wwPDB X-ray Structure Validation Summary Report (i)

Jan 13, 2024 – 04:43 pm GMT

PDB ID	:	6RUA
Title	:	Structure of recombinant human butyrylcholinesterase in complex with a coumarin-based fluorescent probe linked to sulfonamide type inhibitor.
Authors	:	Coquelle, N.; Knez, D.; Brus, B.; Gobec, S.; Colletier, J.P.
Deposited on	:	2019-05-27
Resolution	:	2.75 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:

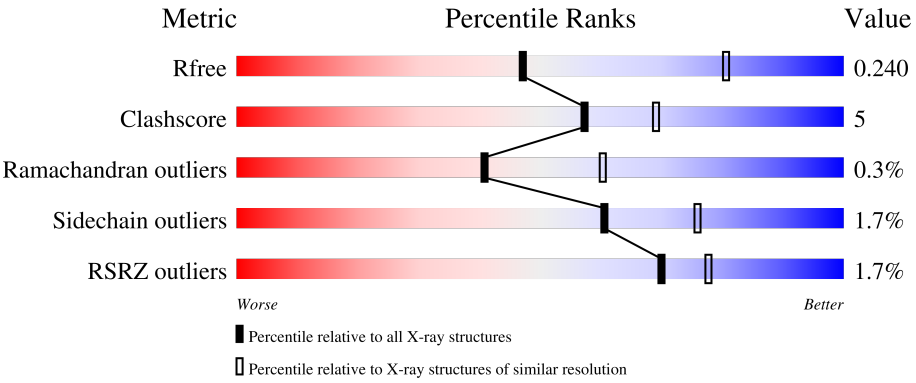
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	574	83% 9% 8%
1	B	574	2% 77% 14% 8%
2	C	2	50% 50%
2	E	2	100%
2	G	2	50% 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	2	100%
4	F	3	33% 67%
5	H	3	33% 67%
5	I	3	67% 33%
5	J	3	100%
5	K	3	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:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
11	PGE	B	623	-	-	X	-
2	NAG	C	1	X	-	-	-
2	NAG	C	2	-	-	-	X
2	NAG	E	2	X	-	-	-
4	NAG	F	2	-	-	-	X
4	FUC	F	3	X	-	-	-
5	NAG	H	1	X	-	-	-
5	NAG	H	2	-	-	-	X
5	NAG	J	2	-	-	-	X
5	FUL	J	3	-	-	-	X
5	NAG	K	2	-	-	-	X
5	FUL	K	3	-	-	-	X
6	NAG	A	615	X	-	-	-
6	NAG	B	614	X	-	-	-
8	CL	A	620	-	-	-	X

2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 9027 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 Cholinesterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	527	Total	C	N	O	S	0	3	0
			4210	2715	714	766	15			
1	B	526	Total	C	N	O	S	0	1	0
			4173	2693	705	760	15			

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

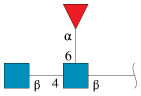
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	G	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 3 is an oligosaccharide called alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	2	Total	C	N	O	0	0	0
			24	14	1	9			

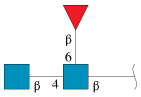
- Molecule 4 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[al

pha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



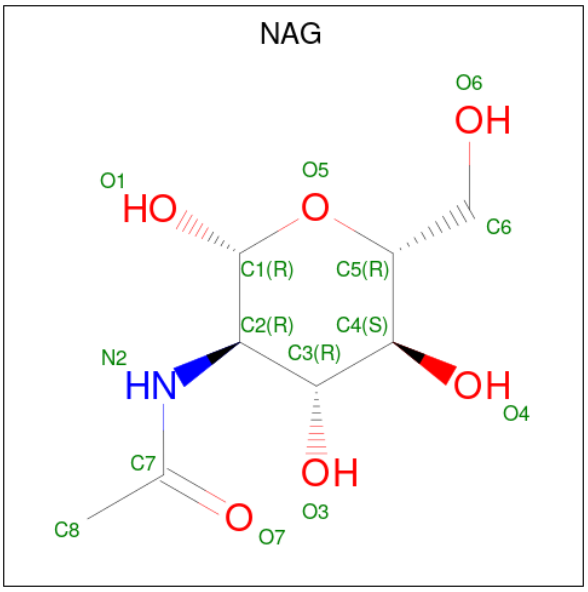
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	F	3	Total	C	N	O	0	0	0
			38	22	2	14			

- Molecule 5 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[be ta-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



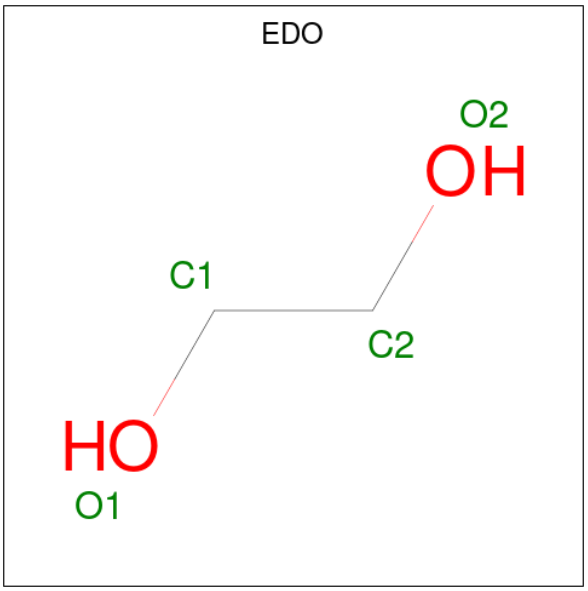
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	H	3	Total	C	N	O	0	0	0
			38	22	2	14			
5	I	3	Total	C	N	O	0	0	0
			38	22	2	14			
5	J	3	Total	C	N	O	0	0	0
			38	22	2	14			
5	K	3	Total	C	N	O	0	0	0
			38	22	2	14			

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

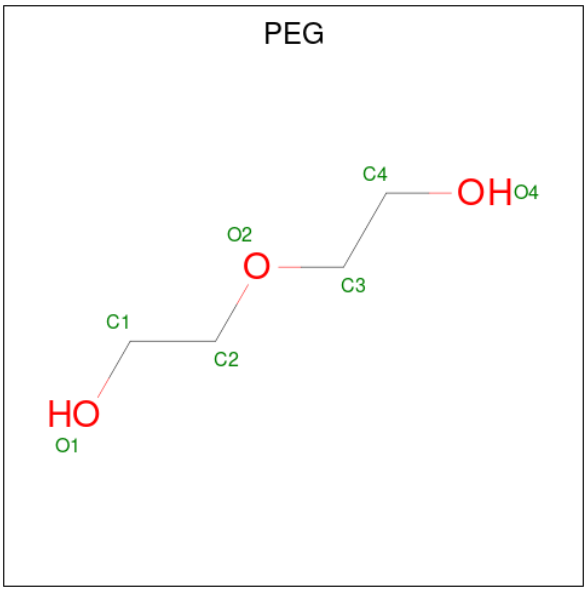


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

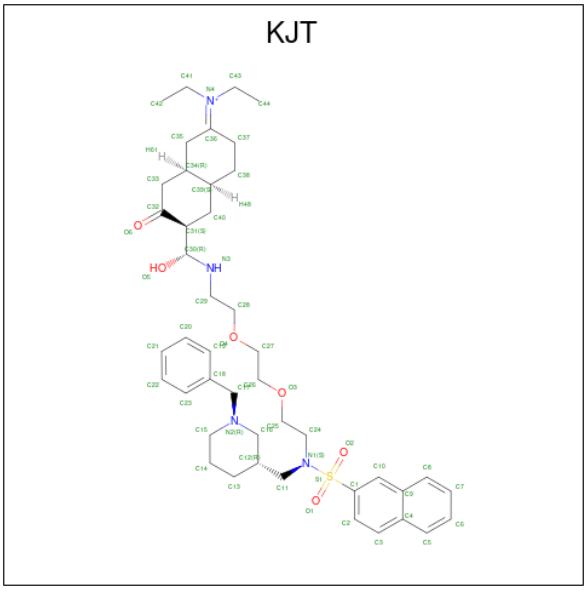
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	2	Total Cl 2 2	0	0
8	B	3	Total Cl 3 3	0	0

- Molecule 9 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



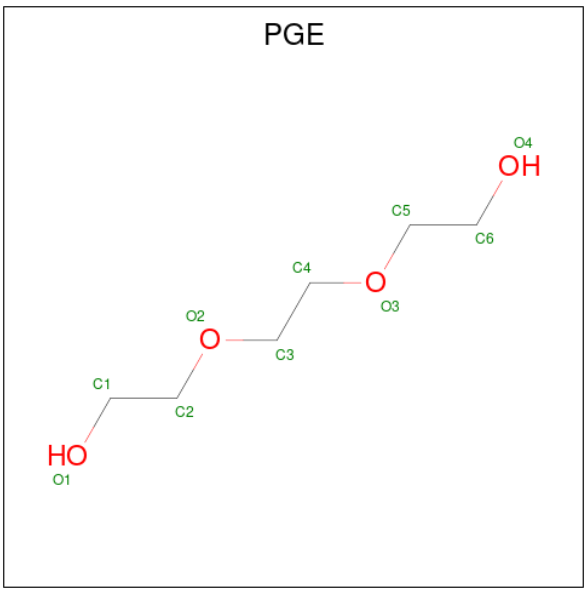
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
9	B	1	7	4	3	0	0

- Molecule 10 is diethyl-[6-[[2-[2-[2-[naphthalen-2-ylsulfonyl-[[(3 {R})-1-(phenylmethyl)piperidin-3-yl]methyl]amino]ethoxy]ethoxy]ethylamino]-oxidanyl-methyl]-7-oxidanylidene-1,3,4,4{a},5,6,8,8 {a}-octahydronaphthalen-2-ylidene]azanium (three-letter code: KJT) (formula: $C_{44}H_{63}N_4O_6S$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
10	B	1	29	24	2	2	1	0	0

- Molecule 11 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



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

- Molecule 12 is water.

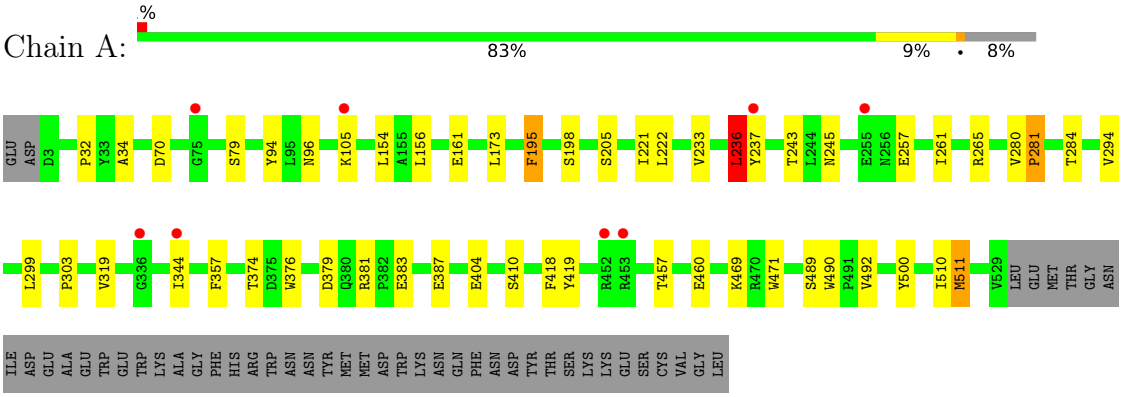
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	A	83	Total O 83 83	0	0
12	B	92	Total O 92 92	0	0

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.

