Full wwPDB X-ray Structure Validation Report (i)

Oct 31, 2023 – 04:37 PM EDT

PDB ID	:	3I27
Title	:	Structure of bovine torovirus Hemagglutinin-Esterase in complex with receptor
Authors	:	Zeng, Q.H.; Huizinga, E.G.
Deposited on	:	2009-06-29
Resolution	:	2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

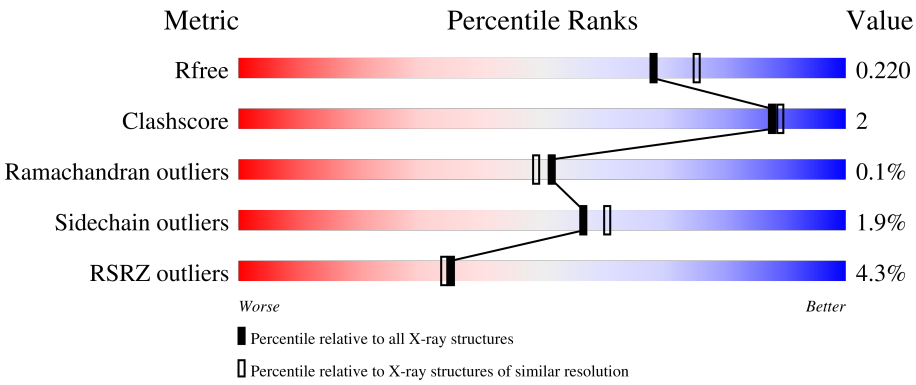
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

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}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 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% 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	384	.% 89% 5% 5%
1	B	384	90% 5% 5%
1	C	384	13% 84% 9% • 7%
1	D	384	2% 93% • •
2	E	3	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	3	100%
2	J	3	33% 67%
3	F	2	100%
3	G	2	50% 50%
3	I	2	50% 50%
3	K	2	50% 50%
3	L	2	50% 50%
4	M	7	43% 57%

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	MAN	M	7	-	-	-	X
5	NAG	B	5154	-	-	-	X

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 12866 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 Hemagglutinin-esterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	365	2862	1821	455	559	27	0	0	0
1	B	364	2855	1819	454	555	27	0	0	0
1	C	358	2813	1795	447	544	27	0	0	0
1	D	369	2889	1837	460	565	27	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

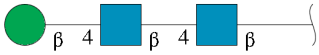
Chain	Residue	Modelled	Actual	Comment	Reference
A	393	ASP	-	expression tag	UNP P0C0V9
A	394	PRO	-	expression tag	UNP P0C0V9
A	395	LEU	-	expression tag	UNP P0C0V9
A	396	VAL	-	expression tag	UNP P0C0V9
A	397	PRO	-	expression tag	UNP P0C0V9
A	398	ARG	-	expression tag	UNP P0C0V9
B	393	ASP	-	expression tag	UNP P0C0V9
B	394	PRO	-	expression tag	UNP P0C0V9
B	395	LEU	-	expression tag	UNP P0C0V9
B	396	VAL	-	expression tag	UNP P0C0V9
B	397	PRO	-	expression tag	UNP P0C0V9
B	398	ARG	-	expression tag	UNP P0C0V9
C	393	ASP	-	expression tag	UNP P0C0V9
C	394	PRO	-	expression tag	UNP P0C0V9
C	395	LEU	-	expression tag	UNP P0C0V9
C	396	VAL	-	expression tag	UNP P0C0V9
C	397	PRO	-	expression tag	UNP P0C0V9
C	398	ARG	-	expression tag	UNP P0C0V9
D	393	ASP	-	expression tag	UNP P0C0V9
D	394	PRO	-	expression tag	UNP P0C0V9
D	395	LEU	-	expression tag	UNP P0C0V9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	396	VAL	-	expression tag	UNP P0C0V9
D	397	PRO	-	expression tag	UNP P0C0V9
D	398	ARG	-	expression tag	UNP P0C0V9

- Molecule 2 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

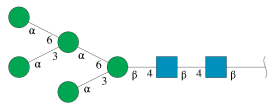


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	E	3	39	22	2	15	0	0	0
2	H	3	39	22	2	15	0	0	0
2	J	3	39	22	2	15	0	0	0

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

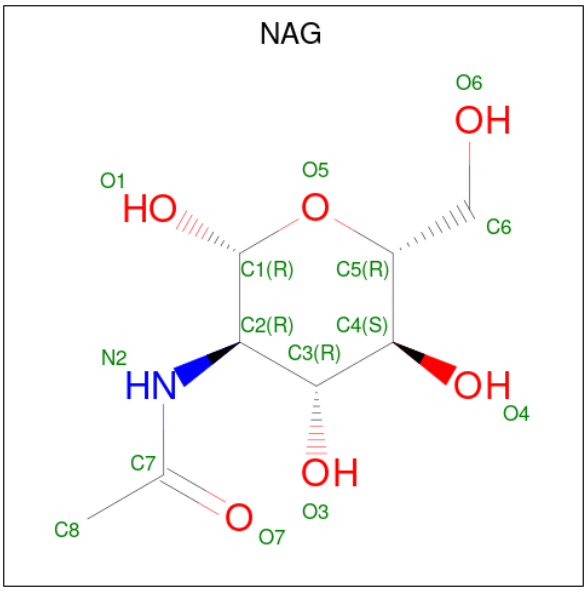
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	F	2	28	16	2	10	0	0	0
3	G	2	28	16	2	10	0	0	0
3	I	2	28	16	2	10	0	0	0
3	K	2	28	16	2	10	0	0	0
3	L	2	28	16	2	10	0	0	0

- Molecule 4 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



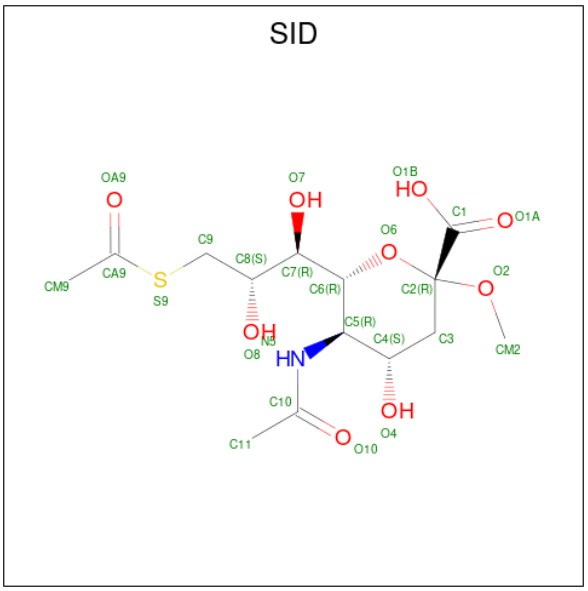
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	M	7	83	46	2	35	0	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 6 is methyl 9-S-acetyl-5-acetamido-3,5-dideoxy-9-thio-D-glycero-alpha-D-galacto-non-2-ulopyranosidonic acid (three-letter code: SID) (formula: $C_{14}H_{23}NO_9S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
6	D	1	25	14	1	9	1	0	0

- Molecule 7 is water.

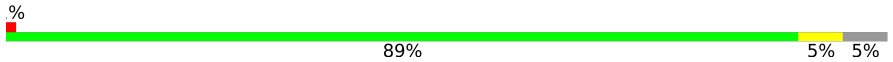
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	347	347	347	0	0
			Total	O		
7	B	338	338	338	0	0
			Total	O		
7	C	65	65	65	0	0
			Total	O		
7	D	262	262	262	0	0

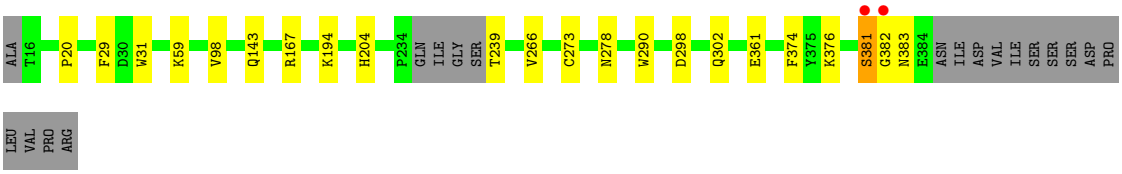
3 Residue-property plots (i)

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 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: Hemagglutinin-esterase

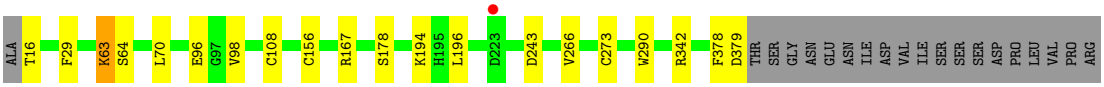
Chain A:





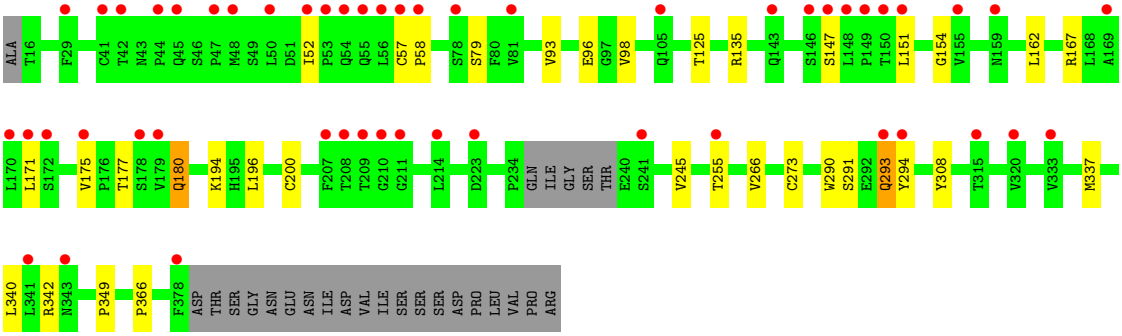
• Molecule 1: Hemagglutinin-esterase

Chain B:



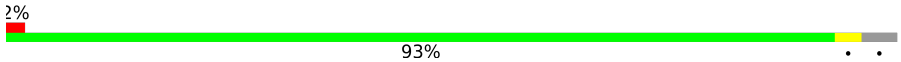
• Molecule 1: Hemagglutinin-esterase

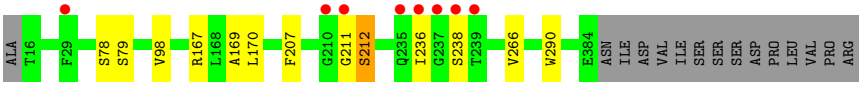
Chain C:



• Molecule 1: Hemagglutinin-esterase

Chain D:





• Molecule 2: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E: 100%

NAG1 NAG2 BMA3

• Molecule 2: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain H: 100%

NAG1 NAG2 BMA3

• Molecule 2: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain J: 33% 67%

NAG1 NAG2 BMA3

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F: 100%

NAG1 NAG2

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G: 50% 50%

NAG1 NAG2

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain I: 50% 50%

NAG1 NAG2

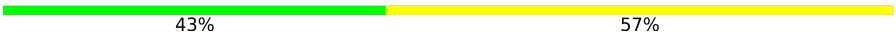
• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain K: 50% 50%

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain L:

• Molecule 4: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain M: 

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	67.63Å 112.89Å 273.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.92 – 2.00 48.92 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.4 (48.92-2.00) 99.4 (48.92-2.00)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.67 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R_{free}	0.179 , 0.210 0.190 , 0.220	Depositor DCC
R_{free} test set	7106 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	22.8	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 49.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	12866	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MAN, NAG, BMA, SID

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	0.68	0/2949	0.70	0/4021
1	B	0.68	0/2943	0.69	1/4014 (0.0%)
1	C	0.48	0/2900	0.57	0/3954
1	D	0.63	0/2977	0.69	0/4060
All	All	0.62	0/11769	0.66	1/16049 (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

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	243	ASP	CB-CG-OD1	5.15	122.93	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	381	SER	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	2862	0	2669	11	0
1	B	2855	0	2672	9	0
1	C	2813	0	2635	14	0
1	D	2889	0	2697	11	0
2	E	39	0	34	0	0
2	H	39	0	34	0	0
2	J	39	0	34	0	0
3	F	28	0	25	0	0
3	G	28	0	25	0	0
3	I	28	0	25	0	0
3	K	28	0	25	0	0
3	L	28	0	25	0	0
4	M	83	0	70	0	0
5	A	28	0	26	0	0
5	B	14	0	13	1	0
5	D	28	0	26	0	0
6	D	25	0	22	2	0
7	A	347	0	0	3	0
7	B	338	0	0	0	0
7	C	65	0	0	0	0
7	D	262	0	0	0	0
All	All	12866	0	11057	45	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 2.

