wwPDB X-ray Structure Validation Summary Report (i)

Dec 4, 2024 – 04:24 am GMT

PDB ID	:	9FYZ
Title	:	Crystal structure of SusA amylase from Bacteroides thetaiotaomicron covalently bound to alpha-1,6 branched pseudo-trisaccharide activity-based probe
Authors	:	Pickles, I.B.; Moroz, O.; Davies, G.
Deposited on	:	2024-07-04
Resolution	:	2.43 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:

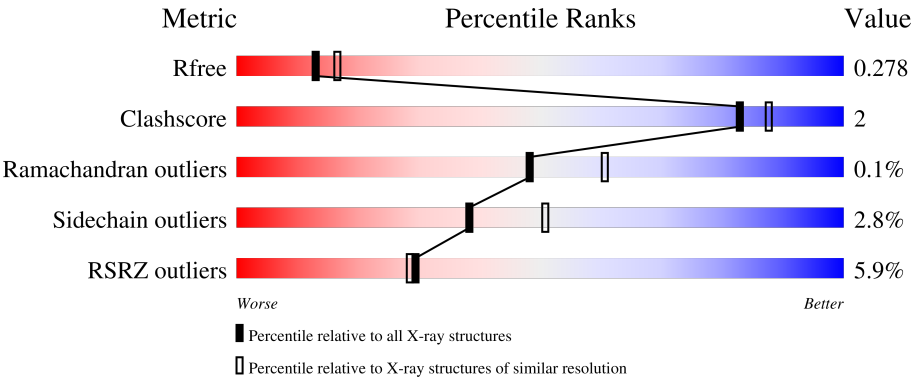
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 2.43 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}	164625	2124 (2.46-2.42)
Clashscore	180529	2259 (2.46-2.42)
Ramachandran outliers	177936	2244 (2.46-2.42)
Sidechain outliers	177891	2244 (2.46-2.42)
RSRZ outliers	164620	2124 (2.46-2.42)

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	610	88% 9%
1	B	610	4% 87% 9%
1	C	610	6% 89% 7%
1	D	610	6% 90% 7%
1	E	610	6% 89% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	610	13% 89% 7%
2	G	2	100%
2	H	2	100%
2	I	2	100%
2	J	2	100%
2	K	2	100%
2	L	2	50% 50%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 28807 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 Neopullulanase SusA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	594	Total	C	N	O	S	0	0	0
			4797	3060	799	917	21			
1	B	591	Total	C	N	O	S	0	0	0
			4753	3033	793	906	21			
1	C	590	Total	C	N	O	S	0	0	0
			4621	2936	771	893	21			
1	D	594	Total	C	N	O	S	0	0	0
			4624	2942	778	883	21			
1	E	593	Total	C	N	O	S	0	0	0
			4512	2857	767	867	21			
1	F	589	Total	C	N	O	S	0	0	0
			4495	2837	763	874	21			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	MET	-	initiating methionine	UNP Q8A1G0
A	9	HIS	-	expression tag	UNP Q8A1G0
A	10	HIS	-	expression tag	UNP Q8A1G0
A	11	HIS	-	expression tag	UNP Q8A1G0
A	12	HIS	-	expression tag	UNP Q8A1G0
A	13	HIS	-	expression tag	UNP Q8A1G0
A	14	HIS	-	expression tag	UNP Q8A1G0
A	15	GLU	-	expression tag	UNP Q8A1G0
A	16	ASN	-	expression tag	UNP Q8A1G0
A	17	LEU	-	expression tag	UNP Q8A1G0
A	18	TYR	-	expression tag	UNP Q8A1G0
A	19	PHE	-	expression tag	UNP Q8A1G0
A	20	GLN	-	expression tag	UNP Q8A1G0
A	21	GLY	-	expression tag	UNP Q8A1G0
B	8	MET	-	initiating methionine	UNP Q8A1G0
B	9	HIS	-	expression tag	UNP Q8A1G0
B	10	HIS	-	expression tag	UNP Q8A1G0

Continued on next page...

Continued from previous page...

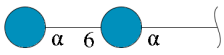
Chain	Residue	Modelled	Actual	Comment	Reference
B	11	HIS	-	expression tag	UNP Q8A1G0
B	12	HIS	-	expression tag	UNP Q8A1G0
B	13	HIS	-	expression tag	UNP Q8A1G0
B	14	HIS	-	expression tag	UNP Q8A1G0
B	15	GLU	-	expression tag	UNP Q8A1G0
B	16	ASN	-	expression tag	UNP Q8A1G0
B	17	LEU	-	expression tag	UNP Q8A1G0
B	18	TYR	-	expression tag	UNP Q8A1G0
B	19	PHE	-	expression tag	UNP Q8A1G0
B	20	GLN	-	expression tag	UNP Q8A1G0
B	21	GLY	-	expression tag	UNP Q8A1G0
C	8	MET	-	initiating methionine	UNP Q8A1G0
C	9	HIS	-	expression tag	UNP Q8A1G0
C	10	HIS	-	expression tag	UNP Q8A1G0
C	11	HIS	-	expression tag	UNP Q8A1G0
C	12	HIS	-	expression tag	UNP Q8A1G0
C	13	HIS	-	expression tag	UNP Q8A1G0
C	14	HIS	-	expression tag	UNP Q8A1G0
C	15	GLU	-	expression tag	UNP Q8A1G0
C	16	ASN	-	expression tag	UNP Q8A1G0
C	17	LEU	-	expression tag	UNP Q8A1G0
C	18	TYR	-	expression tag	UNP Q8A1G0
C	19	PHE	-	expression tag	UNP Q8A1G0
C	20	GLN	-	expression tag	UNP Q8A1G0
C	21	GLY	-	expression tag	UNP Q8A1G0
D	8	MET	-	initiating methionine	UNP Q8A1G0
D	9	HIS	-	expression tag	UNP Q8A1G0
D	10	HIS	-	expression tag	UNP Q8A1G0
D	11	HIS	-	expression tag	UNP Q8A1G0
D	12	HIS	-	expression tag	UNP Q8A1G0
D	13	HIS	-	expression tag	UNP Q8A1G0
D	14	HIS	-	expression tag	UNP Q8A1G0
D	15	GLU	-	expression tag	UNP Q8A1G0
D	16	ASN	-	expression tag	UNP Q8A1G0
D	17	LEU	-	expression tag	UNP Q8A1G0
D	18	TYR	-	expression tag	UNP Q8A1G0
D	19	PHE	-	expression tag	UNP Q8A1G0
D	20	GLN	-	expression tag	UNP Q8A1G0
D	21	GLY	-	expression tag	UNP Q8A1G0
E	8	MET	-	initiating methionine	UNP Q8A1G0
E	9	HIS	-	expression tag	UNP Q8A1G0
E	10	HIS	-	expression tag	UNP Q8A1G0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	11	HIS	-	expression tag	UNP Q8A1G0
E	12	HIS	-	expression tag	UNP Q8A1G0
E	13	HIS	-	expression tag	UNP Q8A1G0
E	14	HIS	-	expression tag	UNP Q8A1G0
E	15	GLU	-	expression tag	UNP Q8A1G0
E	16	ASN	-	expression tag	UNP Q8A1G0
E	17	LEU	-	expression tag	UNP Q8A1G0
E	18	TYR	-	expression tag	UNP Q8A1G0
E	19	PHE	-	expression tag	UNP Q8A1G0
E	20	GLN	-	expression tag	UNP Q8A1G0
E	21	GLY	-	expression tag	UNP Q8A1G0
F	8	MET	-	initiating methionine	UNP Q8A1G0
F	9	HIS	-	expression tag	UNP Q8A1G0
F	10	HIS	-	expression tag	UNP Q8A1G0
F	11	HIS	-	expression tag	UNP Q8A1G0
F	12	HIS	-	expression tag	UNP Q8A1G0
F	13	HIS	-	expression tag	UNP Q8A1G0
F	14	HIS	-	expression tag	UNP Q8A1G0
F	15	GLU	-	expression tag	UNP Q8A1G0
F	16	ASN	-	expression tag	UNP Q8A1G0
F	17	LEU	-	expression tag	UNP Q8A1G0
F	18	TYR	-	expression tag	UNP Q8A1G0
F	19	PHE	-	expression tag	UNP Q8A1G0
F	20	GLN	-	expression tag	UNP Q8A1G0
F	21	GLY	-	expression tag	UNP Q8A1G0

- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose.