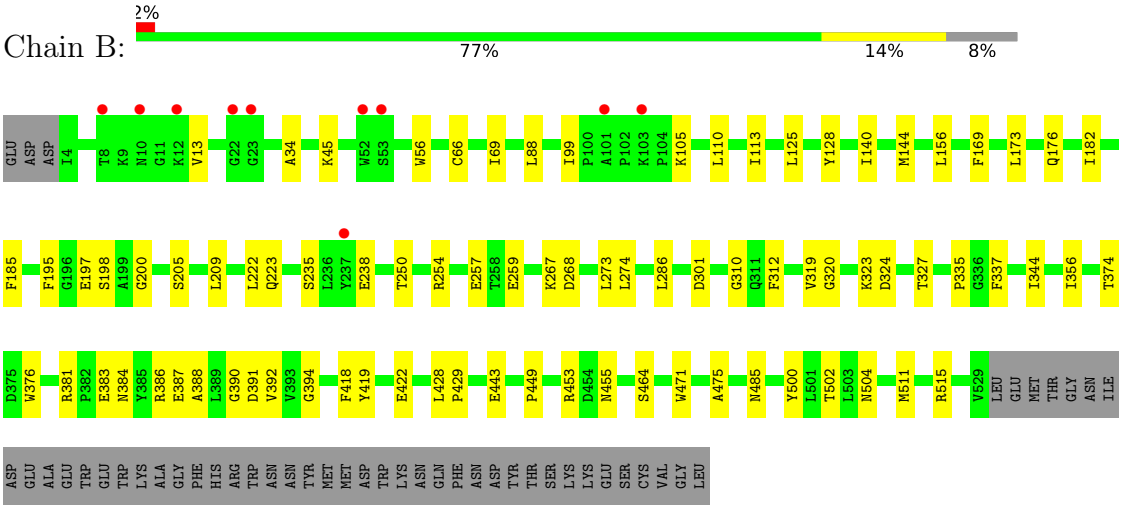
• Molecule 1: Cholinesterase

Chain A:



• Molecule 1: Cholinesterase

Chain B:

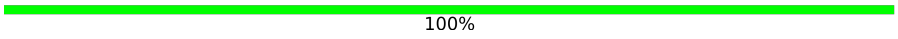


• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:

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

Chain E:

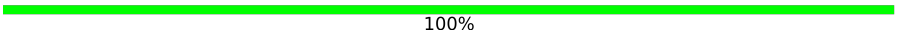


• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:

• Molecule 3: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:

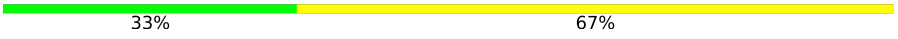


• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:

• Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[beta-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain H:

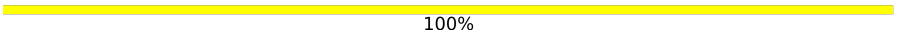


• Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[beta-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain I:

• Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[beta-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain J:



• Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[beta-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain K:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	76.29Å 80.20Å 231.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.93 – 2.75 46.91 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.1 (44.93-2.75) 99.2 (46.91-2.75)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 2.77Å)	Xtriage
Refinement program	PHENIX (dev_3139: ???)	Depositor
R, R_{free}	0.181 , 0.237 0.183 , 0.240	Depositor DCC
R_{free} test set	1122 reflections (2.99%)	wwPDB-VP
Wilson B-factor (Å^2)	58.5	Xtriage
Anisotropy	0.260	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 39.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.029 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9027	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% 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: PGE, FUL, KJT, PEG, FUC, EDO, NAG, CL

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	0.34	0/4338	0.51	1/5890 (0.0%)
1	B	0.34	0/4294	0.51	0/5833
All	All	0.34	0/8632	0.51	1/11723 (0.0%)

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	A	236	LEU	CB-CG-CD2	-7.41	98.40	111.00

There are no chirality outliers.

There are no planarity outliers.

CLOSE-CONTACTS INFOmissingINFO

5.2 Torsion angles

5.2.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	528/574 (92%)	507 (96%)	19 (4%)	2 (0%)	34	53
1	B	525/574 (92%)	497 (95%)	27 (5%)	1 (0%)	47	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1053/1148 (92%)	1004 (95%)	46 (4%)	3 (0%)	41 60

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	236	LEU
1	B	455	ASN
1	A	281	PRO

5.2.2 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	453/494 (92%)	447 (99%)	6 (1%)	69 81
1	B	444/494 (90%)	435 (98%)	9 (2%)	55 72
All	All	897/988 (91%)	882 (98%)	15 (2%)	60 76

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	176	GLN
1	B	504	ASN
1	B	195	PHE
1	B	515	ARG
1	B	301	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	311	GLN

5.2.3 RNA

There are no RNA molecules in this entry.

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

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

5.4 Carbohydrates

23 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	C	1	1,2	14,14,15	1.08	1 (7%)	17,19,21	0.89	0
2	NAG	C	2	2	14,14,15	0.62	0	17,19,21	0.36	0
3	NAG	D	1	3,1	14,14,15	0.35	0	17,19,21	0.46	0
3	FUC	D	2	3	10,10,11	0.82	0	14,14,16	0.84	0
2	NAG	E	1	1,2	14,14,15	0.29	0	17,19,21	0.62	0
2	NAG	E	2	2	14,14,15	0.45	0	17,19,21	0.42	0
4	NAG	F	1	4,1	14,14,15	0.47	0	17,19,21	0.55	0
4	NAG	F	2	4	14,14,15	0.64	1 (7%)	17,19,21	0.77	1 (5%)
4	FUC	F	3	4	10,10,11	1.00	0	14,14,16	0.95	0
2	NAG	G	1	1,2	14,14,15	0.38	0	17,19,21	0.59	0
2	NAG	G	2	2	14,14,15	0.71	1 (7%)	17,19,21	0.67	1 (5%)
5	NAG	H	1	1,5	14,14,15	1.21	2 (14%)	17,19,21	1.02	1 (5%)
5	NAG	H	2	5	14,14,15	0.36	0	17,19,21	0.50	0
5	FUL	H	3	5	10,10,11	1.89	3 (30%)	14,14,16	1.16	0
5	NAG	I	1	1,5	14,14,15	0.39	0	17,19,21	0.46	0
5	NAG	I	2	5	14,14,15	0.19	0	17,19,21	0.49	0
5	FUL	I	3	5	10,10,11	1.92	3 (30%)	14,14,16	1.84	3 (21%)
5	NAG	J	1	1,5	14,14,15	0.33	0	17,19,21	0.44	0
5	NAG	J	2	5	14,14,15	0.35	0	17,19,21	0.47	0
5	FUL	J	3	5	10,10,11	1.79	3 (30%)	14,14,16	1.22	3 (21%)
5	NAG	K	1	1,5	14,14,15	0.27	0	17,19,21	0.81	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	NAG	K	2	5	14,14,15	0.58	0	17,19,21	1.27	1 (5%)
5	FUL	K	3	5	10,10,11	1.86	3 (30%)	14,14,16	1.62	3 (21%)

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	C	1	1,2	1/1/5/7	4/6/23/26	0/1/1/1
2	NAG	C	2	2	-	2/6/23/26	0/1/1/1
3	NAG	D	1	3,1	-	2/6/23/26	0/1/1/1
3	FUC	D	2	3	-	-	0/1/1/1
2	NAG	E	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	E	2	2	1/1/5/7	4/6/23/26	0/1/1/1
4	NAG	F	1	4,1	-	4/6/23/26	0/1/1/1
4	NAG	F	2	4	-	2/6/23/26	0/1/1/1
4	FUC	F	3	4	1/1/4/5	-	0/1/1/1
2	NAG	G	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	G	2	2	-	0/6/23/26	0/1/1/1
5	NAG	H	1	1,5	1/1/5/7	2/6/23/26	0/1/1/1
5	NAG	H	2	5	-	2/6/23/26	0/1/1/1
5	FUL	H	3	5	-	-	0/1/1/1
5	NAG	I	1	1,5	-	2/6/23/26	0/1/1/1
5	NAG	I	2	5	-	1/6/23/26	0/1/1/1
5	FUL	I	3	5	-	-	0/1/1/1
5	NAG	J	1	1,5	-	2/6/23/26	0/1/1/1
5	NAG	J	2	5	-	2/6/23/26	0/1/1/1
5	FUL	J	3	5	-	-	0/1/1/1
5	NAG	K	1	1,5	-	2/6/23/26	0/1/1/1
5	NAG	K	2	5	-	3/6/23/26	0/1/1/1
5	FUL	K	3	5	-	-	0/1/1/1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	I	3	FUL	O5-C1	4.94	1.51	1.43
5	H	3	FUL	O5-C1	4.76	1.51	1.43
5	K	3	FUL	O5-C1	4.69	1.51	1.43

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	J	3	FUL	O5-C1	4.31	1.50	1.43
2	C	1	NAG	O5-C1	-3.92	1.37	1.43

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	I	3	FUL	C1-C2-C3	5.08	115.91	109.67
5	K	3	FUL	C1-C2-C3	4.24	114.88	109.67
5	K	2	NAG	C2-N2-C7	4.21	128.90	122.90
5	H	1	NAG	C4-C3-C2	3.22	115.74	111.02
4	F	2	NAG	C1-O5-C5	2.86	116.07	112.19

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	1	NAG	C1
2	E	2	NAG	C1
4	F	3	FUC	C1
5	H	1	NAG	C1

5 of 36 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	F	2	NAG	C4-C5-C6-O6
2	E	2	NAG	O5-C5-C6-O6
5	K	1	NAG	O5-C5-C6-O6
5	I	1	NAG	C4-C5-C6-O6
5	J	1	NAG	O5-C5-C6-O6

There are no ring outliers.

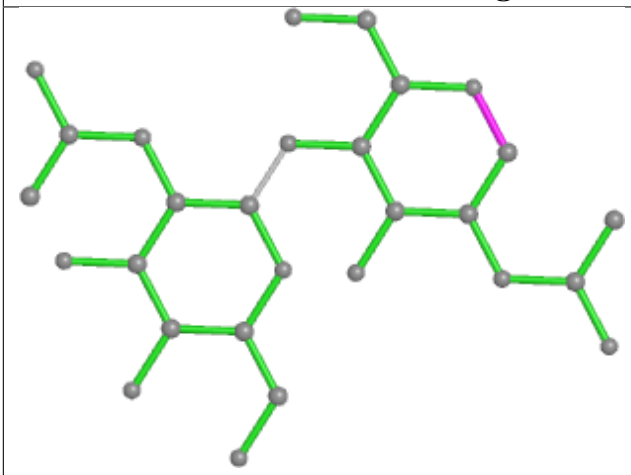
6 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	J	1	NAG	1	0
4	F	3	FUC	1	0
5	J	2	NAG	1	0
5	K	3	FUL	1	0
5	K	1	NAG	1	0
5	K	2	NAG	1	0

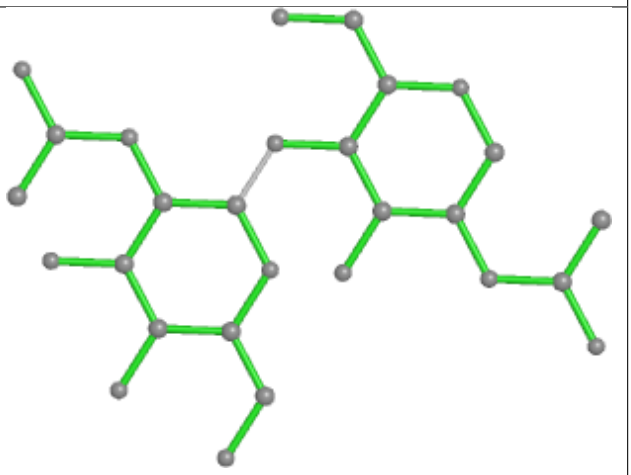
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

bond angles, torsion angles, and ring geometry for oligosaccharide.