All (45) 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:D:211:GLY:HA3	1:D:212:SER:CB	2.15	0.76
1:D:211:GLY:HA3	1:D:212:SER:HB2	1.78	0.64
1:C:98:VAL:HG22	1:C:290:TRP:CE2	2.33	0.64
1:A:381:SER:HB2	1:A:383:ASN:H	1.64	0.62
1:B:63:LYS:HD3	1:B:64:SER:N	2.20	0.56
1:D:211:GLY:HA3	1:D:212:SER:HB3	1.86	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:29:PHE:HB3	5:B:5154:NAG:H5	1.87	0.55
1:D:167:ARG:O	1:D:266:VAL:HA	2.07	0.55
1:A:204:HIS:ND1	7:A:671:HOH:O	2.32	0.54
1:D:78:SER:O	1:D:79:SER:CB	2.56	0.53
1:B:378:PHE:O	1:B:379:ASP:HB2	2.10	0.52
1:D:207:PHE:CE2	6:D:1001:SID:HM9B	2.48	0.48
1:D:207:PHE:CZ	6:D:1001:SID:HM9B	2.49	0.48
1:C:154:GLY:O	1:C:162:LEU:HD22	2.13	0.48
1:D:78:SER:O	1:D:79:SER:HB3	2.14	0.48
1:A:98:VAL:HG22	1:A:290:TRP:CE2	2.48	0.47
1:C:93:VAL:HG21	1:C:340:LEU:HD21	1.95	0.47
1:D:98:VAL:HG22	1:D:290:TRP:CE2	2.49	0.47
1:C:349:PRO:HD3	1:C:366:PRO:O	2.16	0.46
1:B:63:LYS:HD3	1:B:63:LYS:C	2.36	0.46
1:B:70:LEU:C	1:B:70:LEU:HD23	2.37	0.45
1:B:108:CYS:O	1:B:156:CYS:HB3	2.17	0.45
1:A:381:SER:CB	1:A:383:ASN:H	2.29	0.45
1:C:167:ARG:O	1:C:266:VAL:HA	2.16	0.45
1:A:167:ARG:O	1:A:266:VAL:HA	2.18	0.44
1:D:211:GLY:CA	1:D:212:SER:CB	2.93	0.44
1:B:167:ARG:O	1:B:266:VAL:HA	2.19	0.43
1:C:180:GLN:HG2	1:C:255:THR:CG2	2.48	0.43
1:A:278:ASN:ND2	7:A:1019:HOH:O	2.52	0.43
1:A:382:GLY:HA2	7:A:959:HOH:O	2.18	0.43
1:C:196:LEU:C	1:C:196:LEU:HD23	2.38	0.43
1:B:98:VAL:HG22	1:B:290:TRP:CE2	2.54	0.42
1:C:52:ILE:HG22	1:C:337:MET:HG3	2.01	0.42
1:C:291:SER:OG	1:C:293:GLN:NE2	2.47	0.42
1:C:125:THR:HG23	1:C:308:TYR:OH	2.19	0.42
1:D:169:ALA:C	1:D:170:LEU:HD12	2.41	0.41
1:C:293:GLN:HG2	1:C:294:TYR:CD1	2.55	0.41
1:A:298:ASP:O	1:A:302:GLN:HG2	2.19	0.41
1:A:374:PHE:CZ	1:A:376:LYS:HD2	2.54	0.41
1:B:196:LEU:C	1:B:196:LEU:HD23	2.40	0.41
1:C:57:CYS:N	1:C:58:PRO:CD	2.84	0.41
1:A:20:PRO:HD2	1:A:361:GLU:HG2	2.04	0.40
1:C:151:LEU:HD23	1:C:171:LEU:HD23	2.02	0.40
1:A:31:TRP:CZ2	1:A:59:LYS:HG3	2.56	0.40
1:C:200:CYS:HA	1:C:245:VAL:O	2.22	0.40

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	361/384 (94%)	345 (96%)	16 (4%)	0	100	100
1	B	362/384 (94%)	347 (96%)	15 (4%)	0	100	100
1	C	354/384 (92%)	341 (96%)	13 (4%)	0	100	100
1	D	367/384 (96%)	349 (95%)	17 (5%)	1 (0%)	41	37
All	All	1444/1536 (94%)	1382 (96%)	61 (4%)	1 (0%)	51	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	212	SER

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	A	327/344 (95%)	322 (98%)	5 (2%)	65	69
1	B	326/344 (95%)	319 (98%)	7 (2%)	53	57
1	C	321/344 (93%)	310 (97%)	11 (3%)	37	36
1	D	330/344 (96%)	328 (99%)	2 (1%)	86	90
All	All	1304/1376 (95%)	1279 (98%)	25 (2%)	57	61

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	PHE
1	A	143	GLN
1	A	194	LYS
1	A	239	THR
1	A	273	CYS
1	B	16	THR
1	B	63	LYS
1	B	96	GLU
1	B	178	SER
1	B	194	LYS
1	B	273	CYS
1	B	342	ARG
1	C	79	SER
1	C	96	GLU
1	C	135	ARG
1	C	147	SER
1	C	175	VAL
1	C	177	THR
1	C	180	GLN
1	C	194	LYS
1	C	273	CYS
1	C	293	GLN
1	C	342	ARG
1	D	236	ILE
1	D	238	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	159	ASN
1	A	278	ASN
1	C	180	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)

26 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NAG	E	1	1,2	14,14,15	0.83	1 (7%)	17,19,21	1.23	1 (5%)
2	NAG	E	2	2	14,14,15	0.78	1 (7%)	17,19,21	1.17	2 (11%)
2	BMA	E	3	2	11,11,12	0.53	0	15,15,17	1.20	2 (13%)
3	NAG	F	1	1,3	14,14,15	0.56	0	17,19,21	1.03	1 (5%)
3	NAG	F	2	3	14,14,15	0.63	0	17,19,21	1.31	2 (11%)
3	NAG	G	1	1,3	14,14,15	0.70	0	17,19,21	0.79	0
3	NAG	G	2	3	14,14,15	0.58	0	17,19,21	1.06	2 (11%)
2	NAG	H	1	1,2	14,14,15	0.63	0	17,19,21	1.42	3 (17%)
2	NAG	H	2	2	14,14,15	0.60	0	17,19,21	1.35	3 (17%)
2	BMA	H	3	2	11,11,12	0.51	0	15,15,17	1.53	3 (20%)
3	NAG	I	1	1,3	14,14,15	0.54	0	17,19,21	1.43	3 (17%)
3	NAG	I	2	3	14,14,15	0.69	0	17,19,21	0.69	0
2	NAG	J	1	1,2	14,14,15	0.55	0	17,19,21	1.07	1 (5%)
2	NAG	J	2	2	14,14,15	0.53	0	17,19,21	1.12	1 (5%)
2	BMA	J	3	2	11,11,12	0.45	0	15,15,17	0.78	0
3	NAG	K	1	1,3	14,14,15	0.96	0	17,19,21	1.57	3 (17%)
3	NAG	K	2	3	14,14,15	0.58	0	17,19,21	0.95	0
3	NAG	L	1	1,3	14,14,15	0.86	0	17,19,21	1.23	1 (5%)
3	NAG	L	2	3	14,14,15	0.50	0	17,19,21	0.84	0
4	NAG	M	1	1,4	14,14,15	0.81	0	17,19,21	1.21	1 (5%)
4	NAG	M	2	4	14,14,15	0.80	0	17,19,21	0.99	0
4	BMA	M	3	4	11,11,12	0.27	0	15,15,17	1.06	1 (6%)
4	MAN	M	4	4	11,11,12	0.70	0	15,15,17	1.32	1 (6%)
4	MAN	M	5	4	11,11,12	0.64	0	15,15,17	0.84	0
4	MAN	M	6	4	11,11,12	0.54	0	15,15,17	1.07	0
4	MAN	M	7	4	11,11,12	0.55	0	15,15,17	0.93	1 (6%)

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	NAG	E	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	E	2	2	-	2/6/23/26	0/1/1/1
2	BMA	E	3	2	-	2/2/19/22	0/1/1/1
3	NAG	F	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	0/6/23/26	0/1/1/1
3	NAG	G	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	G	2	3	-	1/6/23/26	0/1/1/1
2	NAG	H	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	H	2	2	-	0/6/23/26	0/1/1/1
2	BMA	H	3	2	-	0/2/19/22	0/1/1/1
3	NAG	I	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	I	2	3	-	0/6/23/26	0/1/1/1
2	NAG	J	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	J	2	2	-	0/6/23/26	0/1/1/1
2	BMA	J	3	2	-	0/2/19/22	0/1/1/1
3	NAG	K	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	K	2	3	-	0/6/23/26	0/1/1/1
3	NAG	L	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	L	2	3	-	2/6/23/26	0/1/1/1
4	NAG	M	1	1,4	-	0/6/23/26	0/1/1/1
4	NAG	M	2	4	-	0/6/23/26	0/1/1/1
4	BMA	M	3	4	-	0/2/19/22	0/1/1/1
4	MAN	M	4	4	-	0/2/19/22	0/1/1/1
4	MAN	M	5	4	-	2/2/19/22	0/1/1/1
4	MAN	M	6	4	-	0/2/19/22	0/1/1/1
4	MAN	M	7	4	-	1/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	1	NAG	O5-C1	-2.48	1.39	1.43
2	E	2	NAG	O5-C1	-2.23	1.40	1.43

