



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 28, 2017 – 02:54 PM EDT

PDB ID : 2NZX  
Title : Crystal Structure of alpha1,3-Fucosyltransferase with GDP  
Authors : Sun, H.Y.; Ko, T.P.  
Deposited on : 2006-11-27  
Resolution : 1.90 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-20029077
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-20029077

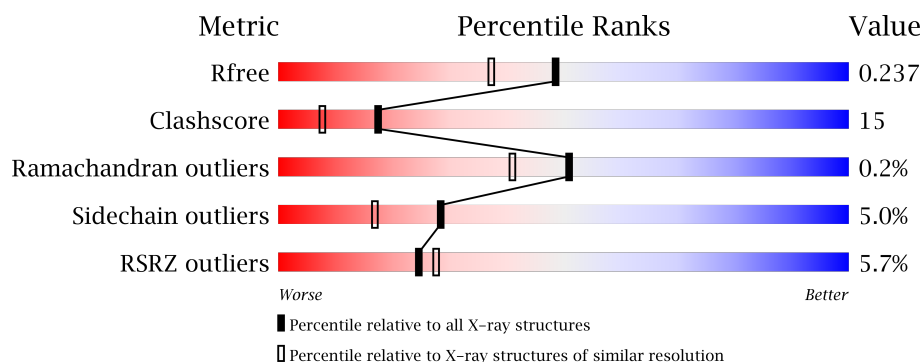
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	5047 (1.90-1.90)
Clashscore	112137	5731 (1.90-1.90)
Ramachandran outliers	110173	5669 (1.90-1.90)
Sidechain outliers	110143	5670 (1.90-1.90)
RSRZ outliers	101464	5100 (1.90-1.90)

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	371	<div> <div>5%</div> <div> <div></div> <div>68%</div> <div>21%</div> <div>• 8%</div> </div> </div>
1	B	371	<div> <div>6%</div> <div> <div></div> <div>66%</div> <div>22%</div> <div>• 9%</div> </div> </div>
1	C	371	<div> <div>4%</div> <div> <div></div> <div>74%</div> <div>17%</div> <div>• 6%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10752 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 Alpha1,3-fucosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	340	Total	C	N	O	S	0	0	0
			2795	1811	454	523	7			
1	B	338	Total	C	N	O	S	0	0	0
			2776	1797	452	520	7			
1	C	350	Total	C	N	O	S	0	0	0
			2871	1861	468	535	7			

There are 24 discrepancies between the modelled and reference sequences:

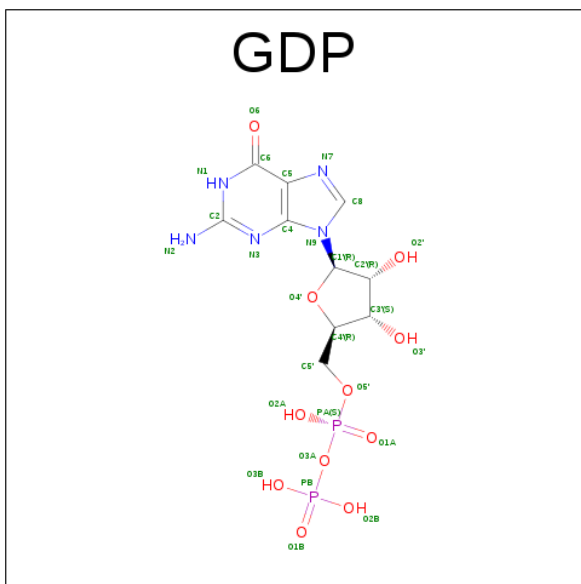
Chain	Residue	Modelled	Actual	Comment	Reference
A	364	LEU	-	EXPRESSION TAG	UNP O30511
A	365	GLU	-	EXPRESSION TAG	UNP O30511
A	366	HIS	-	EXPRESSION TAG	UNP O30511
A	367	HIS	-	EXPRESSION TAG	UNP O30511
A	368	HIS	-	EXPRESSION TAG	UNP O30511
A	369	HIS	-	EXPRESSION TAG	UNP O30511
A	370	HIS	-	EXPRESSION TAG	UNP O30511
A	371	HIS	-	EXPRESSION TAG	UNP O30511
B	364	LEU	-	EXPRESSION TAG	UNP O30511
B	365	GLU	-	EXPRESSION TAG	UNP O30511
B	366	HIS	-	EXPRESSION TAG	UNP O30511
B	367	HIS	-	EXPRESSION TAG	UNP O30511
B	368	HIS	-	EXPRESSION TAG	UNP O30511
B	369	HIS	-	EXPRESSION TAG	UNP O30511
B	370	HIS	-	EXPRESSION TAG	UNP O30511
B	371	HIS	-	EXPRESSION TAG	UNP O30511
C	364	LEU	-	EXPRESSION TAG	UNP O30511
C	365	GLU	-	EXPRESSION TAG	UNP O30511
C	366	HIS	-	EXPRESSION TAG	UNP O30511
C	367	HIS	-	EXPRESSION TAG	UNP O30511
C	368	HIS	-	EXPRESSION TAG	UNP O30511
C	369	HIS	-	EXPRESSION TAG	UNP O30511
C	370	HIS	-	EXPRESSION TAG	UNP O30511

