wwPDB X-ray Structure Validation Summary Report (i)

Dec 18, 2023 – 08:44 am GMT

PDB ID	:	4BQ5
Title	:	Structural analysis of an exo-beta-agarase
Authors	:	Pluvinage, B.; Hehemann, J.H.; Boraston, A.B.
Deposited on	:	2013-05-29
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 (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:

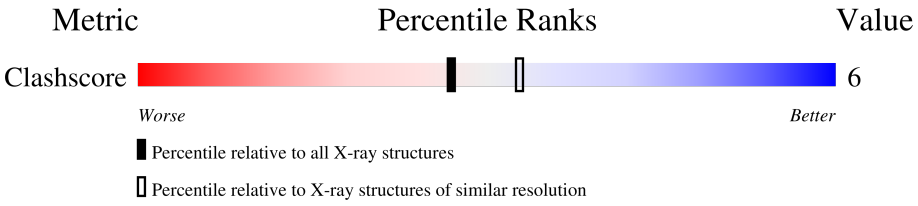
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 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))
Clashscore	141614	5643 (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 <=5%

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	750	87% 13%
1	B	750	85% 15%
2	C	2	50% 50%
2	D	2	100%
2	E	2	50% 50%
2	G	2	100%
2	H	2	50% 50%
3	F	4	50% 50%

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
4	AAL	A	1795	-	-	X	-
5	GAL	A	1796	-	-	X	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 12994 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 B-AGARASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	748	Total	C	N	O	S	3	3	0
			5920	3776	991	1130	23			
1	B	749	Total	C	N	O	S	14	2	0
			5912	3770	997	1122	23			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	44	GLY	-	expression tag	UNP Q21HC5
A	45	SER	-	expression tag	UNP Q21HC5
A	46	HIS	-	expression tag	UNP Q21HC5
A	534	GLN	GLU	engineered mutation	UNP Q21HC5
B	44	GLY	-	expression tag	UNP Q21HC5
B	45	SER	-	expression tag	UNP Q21HC5
B	46	HIS	-	expression tag	UNP Q21HC5
B	534	GLN	GLU	engineered mutation	UNP Q21HC5

- Molecule 2 is an oligosaccharide called 3,6-anhydro-alpha-L-galactopyranose-(1-3)-beta-D-galactopyranose.

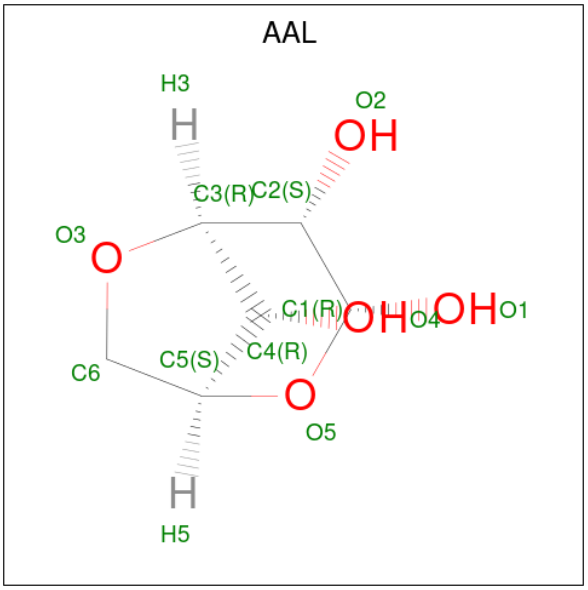
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	C	2	Total	C	O	0	0	0
			21	12	9			
2	D	2	Total	C	O	0	0	0
			21	12	9			
2	E	2	Total	C	O	0	0	0
			22	12	10			
2	G	2	Total	C	O	0	0	0
			21	12	9			
2	H	2	Total	C	O	0	0	0
			22	12	10			

- Molecule 3 is an oligosaccharide called 3,6-anhydro-alpha-L-galactopyranose-(1-3)-beta-D-g

alactopyranose-(1-4)-3,6-anhydro-alpha-L-galactopyranose-(1-3)-beta-D-galactopyranose.