All (32) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	K	1	NAG	O5-C1-C2	-4.02	104.93	111.29
2	H	2	NAG	C1-O5-C5	3.62	117.09	112.19
2	E	1	NAG	C1-O5-C5	3.53	116.98	112.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	I	1	NAG	C1-C2-N2	3.47	116.42	110.49
2	H	1	NAG	C1-O5-C5	3.33	116.71	112.19
3	K	1	NAG	C1-O5-C5	3.32	116.69	112.19
2	E	2	NAG	O5-C5-C6	3.26	112.32	107.20
4	M	4	MAN	C1-O5-C5	3.10	116.39	112.19
2	H	3	BMA	O5-C5-C6	3.09	112.05	107.20
2	H	3	BMA	C1-C2-C3	3.02	113.38	109.67
2	H	1	NAG	C1-C2-N2	-2.99	105.37	110.49
3	F	2	NAG	C4-C3-C2	-2.93	106.72	111.02
2	J	1	NAG	C1-O5-C5	2.85	116.06	112.19
2	J	2	NAG	O5-C5-C6	2.81	111.61	107.20
3	F	2	NAG	C1-C2-N2	2.62	114.96	110.49
4	M	3	BMA	O5-C1-C2	-2.57	106.80	110.77
3	F	1	NAG	O5-C1-C2	-2.56	107.24	111.29
3	G	2	NAG	O6-C6-C5	-2.43	102.95	111.29
4	M	1	NAG	C3-C4-C5	-2.37	106.00	110.24
4	M	7	MAN	C1-C2-C3	2.31	112.50	109.67
3	L	1	NAG	O3-C3-C2	-2.29	104.72	109.47
2	E	2	NAG	O5-C1-C2	-2.29	107.67	111.29
3	K	1	NAG	O4-C4-C5	-2.23	103.75	109.30
2	E	3	BMA	C2-C3-C4	-2.18	107.12	110.89
3	I	1	NAG	O3-C3-C2	-2.16	104.99	109.47
2	H	2	NAG	O7-C7-N2	2.16	125.92	121.95
3	I	1	NAG	C1-O5-C5	2.13	115.08	112.19
2	E	3	BMA	O3-C3-C2	2.11	114.04	109.99
2	H	2	NAG	C2-N2-C7	2.11	125.91	122.90
3	G	2	NAG	C3-C4-C5	-2.07	106.55	110.24
2	H	3	BMA	C3-C4-C5	-2.05	106.58	110.24
2	H	1	NAG	O3-C3-C4	2.05	115.09	110.35

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	M	5	MAN	O5-C5-C6-O6
2	E	2	NAG	C4-C5-C6-O6
2	E	2	NAG	O5-C5-C6-O6
4	M	5	MAN	C4-C5-C6-O6
3	L	2	NAG	O5-C5-C6-O6
4	M	7	MAN	O5-C5-C6-O6
3	G	2	NAG	C4-C5-C6-O6
2	E	3	BMA	C4-C5-C6-O6

Continued on next page...

Continued from previous page...

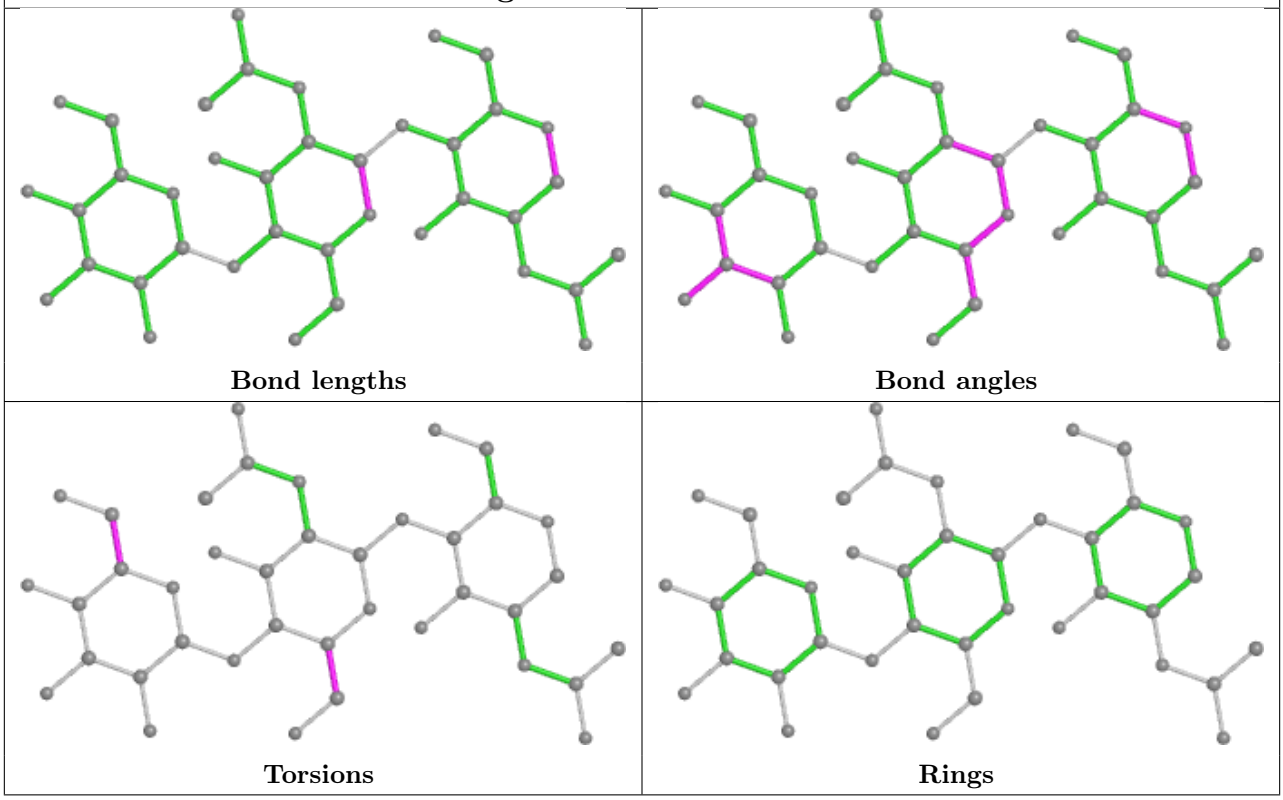
Mol	Chain	Res	Type	Atoms
3	L	2	NAG	C4-C5-C6-O6
2	E	3	BMA	O5-C5-C6-O6

There are no ring outliers.

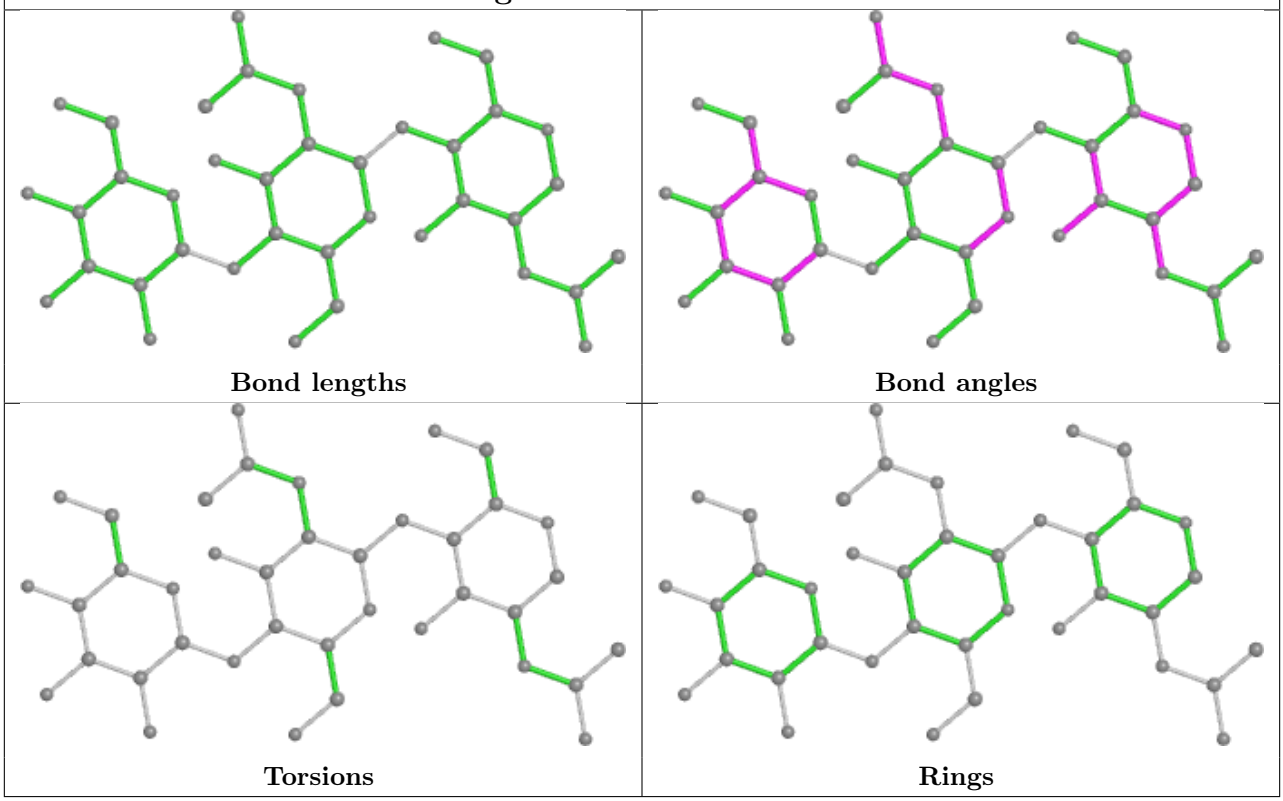
No monomer is involved in short contacts.