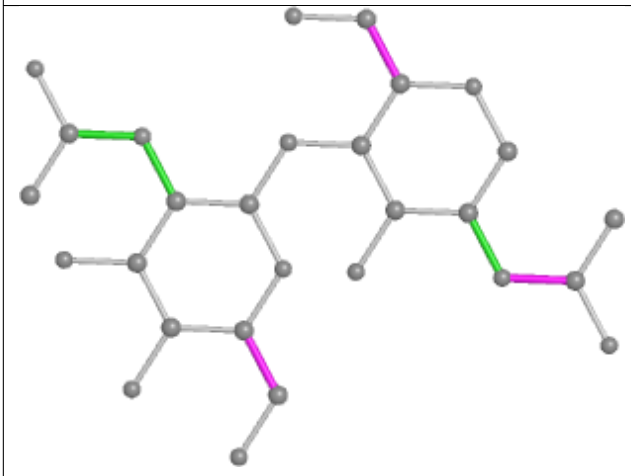
Oligosaccharide Chain C



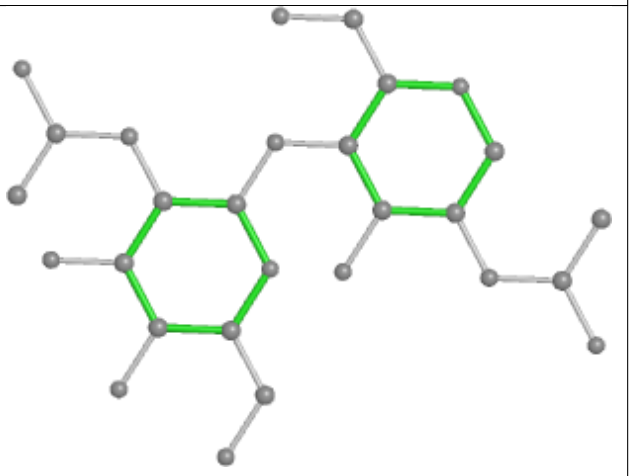
Bond lengths



Bond angles

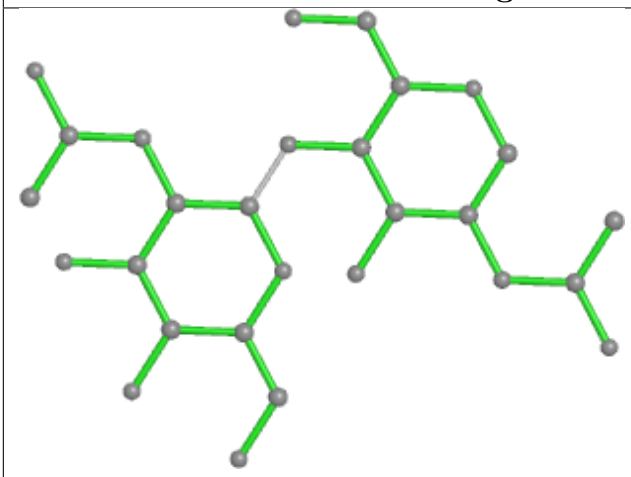


Torsions

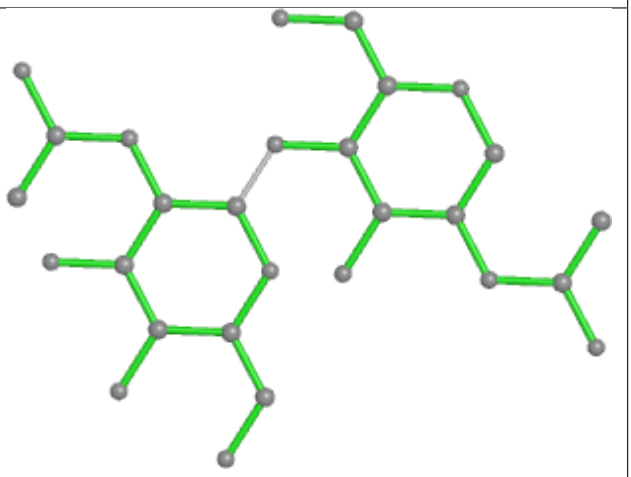


Rings

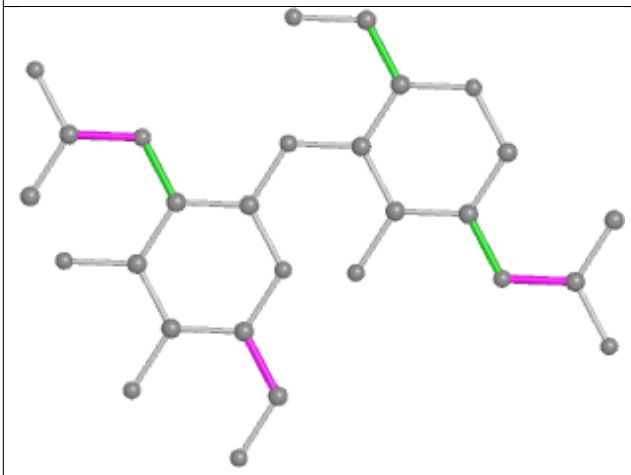
Oligosaccharide Chain E



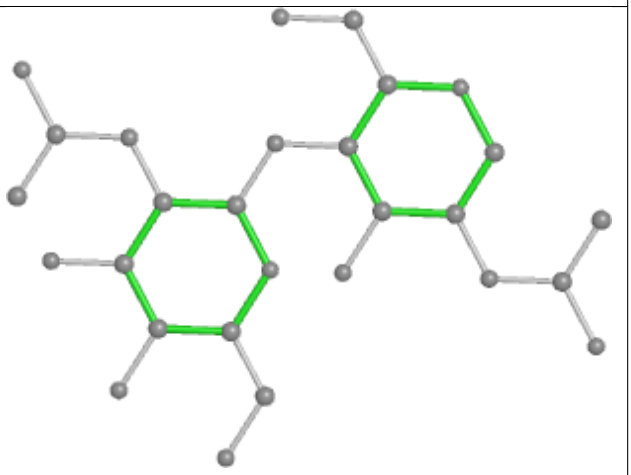
Bond lengths



Bond angles

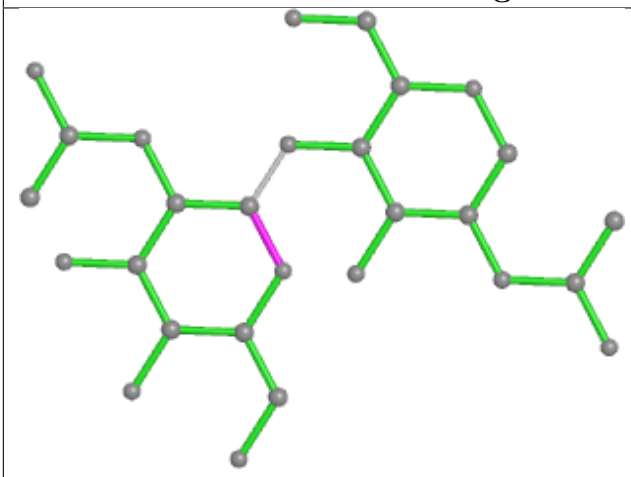


Torsions

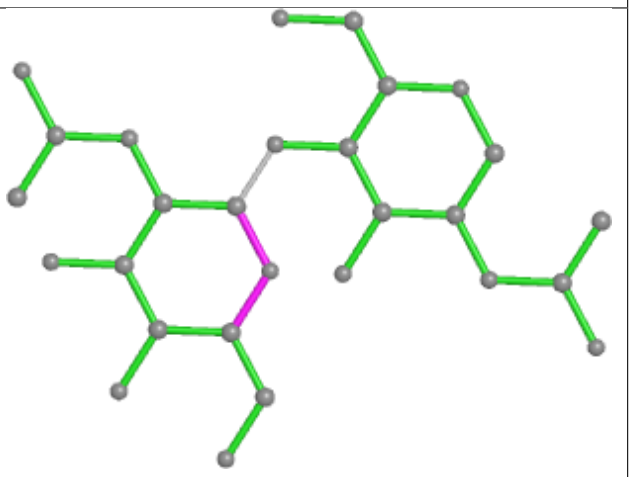


Rings

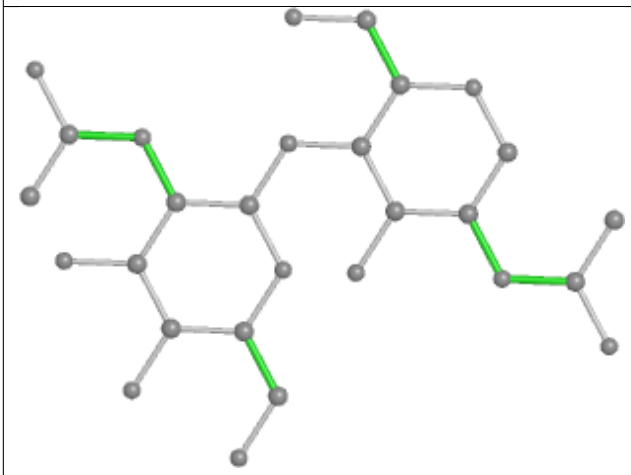
Oligosaccharide Chain G



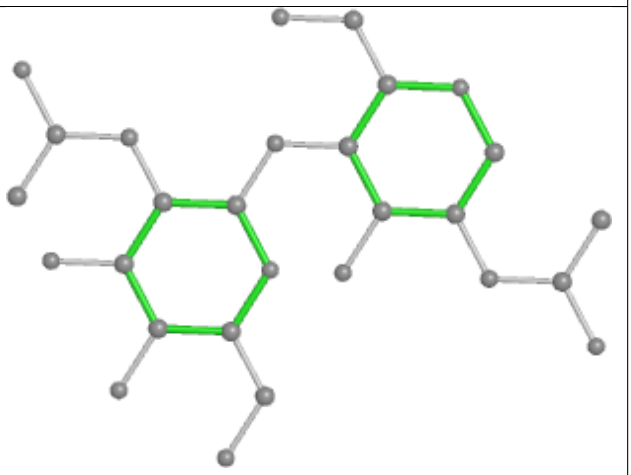
Bond lengths



Bond angles

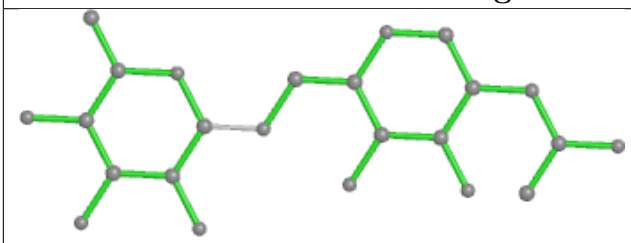


Torsions

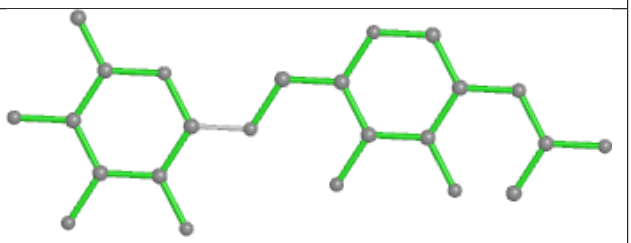


Rings

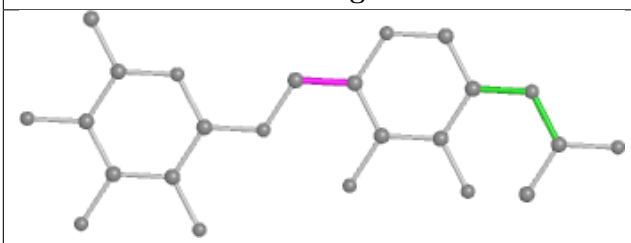
Oligosaccharide Chain D



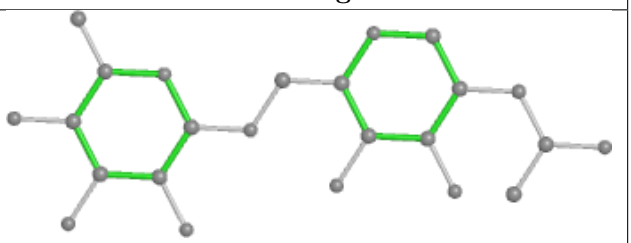
Bond lengths



Bond angles

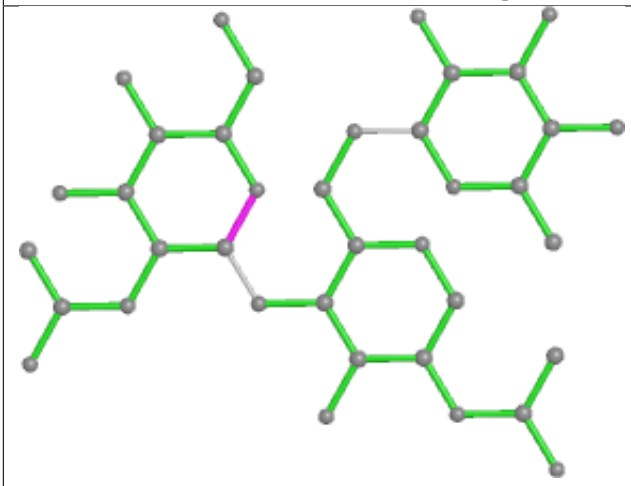


Torsions

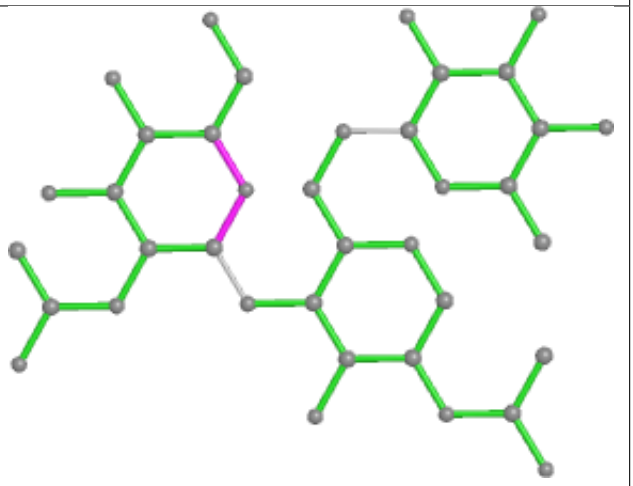


Rings

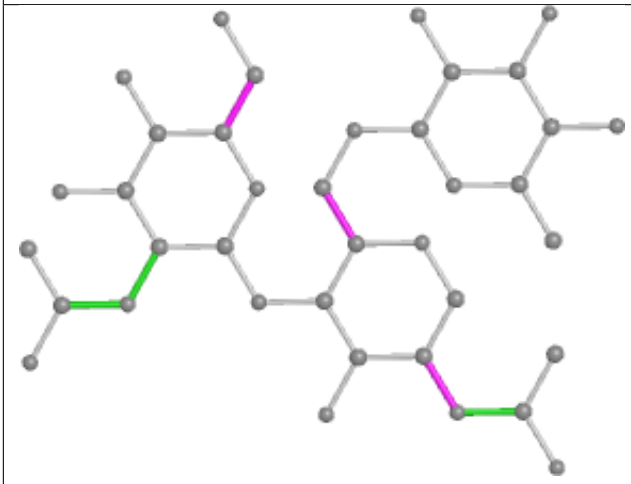
Oligosaccharide Chain F



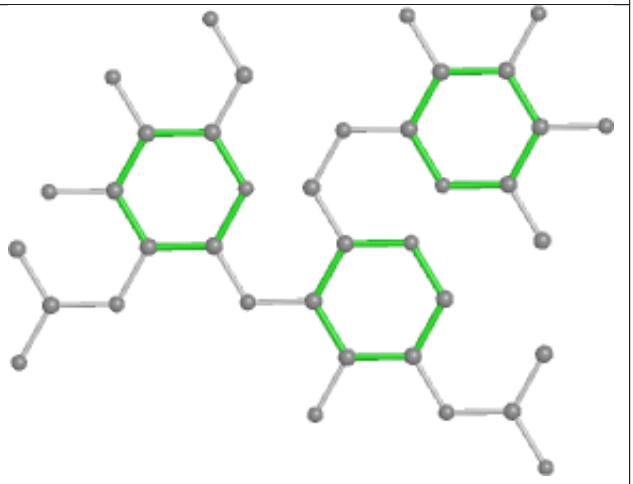
Bond lengths



Bond angles

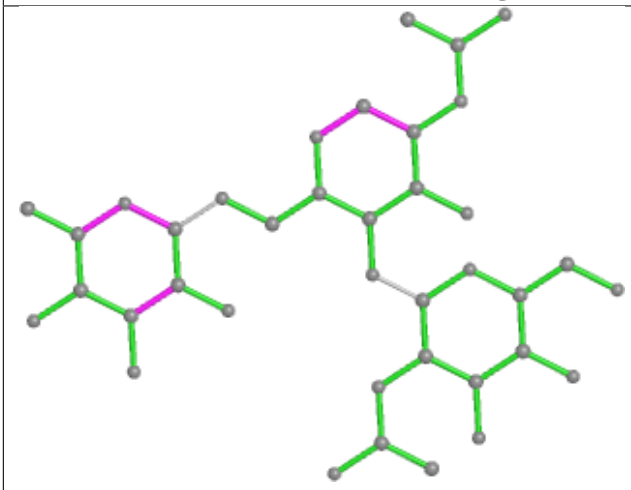


Torsions

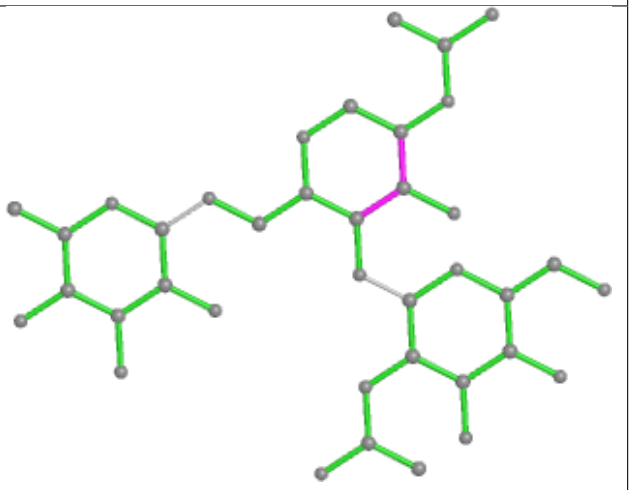


Rings

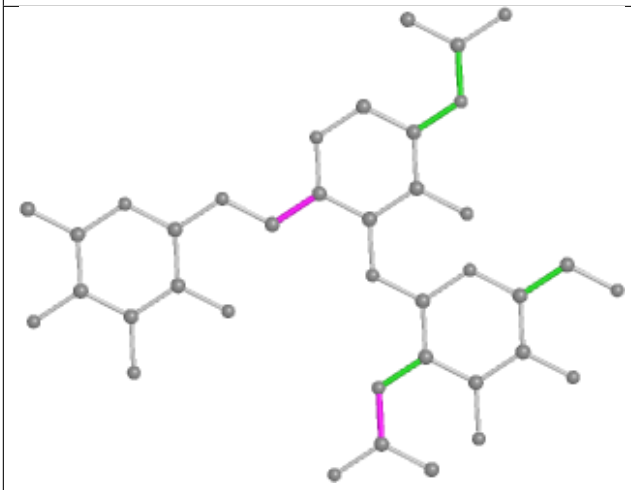
Oligosaccharide Chain H



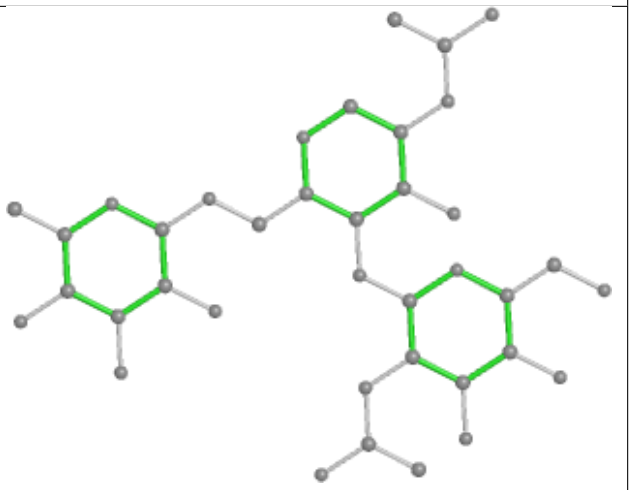
Bond lengths



Bond angles

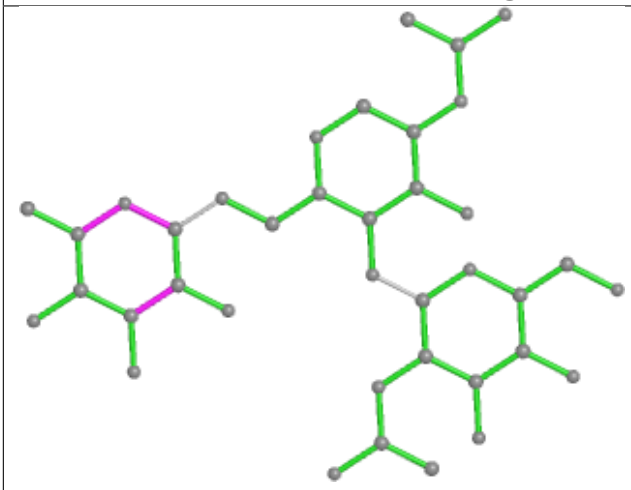


Torsions

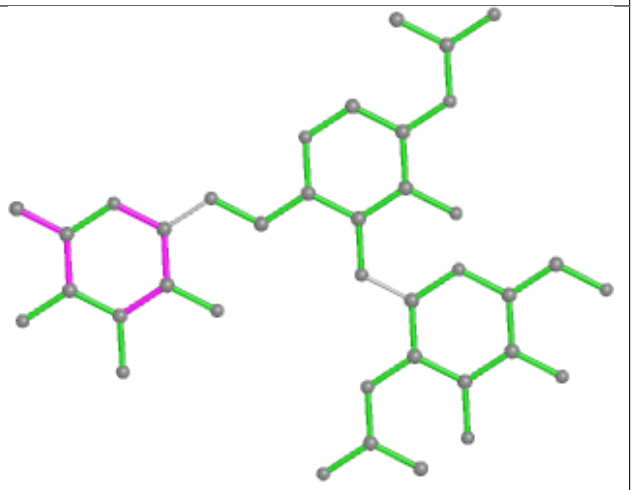


Rings

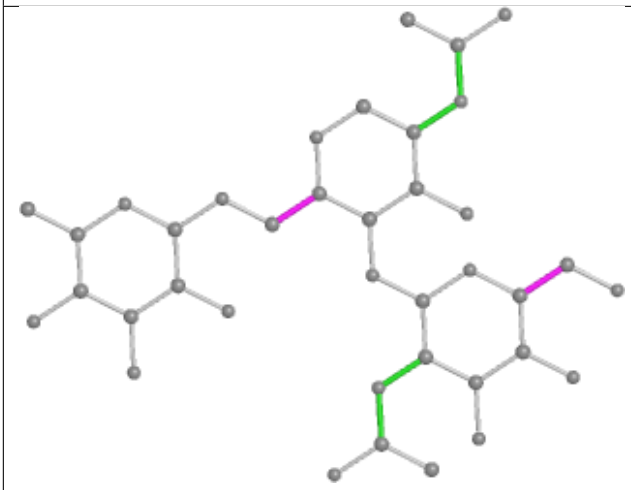
Oligosaccharide Chain I