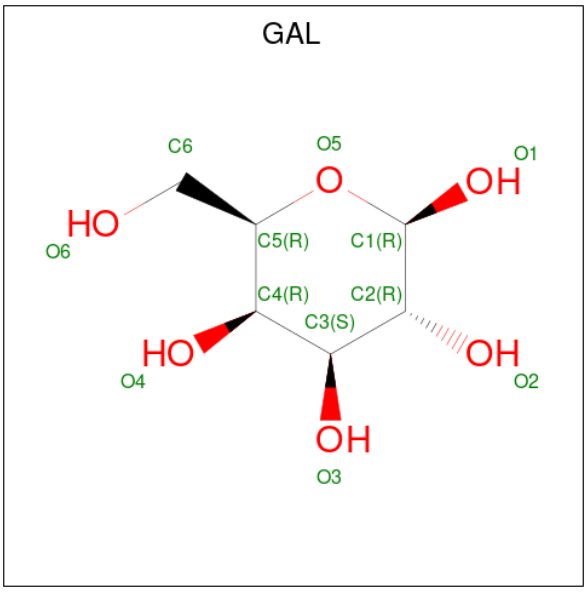
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
3	F	4	Total C O 42 24 18	0	0	0

- Molecule 4 is 3,6-anhydro-alpha-L-galactopyranose (three-letter code: AAL) (formula: $C_6H_{10}O_5$).



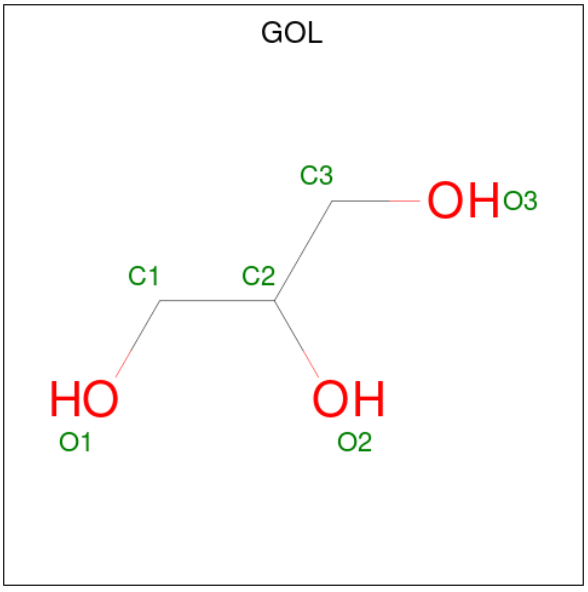
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 10 6 4	0	0

- Molecule 5 is beta-D-galactopyranose (three-letter code: GAL) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 11 6 5	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
6	B	1	Total C O 6 3 3	0	0
6	B	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	3	Total Ca 3 3	0	0
7	B	3	Total Ca 3 3	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	512	Total O 512 512	0	0
8	B	450	Total O 450 450	0	0

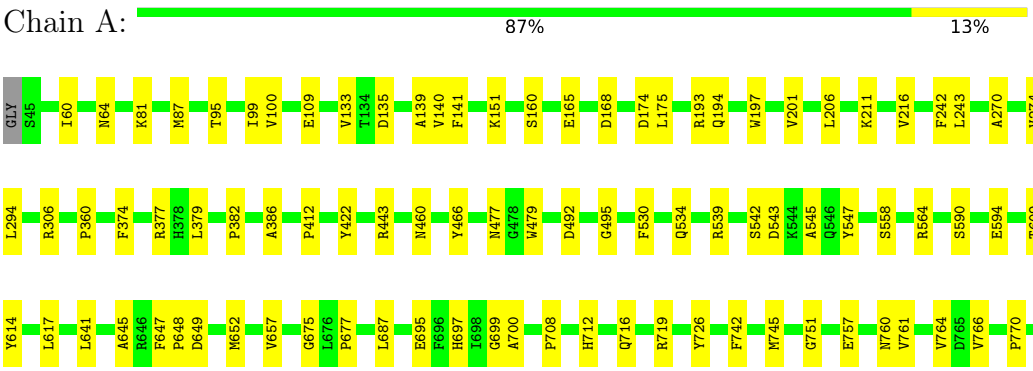
3 Residue-property plots (i)

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 failed to run properly.