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

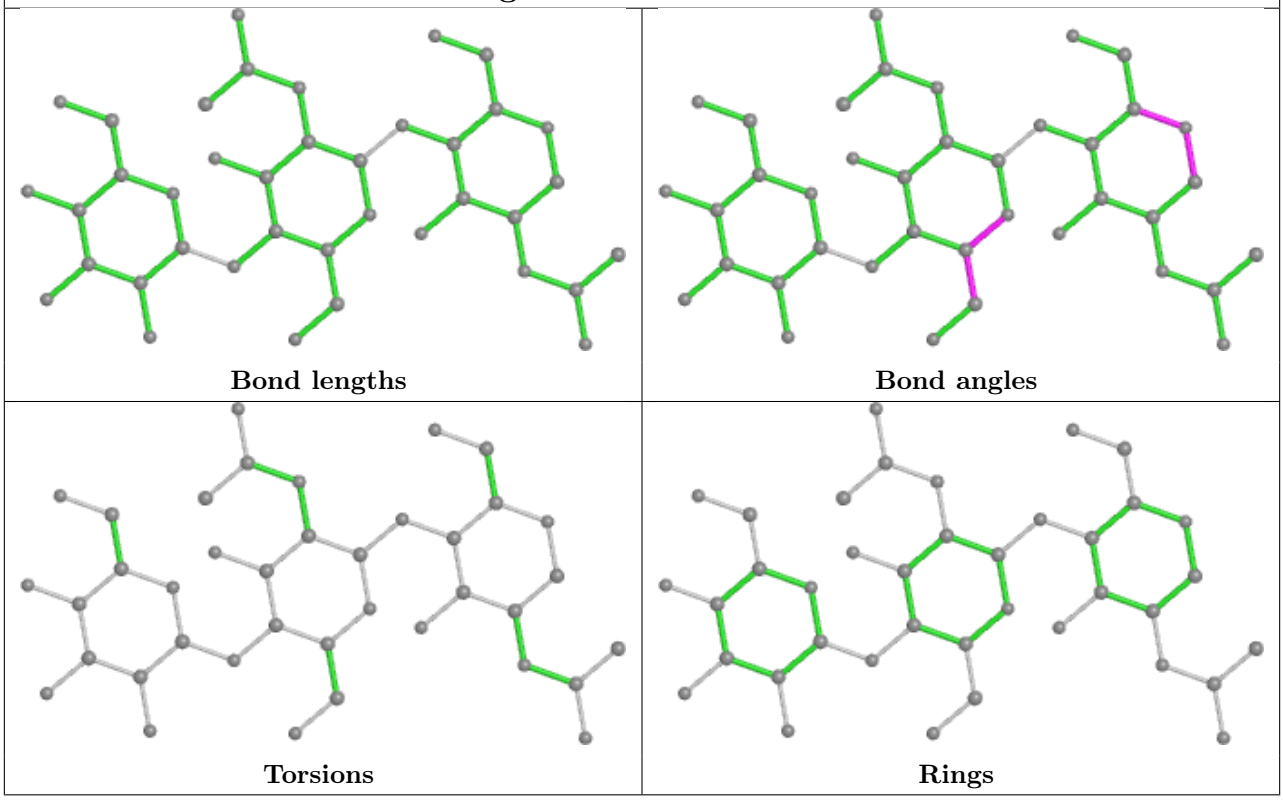
Oligosaccharide Chain E



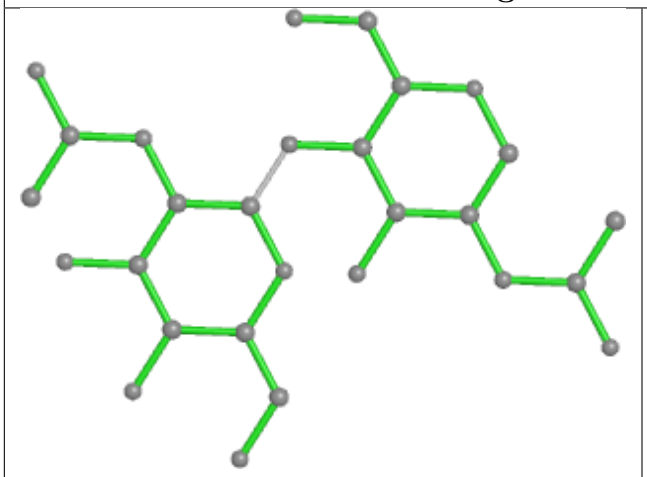
Oligosaccharide Chain H



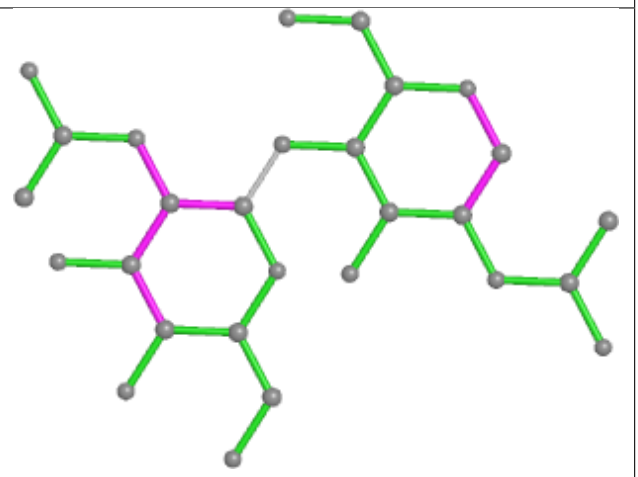
Oligosaccharide Chain J



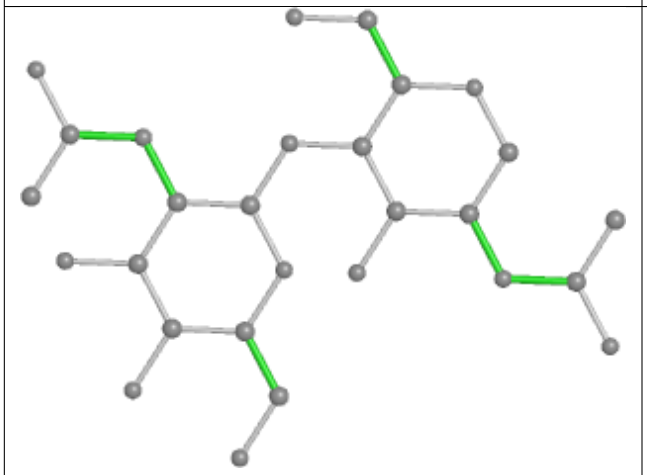
Oligosaccharide Chain F



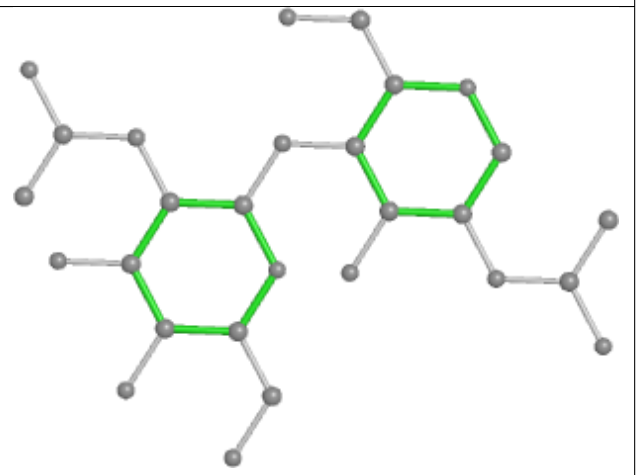
Bond lengths



Bond angles

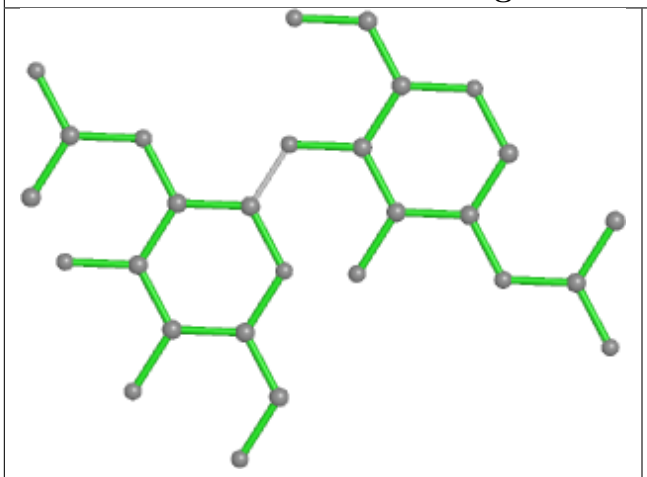


Torsions

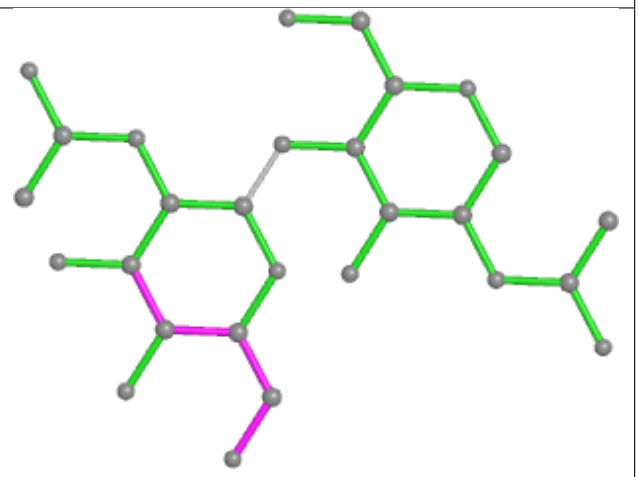


Rings

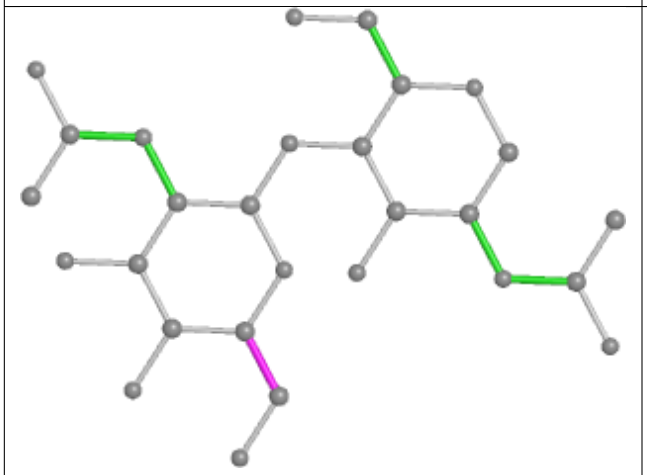
Oligosaccharide Chain G



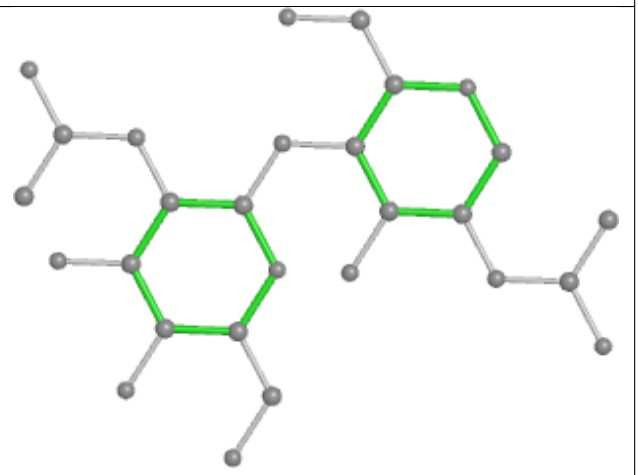
Bond lengths



Bond angles

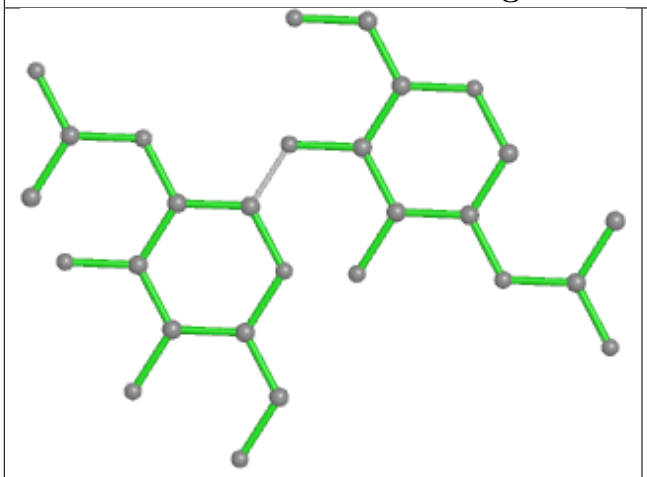


Torsions

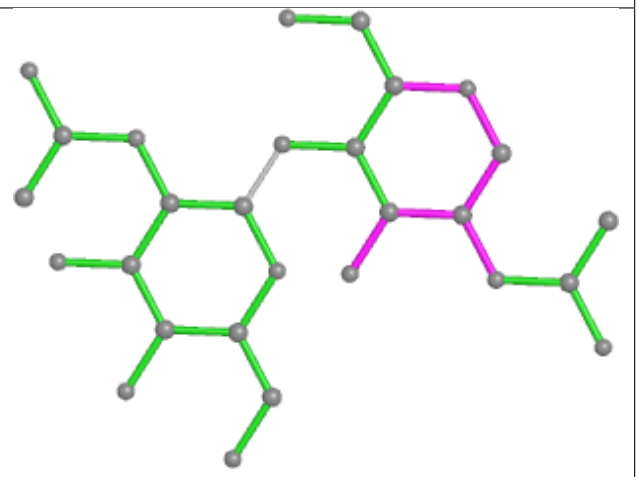


Rings

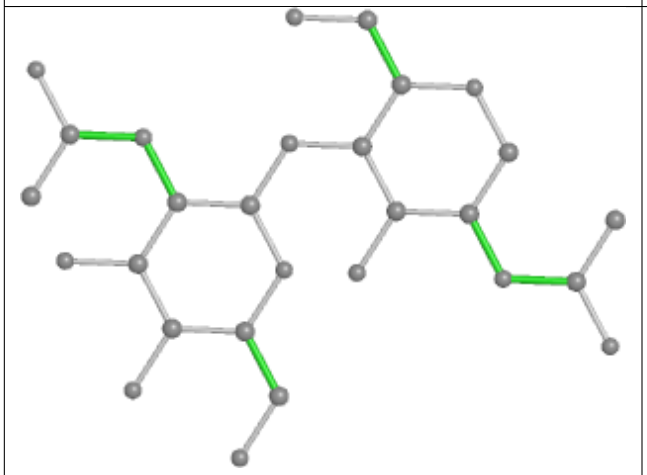
Oligosaccharide Chain I



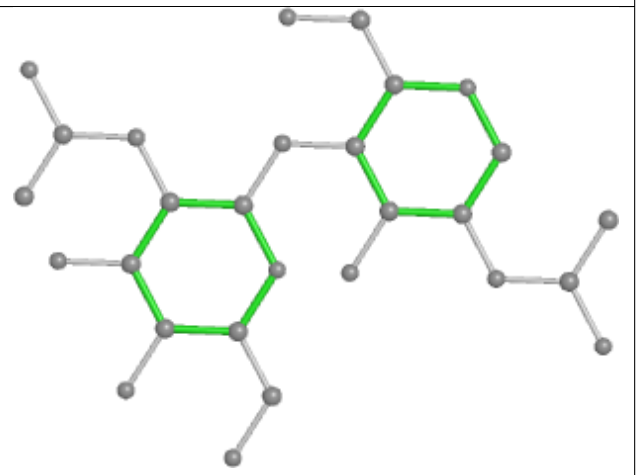
Bond lengths



Bond angles

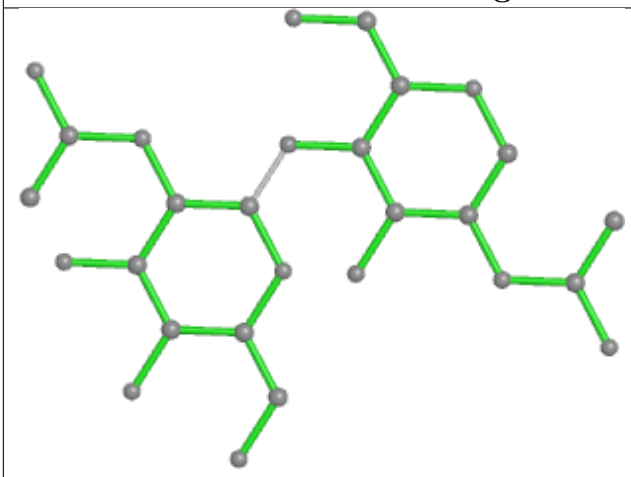


Torsions

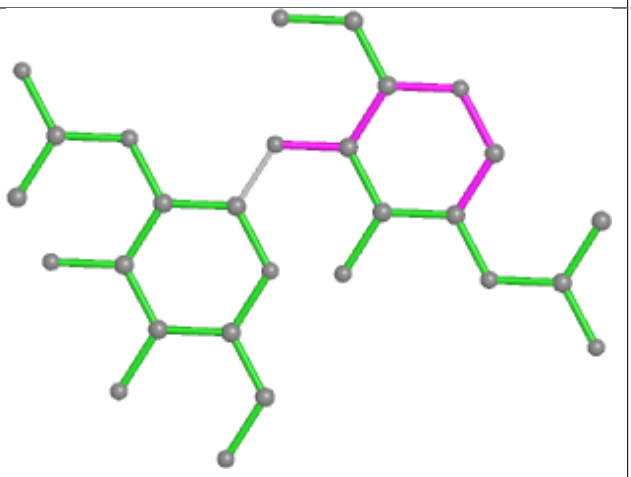


Rings

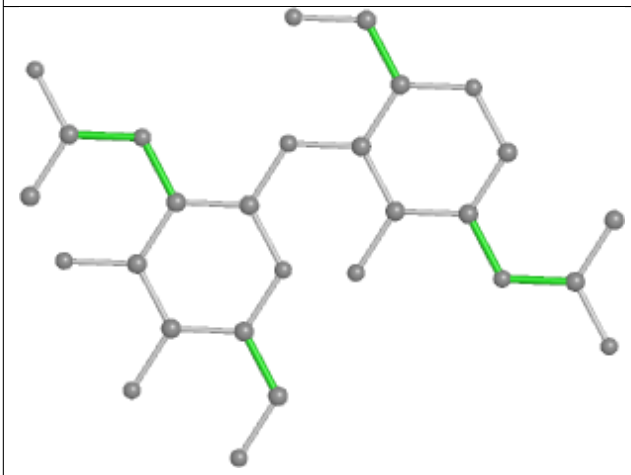
Oligosaccharide Chain K



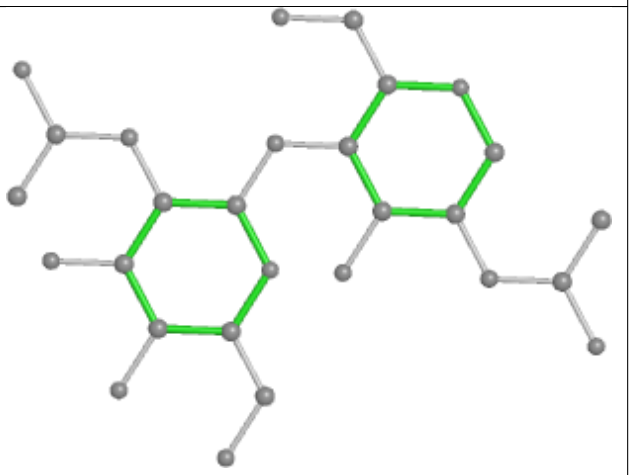
Bond lengths



Bond angles

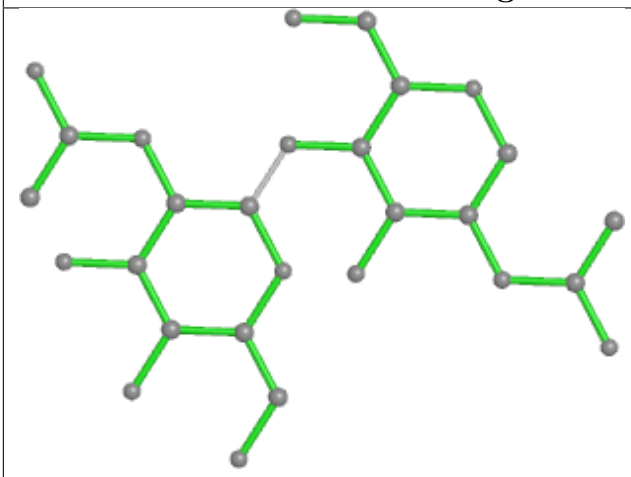


Torsions

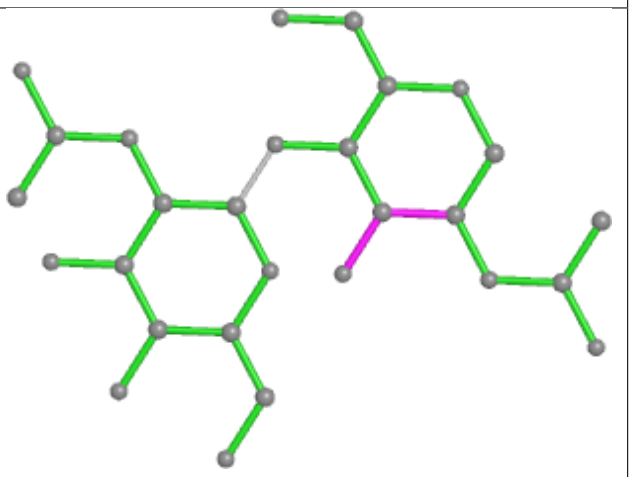


Rings

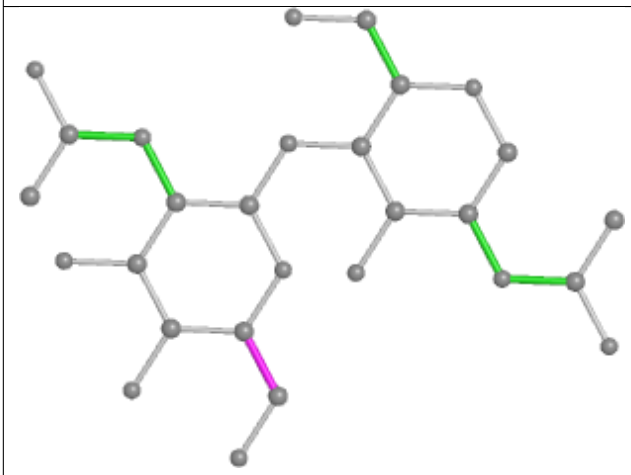
Oligosaccharide Chain L



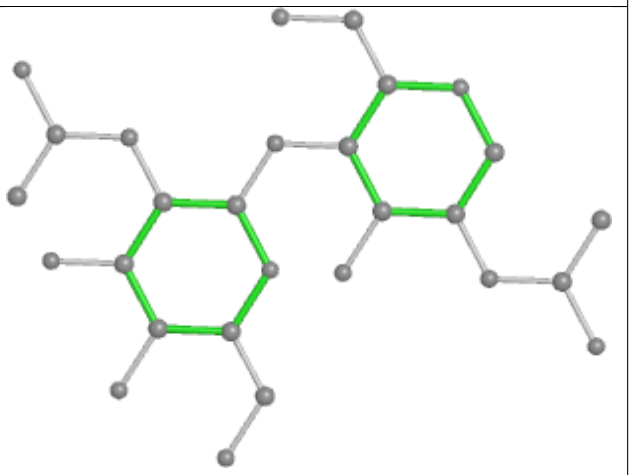
Bond lengths



Bond angles

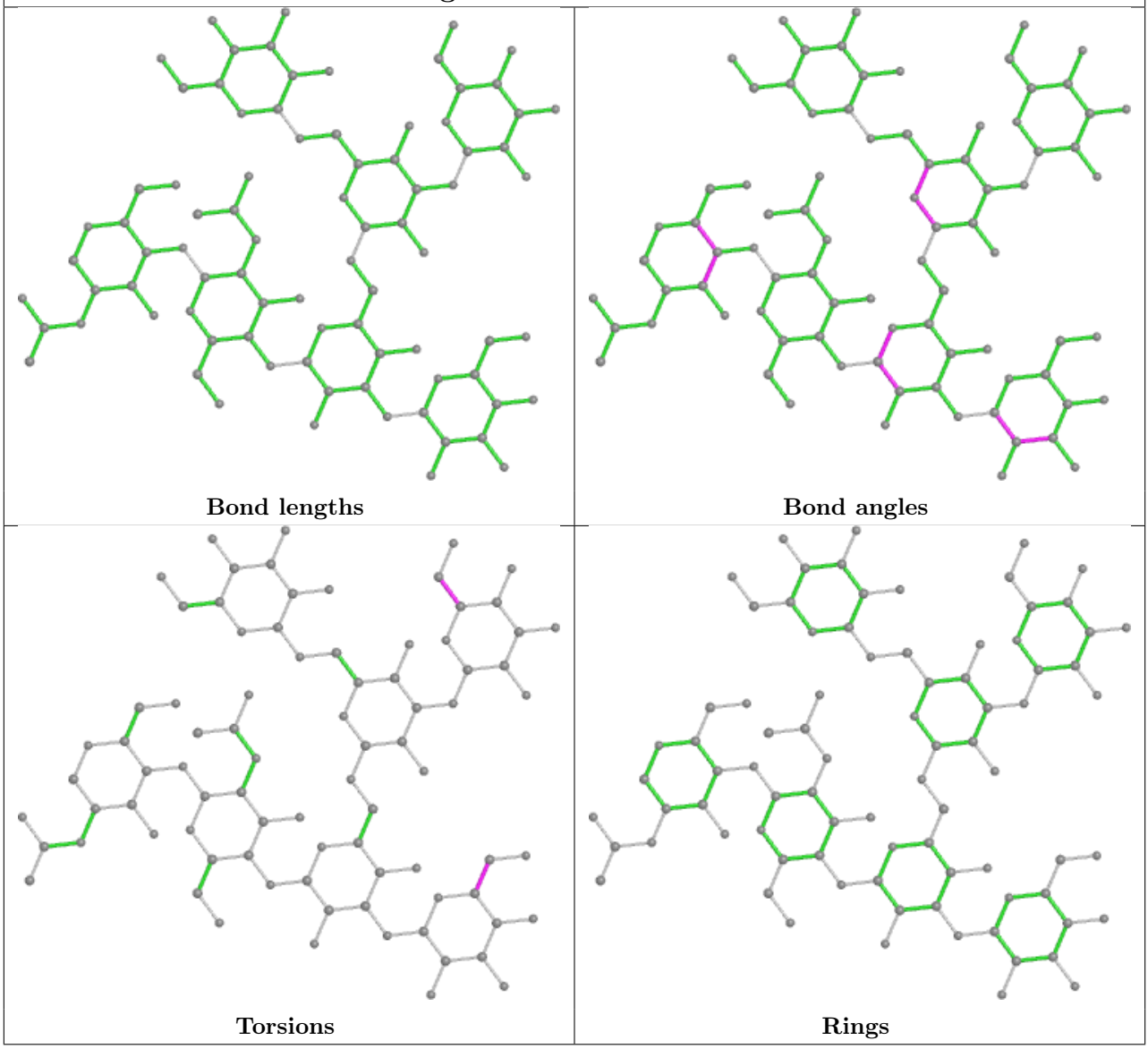


Torsions



Rings

Oligosaccharide Chain M



5.6 Ligand geometry (i)

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	$\#\|Z\| > 2$
5	NAG	A	5156	1	14,14,15	0.78	1 (7%)	17,19,21	1.59	4 (23%)
