



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 18, 2026 – 02:38 AM UTC

PDB ID : 1PZY / pdb_00001pzy
Title : W 3 1 4 A - B E T A 1 , 4 - G A L A C T O S Y L T R A N S F E R A S E - I C O M - P L E X E D W I T H A L P H A - L A C T A L B U M I N I N T H E P R E S E N C E O F N - A C E T Y L G L U C O S A M I N E , U D P A N D M A N G A N E S E
Authors : Ramasamy, V.; Ramakrishnan, B.; Boeggeman, E.; Qasba, P.K.
Deposited on : 2003-07-14
Resolution : 2.30 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 : 4-5-2 with Phenix2.0
Mogul : ?? (???), CSD ??CSD?? (????)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

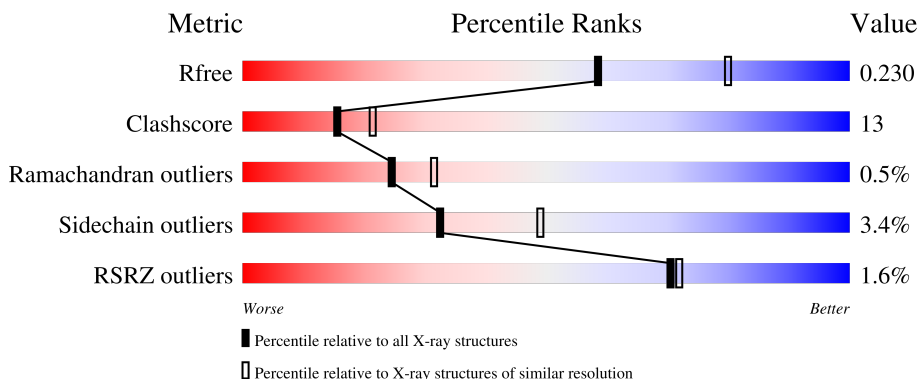
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	180053	6319 (2.30-2.30)
Clashscore	190562	6919 (2.30-2.30)
Ramachandran outliers	187476	6854 (2.30-2.30)
Sidechain outliers	187428	6854 (2.30-2.30)
RSRZ outliers	180081	6325 (2.30-2.30)

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	A	123	3% 74% 24% .
1	C	123	3% 71% 24% 5%
2	B	286	63% 30% 5%
2	D	286	% 66% 25% 5%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6699 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 ALPHA-LACTALBUMIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	123	Total	C	N	O	S	0	0	0
			980	620	156	195	9			
1	C	123	Total	C	N	O	S	0	0	0
			980	620	156	195	9			

- Molecule 2 is a protein called BETA-1,4-GALACTOSYLTRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	271	Total	C	N	O	S	0	0	0
			2202	1411	380	398	13			
2	D	271	Total	C	N	O	S	0	0	0
			2202	1411	380	398	13			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	117	ALA	-	expression tag	UNP P08037
B	118	SER	-	expression tag	UNP P08037
B	119	MET	-	expression tag	UNP P08037
B	120	THR	-	expression tag	UNP P08037
B	121	GLY	-	expression tag	UNP P08037
B	122	GLY	-	expression tag	UNP P08037
B	123	GLN	-	expression tag	UNP P08037
B	124	GLN	-	expression tag	UNP P08037
B	125	MET	-	expression tag	UNP P08037
B	126	GLY	-	expression tag	UNP P08037
B	127	ARG	-	expression tag	UNP P08037
B	128	GLY	-	expression tag	UNP P08037
B	129	SER	-	expression tag	UNP P08037
B	314	ALA	TRP	engineered mutation	UNP P08037
B	342	THR	CYS	engineered mutation	UNP P08037
D	117	ALA	-	expression tag	UNP P08037

Continued on next page...

