



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 12:49 PM GMT

PDB ID : 3L4Z
Title : Crystal complex of N-terminal Human Maltase-Glucoamylase with Salacinol
Authors : Sim, L.; Rose, D.R.
Deposited on : 2009-12-21
Resolution : 2.00 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

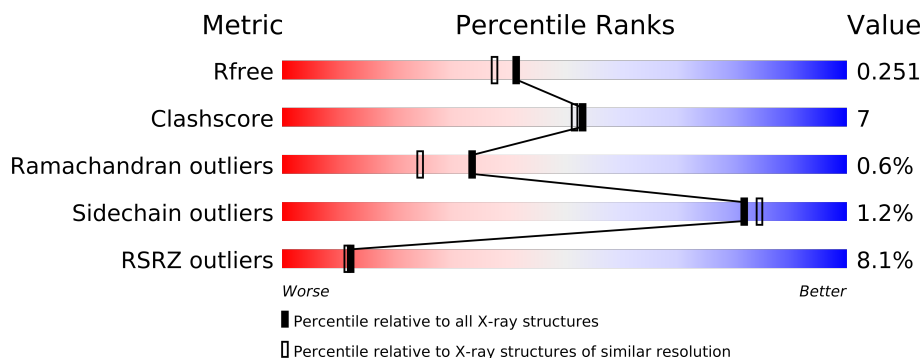
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.00 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}	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	875	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	NAG	A	2005	-	X
5	GOL	A	3001	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7477 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

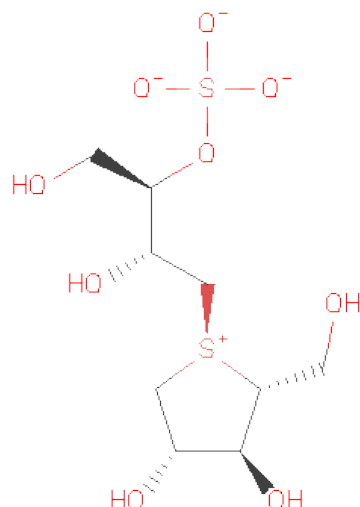
- Molecule 1 is a protein called Maltase-glucoamylase,intestinal.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	863	Total	C	N	O	S	0	4	0
			6934	4420	1174	1313	27			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	772	ASP	ASN	VARIANT	UNP O43451
A	869	ALA	-	EXPRESSION TAG	UNP O43451
A	870	HIS	-	EXPRESSION TAG	UNP O43451
A	871	HIS	-	EXPRESSION TAG	UNP O43451
A	872	HIS	-	EXPRESSION TAG	UNP O43451
A	873	HIS	-	EXPRESSION TAG	UNP O43451
A	874	HIS	-	EXPRESSION TAG	UNP O43451
A	875	HIS	-	EXPRESSION TAG	UNP O43451

- Molecule 2 is 1,4-DIDEOXY-1,4-[[2R,3R)-2,4-DIHYDROXY-3-(SULFOXY)BUTYL]EPIS ULFONIUMYLIDENE]-D-ARABINITOLINNER SALT (three-letter code: SSD) (formula: C₉H₁₈O₉S₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	S	0	0
			20	9	9	2		

- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	A	2	Total	C	N	O	0	0
			28	16	2	10		

There are 16 discrepancies between the modelled and reference sequences:

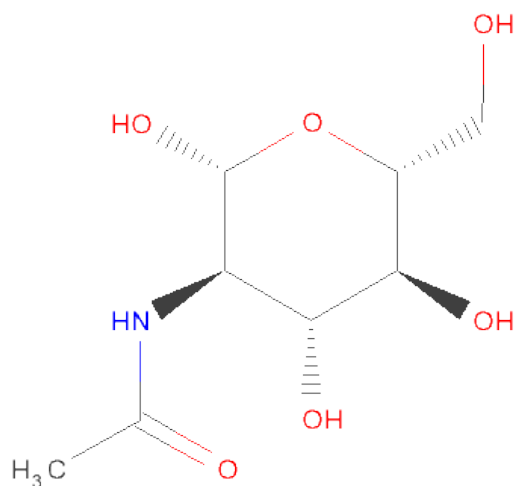
Chain	Residue	Modelled	Actual	Comment	Reference
A	772	ASP	ASN	VARIANT	UNP O43451
A	869	ALA	-	EXPRESSION TAG	UNP O43451
A	870	HIS	-	EXPRESSION TAG	UNP O43451
A	871	HIS	-	EXPRESSION TAG	UNP O43451
A	872	HIS	-	EXPRESSION TAG	UNP O43451
A	873	HIS	-	EXPRESSION TAG	UNP O43451
A	874	HIS	-	EXPRESSION TAG	UNP O43451
A	875	HIS	-	EXPRESSION TAG	UNP O43451
A	772	ASP	ASN	VARIANT	UNP O43451
A	869	ALA	-	EXPRESSION TAG	UNP O43451
A	870	HIS	-	EXPRESSION TAG	UNP O43451
A	871	HIS	-	EXPRESSION TAG	UNP O43451
A	872	HIS	-	EXPRESSION TAG	UNP O43451

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	873	HIS	-	EXPRESSION TAG	UNP O43451
A	874	HIS	-	EXPRESSION TAG	UNP O43451
A	875	HIS	-	EXPRESSION TAG	UNP O43451

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	429	Total	O	0	0
			429	429		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ($\text{RSRZ} > 2$). 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: Maltase-glucoamylase,intestinal



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	92.42Å 108.78Å 111.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.83 – 2.00 19.80 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.2 (19.83-2.00) 98.2 (19.80-2.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.84 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.202 , 0.253 0.202 , 0.251	Depositor DCC
R_{free} test set	3757 reflections (5.30%)	DCC
Wilson B-factor (Å ²)	33.5	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.43 , 50.4	EDS
Estimated twinning fraction	0.010 for -h,l,k	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 74637 reflections (0.001%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7477	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.64% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, NAG, SSD

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	0.77	0/7142	0.76	3/9742 (0.0%)

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	1
3	A	1	0
All	All	1	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	624	ARG	NE-CZ-NH1	5.82	123.21	120.30
1	A	624	ARG	NE-CZ-NH2	-5.73	117.44	120.30
1	A	578	ASP	CB-CG-OD1	5.11	122.89	118.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	2002	NAG	C1

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	121	GLY	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6934	0	6598	96	0
2	A	20	0	18	0	0
3	A	56	0	50	0	0
4	A	14	0	13	0	0
5	A	24	0	32	7	0
6	A	429	0	0	13	0
All	All	7477	0	6711	96	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 7.

The worst 5 of 96 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:98:HIS:HE1	1:A:261:ASP:OD1	1.55	0.88
1:A:479[B]:CYS:SG	6:A:4310:HOH:O	2.31	0.87
1:A:691:VAL:HG23	6:A:4177:HOH:O	1.76	0.86
1:A:493:GLN:NE2	1:A:497:HIS:HD2	1.82	0.76
1:A:779:VAL:HG12	5:A:3003:GOL:H11	1.68	0.74

There are no symmetry-related clashes.

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	863/875 (99%)	819 (95%)	39 (4%)	5 (1%)	33 24

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	793	GLN
1	A	50	HIS
1	A	51	SER
1	A	577	LEU
1	A	405	VAL

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	760/767 (99%)	750 (99%)	10 (1%)	80	82

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	322	ASP
1	A	419	ASN
1	A	707	SER
1	A	149	ARG
1	A	539	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	239	ASN
1	A	357	ASN
1	A	714	GLN
1	A	295	HIS
1	A	464	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