6	SID	D	1001	-	24,25,25	1.24	3 (12%)	28,36,36	1.13	1 (3%)
5	NAG	D	5157	1	14,14,15	0.54	0	17,19,21	0.80	0
5	NAG	A	5157	1	14,14,15	0.49	0	17,19,21	1.06	1 (5%)
5	NAG	D	5156	1	14,14,15	0.44	0	17,19,21	1.61	3 (17%)
5	NAG	B	5154	1	14,14,15	0.66	0	17,19,21	1.25	2 (11%)

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
5	NAG	A	5156	1	-	2/6/23/26	0/1/1/1
6	SID	D	1001	-	-	5/26/44/44	0/1/1/1
5	NAG	D	5157	1	-	2/6/23/26	0/1/1/1
5	NAG	A	5157	1	-	2/6/23/26	0/1/1/1
5	NAG	D	5156	1	-	2/6/23/26	0/1/1/1
5	NAG	B	5154	1	-	0/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	1001	SID	O6-C2	2.93	1.46	1.42
6	D	1001	SID	C3-C2	2.78	1.56	1.52
5	A	5156	NAG	C1-C2	2.35	1.55	1.52
6	D	1001	SID	C3-C4	2.20	1.56	1.53

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	D	5156	NAG	C1-O5-C5	4.49	118.27	112.19
6	D	1001	SID	C8-C7-C6	-3.27	106.83	113.03
5	A	5157	NAG	O5-C5-C6	3.17	112.17	107.20
5	A	5156	NAG	C1-O5-C5	2.82	116.01	112.19
5	A	5156	NAG	C3-C4-C5	-2.65	105.52	110.24
5	B	5154	NAG	O5-C5-C6	2.62	111.31	107.20
5	B	5154	NAG	C1-C2-N2	-2.61	106.03	110.49
5	A	5156	NAG	O7-C7-C8	-2.32	117.75	122.06
5	A	5156	NAG	O5-C5-C6	2.30	110.82	107.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	D	5156	NAG	C3-C4-C5	2.17	114.12	110.24
5	D	5156	NAG	C6-C5-C4	-2.06	108.17	113.00

There are no chirality outliers.

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	D	1001	SID	O1B-C1-C2-O2
6	D	1001	SID	C7-C8-C9-S9
6	D	1001	SID	O8-C8-C9-S9
6	D	1001	SID	OA9-CA9-S9-C9
6	D	1001	SID	CM9-CA9-S9-C9
5	D	5157	NAG	C4-C5-C6-O6
5	D	5157	NAG	O5-C5-C6-O6
5	A	5157	NAG	O5-C5-C6-O6
5	D	5156	NAG	C4-C5-C6-O6
5	A	5156	NAG	O5-C5-C6-O6
5	A	5157	NAG	C4-C5-C6-O6
5	A	5156	NAG	C4-C5-C6-O6