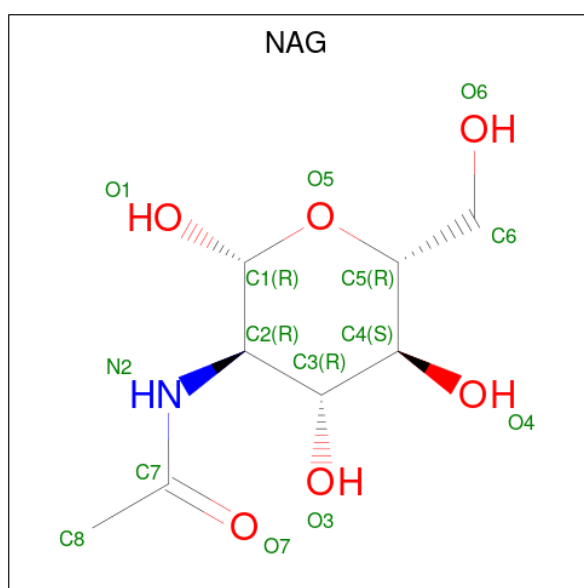
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	118	SER	-	expression tag	UNP P08037
D	119	MET	-	expression tag	UNP P08037
D	120	THR	-	expression tag	UNP P08037
D	121	GLY	-	expression tag	UNP P08037
D	122	GLY	-	expression tag	UNP P08037
D	123	GLN	-	expression tag	UNP P08037
D	124	GLN	-	expression tag	UNP P08037
D	125	MET	-	expression tag	UNP P08037
D	126	GLY	-	expression tag	UNP P08037
D	127	ARG	-	expression tag	UNP P08037
D	128	GLY	-	expression tag	UNP P08037
D	129	SER	-	expression tag	UNP P08037
D	314	ALA	TRP	engineered mutation	UNP P08037
D	342	THR	CYS	engineered mutation	UNP P08037

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Ca 1 1	0	0
3	C	1	Total Ca 1 1	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula: C₈H₁₅NO₆).

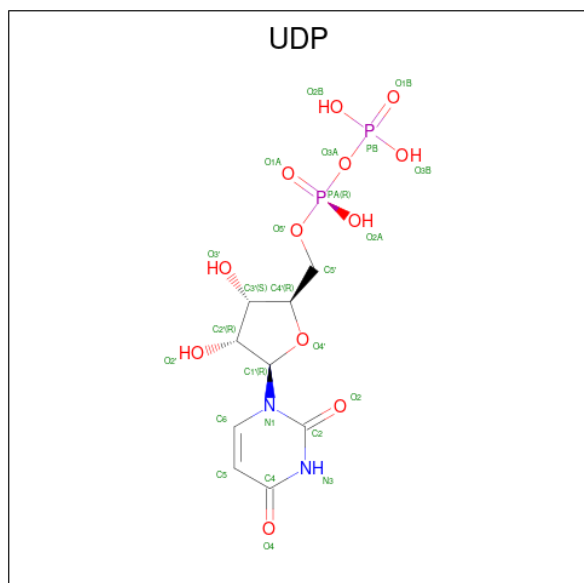


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	N	O	0	0
			15	8	1	6		
4	D	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 5 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Mn	0	0
			1	1		
5	D	1	Total	Mn	0	0
			1	1		

- Molecule 6 is URIDINE-5'-DIPHOSPHATE (CCD ID: UDP) (formula: C₉H₁₄N₂O₁₂P₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
6	D	1	Total	C	N	O	P	0	0
			25	9	2	12	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	46	Total	O	0	0
			46	46		

Continued on next page...

