

wwPDB X-ray Structure Validation Summary Report (i)

Aug 2, 2023 – 08:06 PM EDT

PDB ID : 1GA2

Title: THE CRYSTAL STRUCTURE OF ENDOGLUCANASE 9G FROM

CLOSTRIDIUM CELLULOLYTICUM COMPLEXED WITH CELLOBIOSE

Authors: Mandelman, D.; Belaich, A.; Belaich, J.P.; Aghajari, N.; Driguez, H.; Haser,

R.

Deposited on : 2000-11-29

Resolution : 1.70 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 (i) 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 (i)) were used in the production of this report:

MolProbity : 4.02b-467

Mogul : 1.8.5 (274361), CSD as541be (2020)

Xtriage (Phenix) : NOT EXECUTED EDS : NOT EXECUTED

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

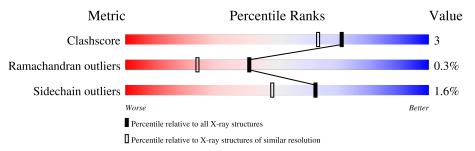
Validation Pipeline (wwPDB-VP) : 2.34

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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	Similar resolution		
Metric	$(\# \mathrm{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$		
Clashscore	141614	4695 (1.70-1.70)		
Ramachandran outliers	138981	4610 (1.70-1.70)		
Sidechain outliers	138945	4610 (1.70-1.70)		

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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain					
1	A	614	92%	7%				
1	В	614	91%	8%				
2	С	3	67%	33%				
3	D	2	100%					
3	Е	2	100%					

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:



N	Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
	7	ACY	A	790	_	_	X	_



2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 10474 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 ENDOGLUCANASE 9G.

Mol	Chain	Residues		Atoms				ZeroOcc	AltConf	Trace
1	A	611	Total 4795	C 3047	N 786	O 943	S 19	0	5	0
1	В	613	Total 4797	C 3050	N 787	O 941	S 19	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	574	THR	ARG	conflict	UNP P37700
A	575	THR	ARG	conflict	UNP P37700
В	574	THR	ARG	conflict	UNP P37700
В	575	THR	ARG	conflict	UNP P37700

• Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace	
2	С	3	Total 34	C 18	O 16	0	0	0

• Molecule 3 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose.





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
3	D	2	Total C O 23 12 11	0	0	0
3	E	2	Total C O 23 12 11	0	0	0

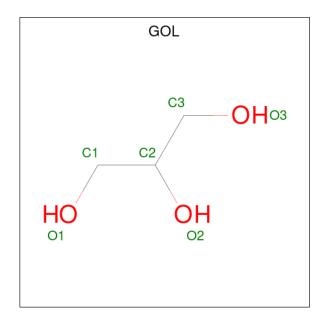
• Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Ca 2 2	0	0
4	В	2	Total Ca 2 2	0	0

• Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	4	Total Mg 4 4	0	0
5	В	3	Total Mg 3 3	0	0

• Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

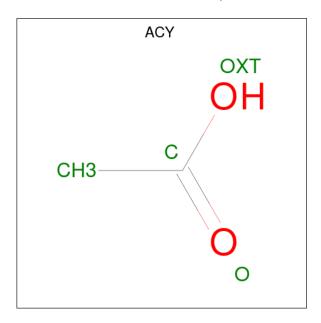
Continued on next page...