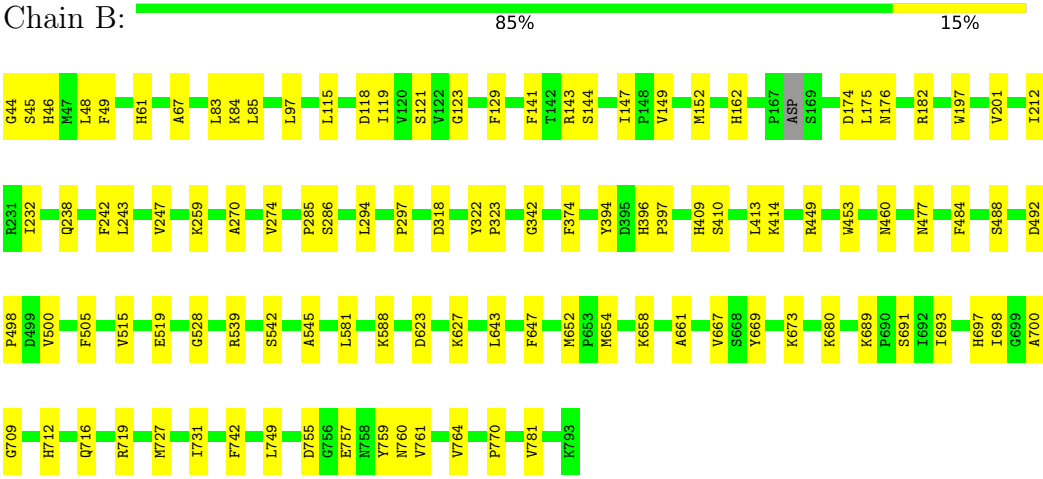
• Molecule 1: B-AGARASE

Chain A:



• Molecule 1: B-AGARASE

Chain B:

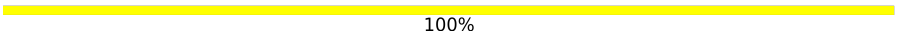


• Molecule 2: 3,6-anhydro-alpha-L-galactopyranose-(1-3)-beta-D-galactopyranose

Chain C:

• Molecule 2: 3,6-anhydro-alpha-L-galactopyranose-(1-3)-beta-D-galactopyranose

Chain D:

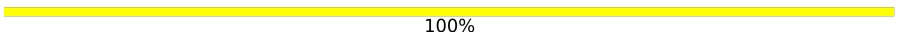


• Molecule 2: 3,6-anhydro-alpha-L-galactopyranose-(1-3)-beta-D-galactopyranose

Chain E:

• Molecule 2: 3,6-anhydro-alpha-L-galactopyranose-(1-3)-beta-D-galactopyranose

Chain G:



• Molecule 2: 3,6-anhydro-alpha-L-galactopyranose-(1-3)-beta-D-galactopyranose

Chain H:

• Molecule 3: 3,6-anhydro-alpha-L-galactopyranose-(1-3)-beta-D-galactopyranose-(1-4)-3,6-anhydro-alpha-L-galactopyranose-(1-3)-beta-D-galactopyranose

Chain F:

4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	69.03Å 116.17Å 207.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.19 – 2.30	Depositor
% Data completeness (in resolution range)	94.1 (39.19-2.30)	Depositor
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.24 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.189 , 0.251	Depositor
Wilson B-factor (Å^2)	17.5	Xtriage
Anisotropy	0.154	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12994	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: GAL, GOL, AAL, CA

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	0.67	1/6094 (0.0%)	0.76	3/8293 (0.0%)
1	B	0.68	4/6085 (0.1%)	0.74	4/8277 (0.0%)
All	All	0.67	5/12179 (0.0%)	0.75	7/16570 (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