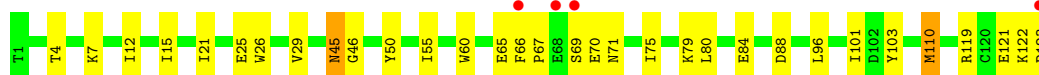
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	85	Total 85	O 85	0	0
7	C	25	Total 25	O 25	0	0
7	D	95	Total 95	O 95	0	0

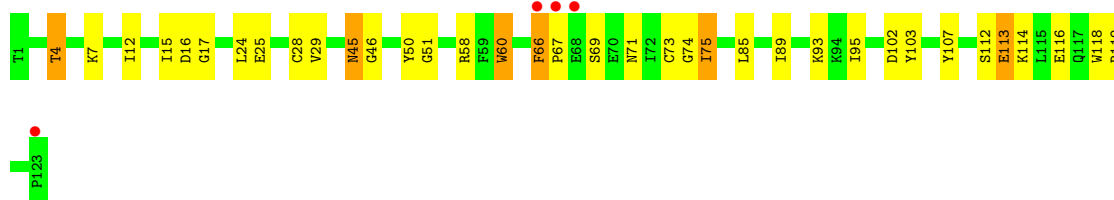
3 Residue-property plots

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.

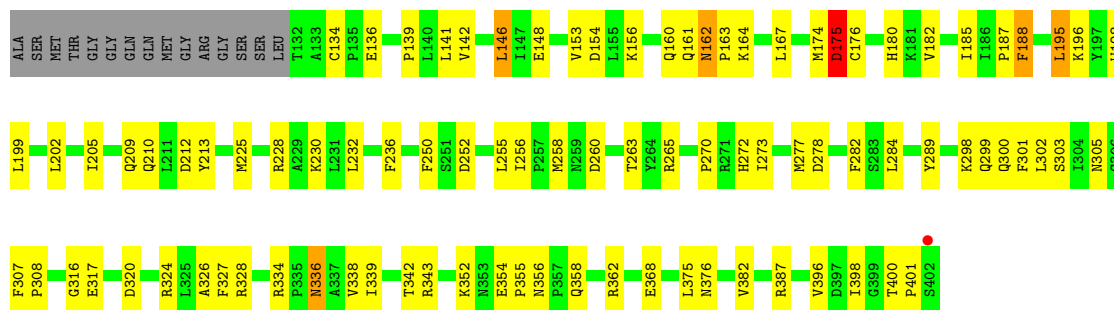
• Molecule 1: ALPHA-LACTALBUMIN



• Molecule 1: ALPHA-LACTALBUMIN

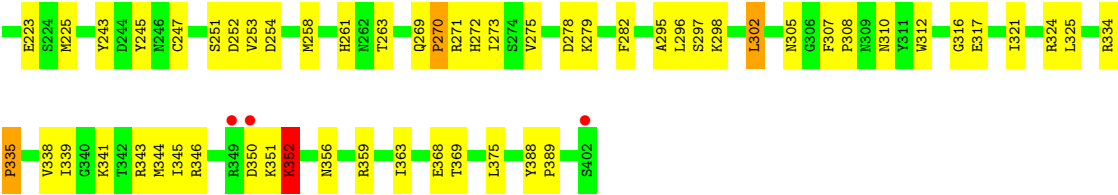


• Molecule 2: BETA-1,4-GALACTOSYLTRANSFERASE



• Molecule 2: BETA-1,4-GALACTOSYLTRANSFERASE





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	55.29Å 98.75Å 102.79Å 90.00° 103.92° 90.00°	Depositor
Resolution (Å)	19.95 – 2.30 19.95 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.1 (19.95-2.30) 97.1 (19.95-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.16 (at 2.30Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.207 , 0.253 (Not available) , 0.230	Depositor DCC
R_{free} test set	4654 reflections (10.05%)	wwPDB-VP
Wilson B-factor (Å ²)	37.2	Xtriage
Anisotropy	0.383	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 32.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.029 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6699	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.60% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, MN, CA, UDP

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	0.42	0/1001	0.95	5/1350 (0.4%)
1	C	0.37	0/1001	0.94	7/1350 (0.5%)
2	B	0.41	0/2260	0.98	10/3060 (0.3%)
2	D	0.43	0/2260	1.01	17/3060 (0.6%)
All	All	0.41	0/6522	0.98	39/8820 (0.4%)

There are no bond length outliers.