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	371	HIS	-	EXPRESSION TAG	UNP O30511

- Molecule 2 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula:  $C_{10}H_{15}N_5O_{11}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
2	B	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
2	C	1	Total	C	N	O	P	0	0
			28	10	5	11	2		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).

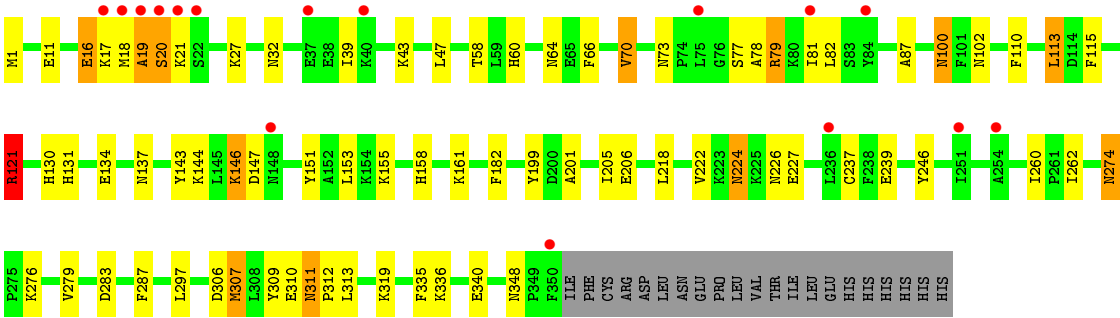


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	797	Total	O	0	0
			797	797		
4	B	650	Total	O	0	0
			650	650		
4	C	774	Total	O	0	0
			774	774		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.15Å 136.25Å 96.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.90 27.81 – 1.90	Depositor EDS
% Data completeness (in resolution range)	89.2 (30.00-1.90) 89.1 (27.81-1.90)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.15 (at 1.89Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.190 , 0.237 0.190 , 0.237	Depositor DCC
$R_{free}$ test set	4859 reflections (5.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	28.4	Xtriage
Anisotropy	0.151	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 64.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10752	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.87% 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: GDP, SO4

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	0.88	0/2880	0.85	5/3910 (0.1%)
1	B	0.94	0/2859	0.91	3/3880 (0.1%)
1	C	1.06	4/2958 (0.1%)	1.18	12/4015 (0.3%)
All	All	0.96	4/8697 (0.0%)	0.99	20/11805 (0.2%)

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	A	0	2
1	B	0	2
All	All	0	4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	237	CYS	CB-SG	-8.70	1.67	1.82
1	C	121	ARG	CG-CD	6.52	1.68	1.51
1	C	307	MET	SD-CE	-6.10	1.43	1.77
1	C	239	GLU	CB-CG	5.29	1.62	1.52

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	121	ARG	NE-CZ-NH1	-32.17	104.22	120.30
1	C	121	ARG	NE-CZ-NH2	28.98	134.79	120.30
1	C	121	ARG	CD-NE-CZ	11.88	140.23	123.60
1	A	113	LEU	CA-CB-CG	7.93	133.54	115.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	113	LEU	CA-CB-CG	7.82	133.28	115.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	119	TYR	Sidechain
1	A	92	TYR	Sidechain
1	B	263	TYR	Sidechain
1	B	9	TYR	Sidechain

## 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	2795	0	2669	86	0
1	B	2776	0	2653	98	0
1	C	2871	0	2756	77	0
2	A	28	0	12	0	0
2	B	28	0	12	1	0
2	C	28	0	12	0	0
3	C	5	0	0	0	0
4	A	797	0	0	22	4
4	B	650	0	0	28	5
4	C	774	0	0	25	3
All	All	10752	0	8114	257	8

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 15.