 $Continued\ from\ previous\ page...$

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

• Molecule 7 is ACETIC ACID (three-letter code: ACY) (formula: C₂H₄O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 4	C 2	O 2	0	0

• Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	401	Total O 401 401	0	0
8	В	356	Total O 356 356	0	0

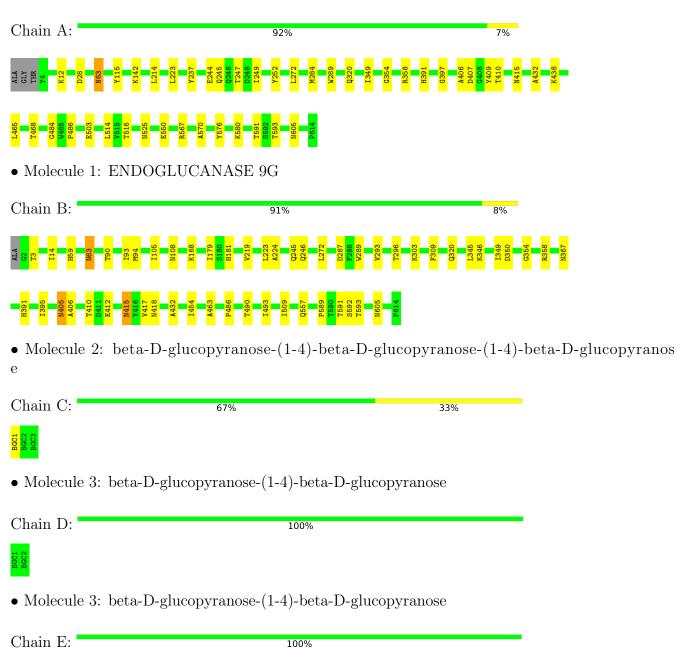


3 Residue-property plots (i)

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. 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. 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.

Note EDS was not executed.











4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source	
Space group	P 1	Depositor	
Cell constants	56.85Å 57.67Å 86.58Å	Depositor	
a, b, c, α , β , γ	93.82° 100.86° 99.46°	Depositor	
Resolution (Å)	43.54 - 1.70	Depositor	
% Data completeness	89.8 (43.54-1.70)	Depositor	
(in resolution range)	03.0 (49.04 1.10)	Deposition	
R_{merge}	(Not available)	Depositor	
R_{sym}	0.04	Depositor	
Refinement program	CNS	Depositor	
R, R_{free}	0.170 , 0.197	Depositor	
Estimated twinning fraction	No twinning to report.	Xtriage	
Total number of atoms	10474	wwPDB-VP	
Average B, all atoms (Å ²)	17.0	wwPDB-VP	



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: BGC, CA, GOL, MG, ACY

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	Mol Chain		Bond lengths		Bond angles		
IVIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5		
1	A	0.39	0/4935	0.64	0/6727		
1	В	0.40	0/4938	0.66	1/6729 (0.0%)		
All	All	0.40	0/9873	0.65	$1/13456 \ (0.0\%)$		

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 maintenain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^{o})$
1	В	415	ASN	N-CA-C	-5.87	95.14	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	454	ILE	СВ

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	4795	0	4473	29	0
1	В	4797	0	4501	36	0
2	С	34	0	30	0	0
3	D	23	0	20	0	0
3	Ε	23	0	21	0	0
4	A	2	0	0	0	0
4	В	2	0	0	0	0
5	A	4	0	0	0	0
5	В	3	0	0	0	0
6	A	24	0	32	0	0
6	В	6	0	8	0	0
7	A	4	0	3	4	0
8	A	401	0	0	2	0
8	В	356	0	0	4	0
All	All	10474	0	9088	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 3.

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	Clash
		$\operatorname{distance}\left(\mathrm{\AA}\right)$	overlap (Å)
1:B:179:ILE:HD11	1:B:224:ALA:HB1	1.58	0.84
1:B:410:THR:HG22	1:B:412:GLU:HG3	1.67	0.76
1:B:345:LEU:O	1:B:349:ILE:HD13	1.85	0.76
1:B:591:THR:O	1:B:593:THR:N	2.17	0.76
1:A:247[A]:THR:HG21	1:A:249:ILE:HD12	1.69	0.75

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	Perce	ntiles
1	A	614/614 (100%)	595 (97%)	17 (3%)	2 (0%)	41	24
1	В	611/614 (100%)	590 (97%)	19 (3%)	2 (0%)	41	24
All	All	1225/1228 (100%)	1185 (97%)	36 (3%)	4 (0%)	41	24

