



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 20, 2024 – 12:39 pm GMT

PDB ID : 7PU1
Title : High resolution X-ray structure of Thermoascus aurantiacus LPMO
Authors : Banerjee, S.; Frandsen, K.E.H.; Singh, R.K.; Bjerrum, M.J.; Lo Leggio, L.
Deposited on : 2021-09-28
Resolution : 1.06 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

<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
Mogul : 1.8.4, CSD as541be (2020)
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

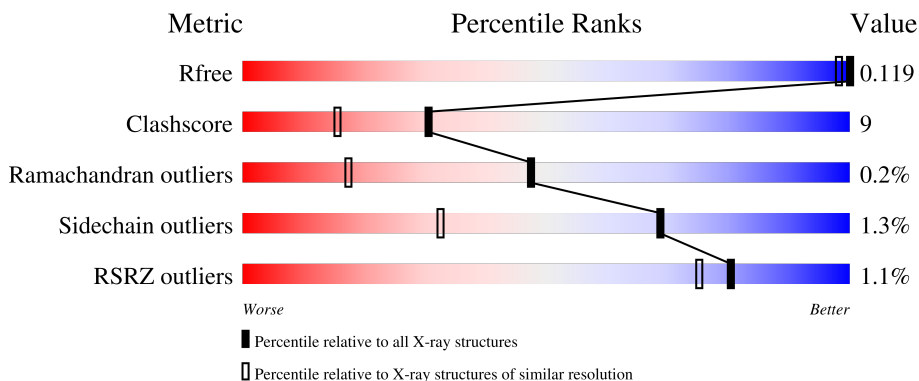
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.06 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	1202 (1.10-1.02)
Clashscore	141614	1252 (1.10-1.02)
Ramachandran outliers	138981	1204 (1.10-1.02)
Sidechain outliers	138945	1202 (1.10-1.02)
RSRZ outliers	127900	1178 (1.10-1.02)

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 ≥ 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	AAA	228	 87% 12% .
1	BBB	228	 86% 12% .

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
3	NAG	AAA	302	-	X	X	-
4	CL	BBB	303	-	-	X	-

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 4345 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.

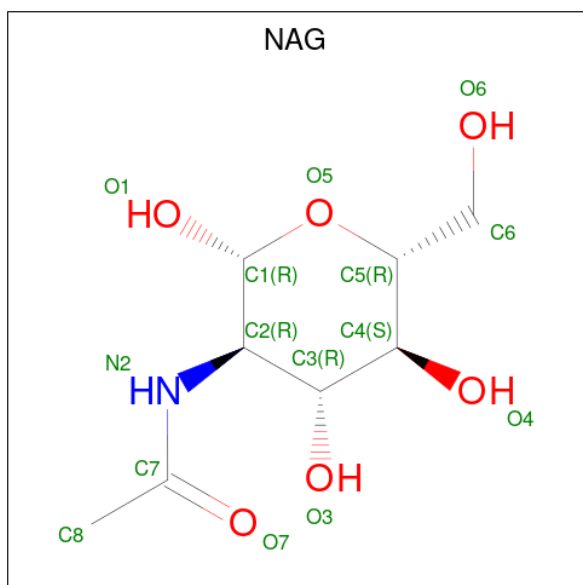
- Molecule 1 is a protein called Gh61 isozyme a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	228	Total 1779	C 1144	N 289	O 341	S 5	0	11	0
1	BBB	228	Total 1786	C 1147	N 291	O 343	S 5	3	13	0

- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	1	Total 1	Cu 1	0	0
2	BBB	1	Total 1	Cu 1	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	AAA	1	Total	C	N	O	1	0
			14	8	1	5		
3	BBB	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	9	Total	Cl	0	0
			9	9		
4	BBB	10	Total	Cl	0	0
			10	10		

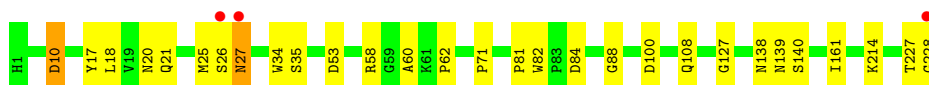
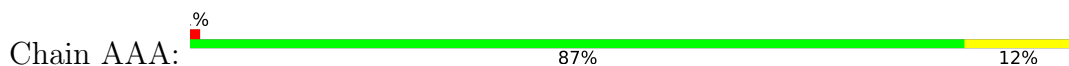
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	356	Total	O	0	11
			369	369		
5	BBB	349	Total	O	0	11
			362	362		