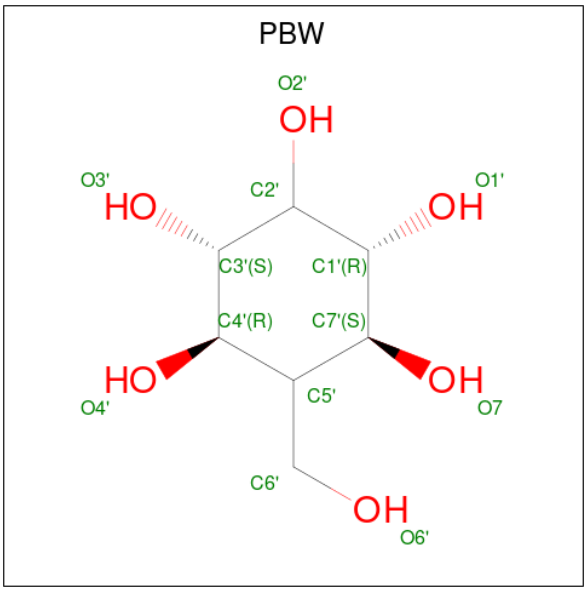
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	G	2	Total C O 22 12 10	0	0	0
2	H	2	Total C O 22 12 10	0	0	0
2	I	2	Total C O 22 12 10	0	0	0
2	J	2	Total C O 22 12 10	0	0	0

Continued on next page...

Continued from previous page...

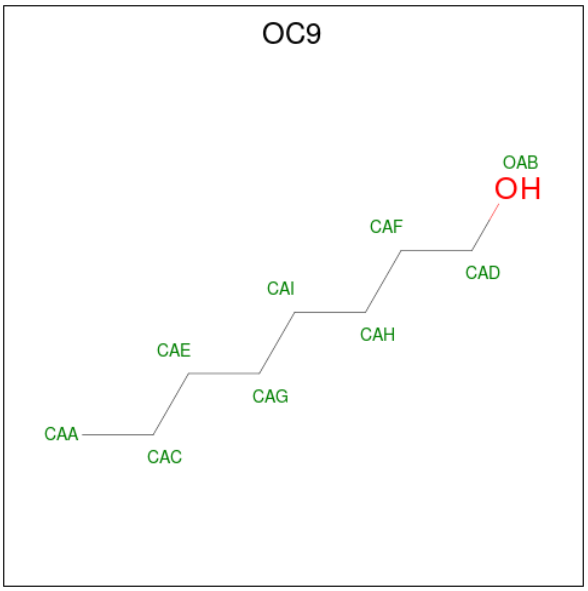
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	K	2	Total C O 22 12 10	0	0	0
2	L	2	Total C O 22 12 10	0	0	0

- Molecule 3 is (1 {S},4 {S},5 {R})-6-(hydroxymethyl)cyclohexane-1,2,3,4,5-pentol (three-letter code: PBW) (formula: $C_7H_{14}O_6$) (labeled as "Ligand of Interest" by depositor).



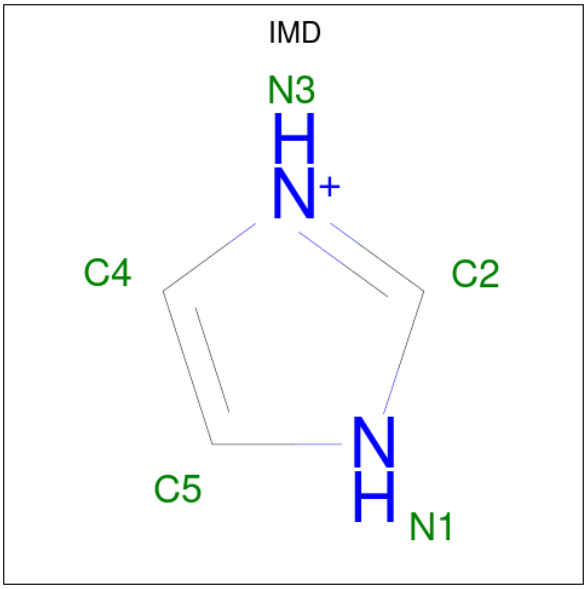
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 12 7 5	0	0
3	B	1	Total C O 12 7 5	0	0
3	C	1	Total C O 12 7 5	0	0
3	D	1	Total C O 12 7 5	0	0
3	E	1	Total C O 12 7 5	0	0
3	F	1	Total C O 12 7 5	0	0

- Molecule 4 is OCTAN-1-OL (three-letter code: OC9) (formula: $C_8H_{18}O$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C 4 4	0	0
4	B	1	Total C 4 4	0	0
4	C	1	Total C 2 2	0	0
4	D	1	Total C 3 3	0	0
4	E	1	Total C 2 2	0	0

- Molecule 5 is IMIDAZOLE (three-letter code: IMD) (formula: $C_3H_5N_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N 5 3 2	0	0
5	A	1	Total C N 5 3 2	0	0
5	A	1	Total C N 5 3 2	0	0
5	B	1	Total C N 5 3 2	0	0
5	C	1	Total C N 5 3 2	0	0
5	C	1	Total C N 5 3 2	0	0
5	C	1	Total C N 5 3 2	0	0
5	D	1	Total C N 5 3 2	0	0
5	D	1	Total C N 5 3 2	0	0
5	D	1	Total C N 5 3 2	0	0
5	E	1	Total C N 5 3 2	0	0
5	F	1	Total C N 5 3 2	0	0

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	2	Total Ca 2 2	0	0
6	B	2	Total Ca 2 2	0	0
6	C	2	Total Ca 2 2	0	0
6	D	2	Total Ca 2 2	0	0
6	E	2	Total Ca 2 2	0	0
6	F	2	Total Ca 2 2	0	0

- Molecule 7 is water.

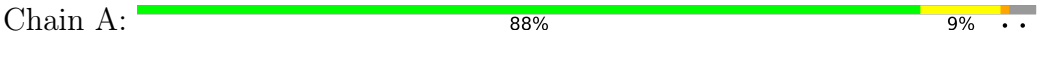
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	295	Total O 295 295	0	0
7	B	180	Total O 180 180	0	0
7	C	84	Total O 84 84	0	0
7	D	68	Total O 68 68	0	0
7	E	42	Total O 42 42	0	0
7	F	45	Total O 45 45	0	0

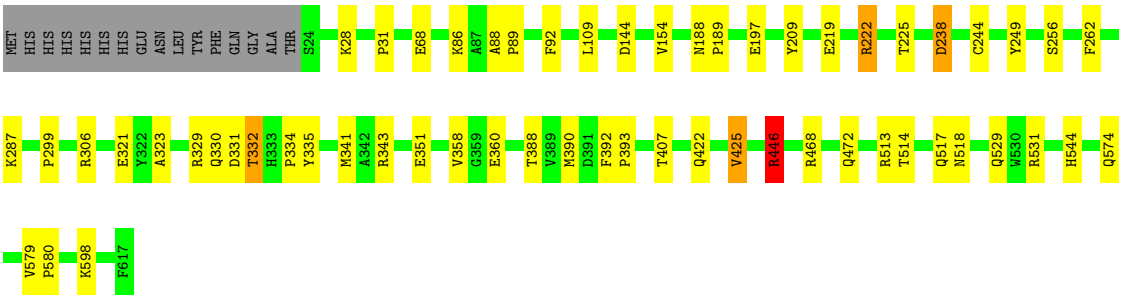
3 Residue-property plots (i)

