	-6.28	86.37	102.48
2	B	182	DUA	C18-C17-N2	-6.12	99.21	108.72
2	C	182	DUA	C3'-C2'-C1'	-5.55	88.24	102.48

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	182	DUA	C1'
2	B	182	DUA	C1'

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	182	DUA	3	0
3	A	183	GOL	6	0
3	B	183	GOL	4	0
2	C	182	DUA	2	0
3	C	183	GOL	5	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, 95th 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 ²)	Q<0.9
1	A	137/181 (75%)	0.11	10 (7%) 16 14	22, 32, 47, 57	18 (13%)
1	B	130/181 (71%)	1.00	27 (20%) 1 1	25, 44, 65, 70	21 (16%)
1	C	143/181 (79%)	0.24	12 (8%) 12 10	23, 32, 52, 68	15 (10%)
All	All	410/543 (75%)	0.44	49 (11%) 5 4	22, 35, 61, 70	54 (13%)

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	63	TYR	5.6
1	B	64	TYR	4.6
1	B	129	HIS	4.6
1	B	81	ILE	4.5
1	B	58	TYR	4.3

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, 95th 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 ²)	Q<0.9
3	GOL	C	183	6/6	0.96	0.30	2.29	44,48,50,54	0
3	GOL	B	183	6/6	0.92	0.30	2.20	57,58,60,62	0
3	GOL	A	183	6/6	0.96	0.22	1.49	44,46,48,51	0
2	DUA	A	182	35/35	0.91	0.17	-0.03	33,39,48,50	0
2	DUA	C	182	35/35	0.95	0.15	-0.22	25,31,36,37	0
2	DUA	B	182	35/35	0.95	0.15	-0.72	25,38,43,46	0

6.5 Other polymers [i](#)

There are no such residues in this entry.