5	D	5156	NAG	O5-C5-C6-O6

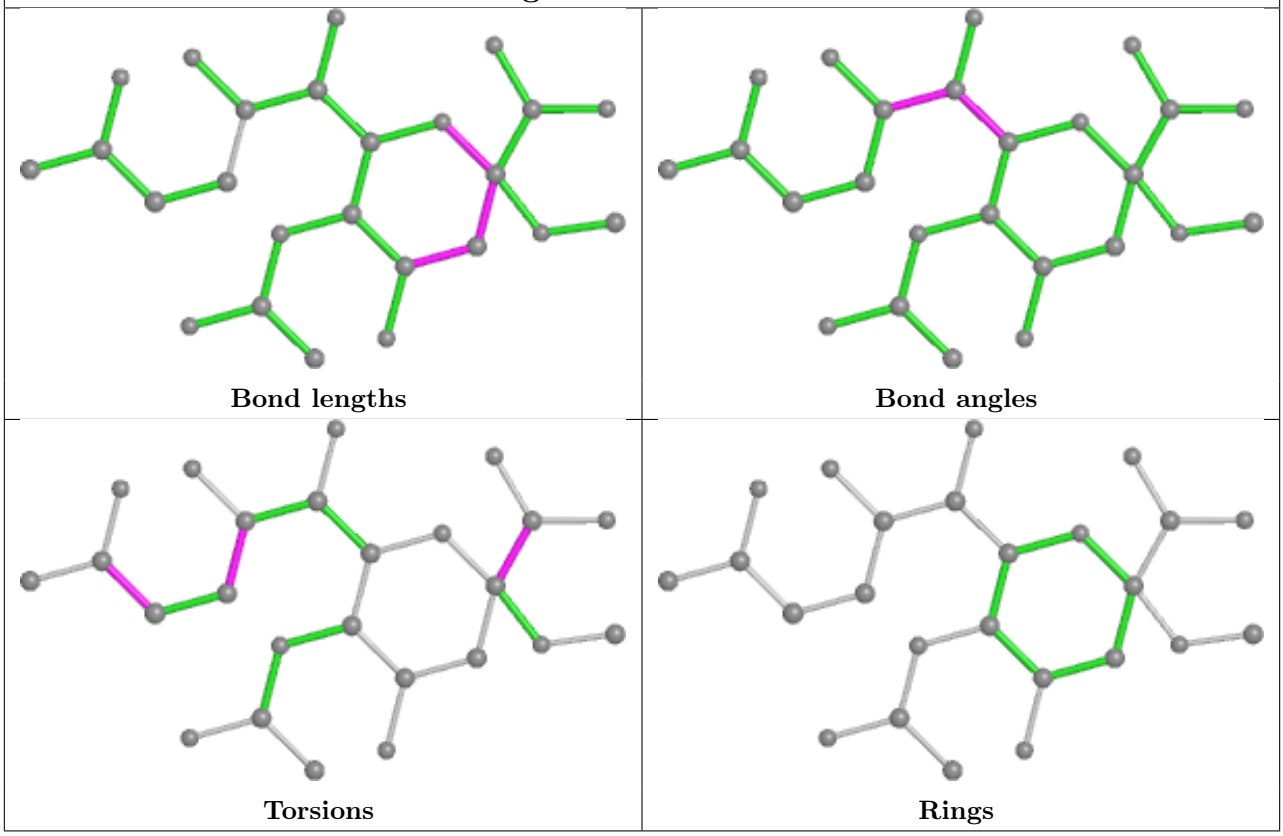
There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	1001	SID	2	0
5	B	5154	NAG	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand SID D 1001



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	365/384 (95%)	-0.24	2 (0%)	91	90	9, 13, 21, 35	0
1	B	364/384 (94%)	-0.20	1 (0%)	94	93	8, 13, 21, 30	0
1	C	358/384 (93%)	0.84	51 (14%)	2	2	15, 25, 32, 37	0
1	D	369/384 (96%)	-0.02	8 (2%)	62	60	10, 17, 28, 45	0
All	All	1456/1536 (94%)	0.09	62 (4%)	35	34	8, 17, 29, 45	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	237	GLY	8.0
1	D	236	ILE	7.9
1	D	238	SER	6.5
1	D	239	THR	6.5
1	C	146	SER	6.0
1	C	147	SER	5.2
1	C	47	PRO	5.1
1	C	148	LEU	5.0
1	C	78	SER	4.8
1	C	179	VAL	4.6
1	C	41	CYS	4.2
1	C	42	THR	4.2
1	C	211	GLY	4.2
1	C	48	MET	3.8
1	D	235	GLN	3.7
1	C	170	LEU	3.5
1	C	209	THR	3.5
1	C	320	VAL	3.4
1	C	210	GLY	3.4
1	C	29	PHE	3.2
1	C	149	PRO	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	57	CYS	3.1
1	C	53	PRO	3.1
1	C	56	LEU	3.0
1	C	58	PRO	2.9
1	D	210	GLY	2.8
1	C	44	PRO	2.8
1	C	143	GLN	2.7
1	C	315	THR	2.7
1	A	382	GLY	2.7
1	C	172	SER	2.6
1	C	378	PHE	2.5
1	C	150	THR	2.5
1	C	343	ASN	2.5
1	C	171	LEU	2.5
1	C	105	GLN	2.5
1	C	208	THR	2.5
1	C	55	GLN	2.5
1	C	178	SER	2.5
1	D	29	PHE	2.4
1	C	54	GLN	2.4
1	C	151	LEU	2.4
1	C	341	LEU	2.4
1	C	175	VAL	2.3
1	A	381	SER	2.3
1	C	241	SER	2.3
1	C	207	PHE	2.3
1	C	50	LEU	2.3
1	C	214	LEU	2.3
1	C	155	VAL	2.2
1	C	293	GLN	2.2
1	C	294	TYR	2.2
1	C	45	GLN	2.2
1	C	81	VAL	2.2
1	C	169	ALA	2.1
1	B	223	ASP	2.1
1	C	223	ASP	2.0
1	C	333	VAL	2.0
1	D	211	GLY	2.0
1	C	159	ASN	2.0
1	C	52	ILE	2.0
1	C	255	THR	2.0