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: Neopullulanase SusA

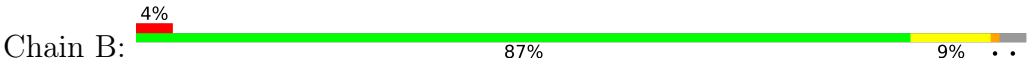
Chain A:

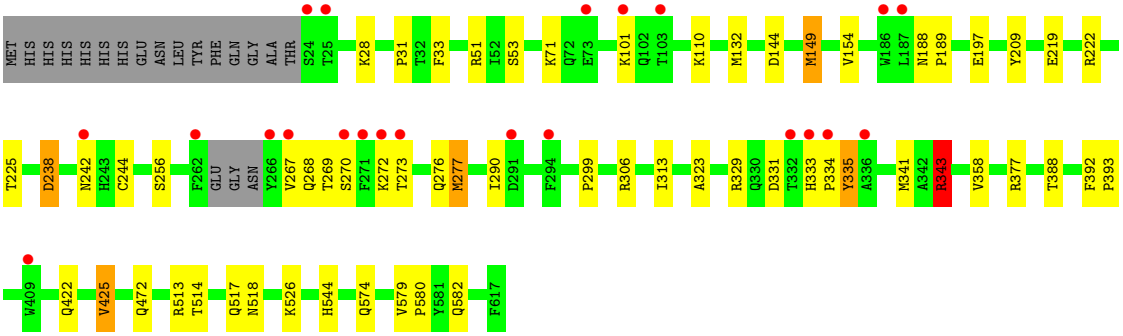




• Molecule 1: Neopullulanase SusA

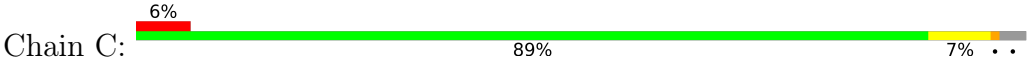
Chain B:

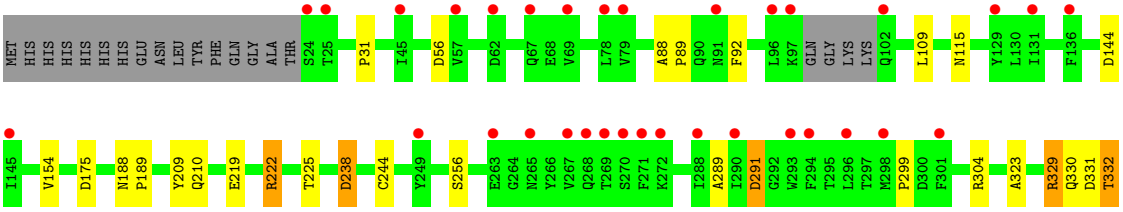


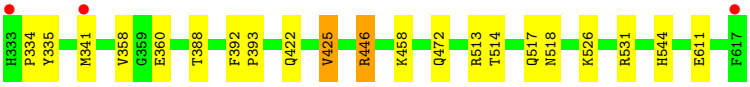


• Molecule 1: Neopullulanase SusA

Chain C:

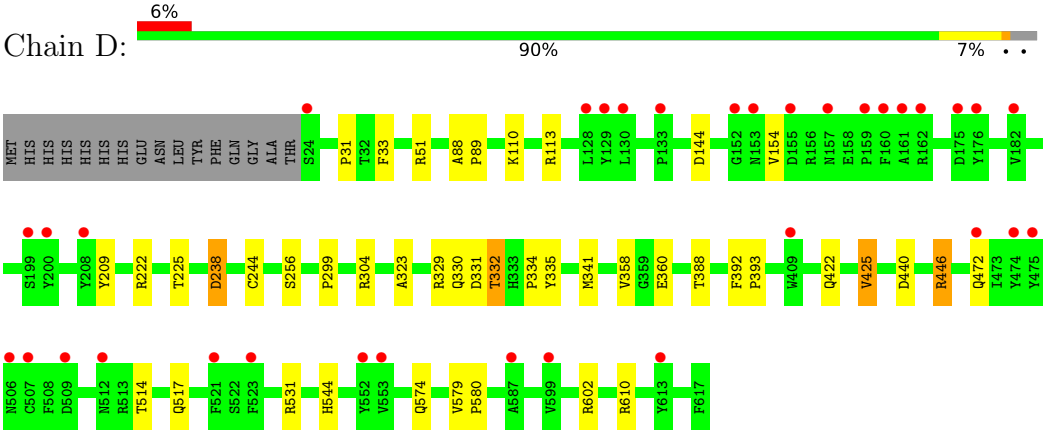






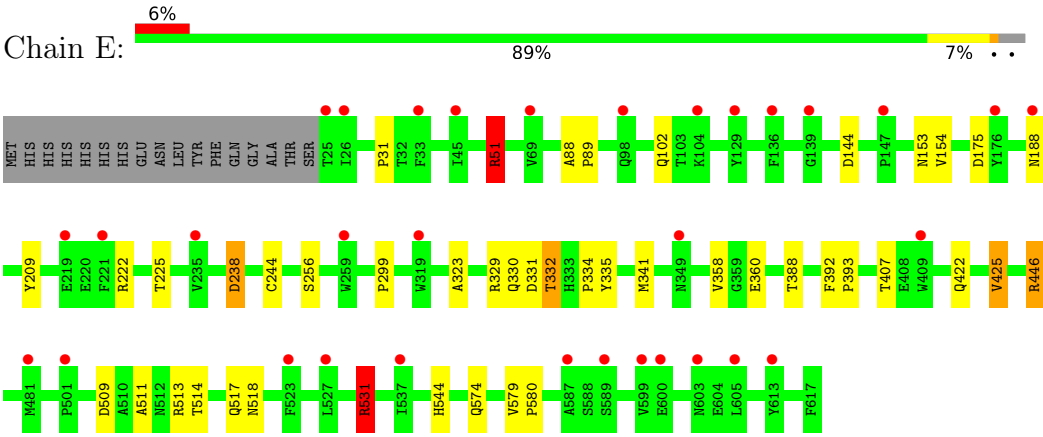
• Molecule 1: Neopullulanase SusA

Chain D:



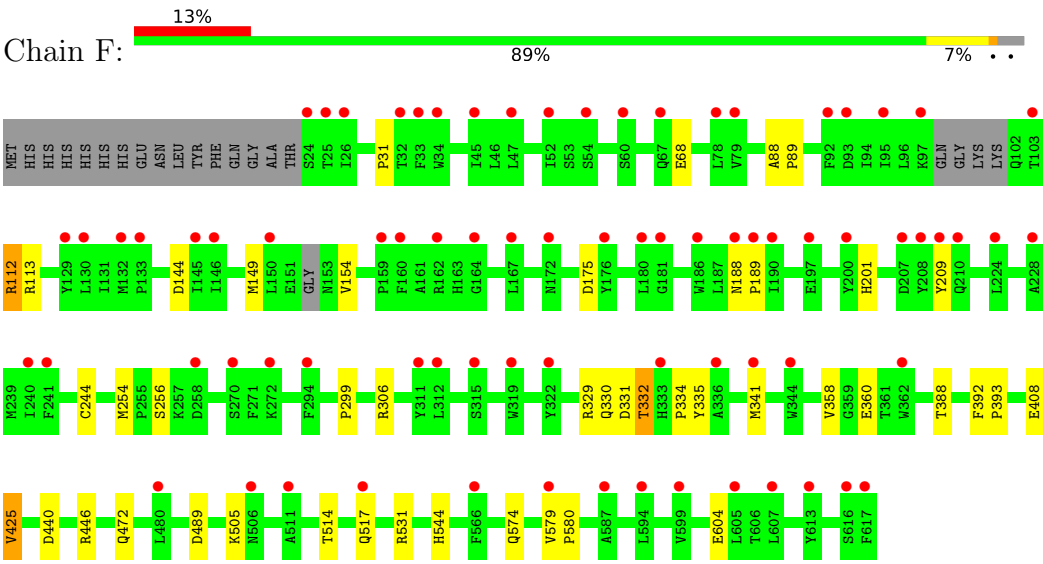
• Molecule 1: Neopullulanase SusA

Chain E:



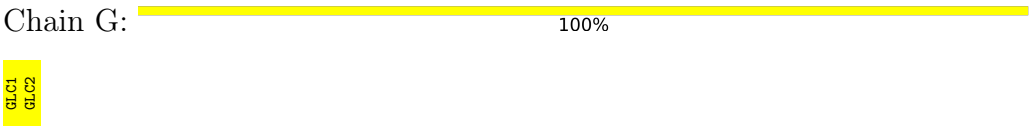
• Molecule 1: Neopullulanase SusA

Chain F:



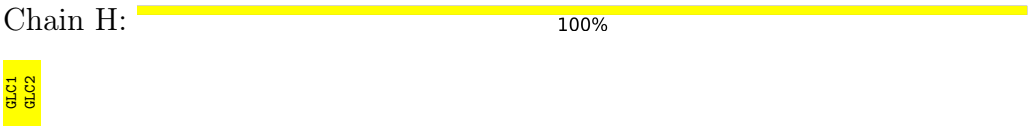
• Molecule 2: alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose

Chain G:



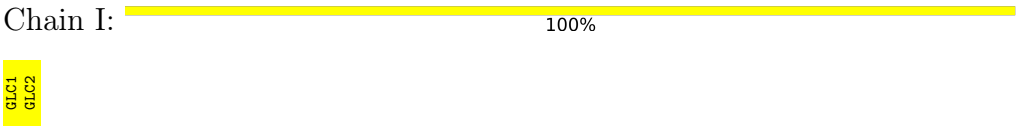
• Molecule 2: alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose

Chain H:



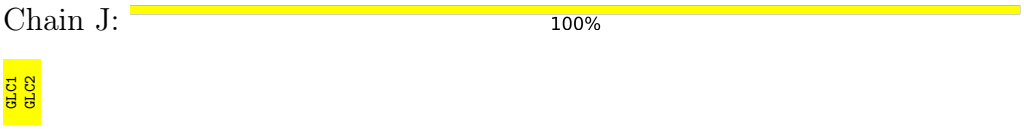
• Molecule 2: alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose

Chain I:



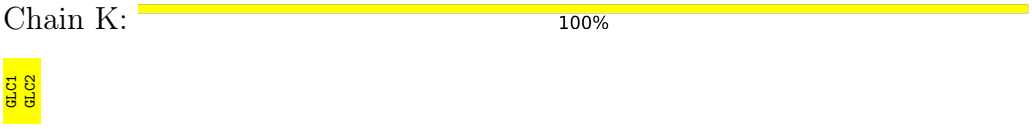
• Molecule 2: alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose

Chain J:



• Molecule 2: alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose

Chain K:



• Molecule 2: alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose

Chain L:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	105.98Å 105.98Å 753.69Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	58.25 – 2.43 58.25 – 2.43	Depositor EDS
% Data completeness (in resolution range)	100.0 (58.25-2.43) 100.0 (58.25-2.43)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.45 (at 2.42Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.230 , 0.281 0.230 , 0.278	Depositor DCC
R_{free} test set	9310 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	58.0	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 57.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.017 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	28807	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.87% 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: OC9, PBW, CA, IMD, GLC

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	0.52	1/4921 (0.0%)	0.82	8/6683 (0.1%)
1	B	0.45	0/4874	0.86	6/6619 (0.1%)
1	C	0.40	0/4743	0.73	3/6464 (0.0%)
1	D	0.38	0/4745	0.73	3/6469 (0.0%)
1	E	0.37	0/4633	0.82	4/6330 (0.1%)
1	F	0.37	0/4613	0.75	10/6297 (0.2%)
All	All	0.42	1/28529 (0.0%)	0.79	34/38862 (0.1%)

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	3
1	B	0	2
1	C	0	4
1	D	0	4
1	E	0	4
1	F	0	2
All	All	0	19

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	351	GLU	CD-OE1	5.17	1.31	1.25

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	531	ARG	NE-CZ-NH1	-31.08	104.76	120.30
1	B	149	MET	CG-SD-CE	-14.85	76.44	100.20
1	B	331	ASP	CB-CG-OD2	-14.85	104.94	118.30
1	B	331	ASP	CB-CG-OD1	12.41	129.47	118.30
1	E	331	ASP	CB-CG-OD2	-10.93	108.47	118.30

There are no chirality outliers.

5 of 19 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	222	ARG	Sidechain
1	A	446	ARG	Sidechain
1	A	468	ARG	Sidechain
1	B	306	ARG	Sidechain
1	B	343	ARG	Sidechain

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	4797	0	4568	24	0
1	B	4753	0	4523	21	0
1	C	4621	0	4192	22	0
1	D	4624	0	4209	18	0
1	E	4512	0	3964	18	0
1	F	4495	0	3917	20	0
2	G	22	0	18	0	0
2	H	22	0	18	0	0
2	I	22	0	18	0	0
2	J	22	0	18	0	0
2	K	22	0	18	0	0
2	L	22	0	19	2	0
3	A	12	0	0	0	0
3	B	12	0	0	0	0
3	C	12	0	0	0	0
3	D	12	0	0	1	0
3	E	12	0	0	0	0
3	F	12	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	4	0	4	0	0
4	B	4	0	4	0	0
4	C	2	0	0	0	0
4	D	3	0	2	0	0
4	E	2	0	0	0	0
5	A	15	0	15	0	0
5	B	5	0	5	0	0
5	C	15	0	15	3	0
5	D	15	0	15	1	0
5	E	5	0	5	0	0
5	F	5	0	5	0	0
6	A	2	0	0	0	0
6	B	2	0	0	0	0
6	C	2	0	0	0	0
6	D	2	0	0	0	0
6	E	2	0	0	0	0
6	F	2	0	0	0	0
7	A	295	0	0	3	1
7	B	180	0	0	0	0
7	C	84	0	0	1	1
7	D	68	0	0	0	0
7	E	42	0	0	0	0
7	F	45	0	0	2	0
All	All	28807	0	25552	123	1

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.