4 carbohydrates 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$
3	NAG	A	2001	1,3	12,14,15	0.94	1 (8%)	15,19,21	1.28	3 (20%)
3	NAG	A	2002	3	12,14,15	0.47	0	15,19,21	2.11	5 (33%)
3	NAG	A	2003	1,3	12,14,15	0.84	0	15,19,21	0.91	0
3	NAG	A	2004	3	12,14,15	0.60	0	15,19,21	0.98	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
3	NAG	A	2001	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	2002	3	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	A	2003	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	2004	3	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2001	NAG	O5-C5	-2.37	1.41	1.45

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	2002	NAG	O5-C5-C6	5.25	112.50	106.98
3	A	2002	NAG	C8-C7-N2	3.15	122.27	116.11

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	2002	NAG	O7-C7-C8	-2.97	116.25	122.04
3	A	2002	NAG	O4-C4-C3	2.36	115.65	110.35
3	A	2001	NAG	C8-C7-N2	2.36	120.73	116.11

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	2002	NAG	C1

There are no torsion outliers.

There are no ring outliers.

5.6 Ligand geometry ⓘ

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	SSD	A	1001	-	19,20,20	1.09	1 (5%)	26,29,29	3.28	9 (34%)
4	NAG	A	2005	1	12,14,15	0.74	1 (8%)	15,19,21	0.92	0
5	GOL	A	3001	-	5,5,5	0.66	0	5,5,5	0.73	0
5	GOL	A	3002	-	5,5,5	0.39	0	5,5,5	0.34	0
5	GOL	A	3003	-	5,5,5	0.50	0	5,5,5	1.21	1 (20%)
5	GOL	A	3004	-	5,5,5	0.30	0	5,5,5	0.47	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	SSD	A	1001	-	2/2/7/7	0/17/33/33	0/1/1/1
4	NAG	A	2005	1	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	A	3001	-	-	0/4/4/4	0/0/0/0
5	GOL	A	3002	-	-	0/4/4/4	0/0/0/0
5	GOL	A	3003	-	-	0/4/4/4	0/0/0/0
5	GOL	A	3004	-	-	0/4/4/4	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1001	SSD	O12-S9	2.43	1.52	1.43
4	A	2005	NAG	O5-C5	-2.01	1.41	1.45

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	SSD	O12-S9-O11	-11.31	106.17	113.77
2	A	1001	SSD	O13-S9-O11	-7.17	108.96	113.77
2	A	1001	SSD	C6-C5-S4	4.86	114.92	110.08
2	A	1001	SSD	O13-S9-O12	-4.04	111.06	113.77
2	A	1001	SSD	C7-S4-C3	3.12	104.58	101.64

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	1001	SSD	C8
2	A	1001	SSD	C9

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	863/875 (98%)	0.36	71 (8%) 12 11	20, 33, 52, 74	0

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	36	GLN	6.2
1	A	50	HIS	6.1
1	A	838	THR	5.4
1	A	75	SER	5.0
1	A	376	SER	4.6

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

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	RSR	LLDF	B-factors(Å ²)	Q<0.9
3	NAG	A	2002	14/15	0.28	16.78	58,61,63,63	0
3	NAG	A	2001	14/15	0.14	0.86	38,42,49,54	0
3	NAG	A	2003	14/15	0.11	-1.13	42,44,47,48	0
3	NAG	A	2004	14/15	0.20	-	53,57,59,60	0

6.4 Ligands

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	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	NAG	A	2005	14/15	0.43	4.62	58,62,66,66	0
5	GOL	A	3001	6/6	0.21	2.39	36,39,41,43	0
5	GOL	A	3003	6/6	0.29	1.33	51,56,58,59	0
5	GOL	A	3002	6/6	0.11	0.81	55,60,61,62	0
5	GOL	A	3004	6/6	0.16	0.39	49,55,56,58	0
2	SSD	A	1001	20/20	0.09	-0.75	33,39,44,45	0

6.5 Other polymers

There are no such residues in this entry.