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)

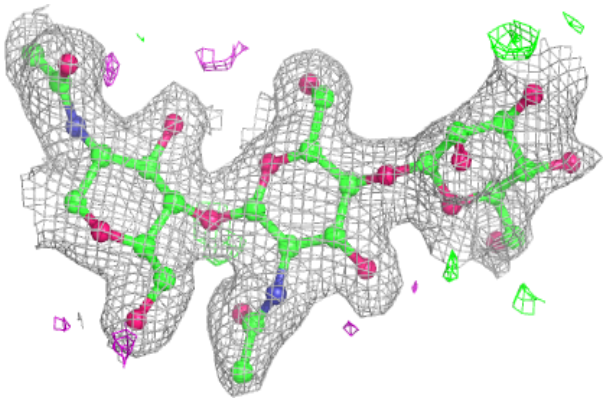
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

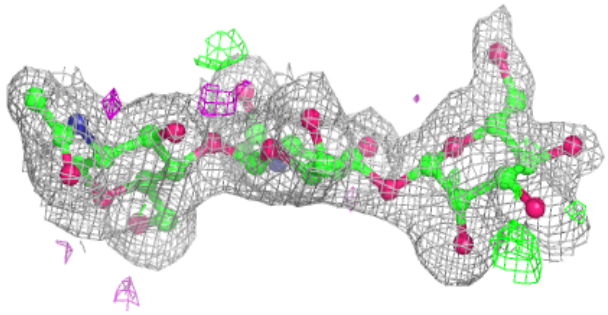
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	BMA	J	3	11/12	0.64	0.36	64,66,67,67	0
4	MAN	M	7	11/12	0.68	0.46	63,66,66,67	0
2	BMA	E	3	11/12	0.79	0.20	41,47,50,50	0
2	NAG	J	2	14/15	0.82	0.30	55,57,60,63	0
4	BMA	M	3	11/12	0.83	0.26	49,51,54,59	0
2	BMA	H	3	11/12	0.83	0.16	44,48,49,52	0
3	NAG	F	2	14/15	0.85	0.23	38,45,49,49	0
4	MAN	M	5	11/12	0.87	0.26	36,43,45,48	0
4	MAN	M	4	11/12	0.88	0.23	46,47,51,53	0
2	NAG	J	1	14/15	0.89	0.20	38,42,46,50	0
4	MAN	M	6	11/12	0.89	0.26	53,55,56,57	0
4	NAG	M	2	14/15	0.89	0.15	31,36,40,46	0
3	NAG	G	2	14/15	0.90	0.14	30,35,39,41	0
3	NAG	L	2	14/15	0.90	0.21	40,42,47,49	0
3	NAG	I	2	14/15	0.92	0.13	34,36,42,42	0
2	NAG	H	2	14/15	0.92	0.15	31,34,41,42	0
3	NAG	L	1	14/15	0.94	0.21	33,36,37,37	0
2	NAG	H	1	14/15	0.95	0.12	22,24,26,29	0
4	NAG	M	1	14/15	0.95	0.10	22,25,29,32	0
3	NAG	K	2	14/15	0.95	0.09	29,31,34,35	0
2	NAG	E	2	14/15	0.95	0.11	27,32,35,40	0
3	NAG	F	1	14/15	0.96	0.10	20,24,27,31	0
3	NAG	G	1	14/15	0.97	0.10	19,22,25,27	0
3	NAG	K	1	14/15	0.97	0.08	15,18,25,27	0
2	NAG	E	1	14/15	0.97	0.09	17,20,26,28	0
3	NAG	I	1	14/15	0.97	0.09	19,22,27,30	0

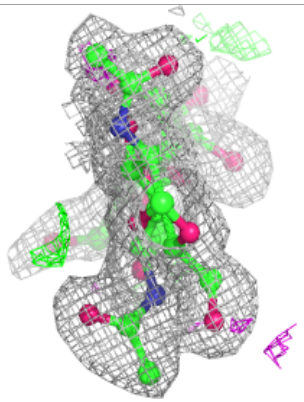
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain E:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

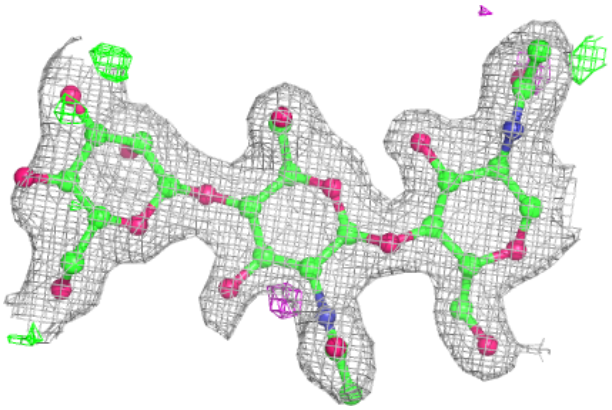


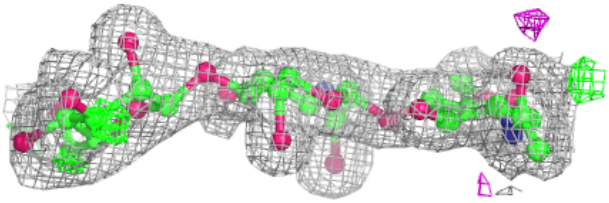




Electron density around Chain H:

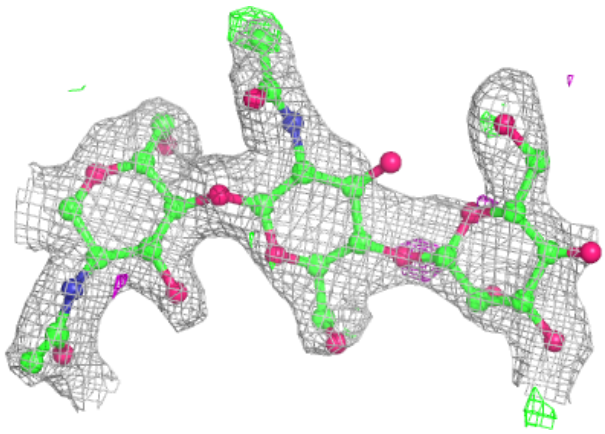
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

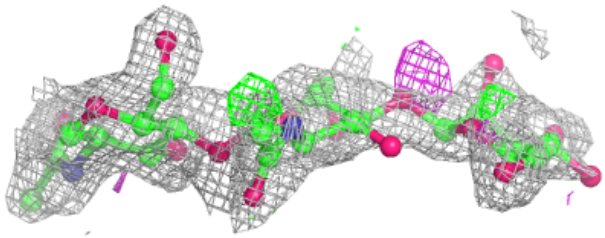


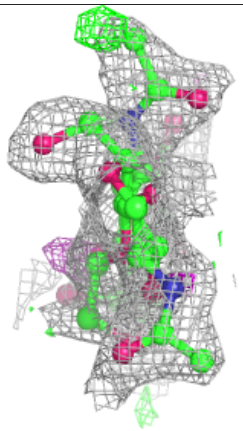


Electron density around Chain J:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

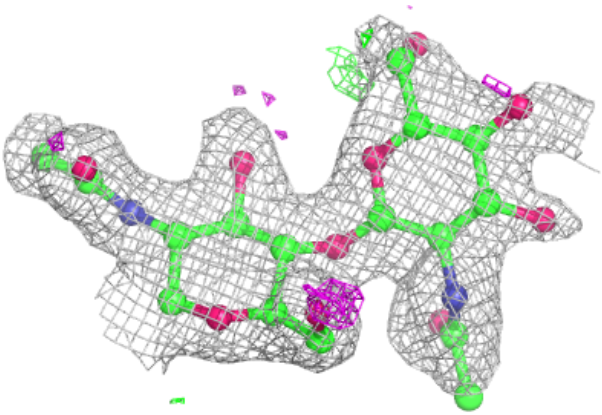


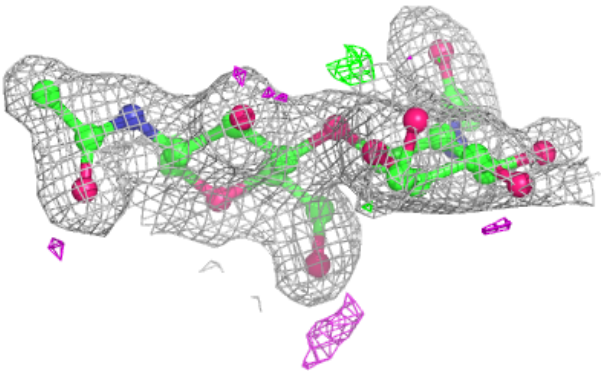


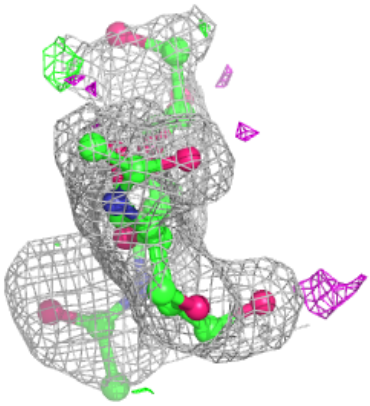


Electron density around Chain F:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

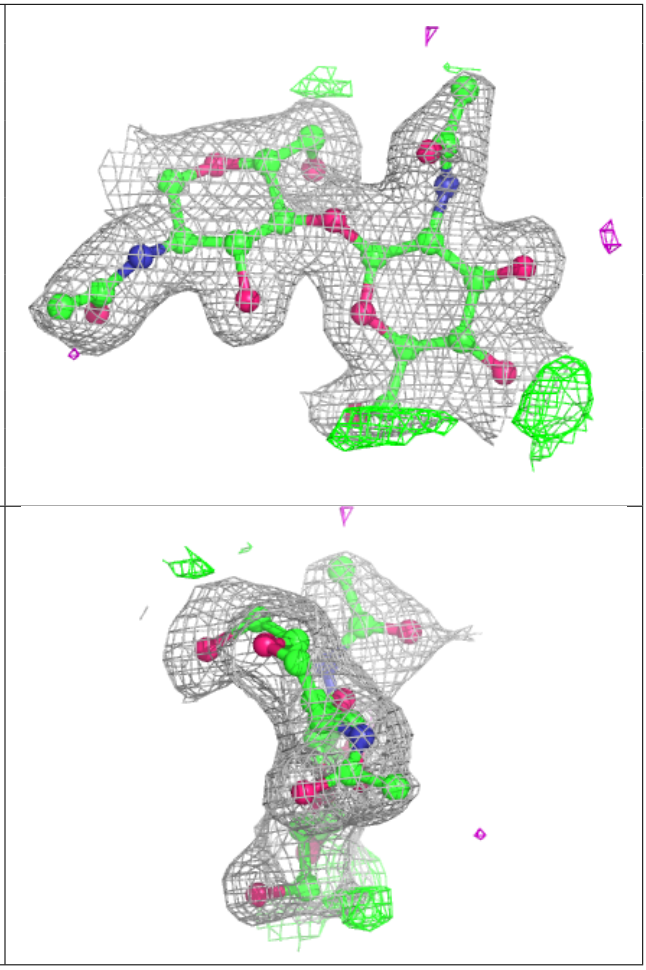






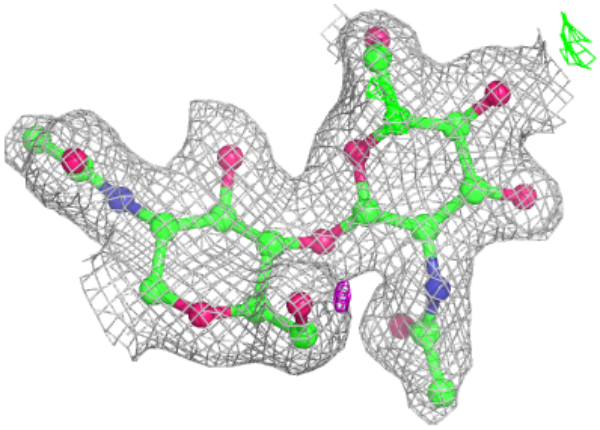
Electron density around Chain G:

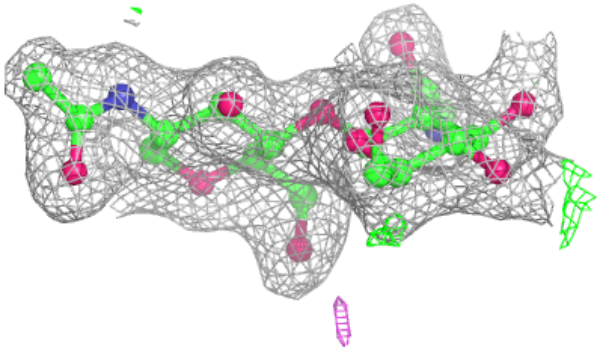
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

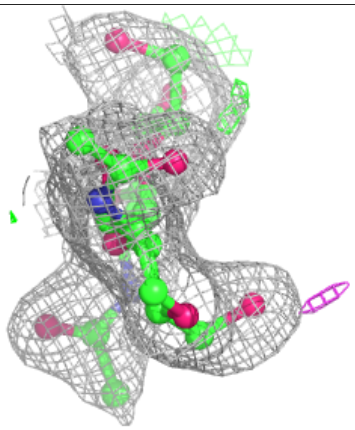


Electron density around Chain I:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

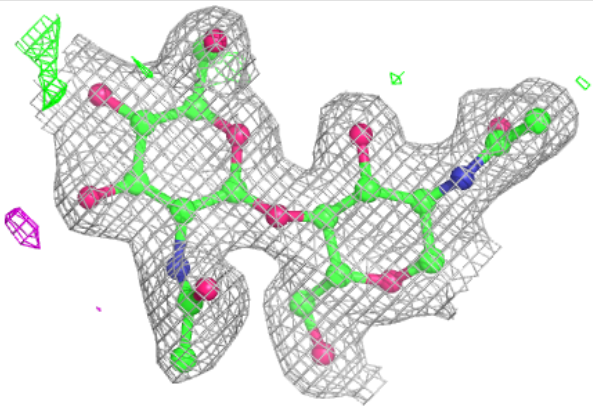


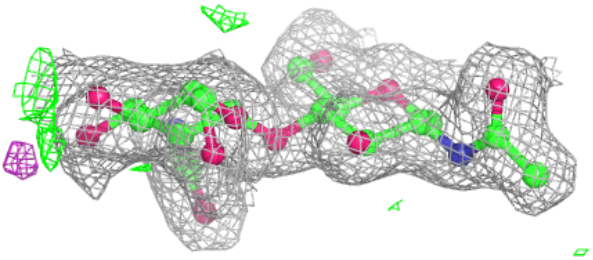


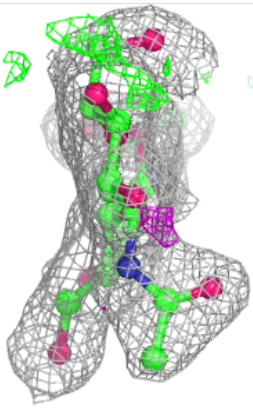


Electron density around Chain K:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

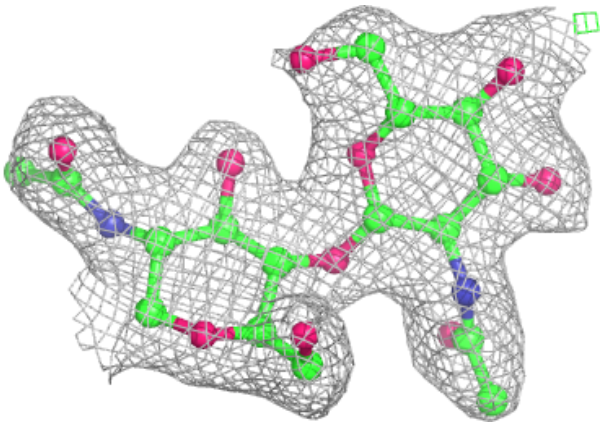


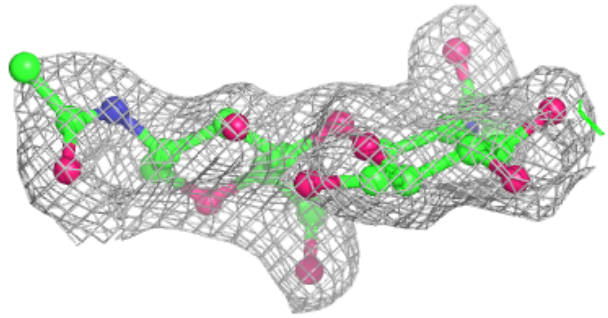


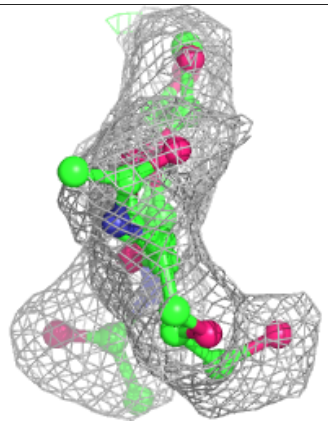


Electron density around Chain L:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

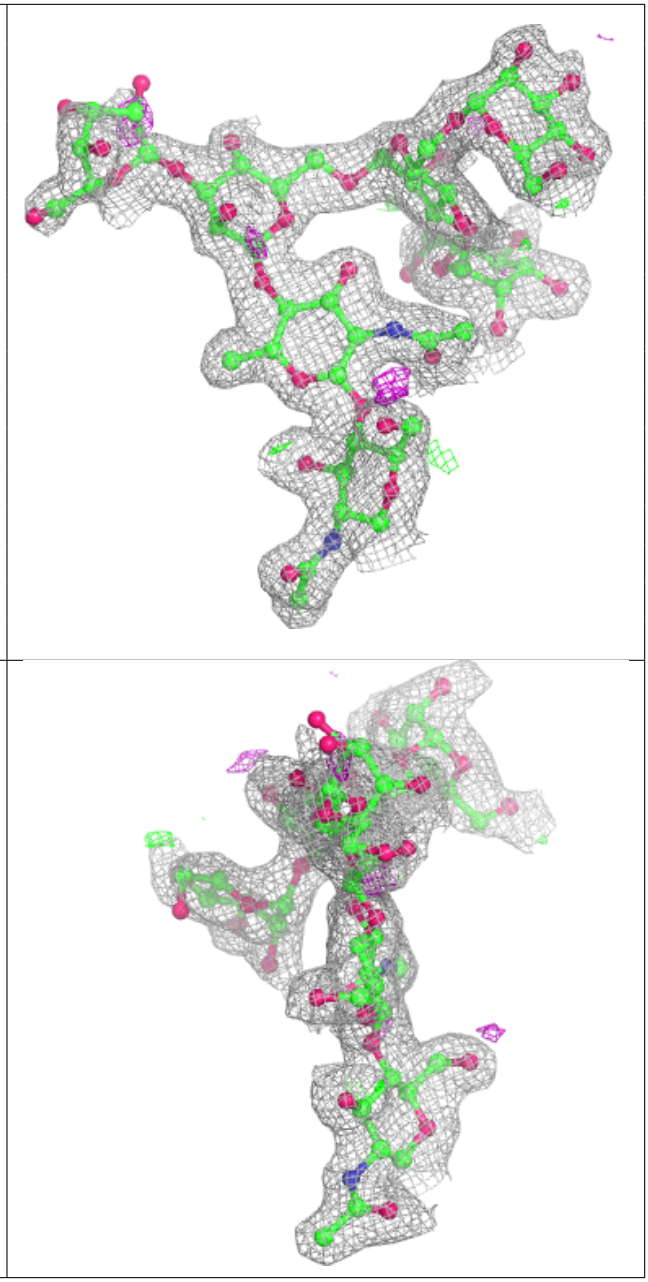






Electron density around Chain M:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



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
5	NAG	B	5154	14/15	0.72	0.42	50,55,57,57	0
5	NAG	A	5157	14/15	0.84	0.38	47,54,56,57	0
5	NAG	D	5157	14/15	0.85	0.23	41,47,50,52	0

Continued on next page...

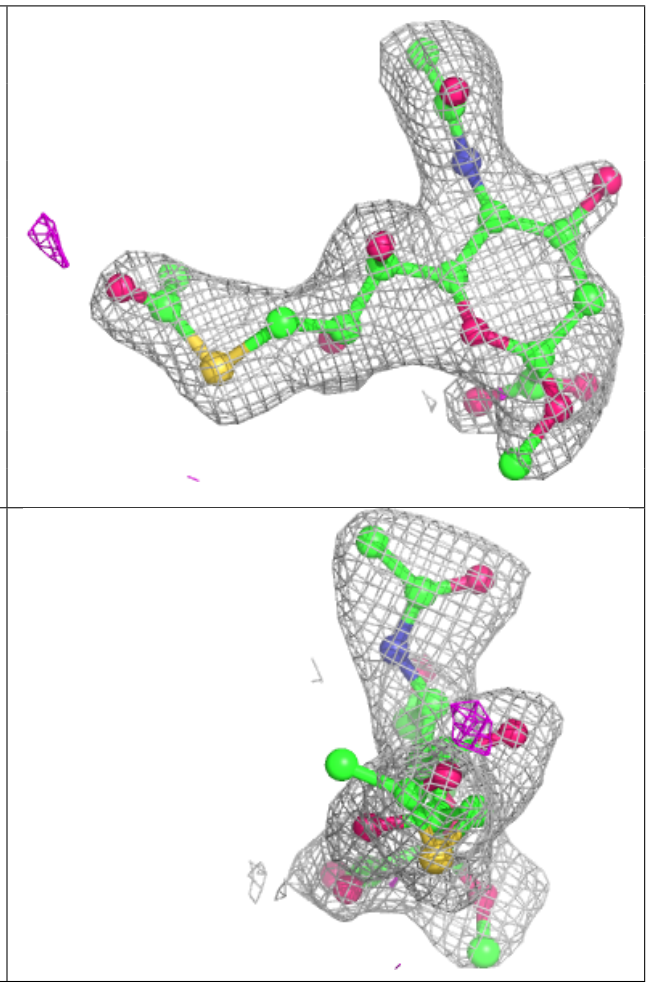
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	NAG	A	5156	14/15	0.86	0.23	32,36,40,40	0
6	SID	D	1001	25/25	0.87	0.28	49,55,62,62	0
5	NAG	D	5156	14/15	0.88	0.17	36,42,45,46	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around SID D 1001:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.