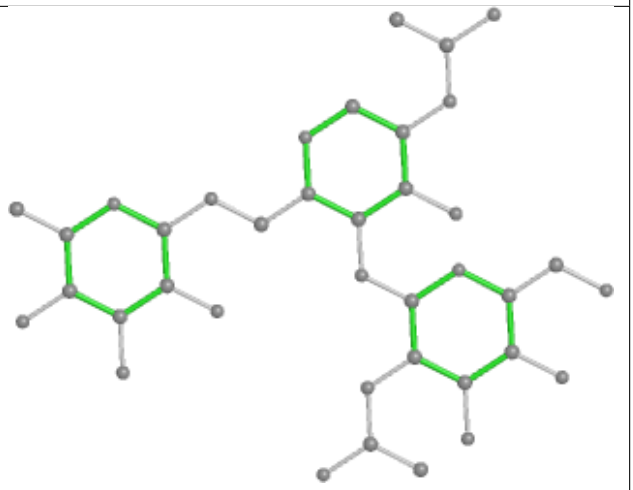
Bond lengths



Bond angles

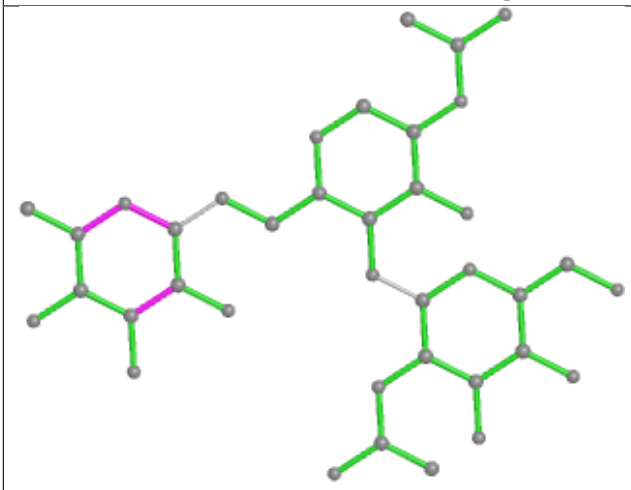


Torsions

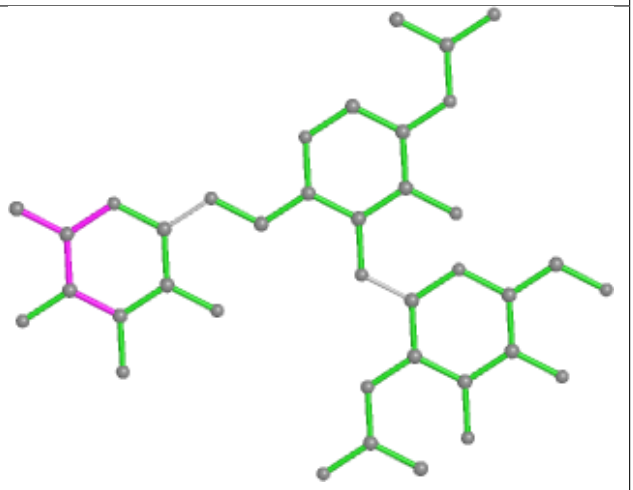


Rings

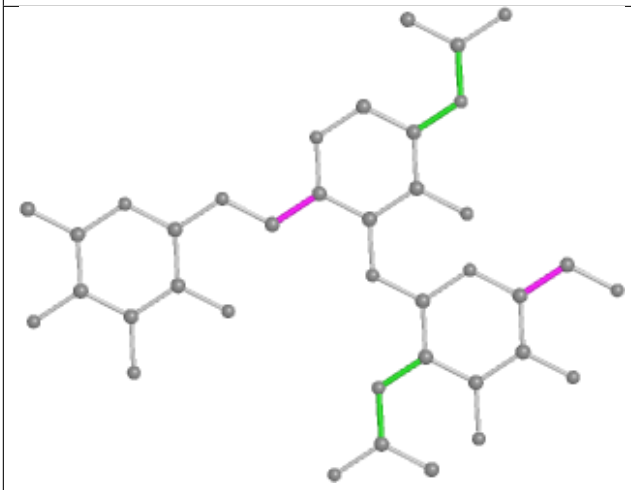
Oligosaccharide Chain J



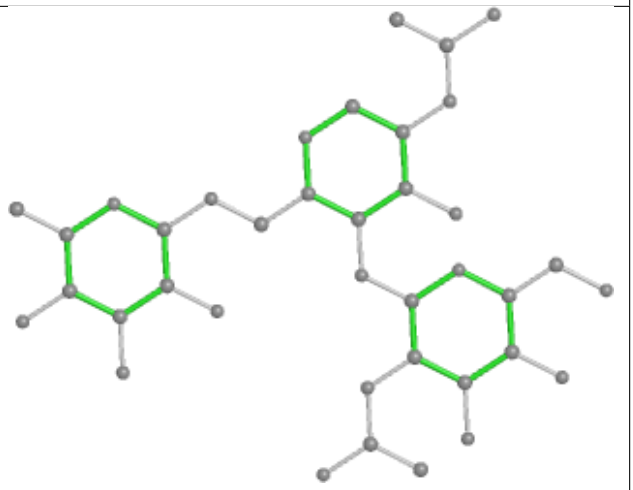
Bond lengths



Bond angles

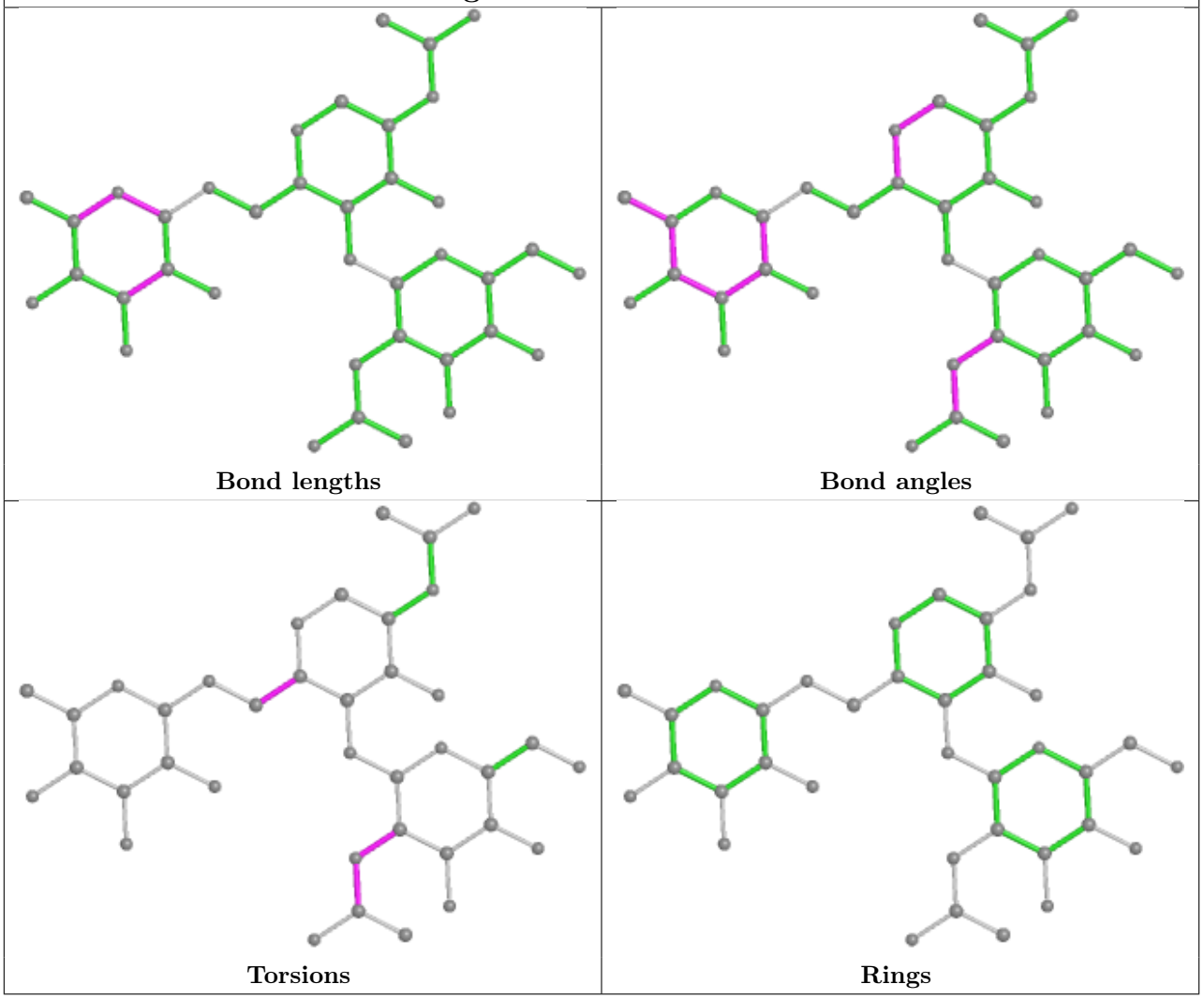


Torsions



Rings

Oligosaccharide Chain K



5.5 Ligand geometry (i)

Of 23 ligands modelled in this entry, 5 are monoatomic - leaving 18 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
6	NAG	B	603	1	14,14,15	0.39	0	17,19,21	0.77	0
7	EDO	A	617	-	3,3,3	0.54	0	2,2,2	0.19	0
6	NAG	B	610	1	14,14,15	0.36	0	17,19,21	0.49	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
7	EDO	A	616	-	3,3,3	0.51	0	2,2,2	0.35	0
7	EDO	A	618	-	3,3,3	0.58	0	2,2,2	0.17	0
7	EDO	A	621	-	3,3,3	0.56	0	2,2,2	0.05	0
7	EDO	B	616	-	3,3,3	0.55	0	2,2,2	0.34	0
6	NAG	B	602	1	14,14,15	0.47	0	17,19,21	0.64	1 (5%)
7	EDO	B	618	-	3,3,3	0.50	0	2,2,2	0.22	0
6	NAG	A	615	1	14,14,15	0.66	1 (7%)	17,19,21	1.00	1 (5%)
6	NAG	B	614	1	14,14,15	0.50	0	17,19,21	0.80	1 (5%)
7	EDO	A	619	-	3,3,3	0.49	0	2,2,2	0.30	0
9	PEG	B	615	-	6,6,6	0.48	0	5,5,5	0.22	0
7	EDO	B	617	-	3,3,3	0.55	0	2,2,2	0.30	0
7	EDO	A	623	-	3,3,3	0.52	0	2,2,2	0.17	0
10	KJT	B	620	-	31,32,60	3.13	5 (16%)	41,45,82	1.83	7 (17%)
11	PGE	B	623	-	9,9,9	0.38	0	8,8,8	0.63	0
6	NAG	B	601	1	14,14,15	0.81	1 (7%)	17,19,21	0.85	1 (5%)

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	NAG	B	603	1	-	2/6/23/26	0/1/1/1
7	EDO	A	617	-	-	1/1/1/1	-