The worst 5 of 123 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:B:290:ILE:HG12	1:B:335:TYR:CZ	2.38	0.58
1:B:269:THR:HG21	1:B:334:PRO:HB3	1.86	0.57
1:C:289:ALA:O	1:C:291:ASP:N	2.34	0.51
1:D:440:ASP:OD1	3:D:701:PBW:O3'	2.30	0.49
1:E:509:ASP:O	1:E:511:ALA:O	2.31	0.49

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:1057:HOH:O	7:C:849:HOH:O[4_665]	2.06	0.14

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	592/610 (97%)	575 (97%)	17 (3%)	0	100 100
1	B	587/610 (96%)	564 (96%)	21 (4%)	2 (0%)	37 44
1	C	586/610 (96%)	567 (97%)	19 (3%)	0	100 100
1	D	592/610 (97%)	573 (97%)	19 (3%)	0	100 100
1	E	591/610 (97%)	571 (97%)	20 (3%)	0	100 100
1	F	583/610 (96%)	565 (97%)	18 (3%)	0	100 100
All	All	3531/3660 (96%)	3415 (97%)	114 (3%)	2 (0%)	48 60

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	335	TYR
1	B	268	GLN

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	513/537 (96%)	500 (98%)	13 (2%)	42 55

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	506/537 (94%)	486 (96%)	20 (4%)	27	36
1	C	467/537 (87%)	456 (98%)	11 (2%)	44	57
1	D	462/537 (86%)	452 (98%)	10 (2%)	47	60
1	E	432/537 (80%)	417 (96%)	15 (4%)	31	42
1	F	430/537 (80%)	420 (98%)	10 (2%)	45	58
All	All	2810/3222 (87%)	2731 (97%)	79 (3%)	38	51

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	175	ASP
1	F	256	SER
1	E	238	ASP
1	E	472	GLN
1	F	505	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

12 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	GLC	G	1	3,2,4	11,11,12	1.44	2 (18%)	15,15,17	1.65	4 (26%)
2	GLC	G	2	2	11,11,12	0.74	0	15,15,17	1.35	2 (13%)
2	GLC	H	1	3,2,4	11,11,12	1.55	3 (27%)	15,15,17	1.11	2 (13%)
2	GLC	H	2	2	11,11,12	0.86	1 (9%)	15,15,17	1.19	1 (6%)
2	GLC	I	1	3,2,4	11,11,12	1.16	0	15,15,17	1.41	3 (20%)
2	GLC	I	2	2	11,11,12	0.60	0	15,15,17	0.90	1 (6%)
2	GLC	J	1	3,2,4	11,11,12	0.93	0	15,15,17	1.42	1 (6%)
2	GLC	J	2	2	11,11,12	0.50	0	15,15,17	1.28	2 (13%)
2	GLC	K	1	3,2,4	11,11,12	1.25	1 (9%)	15,15,17	1.29	2 (13%)
2	GLC	K	2	2	11,11,12	0.64	0	15,15,17	1.42	3 (20%)
2	GLC	L	1	3,2	11,11,12	1.09	1 (9%)	15,15,17	1.69	2 (13%)
2	GLC	L	2	2	11,11,12	0.76	1 (9%)	15,15,17	1.26	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	GLC	G	1	3,2,4	-	0/2/19/22	0/1/1/1
2	GLC	G	2	2	-	2/2/19/22	0/1/1/1
2	GLC	H	1	3,2,4	-	0/2/19/22	0/1/1/1
2	GLC	H	2	2	-	0/2/19/22	0/1/1/1
2	GLC	I	1	3,2,4	-	0/2/19/22	0/1/1/1
2	GLC	I	2	2	-	2/2/19/22	0/1/1/1
2	GLC	J	1	3,2,4	-	0/2/19/22	0/1/1/1
2	GLC	J	2	2	-	2/2/19/22	0/1/1/1
2	GLC	K	1	3,2,4	-	0/2/19/22	0/1/1/1
2	GLC	K	2	2	-	2/2/19/22	0/1/1/1
2	GLC	L	1	3,2	-	0/2/19/22	0/1/1/1
2	GLC	L	2	2	-	0/2/19/22	0/1/1/1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	1	GLC	O5-C5	3.14	1.49	1.43
2	G	1	GLC	C1-C2	-2.95	1.45	1.52
2	H	2	GLC	O5-C5	2.65	1.48	1.43
2	K	1	GLC	O5-C1	-2.65	1.39	1.43

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	1	GLC	C2-C3	-2.57	1.48	1.52

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	L	1	GLC	C1-C2-C3	-5.17	103.31	109.67
2	J	1	GLC	C1-C2-C3	-4.32	104.35	109.67
2	K	1	GLC	C1-C2-C3	-3.82	104.97	109.67
2	G	1	GLC	C1-C2-C3	-3.49	105.37	109.67
2	K	2	GLC	C3-C4-C5	3.46	116.41	110.24

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	I	2	GLC	O5-C5-C6-O6
2	K	2	GLC	O5-C5-C6-O6
2	J	2	GLC	O5-C5-C6-O6
2	J	2	GLC	C4-C5-C6-O6
2	I	2	GLC	C4-C5-C6-O6

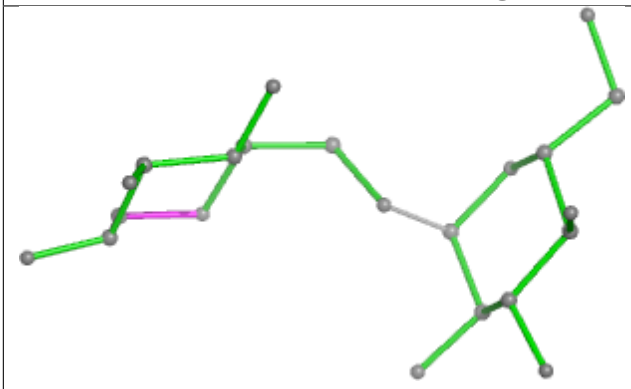
There are no ring outliers.

1 monomer is involved in 2 short contacts:

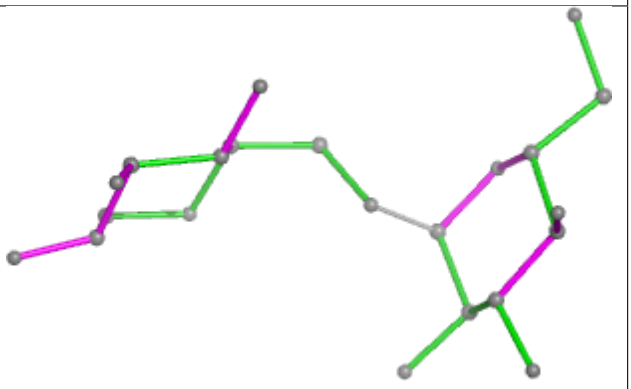
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	L	1	GLC	2	0