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	165	GLU	CG-CD	-20.29	1.21	1.51
1	B	84	LYS	CE-NZ	-15.21	1.11	1.49
1	B	259	LYS	CG-CD	-15.01	1.01	1.52
1	B	414	LYS	CD-CE	-12.48	1.20	1.51
1	B	588	LYS	CG-CD	7.52	1.78	1.52

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	259	LYS	CB-CG-CD	11.21	140.74	111.60
1	B	588	LYS	CB-CG-CD	-7.41	92.32	111.60
1	B	581	LEU	CA-CB-CG	6.19	129.53	115.30
1	A	539	ARG	NE-CZ-NH2	-5.72	117.44	120.30
1	A	193	ARG	NE-CZ-NH2	-5.35	117.62	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	285	PRO	Peptide

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	5920	0	5586	60	0
1	B	5912	0	5576	67	0
2	C	21	0	17	5	0
2	D	21	0	17	8	0
2	E	22	0	19	3	0
2	G	21	0	17	9	0
2	H	22	0	19	4	0
3	F	42	0	33	7	0
4	A	10	0	8	7	0
5	A	11	0	10	8	0
6	A	6	0	8	0	0
6	B	18	0	24	1	0
7	A	3	0	0	0	0
7	B	3	0	0	0	0
8	A	512	0	0	11	0
8	B	450	0	0	5	0
All	All	12994	0	11334	151	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 6.

The worst 5 of 151 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)
3:F:1:GAL:C1	2:G:2:AAL:O4	1.69	1.40
4:A:1795:AAL:C1	5:A:1796:GAL:O3	1.72	1.38
5:A:1796:GAL:C1	2:D:2:AAL:O4	1.75	1.34
4:A:1795:AAL:O4	2:C:1:GAL:C1	1.90	1.20

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:1:GAL:C1	2:G:2:AAL:HO4	1.47	1.18

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

There are no protein backbone outliers to report in this entry.

5.3.2 Protein sidechains

There are no protein residues with a non-rotameric sidechain to report in this entry.

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

14 monosaccharides are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

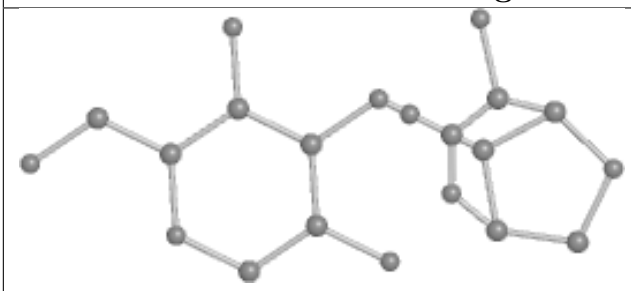
There are no torsion outliers.

There are no ring outliers.

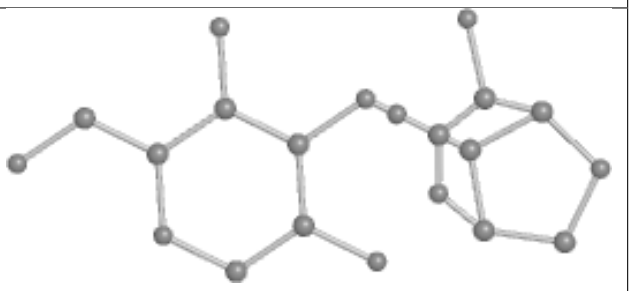
No monomer is involved in short contacts.