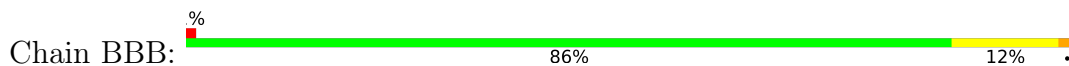
3 Residue-property plots [i](#)

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: Gh61 isozyme a



- Molecule 1: Gh61 isozyme a



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	37.65Å 89.05Å 70.47Å 90.00° 103.33° 90.00°	Depositor
Resolution (Å)	37.37 – 1.06 37.34 – 1.06	Depositor EDS
% Data completeness (in resolution range)	84.2 (37.37-1.06) 84.3 (37.34-1.06)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.36 (at 1.06Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.117 , 0.119 0.121 , 0.119	Depositor DCC
R_{free} test set	8495 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å ²)	10.7	Xtrriage
Anisotropy	0.111	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 45.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.004 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	4345	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 29.53 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.5250e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: HIC, NAG, CL, CU

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	AAA	0.83	2/1857 (0.1%)	1.03	9/2556 (0.4%)
1	BBB	0.87	4/1865 (0.2%)	1.04	7/2568 (0.3%)
All	All	0.85	6/3722 (0.2%)	1.03	16/5124 (0.3%)

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	BBB	0	1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	BBB	58	ARG	NE-CZ	14.56	1.51	1.33
1	AAA	140	SER	CB-OG	-12.72	1.25	1.42
1	BBB	58	ARG	CZ-NH2	6.93	1.42	1.33
1	BBB	53	ASP	CG-OD1	6.58	1.40	1.25
1	BBB	30	GLU	CD-OE1	-6.12	1.19	1.25

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	BBB	58	ARG	NE-CZ-NH2	-16.92	111.84	120.30
1	AAA	58	ARG	NE-CZ-NH1	12.80	126.70	120.30
1	AAA	58	ARG	NE-CZ-NH2	-11.95	114.32	120.30
1	BBB	58	ARG	NH1-CZ-NH2	9.33	129.67	119.40
1	AAA	25	MET	CA-CB-CG	8.42	127.61	113.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	BBB	53	ASP	Sidechain

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	AAA	1779	0	1727	28	0
1	BBB	1786	0	1739	33	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0
3	AAA	14	0	13	12	0
3	BBB	14	0	12	2	0
4	AAA	9	0	0	4	0
4	BBB	10	0	0	7	0
5	AAA	369	0	0	10	0
5	BBB	362	0	0	13	0
All	All	4345	0	3491	65	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 9.

The worst 5 of 65 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:AAA:10:ASP:HB3	5:AAA:569:HOH:O	1.10	1.21
1:AAA:138[A]:ASN:HD21	3:AAA:302:NAG:C1	1.64	1.10
3:AAA:302:NAG:H2	1:BBB:219[A]:ILE:CD1	1.86	1.06
1:BBB:80[A]:THR:HG23	5:BBB:436:HOH:O	1.55	1.05
3:AAA:302:NAG:H2	1:BBB:219[A]:ILE:HD11	1.02	1.01

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	AAA	237/228 (104%)	232 (98%)	5 (2%)	0	100	100
1	BBB	239/228 (105%)	235 (98%)	3 (1%)	1 (0%)	34	11
All	All	476/456 (104%)	467 (98%)	8 (2%)	1 (0%)	47	17

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	BBB	57	HIS

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	AAA	197/186 (106%)	195 (99%)	2 (1%)	76	44
1	BBB	198/186 (106%)	194 (98%)	4 (2%)	55	17
All	All	395/372 (106%)	389 (98%)	6 (2%)	69	26

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	80[A]	THR
1	BBB	80[B]	THR
1	BBB	174	ASN
1	AAA	27	ASN
1	AAA	20	ASN

Sometimes 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](#)

2 non-standard protein/DNA/RNA residues 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	$\# Z > 2$
1	HIC	BBB	1	1,2	8,11,12	1.06	0	6,14,16	0.58	0
1	HIC	AAA	1	1,2	8,11,12	0.83	0	6,14,16	0.50	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
1	HIC	BBB	1	1,2	-	0/5/6/8	0/1/1/1
1	HIC	AAA	1	1,2	-	0/5/6/8	0/1/1/1

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.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 23 ligands modelled in this entry, 21 are monoatomic - leaving 2 for Mogul analysis.