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

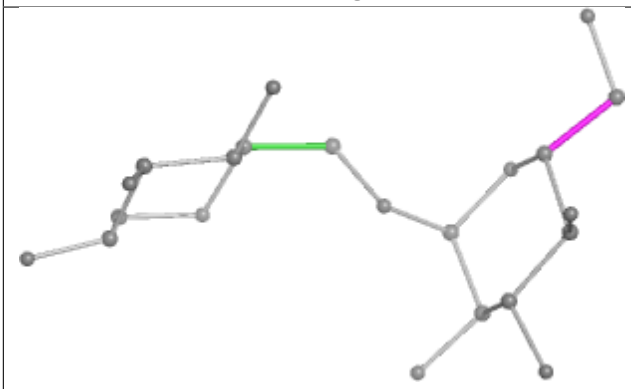
Oligosaccharide Chain G



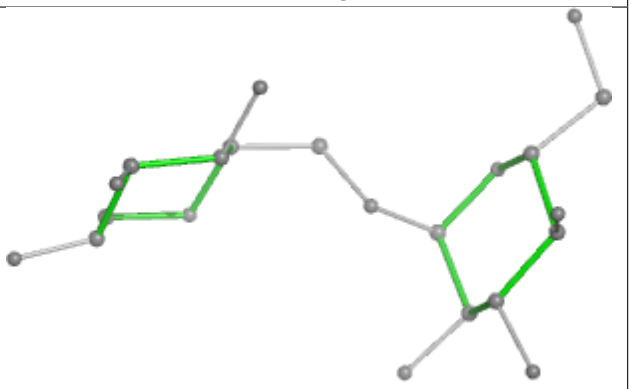
Bond lengths



Bond angles

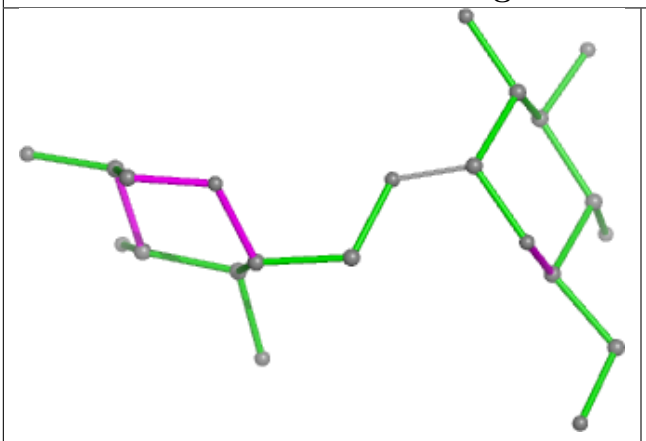


Torsions

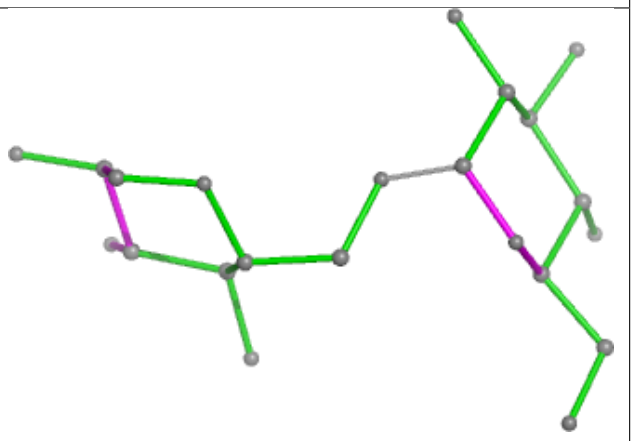


Rings

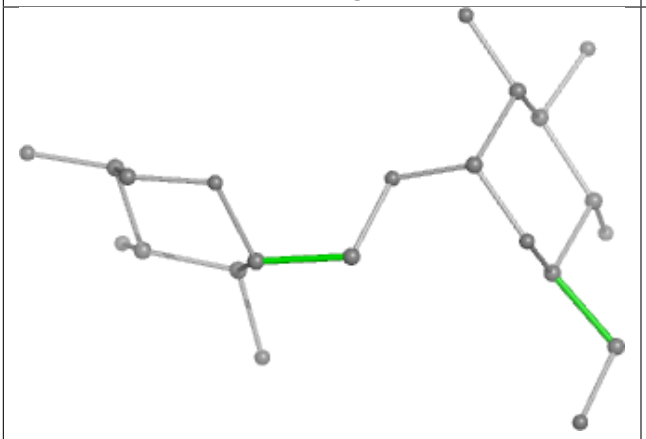
Oligosaccharide Chain H



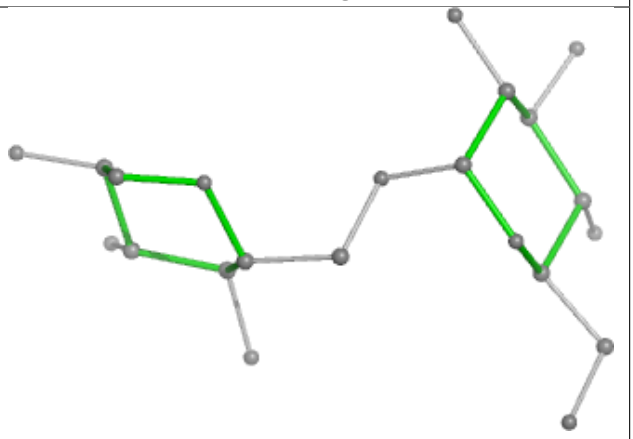
Bond lengths



Bond angles

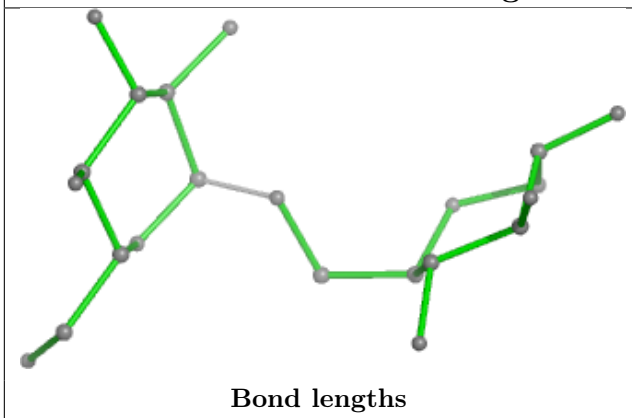


Torsions

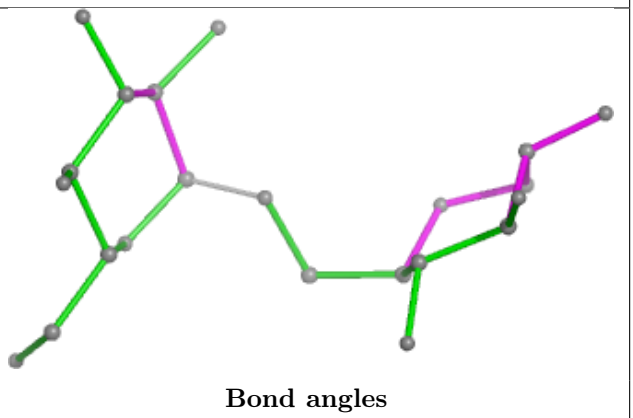


Rings

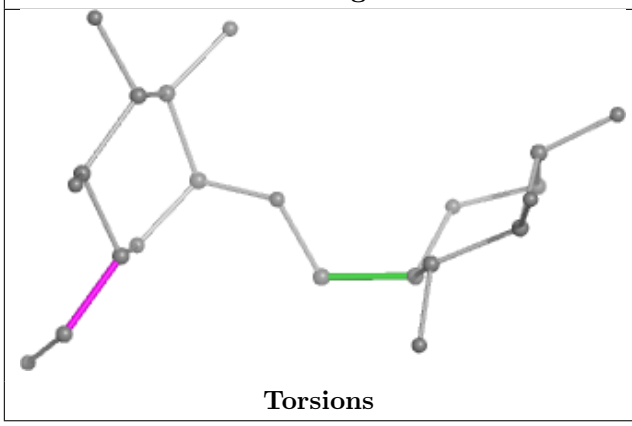
Oligosaccharide Chain I



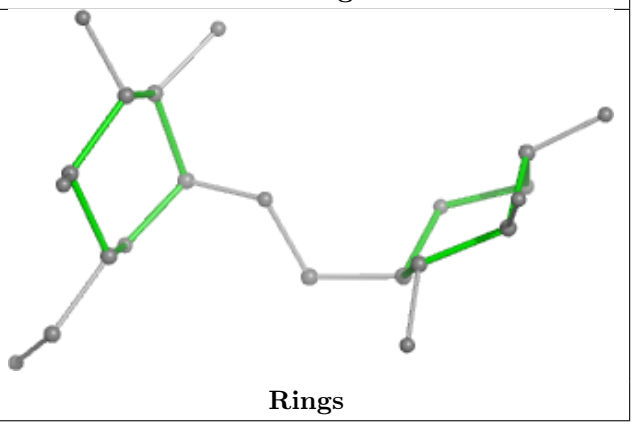
Bond lengths



Bond angles

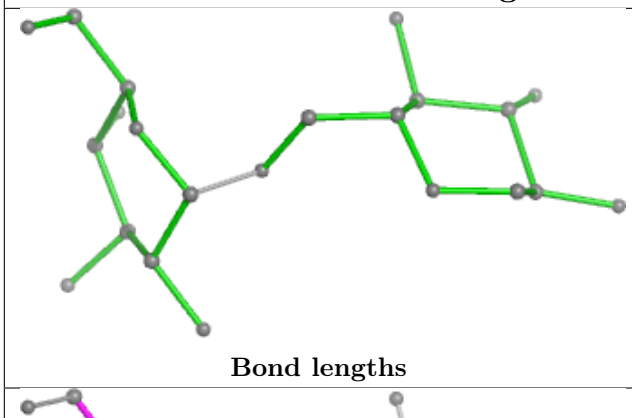


Torsions

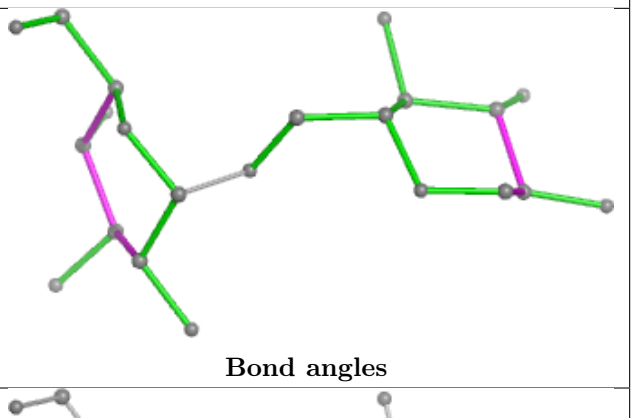


Rings

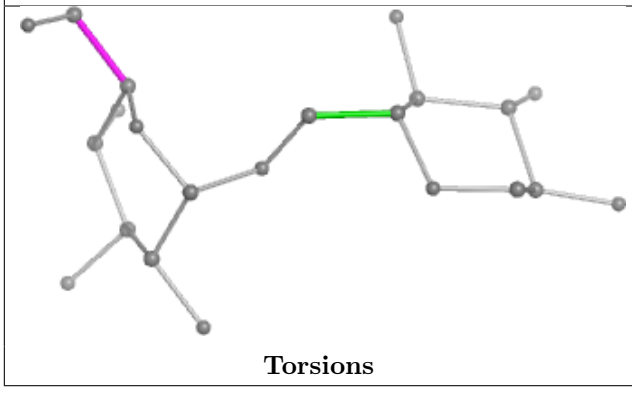
Oligosaccharide Chain J



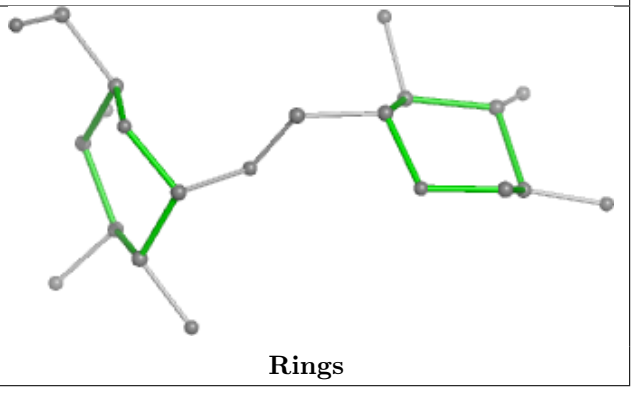
Bond lengths



Bond angles

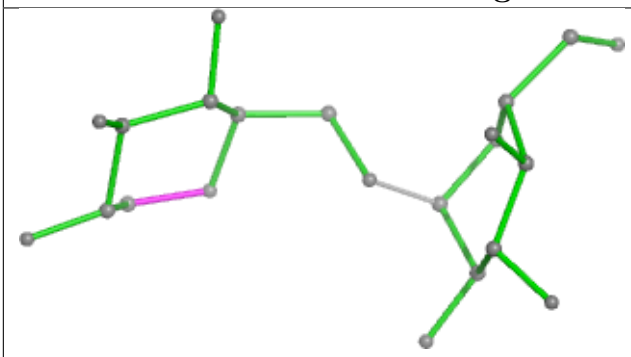


Torsions

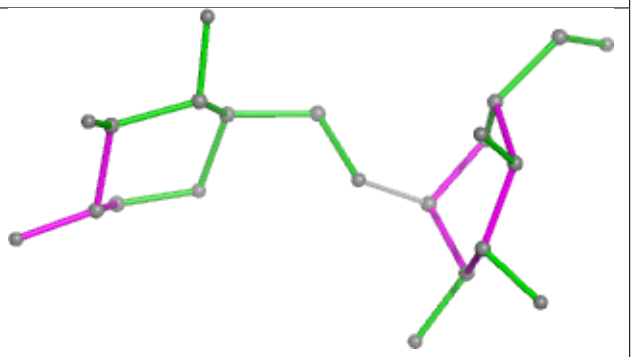


Rings

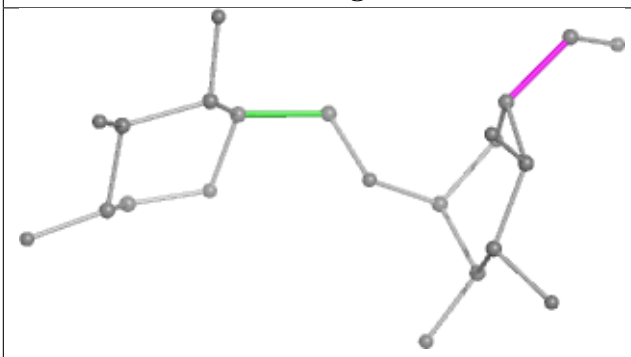
Oligosaccharide Chain K



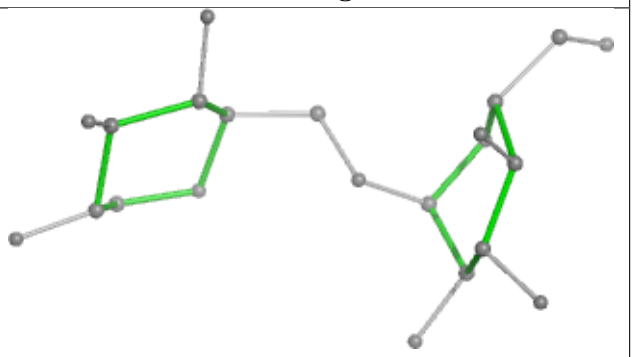
Bond lengths



Bond angles

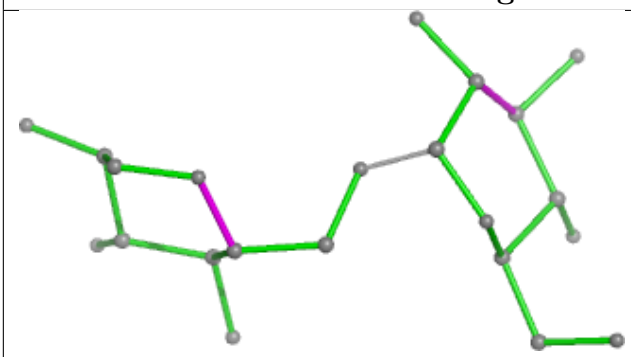


Torsions

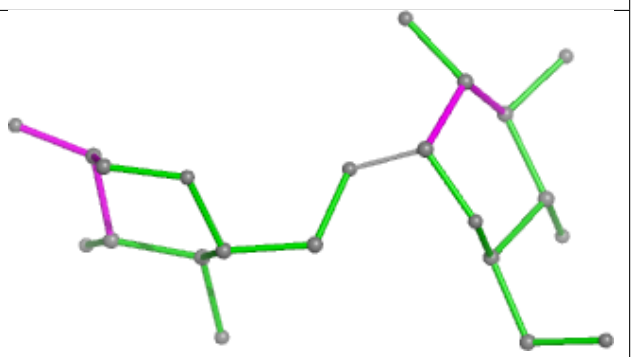


Rings

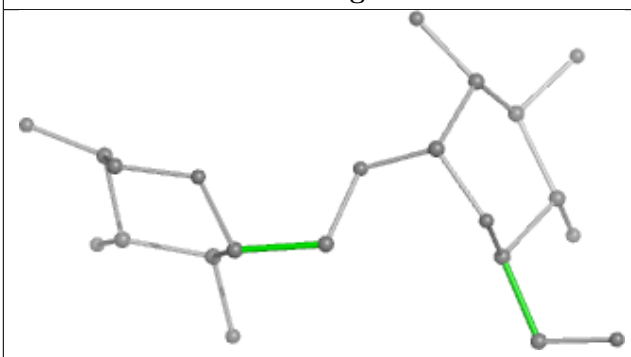