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

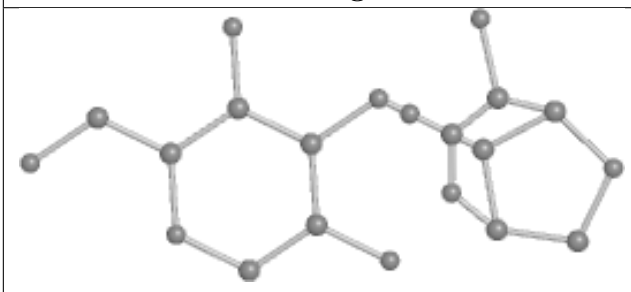
Oligosaccharide Chain C



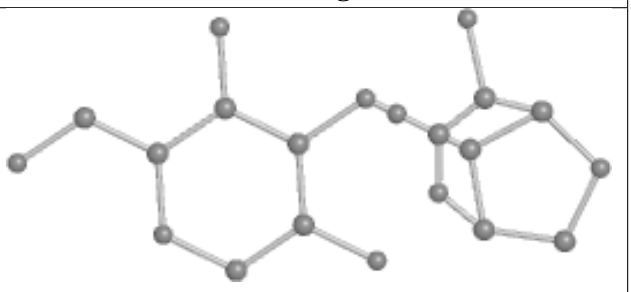
Bond lengths



Bond angles

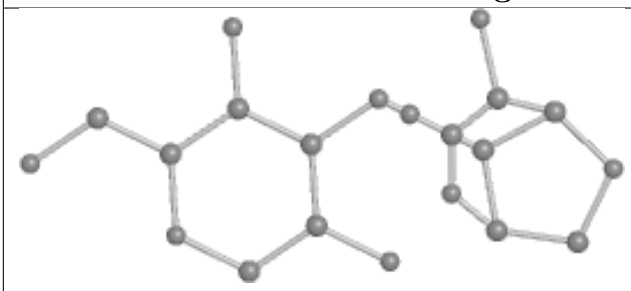


Torsions

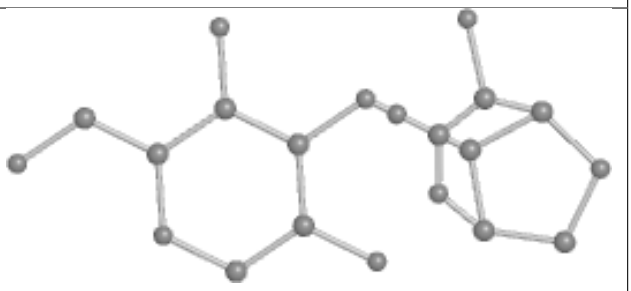


Rings

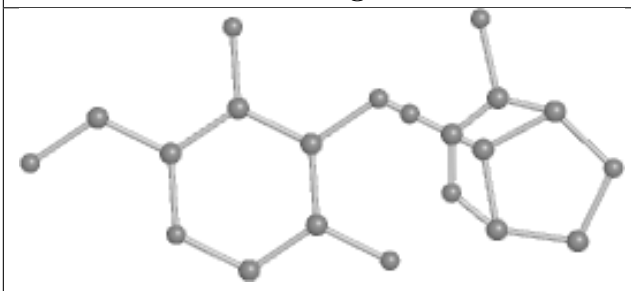
Oligosaccharide Chain D