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	В	592	SER
1	В	354	GLY
1	A	354	GLY
1	A	484	GLY

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	Perce	ntiles
1	A	503/507~(99%)	497 (99%)	6 (1%)	71	59
1	В	506/507 (100%)	496 (98%)	10 (2%)	55	38
All	All	1009/1014 (100%)	993 (98%)	16 (2%)	62	48

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	В	557	GLN
1	В	486	PRO
1	В	296	THR
1	В	405	ASN
1	В	63	ASN

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



Mol	Chain	Res	Type
1	В	245	GLN
1	В	367	ASN
1	В	605	ASN
1	В	294	ASN
1	В	405	ASN

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)

7 monosaccharides 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	Trno	Chain	Res	Link	Вс	Bond lengths			Bond angles		
MIOI	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2	
2	BGC	С	1	2	12,12,12	0.55	0	17,17,17	0.74	1 (5%)	
2	BGC	С	2	2	11,11,12	0.54	0	15,15,17	0.47	0	
2	BGC	С	3	2	11,11,12	0.44	0	15,15,17	0.33	0	
3	BGC	D	1	3	12,12,12	0.43	0	17,17,17	0.35	0	
3	BGC	D	2	3,5	11,11,12	0.53	0	15,15,17	0.46	0	
3	BGC	Е	1	3	12,12,12	0.51	0	17,17,17	0.37	0	
3	BGC	Е	2	3	11,11,12	0.48	0	15,15,17	0.48	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	BGC	С	1	2	-	0/2/22/22	0/1/1/1
2	BGC	С	2	2	-	0/2/19/22	0/1/1/1
2	BGC	С	3	2	-	1/2/19/22	0/1/1/1
3	BGC	D	1	3	-	0/2/22/22	0/1/1/1
3	BGC	D	2	3,5	-	0/2/19/22	0/1/1/1
3	BGC	Е	1	3	-	1/2/22/22	0/1/1/1
3	BGC	Е	2	3	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
2	С	1	BGC	C4-C3-C2	-2.12	107.13	110.82

There are no chirality outliers.

All (2) torsion outliers are listed below:

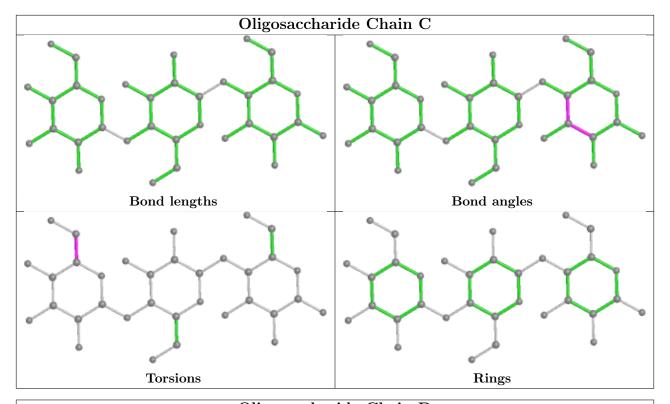
Mol	Chain	Res	Type	Atoms
2	С	3	BGC	O5-C5-C6-O6
3	Е	1	BGC	C4-C5-C6-O6

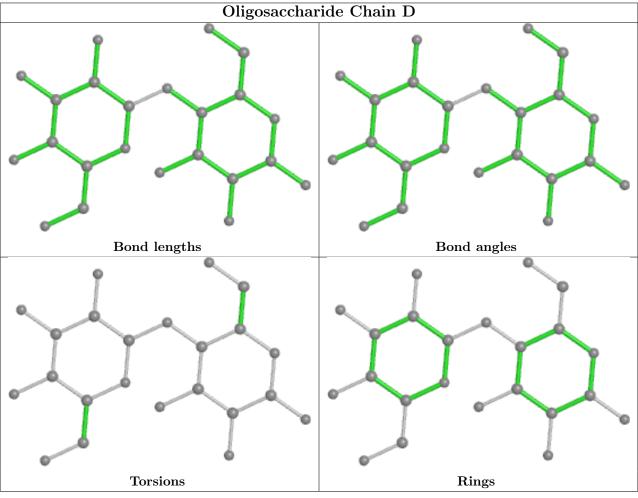
There are no ring outliers.