Oligosaccharide Chain L



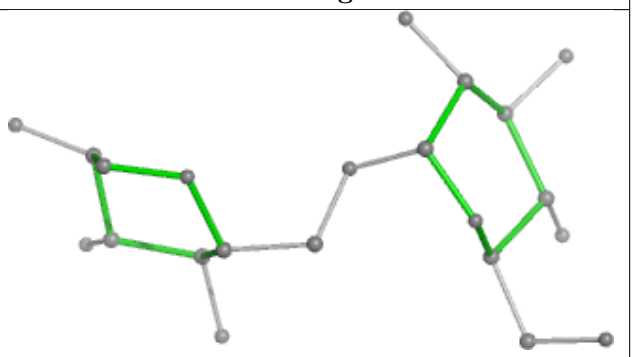
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

Of 35 ligands modelled in this entry, 12 are monoatomic - leaving 23 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	$\#\|Z\| > 2$
5	IMD	D	705	-	3,5,5	0.23	0	4,5,5	0.64	0
3	PBW	B	701	2,1	12,12,13	0.58	0	14,17,19	0.73	0
3	PBW	E	701	2,1	12,12,13	0.58	0	14,17,19	0.87	0
5	IMD	A	704	-	3,5,5	0.35	0	4,5,5	0.80	0
5	IMD	D	704	-	3,5,5	0.27	0	4,5,5	0.56	0
4	OC9	B	702	2	3,3,8	0.56	0	2,2,7	0.16	0
5	IMD	A	705	-	3,5,5	0.47	0	4,5,5	0.54	0
4	OC9	E	702	2	1,1,8	0.54	0	-		
5	IMD	A	703	-	3,5,5	0.61	0	4,5,5	0.81	0
4	OC9	A	702	2	3,3,8	0.47	0	2,2,7	0.23	0
3	PBW	D	701	2,1	12,12,13	0.66	0	14,17,19	0.78	0
4	OC9	D	702	2	2,2,8	0.50	0	0,1,7	-	-
5	IMD	E	703	-	3,5,5	0.25	0	4,5,5	0.75	0
5	IMD	B	703	-	3,5,5	0.42	0	4,5,5	0.69	0
5	IMD	C	703	-	3,5,5	0.26	0	4,5,5	0.64	0
3	PBW	C	701	2,1	12,12,13	0.50	0	14,17,19	0.78	1 (7%)
5	IMD	D	703	-	3,5,5	0.19	0	4,5,5	0.69	0
3	PBW	A	701	2,1	12,12,13	1.85	2 (16%)	14,17,19	1.00	0
4	OC9	C	702	2	1,1,8	0.61	0	-		
5	IMD	C	704	-	3,5,5	0.28	0	4,5,5	0.64	0
5	IMD	F	702	-	3,5,5	0.18	0	4,5,5	0.56	0
5	IMD	C	705	-	3,5,5	0.14	0	4,5,5	0.76	0
3	PBW	F	701	2,1	12,12,13	0.78	1 (8%)	14,17,19	1.05	1 (7%)