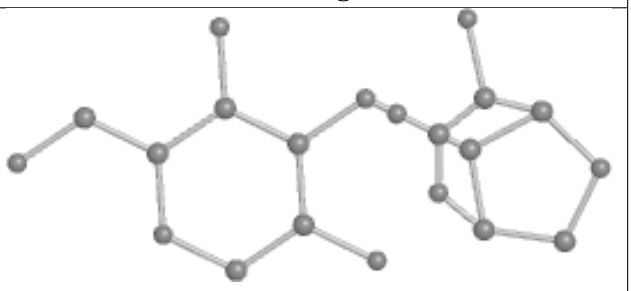
Bond lengths



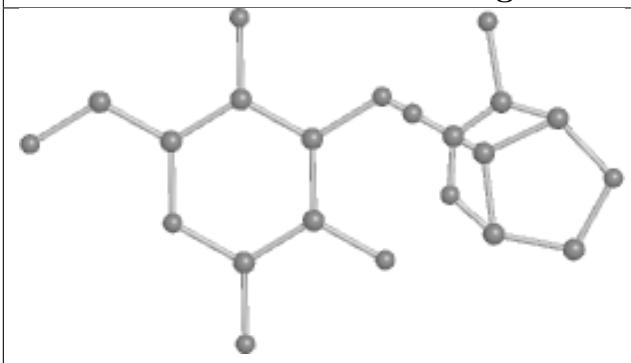
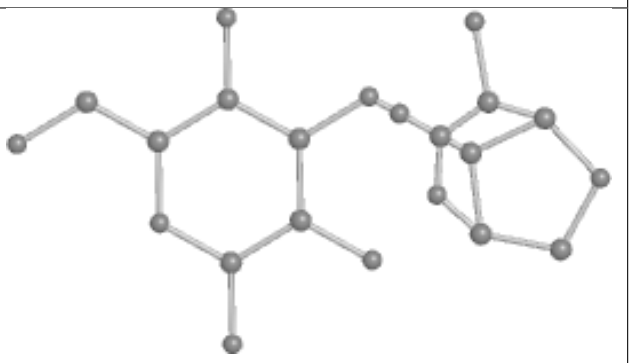
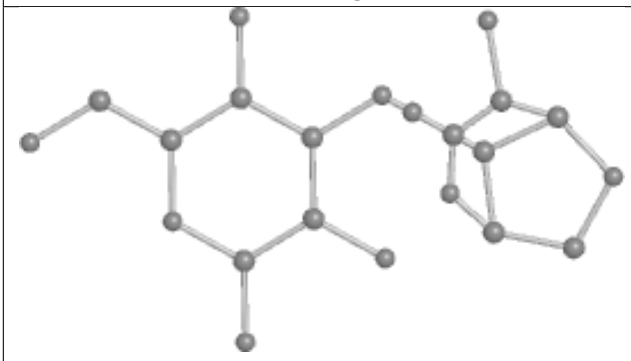
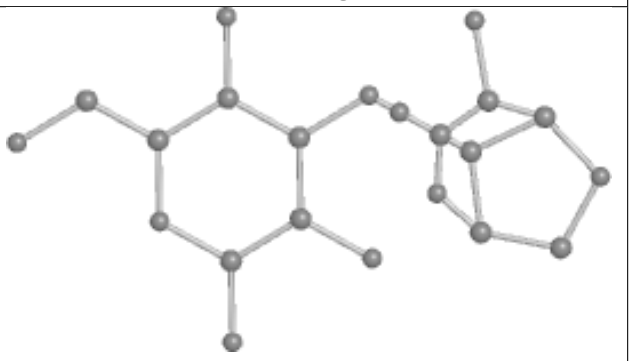
Bond angles

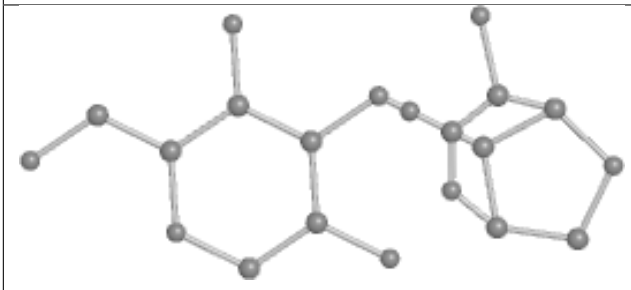
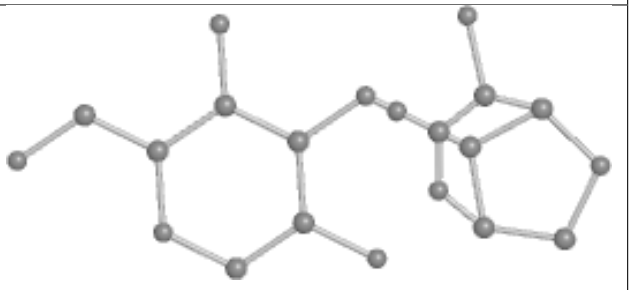
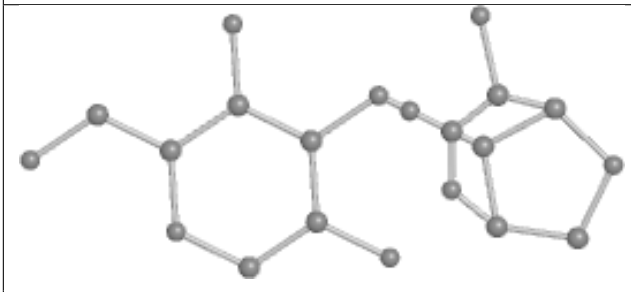
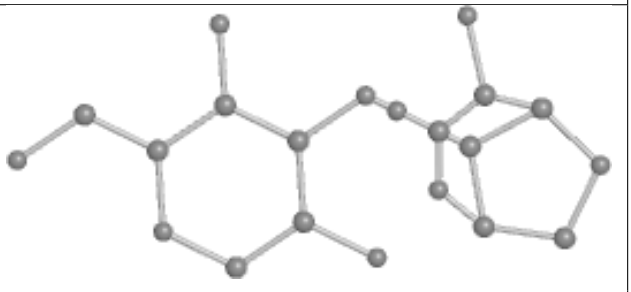