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
3	NAG	BBB	301	1	14,14,15	0.93	1 (7%)	17,19,21	1.52	4 (23%)
3	NAG	AAA	302	1	14,14,15	4.80	8 (57%)	17,19,21	7.74	13 (76%)

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
3	NAG	BBB	301	1	-	0/6/23/26	0/1/1/1
3	NAG	AAA	302	1	-	3/6/23/26	0/1/1/1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	AAA	302	NAG	O4-C4	14.99	1.78	1.43
3	AAA	302	NAG	C1-C2	5.69	1.60	1.52
3	AAA	302	NAG	C6-C5	4.40	1.66	1.51
3	AAA	302	NAG	O6-C6	3.37	1.56	1.42
3	AAA	302	NAG	O5-C1	2.99	1.48	1.43

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	AAA	302	NAG	O5-C5-C6	15.72	131.84	107.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	AAA	302	NAG	C1-O5-C5	-14.68	92.31	112.19
3	AAA	302	NAG	O4-C4-C3	-11.36	84.08	110.35
3	AAA	302	NAG	C8-C7-N2	10.96	134.65	116.10
3	AAA	302	NAG	O6-C6-C5	-8.54	82.00	111.29

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	AAA	302	NAG	C8-C7-N2-C2
3	AAA	302	NAG	O7-C7-N2-C2
3	AAA	302	NAG	C4-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	BBB	301	NAG	2	0
3	AAA	302	NAG	12	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	AAA	227/228 (99%)	-0.14	3 (1%) 77 71	8, 11, 19, 33	0
1	BBB	227/228 (99%)	-0.12	2 (0%) 84 80	8, 12, 19, 48	1 (0%)
All	All	454/456 (99%)	-0.13	5 (1%) 80 75	8, 12, 20, 48	1 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	228	GLY	3.6
1	AAA	27	ASN	2.6
1	BBB	227	THR	2.5
1	BBB	228	GLY	2.4
1	AAA	26	SER	2.1

6.2 Non-standard residues in protein, DNA, RNA chains [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. 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	B-factors(Å ²)	Q<0.9
1	HIC	AAA	1	11/12	0.97	0.08	9,11,12,14	0
1	HIC	BBB	1	11/12	0.98	0.07	9,11,12,15	0

6.3 Carbohydrates [i](#)

There are no monosaccharides 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. 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	B-factors(Å ²)	Q<0.9
3	NAG	AAA	302	14/15	0.56	0.25	26,35,48,74	14
4	CL	AAA	308	1/1	0.75	0.31	65,65,65,65	0
4	CL	BBB	303	1/1	0.84	0.14	44,44,44,44	1
4	CL	BBB	310	1/1	0.84	0.15	61,61,61,61	0
4	CL	BBB	309	1/1	0.90	0.13	52,52,52,52	1
4	CL	AAA	311	1/1	0.90	0.20	53,53,53,53	0
4	CL	AAA	304	1/1	0.92	0.26	53,53,53,53	1
4	CL	AAA	310	1/1	0.92	0.26	38,38,38,38	1
4	CL	AAA	306	1/1	0.93	0.16	51,51,51,51	1
4	CL	BBB	307	1/1	0.94	0.19	46,46,46,46	1
3	NAG	BBB	301	14/15	0.95	0.09	13,15,18,20	3
4	CL	AAA	309	1/1	0.95	0.12	49,49,49,49	0
4	CL	BBB	312	1/1	0.95	0.16	36,36,36,36	0
4	CL	BBB	306	1/1	0.98	0.15	26,26,26,26	1
4	CL	AAA	303	1/1	0.98	0.11	30,30,30,30	1
4	CL	AAA	305	1/1	0.98	0.15	36,36,36,36	1
4	CL	BBB	304	1/1	0.98	0.05	28,28,28,28	1
4	CL	BBB	305	1/1	0.98	0.19	33,33,33,33	0
4	CL	BBB	308	1/1	0.99	0.15	29,29,29,29	1
4	CL	BBB	311	1/1	0.99	0.05	26,26,26,26	1
4	CL	AAA	307	1/1	0.99	0.15	24,24,24,24	1
2	CU	BBB	302	1/1	1.00	0.04	9,9,9,9	1
2	CU	AAA	301	1/1	1.00	0.05	9,9,9,9	1

6.5 Other polymers [i](#)

There are no such residues in this entry.