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	182	VAL	N-CA-C	8.86	120.58	108.17
1	A	60	TRP	N-CA-C	8.79	122.12	111.40
2	D	202	LEU	N-CA-C	7.54	119.50	111.28
2	B	162	ASN	CA-C-N	7.39	126.93	119.24
2	B	162	ASN	C-N-CA	7.39	126.93	119.24

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	980	0	936	20	0
1	C	980	0	936	24	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	2202	0	2171	68	0
2	D	2202	0	2171	58	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
4	B	15	0	15	1	0
4	D	15	0	15	0	0
5	B	1	0	0	0	0
5	D	1	0	0	0	0
6	B	25	0	11	1	0
6	D	25	0	11	1	0
7	A	46	0	0	2	0
7	B	85	0	0	0	0
7	C	25	0	0	1	0
7	D	95	0	0	2	0
All	All	6699	0	6266	170	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 13.

The worst 5 of 170 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)
2:D:352:LYS:HE3	2:D:352:LYS:HA	1.47	0.94
2:B:258:MET:HE1	2:B:282:PHE:HE2	1.39	0.88
2:B:336:ASN:HD22	2:B:336:ASN:C	1.80	0.87
2:B:336:ASN:ND2	2:B:339:ILE:H	1.76	0.84
1:A:66:PHE:HB3	1:A:69:SER:HB2	1.59	0.84

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	121/123 (98%)	111 (92%)	10 (8%)	0	100	100
1	C	121/123 (98%)	108 (89%)	12 (10%)	1 (1%)	16	20
2	B	269/286 (94%)	258 (96%)	9 (3%)	2 (1%)	18	23
2	D	269/286 (94%)	257 (96%)	11 (4%)	1 (0%)	30	38
All	All	780/818 (95%)	734 (94%)	42 (5%)	4 (0%)	24	31

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	175	ASP
1	C	113	GLU
2	D	352	LYS
2	B	327	PHE

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	109/109 (100%)	107 (98%)	2 (2%)	51	70
1	C	109/109 (100%)	105 (96%)	4 (4%)	30	45
2	B	243/253 (96%)	234 (96%)	9 (4%)	30	45
2	D	243/253 (96%)	234 (96%)	9 (4%)	30	45
All	All	704/724 (97%)	680 (97%)	24 (3%)	32	49

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	116	GLU
2	D	175	ASP
2	D	153	VAL
2	D	195	LEU
2	B	175	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22

such sidechains are listed below:

Mol	Chain	Res	Type
2	D	207	GLN
2	D	269	GLN
2	D	261	HIS
2	D	299	GLN
2	B	299	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 for Mogul analysis.

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.

5.7 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	123/123 (100%)	-0.08	4 (3%)	49	51	24, 36, 68, 105	0
1	C	123/123 (100%)	0.30	4 (3%)	49	51	31, 48, 70, 91	0
2	B	271/286 (94%)	-0.19	1 (0%)	88	89	25, 39, 60, 82	0
2	D	271/286 (94%)	-0.20	4 (1%)	72	73	22, 37, 61, 84	0
All	All	788/818 (96%)	-0.10	13 (1%)	70	72	22, 39, 68, 105	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	123	PRO	3.8
1	C	66	PHE	3.4
1	C	123	PRO	3.0
2	D	132	THR	2.9
2	D	350	ASP	2.6

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
6	UDP	D	409	25/25	0.87	0.11	49,57,85,86	4
5	MN	D	410	1/1	0.88	0.08	79,79,79,79	1
6	UDP	B	407	25/25	0.91	0.09	49,54,79,81	4
4	NAG	D	406	15/15	0.94	0.07	40,42,46,46	0
4	NAG	B	405	15/15	0.94	0.07	34,36,38,39	0
3	CA	C	404	1/1	0.96	0.04	49,49,49,49	0
5	MN	B	408	1/1	0.98	0.04	66,66,66,66	1
3	CA	A	403	1/1	0.99	0.03	36,36,36,36	0

6.5 Other polymers [i](#)

There are no such residues in this entry.