6	NAG	B	610	1	-	2/6/23/26	0/1/1/1
7	EDO	A	616	-	-	0/1/1/1	-
7	EDO	A	618	-	-	1/1/1/1	-
7	EDO	A	621	-	-	0/1/1/1	-
7	EDO	B	616	-	-	0/1/1/1	-
6	NAG	B	602	1	-	4/6/23/26	0/1/1/1
7	EDO	B	618	-	-	0/1/1/1	-
6	NAG	B	614	1	1/1/5/7	2/6/23/26	0/1/1/1
6	NAG	A	615	1	1/1/5/7	2/6/23/26	0/1/1/1
7	EDO	A	619	-	-	1/1/1/1	-
9	PEG	B	615	-	-	1/4/4/4	-
7	EDO	B	617	-	-	0/1/1/1	-
7	EDO	A	623	-	-	0/1/1/1	-
10	KJT	B	620	-	-	14/20/30/79	0/4/4/6
11	PGE	B	623	-	-	6/7/7/7	-
6	NAG	B	601	1	-	3/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	B	620	KJT	O1-S1	11.09	1.55	1.43
10	B	620	KJT	O2-S1	10.26	1.54	1.43
10	B	620	KJT	C17-C18	-6.22	1.40	1.51
10	B	620	KJT	S1-N1	3.75	1.73	1.63
6	B	601	NAG	O5-C1	2.47	1.47	1.43

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
10	B	620	KJT	O2-S1-N1	6.89	115.64	107.05
10	B	620	KJT	O1-S1-O2	-5.11	111.24	119.52
6	A	615	NAG	C1-O5-C5	3.88	117.45	112.19
6	B	601	NAG	C1-O5-C5	3.20	116.53	112.19
10	B	620	KJT	O1-S1-C1	-3.10	104.12	108.05

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	A	615	NAG	C1
6	B	614	NAG	C1

5 of 39 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
10	B	620	KJT	N1-C11-C12-C13
10	B	620	KJT	N1-C11-C12-C16
10	B	620	KJT	C12-C11-N1-S1
6	B	602	NAG	C4-C5-C6-O6
6	B	614	NAG	C4-C5-C6-O6

There are no ring outliers.

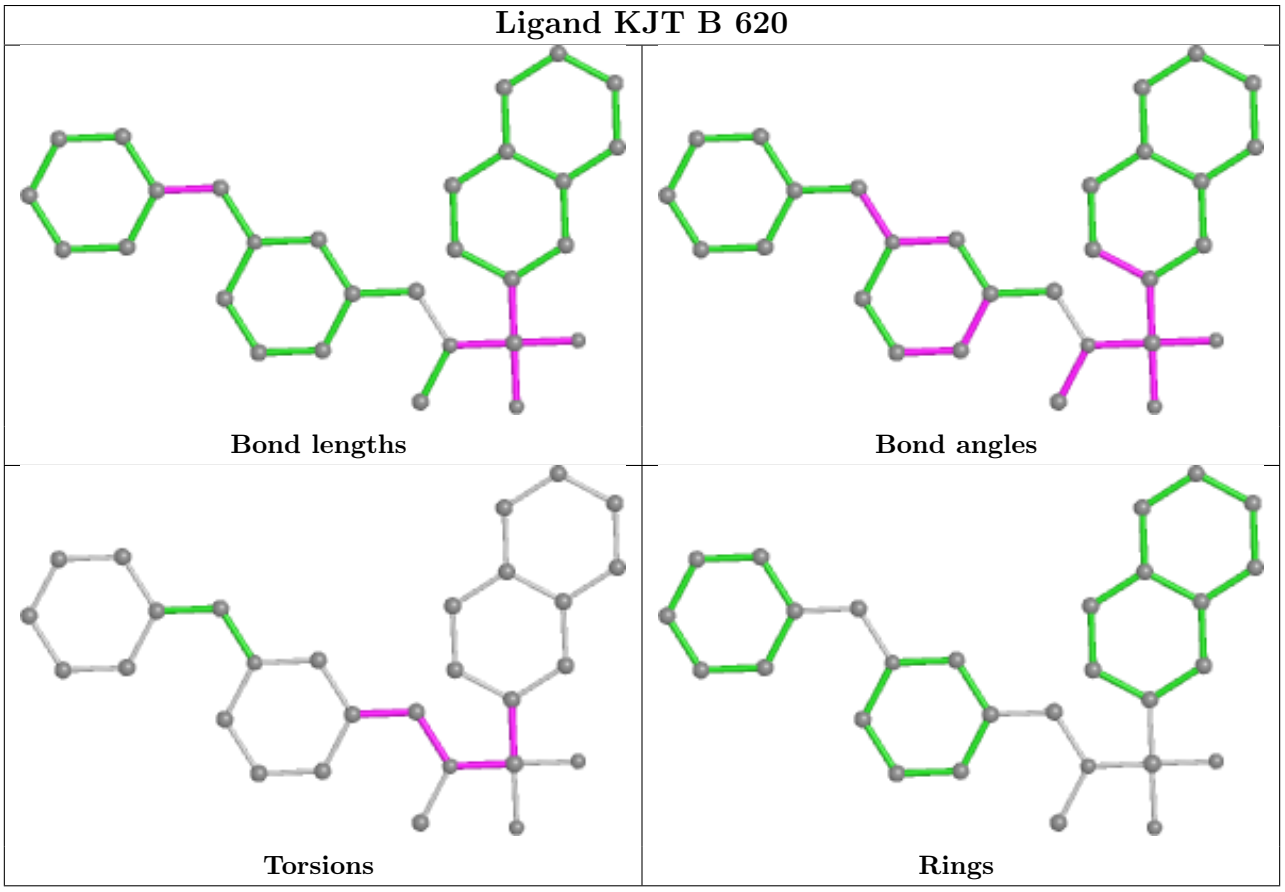
3 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	B	615	PEG	1	0
7	B	617	EDO	1	0
11	B	623	PGE	13	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 KJT B 620



5.6 Other polymers (i)

There are no such residues in this entry.