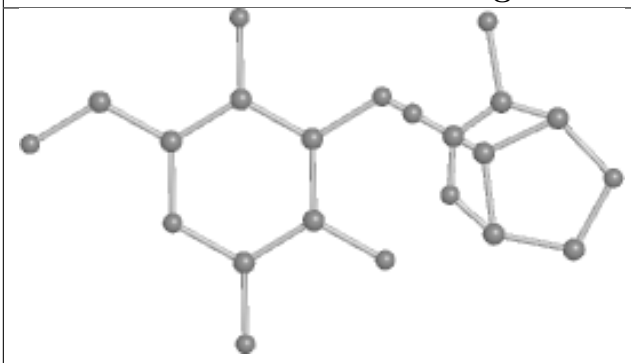
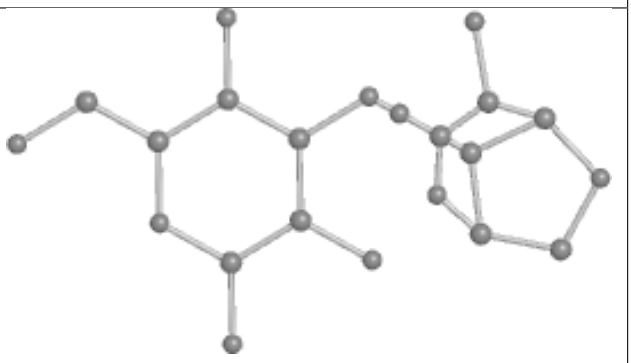
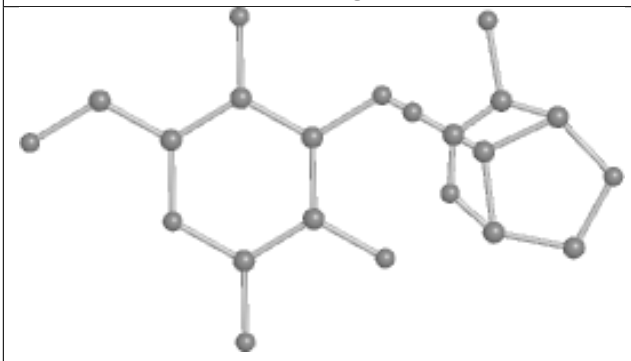
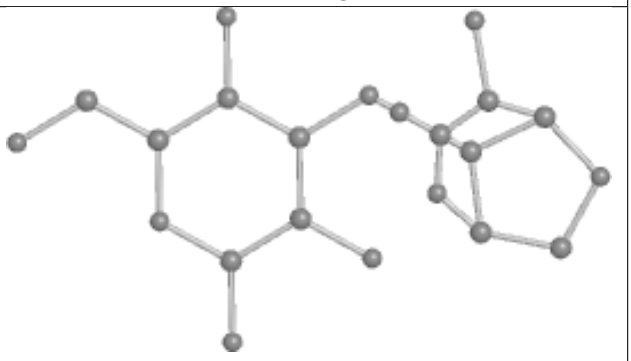
Torsions