The worst 5 of 257 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:260:ILE:CG2	1:C:307:MET:HE3	1.81	1.10
1:C:260:ILE:HG21	1:C:307:MET:HE3	1.21	1.09

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:287:PHE:HD1	4:C:3726:HOH:O	1.37	1.05
1:B:319:LYS:HD2	4:B:3651:HOH:O	1.55	1.05
1:A:349:PRO:HD2	4:A:3750:HOH:O	1.58	1.02

The worst 5 of 8 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)
4:B:3154:HOH:O	4:B:3154:HOH:O[2_565]	1.81	0.39
4:B:3459:HOH:O	4:B:3459:HOH:O[2_565]	1.90	0.30
4:A:3568:HOH:O	4:A:3568:HOH:O[2_565]	2.06	0.14
4:A:3784:HOH:O	4:C:3457:HOH:O[3_557]	2.10	0.10
4:A:3726:HOH:O	4:B:3497:HOH:O[4_466]	2.10	0.10

## 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	336/371 (91%)	321 (96%)	15 (4%)	0	100	100
1	B	334/371 (90%)	326 (98%)	8 (2%)	0	100	100
1	C	348/371 (94%)	336 (97%)	10 (3%)	2 (1%)	28	16
All	All	1018/1113 (92%)	983 (97%)	33 (3%)	2 (0%)	51	41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	19	ALA
1	C	20	SER

### 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	305/334 (91%)	290 (95%)	15 (5%)	29	17
1	B	303/334 (91%)	286 (94%)	17 (6%)	25	13
1	C	313/334 (94%)	299 (96%)	14 (4%)	32	21
All	All	921/1002 (92%)	875 (95%)	46 (5%)	28	17

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	100	ASN
1	B	166	ASN
1	C	155	LYS
1	B	113	LEU
1	B	144	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	131	HIS
1	B	231	GLN
1	C	224	ASN
1	B	172	ASN
1	B	196	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 ⓘ

4 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	GDP	A	3001	-	25,30,30	1.65	5 (20%)	26,47,47	1.95	6 (23%)
2	GDP	B	3002	-	25,30,30	1.48	2 (8%)	26,47,47	1.99	5 (19%)
2	GDP	C	3003	-	25,30,30	2.00	5 (20%)	26,47,47	1.93	5 (19%)
3	SO4	C	3004	-	4,4,4	0.43	0	6,6,6	0.23	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	GDP	A	3001	-	-	0/12/32/32	0/3/3/3
2	GDP	B	3002	-	-	0/12/32/32	0/3/3/3
2	GDP	C	3003	-	-	0/12/32/32	0/3/3/3
3	SO4	C	3004	-	-	0/0/0/0	0/0/0/0

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	3001	GDP	C8-N7	-2.72	1.29	1.34
2	A	3001	GDP	PB-O3A	2.15	1.63	1.60
2	C	3003	GDP	O2'-C2'	2.35	1.48	1.43
2	C	3003	GDP	C3'-C4'	2.80	1.60	1.53
2	A	3001	GDP	C2-N1	2.88	1.40	1.35

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	A	3001	GDP	N3-C2-N1	-4.40	121.04	127.46
2	C	3003	GDP	N3-C2-N1	-4.28	121.21	127.46
2	B	3002	GDP	N3-C2-N1	-4.26	121.24	127.46
2	A	3001	GDP	C5-C6-N1	-4.08	117.67	123.48
2	C	3003	GDP	C5-C6-N1	-3.86	117.99	123.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	3002	GDP	1	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	340/371 (91%)	0.30	19 (5%) 25 28	20, 37, 66, 79	0
1	B	338/371 (91%)	0.28	24 (7%) 17 19	19, 36, 61, 74	0
1	C	350/371 (94%)	0.01	16 (4%) 33 37	18, 28, 50, 77	0
All	All	1028/1113 (92%)	0.20	59 (5%) 24 28	18, 33, 62, 79	0

The worst 5 of 59 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	20	SER	8.4
1	C	19	ALA	8.1
1	A	19	ALA	7.6
1	B	19	ALA	7.4
1	C	20	SER	5.4

### 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
2	GDP	A	3001	28/28	0.95	0.11	-0.56	42,51,56,57	0
2	GDP	B	3002	28/28	0.94	0.11	-0.78	34,37,51,52	0
2	GDP	C	3003	28/28	0.98	0.06	-1.38	21,24,29,35	0
3	SO4	C	3004	5/5	0.94	0.14	-	68,68,69,70	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.