5.7 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

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, 95^{th} 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	527/574 (91%)	-0.22	8 (1%)	73	81	27, 49, 75, 101	0
1	B	526/574 (91%)	-0.14	10 (1%)	66	75	30, 48, 79, 116	0
All	All	1053/1148 (91%)	-0.18	18 (1%)	70	78	27, 48, 78, 116	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	22	GLY	3.5
1	B	53	SER	3.4
1	B	52	TRP	3.1
1	A	336	GLY	3.0
1	B	12	LYS	2.9

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

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

6.3 Carbohydrates

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} 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	K	2	14/15	0.57	0.46	88,134,138,138	0
5	NAG	J	2	14/15	0.61	0.71	108,142,154,156	0
2	NAG	C	2	14/15	0.67	0.55	147,163,168,170	0
5	NAG	H	2	14/15	0.67	0.57	116,130,136,137	0

Continued on next page...

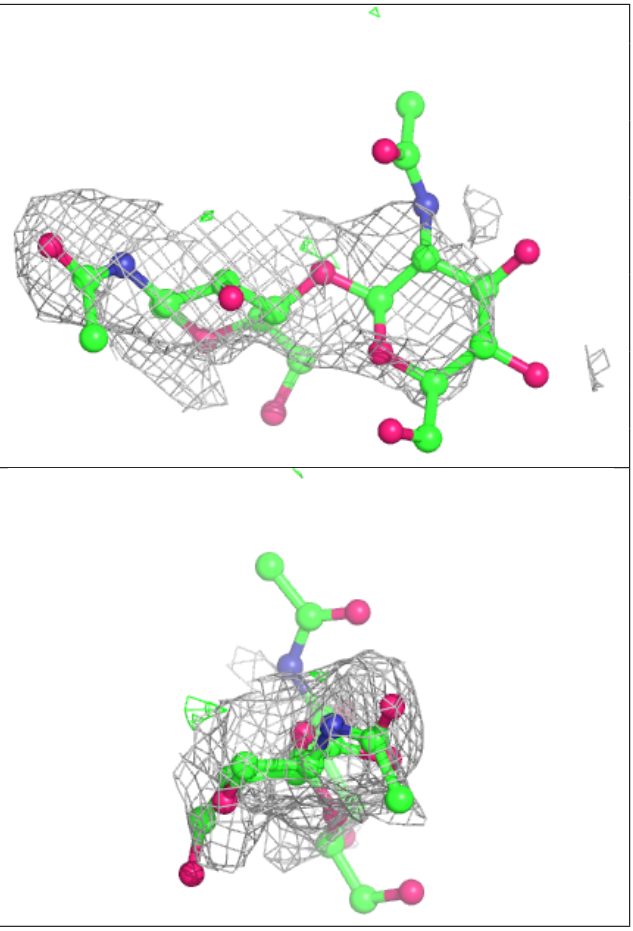
Continued from previous page...

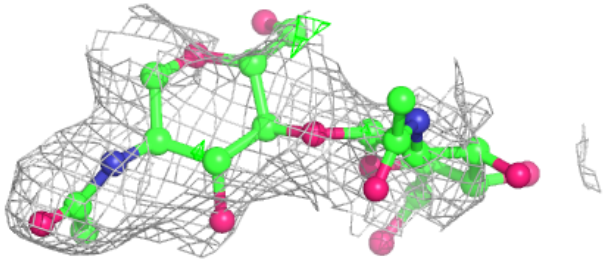
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	FUL	J	3	10/11	0.68	0.67	131,135,142,146	0
2	NAG	C	1	14/15	0.71	0.31	104,124,144,150	0
5	FUL	K	3	10/11	0.73	0.42	123,134,141,145	0
4	NAG	F	2	14/15	0.74	0.53	116,128,138,138	0
5	NAG	H	1	14/15	0.75	0.27	86,118,127,135	0
5	NAG	J	1	14/15	0.75	0.25	70,101,123,132	0
2	NAG	G	2	14/15	0.77	0.23	48,81,90,93	0
3	NAG	D	1	14/15	0.78	0.24	88,107,115,116	0
5	NAG	K	1	14/15	0.78	0.37	108,122,133,134	0
5	FUL	H	3	10/11	0.80	0.28	82,98,109,114	0
4	NAG	F	1	14/15	0.82	0.26	105,119,127,130	0
2	NAG	E	2	14/15	0.85	0.40	91,103,109,110	0
5	NAG	I	2	14/15	0.85	0.45	94,112,123,123	0
3	FUC	D	2	10/11	0.87	0.20	83,104,114,125	0
5	FUL	I	3	10/11	0.89	0.51	73,91,96,98	0
4	FUC	F	3	10/11	0.91	0.29	82,101,106,111	0
5	NAG	I	1	14/15	0.92	0.33	81,100,115,123	0
2	NAG	G	1	14/15	0.93	0.15	60,74,81,83	0
2	NAG	E	1	14/15	0.94	0.23	63,73,87,96	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 C:

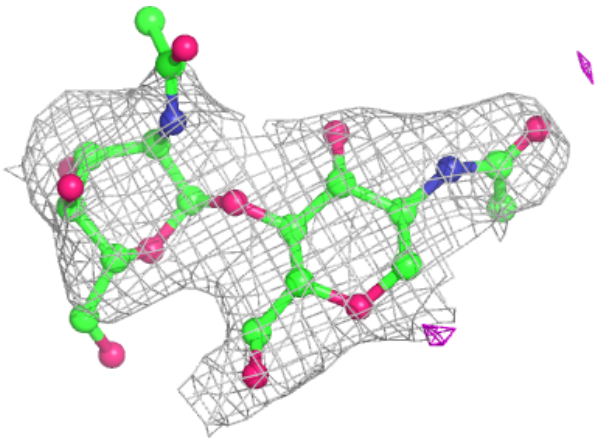
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

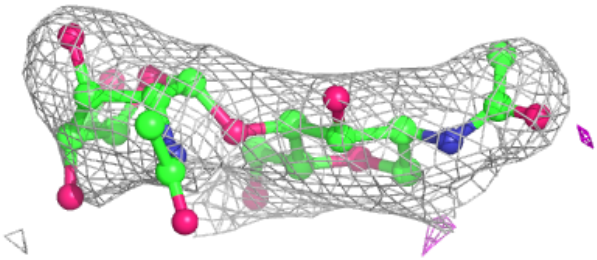


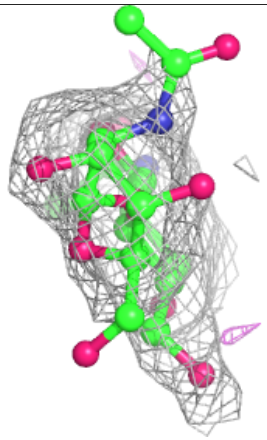


Electron density around Chain E:

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

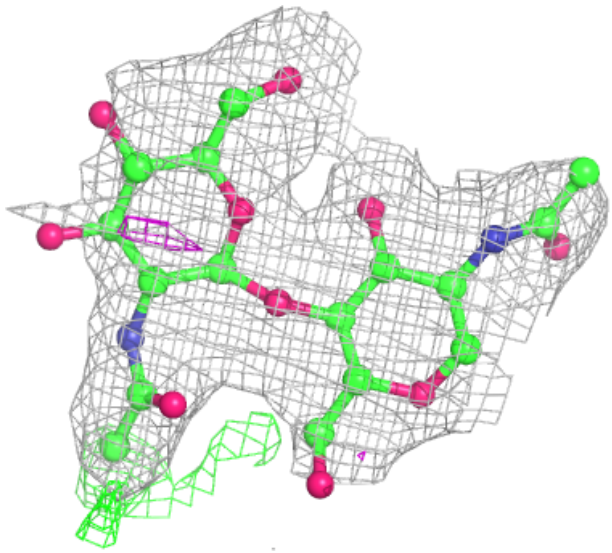


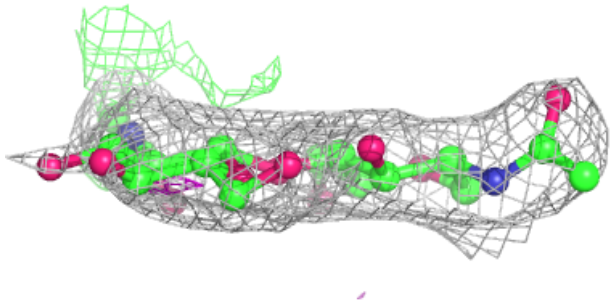


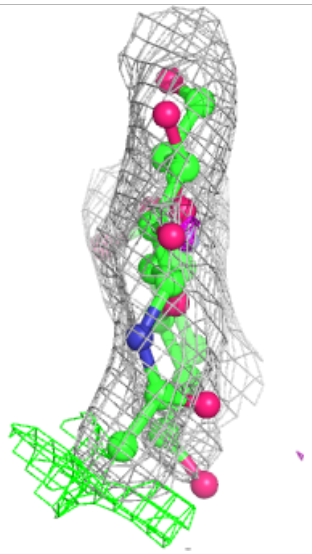


Electron density around Chain G:

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

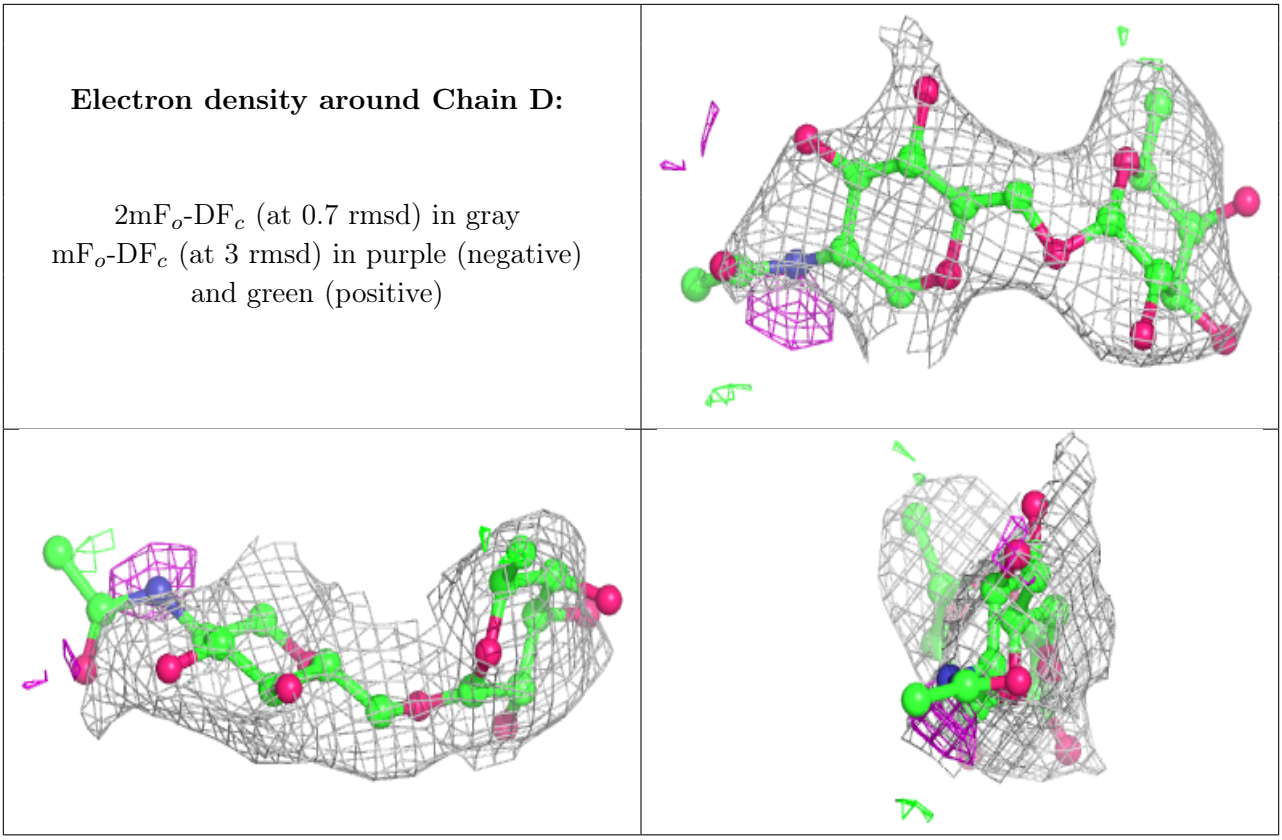






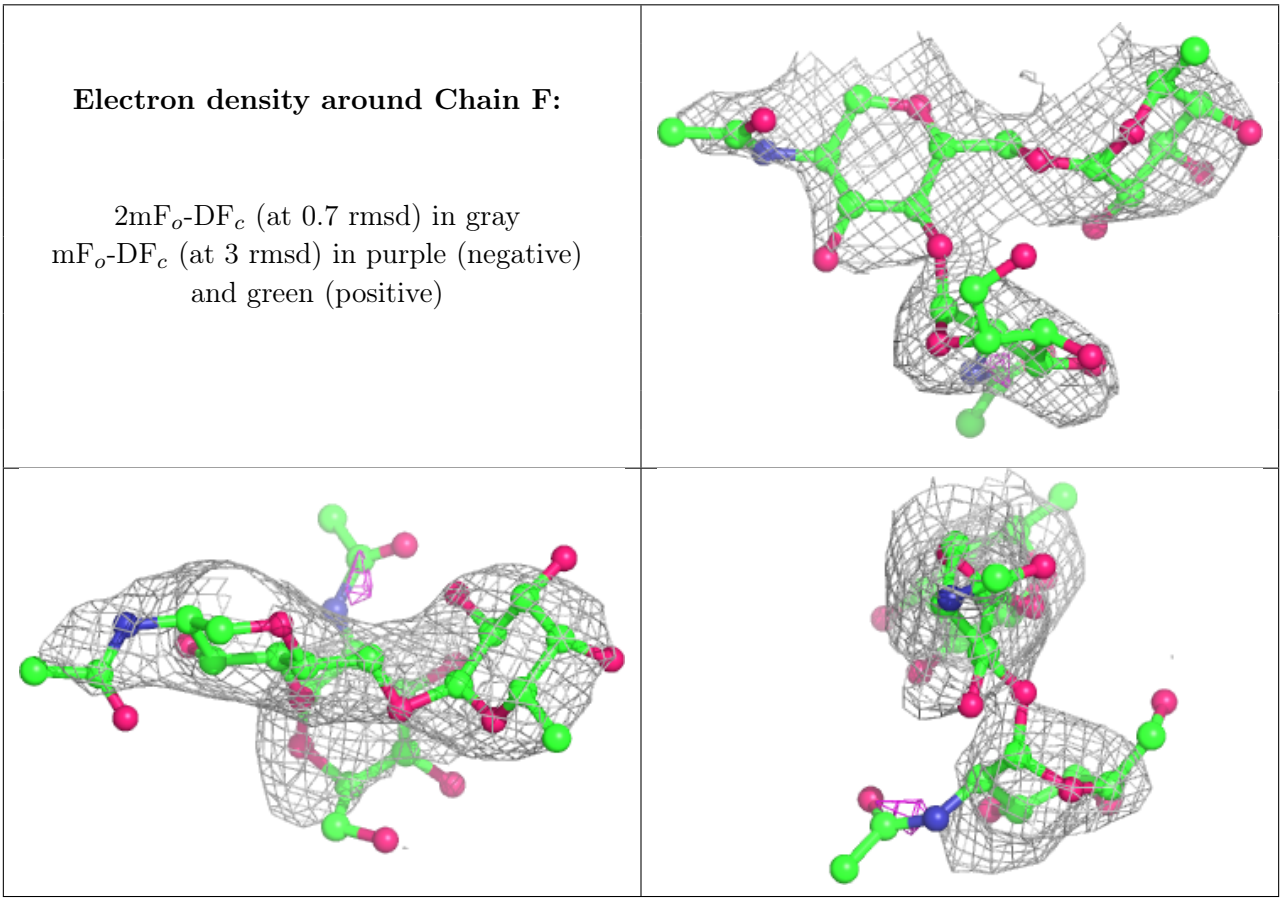
Electron density around Chain D:

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



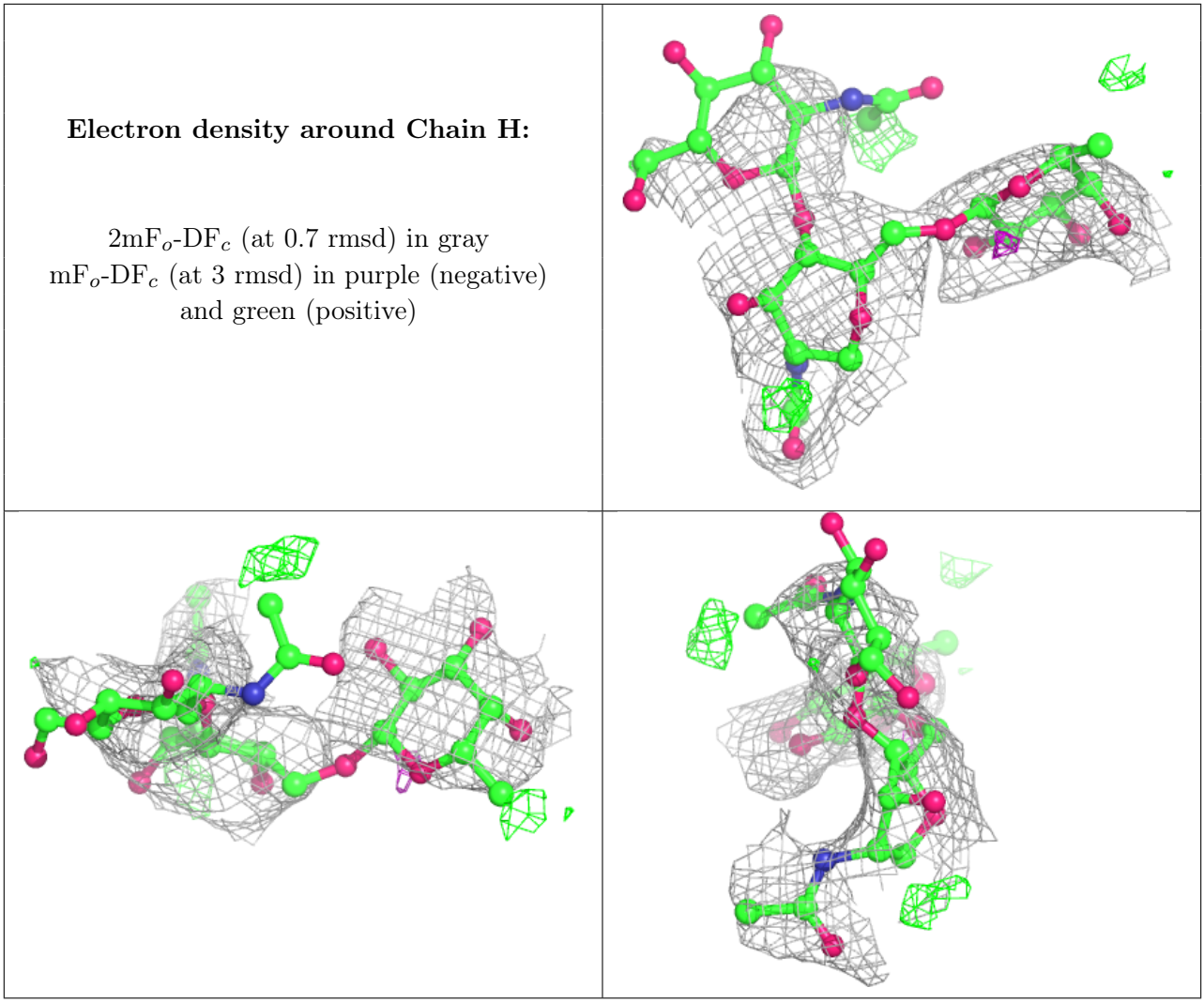
Electron density around Chain F:

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



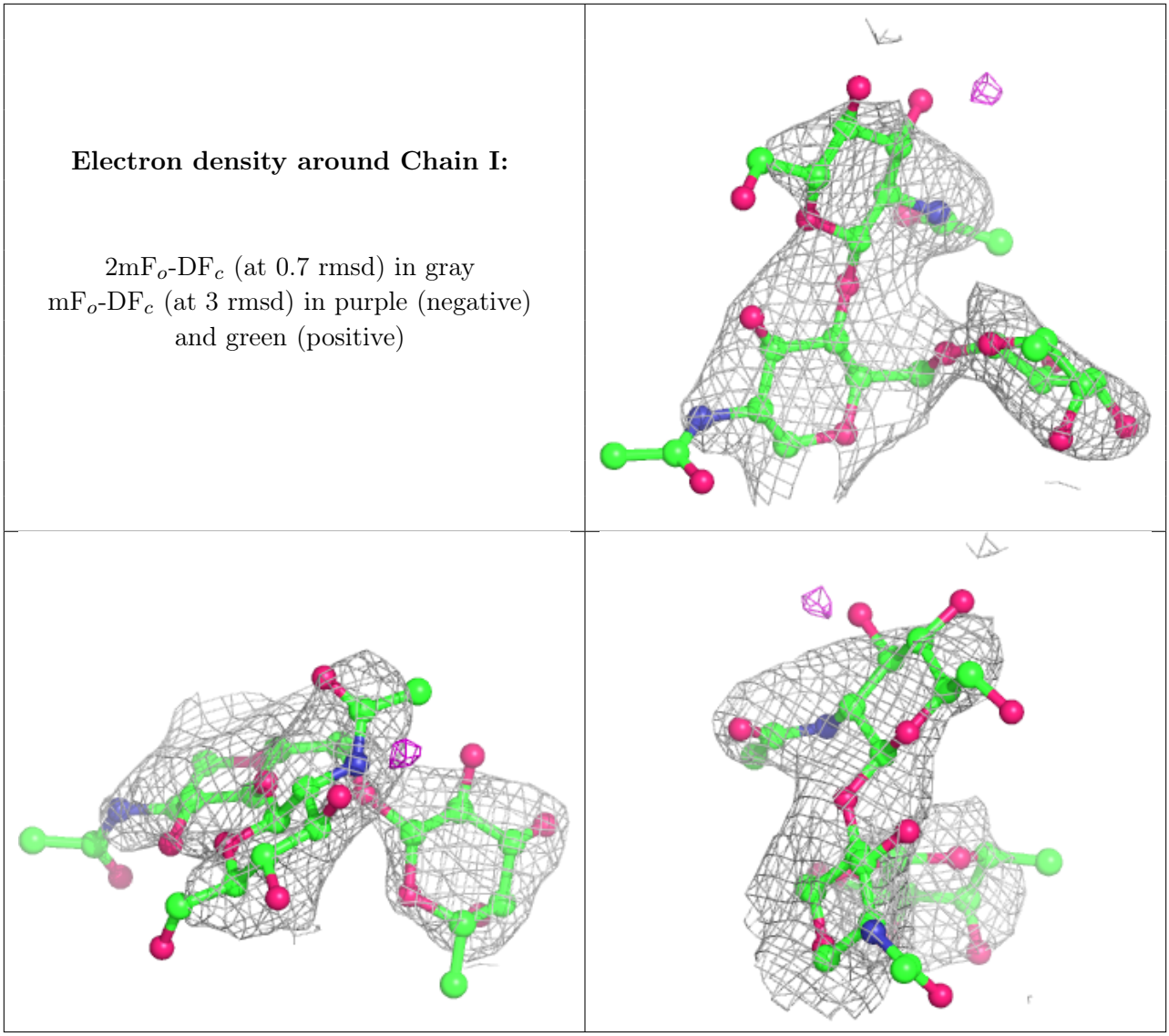
Electron density around Chain H:

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



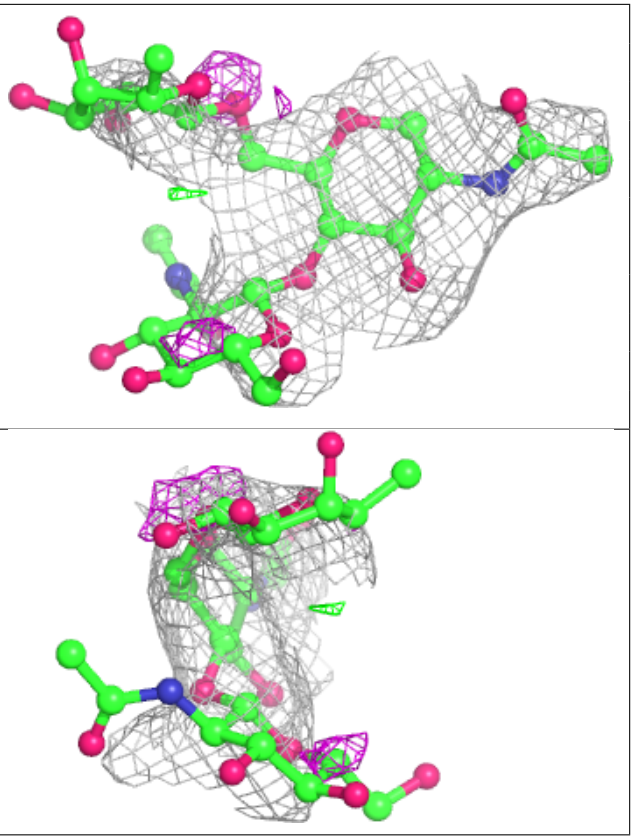
Electron density around Chain I:

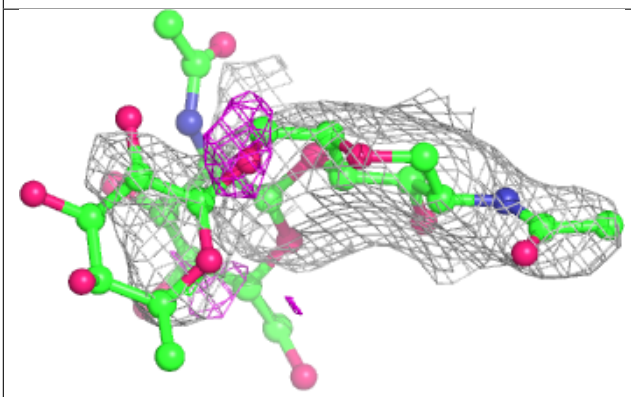
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



Electron density around Chain J:

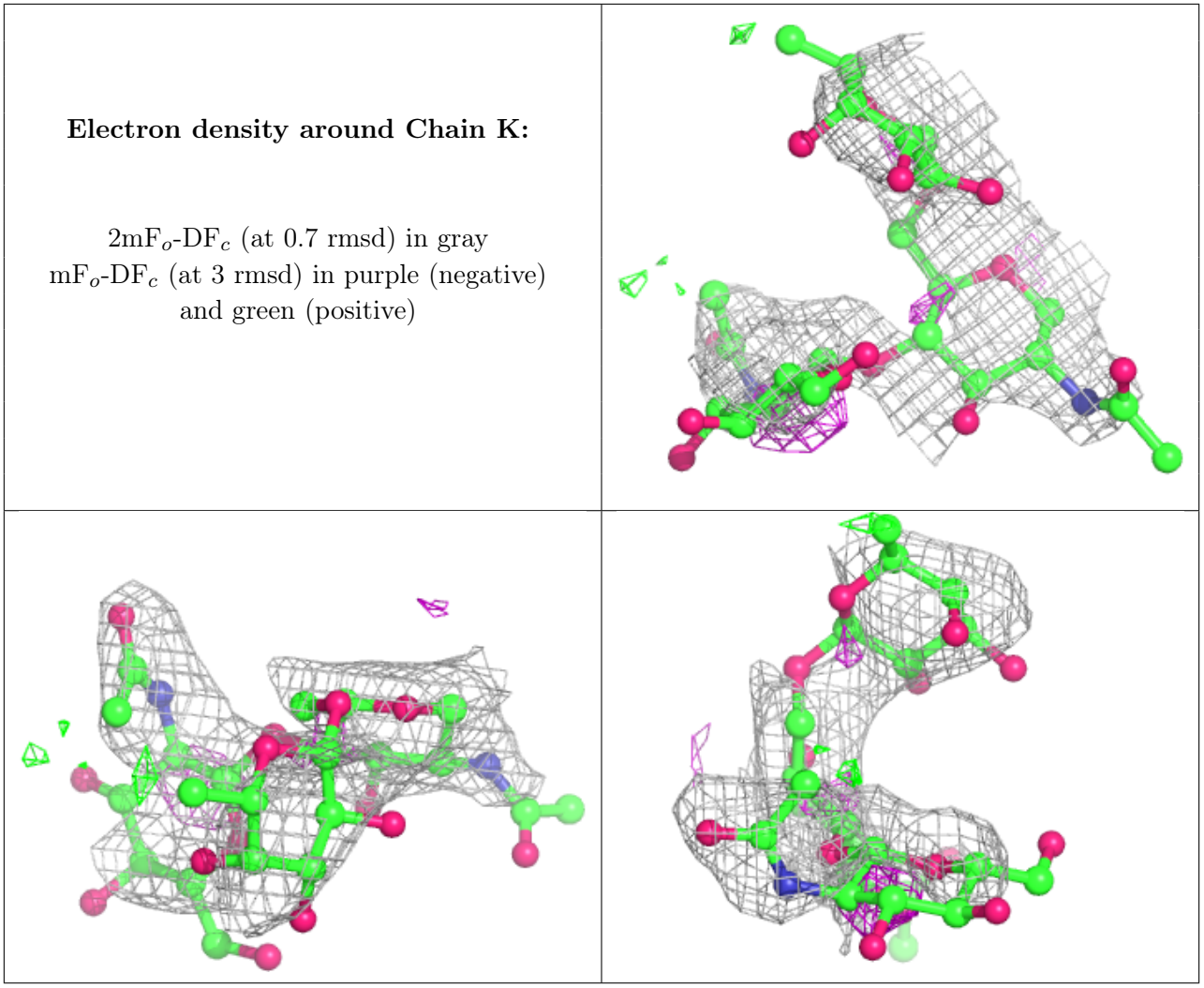
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)





Electron density around Chain K:

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



6.4 Ligands

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, 95^{th} 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
6	NAG	B	602	14/15	0.65	0.34	95,111,120,122	0
8	CL	A	620	1/1	0.67	0.43	93,93,93,93	0
7	EDO	A	621	4/4	0.77	0.30	57,61,67,77	0
6	NAG	B	603	14/15	0.78	0.32	83,95,102,104	0
7	EDO	A	616	4/4	0.79	0.18	72,76,77,83	0
7	EDO	A	618	4/4	0.81	0.43	41,45,60,62	0
6	NAG	B	614	14/15	0.81	0.25	71,104,114,121	0
6	NAG	B	601	14/15	0.81	0.30	98,113,126,126	0
7	EDO	B	617	4/4	0.84	0.50	56,61,62,65	0

Continued on next page...

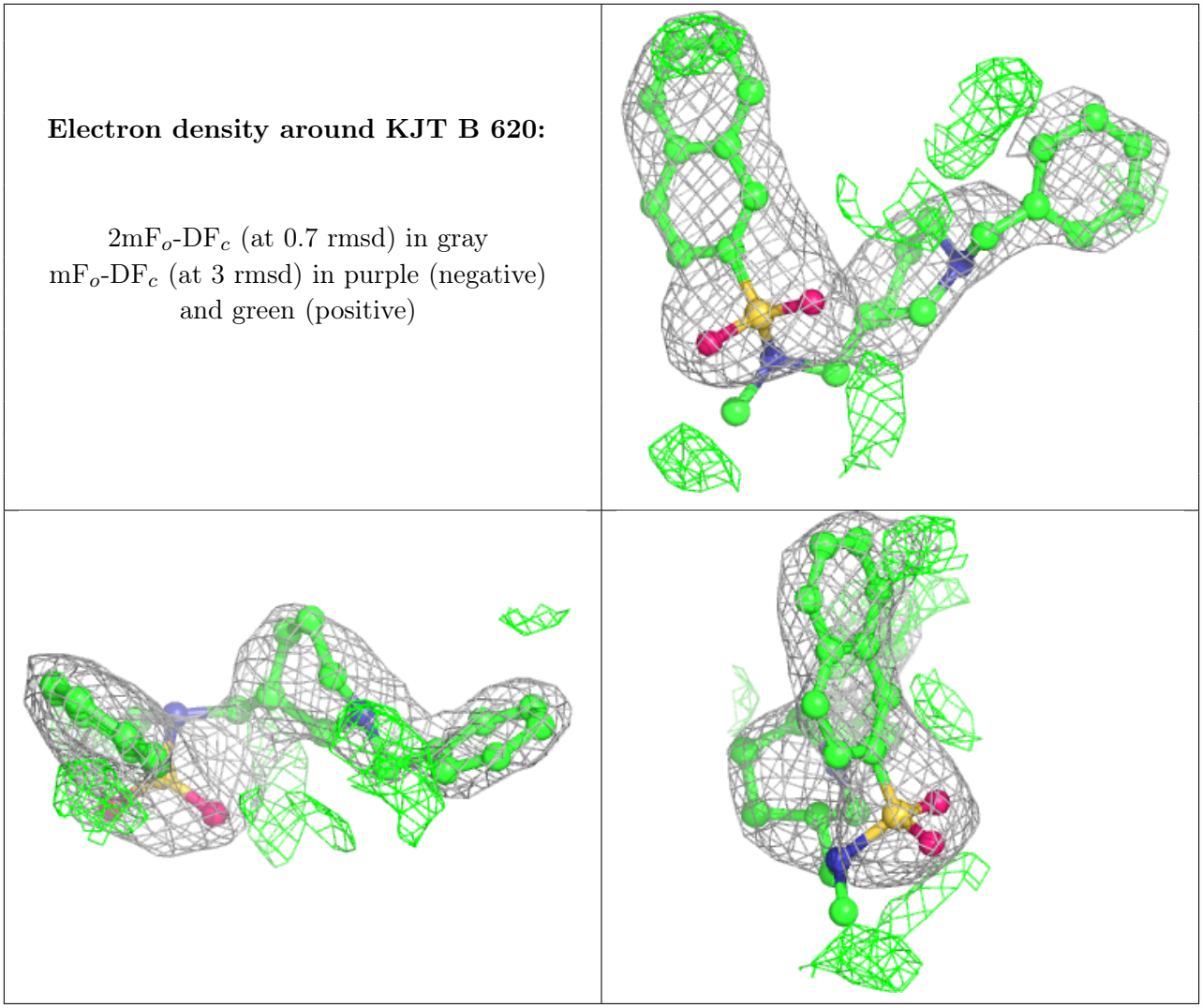
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	EDO	A	617	4/4	0.84	0.58	70,71,77,81	0
6	NAG	A	615	14/15	0.86	0.26	62,72,80,81	0
10	KJT	B	620	29/55	0.87	0.28	40,56,75,83	29
7	EDO	B	616	4/4	0.88	0.42	49,50,56,60	0
11	PGE	B	623	10/10	0.89	0.40	46,54,55,63	0
7	EDO	A	623	4/4	0.92	0.17	52,52,56,58	0
6	NAG	B	610	14/15	0.92	0.22	59,66,72,76	0
8	CL	B	619	1/1	0.93	0.20	84,84,84,84	0
8	CL	B	622	1/1	0.93	0.24	70,70,70,70	0
7	EDO	B	618	4/4	0.94	0.45	57,60,64,69	0
7	EDO	A	619	4/4	0.94	0.09	63,68,69,81	0
9	PEG	B	615	7/7	0.95	0.25	53,58,63,63	0
8	CL	A	622	1/1	0.98	0.44	54,54,54,54	0
8	CL	B	621	1/1	0.98	0.26	51,51,51,51	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 KJT B 620:

$2mF_o$-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

There are no such residues in this entry.