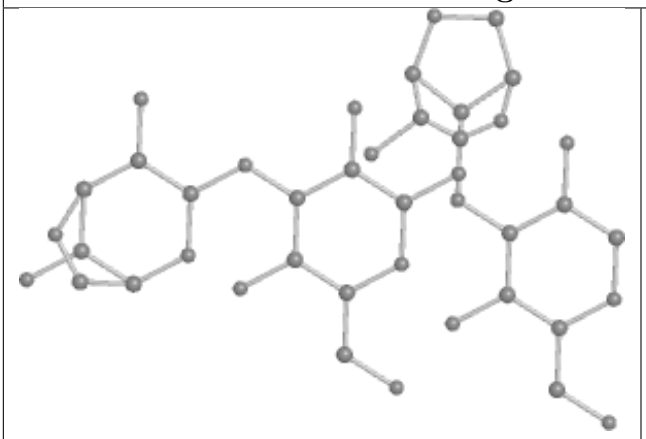
Rings

Oligosaccharide Chain E	
	
Bond lengths	**Bond angles**
	
Torsions	**Rings**

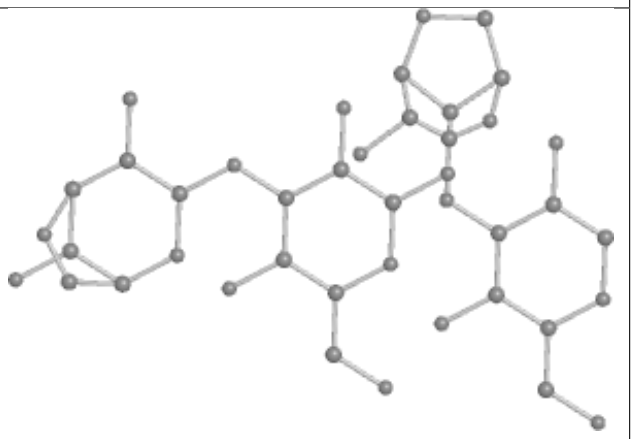
Oligosaccharide Chain G	
	
Bond lengths	**Bond angles**
	
Torsions	**Rings**

Oligosaccharide Chain H	
	
Bond lengths	**Bond angles**
	
Torsions	**Rings**

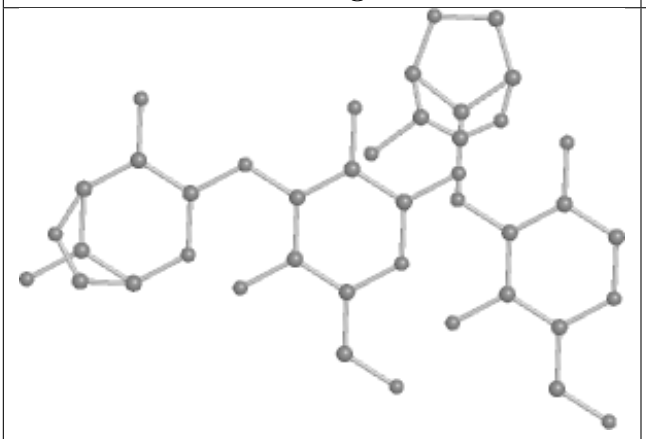
Oligosaccharide Chain F



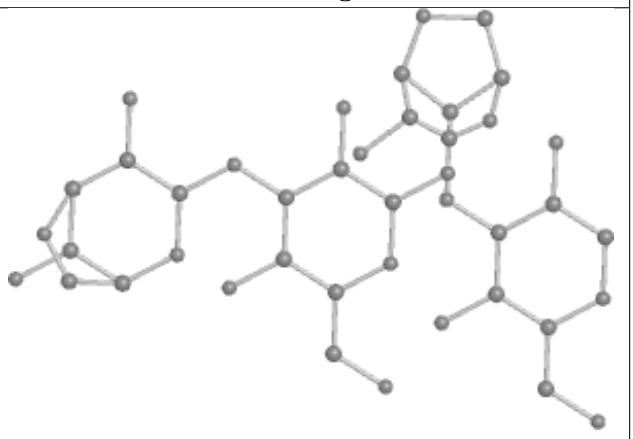
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 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

6.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.