No monomer is involved in short contacts.

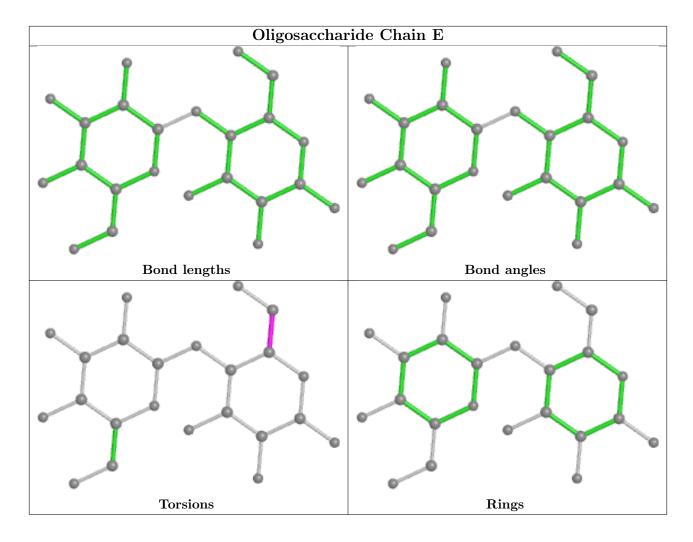
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.











5.6 Ligand geometry (i)

Of 17 ligands modelled in this entry, 11 are monoatomic - leaving 6 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	Tune	Chain	Res	Link	В	Bond lengths			Bond angles		
WIOI	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2	
7	ACY	A	790	-	3,3,3	1.00	0	3,3,3	1.94	2 (66%)	
6	GOL	A	788	-	5,5,5	1.04	0	5,5,5	0.93	0	
6	GOL	В	787	-	5,5,5	1.20	1 (20%)	5,5,5	0.70	0	
6	GOL	A	789	-	5,5,5	0.75	0	5,5,5	2.32	2 (40%)	
6	GOL	A	787	-	5,5,5	1.12	0	5,5,5	0.74	0	



7	Mol	Type	Chain	Res	Link	B	Bond lengths		Bond angles		
1	VIOI	туре	Chain	rtes	Lilik	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z >2
	6	GOL	A	786	-	5,5,5	0.78	0	5,5,5	0.58	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
6	GOL	A	788	-	-	4/4/4/4	-
6	GOL	В	787	-	-	2/4/4/4	-
6	GOL	A	789	-	-	3/4/4/4	-
6	GOL	A	787	-	-	0/4/4/4	-
6	GOL	A	786	-	-	0/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\text{\AA})$	$\operatorname{Ideal}(ext{\AA})$
6	В	787	GOL	O2-C2	2.01	1.49	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^{o})$
6	A	789	GOL	O1-C1-C2	3.46	126.77	110.20
6	A	789	GOL	O3-C3-C2	3.40	126.51	110.20
7	A	790	ACY	O-C-CH3	-2.57	112.31	122.33
7	A	790	ACY	OXT-C-O	2.02	129.50	122.05

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	788	GOL	C1-C2-C3-O3
6	A	788	GOL	O2-C2-C3-O3
6	A	789	GOL	C1-C2-C3-O3
6	В	787	GOL	C1-C2-C3-O3
6	В	787	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 4 short contacts:



Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	790	ACY	4	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)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.