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	IMD	D	705	-	-	-	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PBW	B	701	2,1	-	0/2/22/26	0/1/1/1
3	PBW	E	701	2,1	-	0/2/22/26	0/1/1/1
5	IMD	A	704	-	-	-	0/1/1/1
5	IMD	D	704	-	-	-	0/1/1/1
4	OC9	B	702	2	-	0/1/1/6	-
5	IMD	A	705	-	-	-	0/1/1/1
5	IMD	A	703	-	-	-	0/1/1/1
4	OC9	A	702	2	-	0/1/1/6	-
3	PBW	D	701	2,1	-	2/2/22/26	0/1/1/1
5	IMD	E	703	-	-	-	0/1/1/1
5	IMD	B	703	-	-	-	0/1/1/1
5	IMD	C	703	-	-	-	0/1/1/1
3	PBW	C	701	2,1	-	0/2/22/26	0/1/1/1
5	IMD	D	703	-	-	-	0/1/1/1
3	PBW	A	701	2,1	-	0/2/22/26	0/1/1/1
5	IMD	C	704	-	-	-	0/1/1/1
5	IMD	F	702	-	-	-	0/1/1/1
5	IMD	C	705	-	-	-	0/1/1/1
3	PBW	F	701	2,1	-	0/2/22/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	701	PBW	C2'-C3'	-4.58	1.45	1.52
3	A	701	PBW	C5'-C4'	-3.58	1.49	1.53
3	F	701	PBW	C2'-C3'	-2.04	1.49	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	701	PBW	C2'-C3'-C4'	-2.70	106.22	110.89
3	C	701	PBW	C1'-C2'-C3'	2.07	113.69	110.69

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	701	PBW	C4'-C5'-C6'-O6'
3	D	701	PBW	C7'-C5'-C6'-O6'

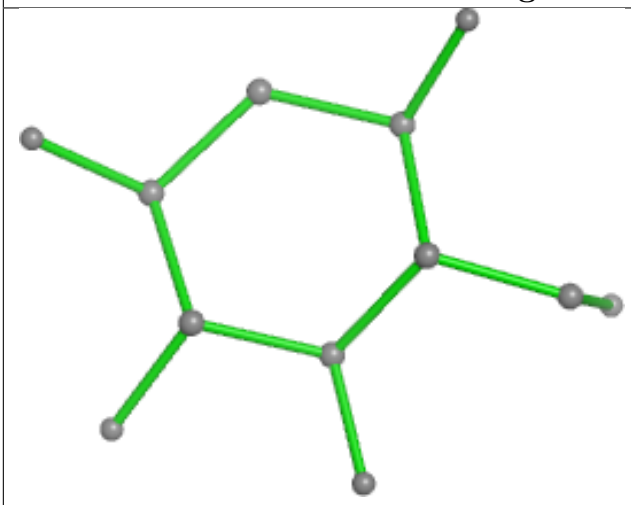
There are no ring outliers.

6 monomers are involved in 6 short contacts:

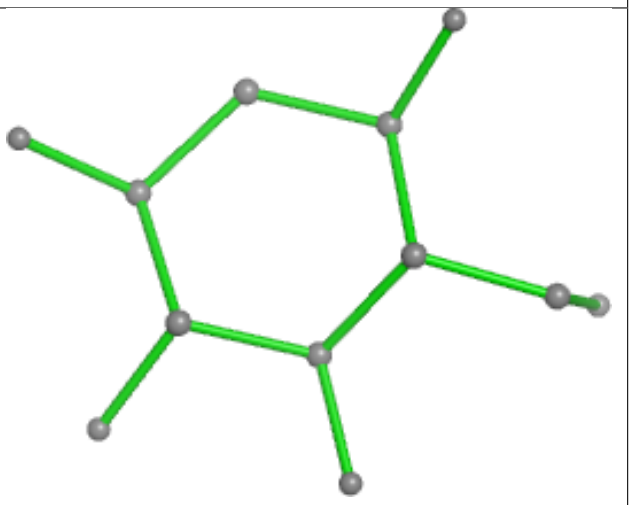
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	704	IMD	1	0
3	D	701	PBW	1	0
5	C	703	IMD	1	0
5	C	704	IMD	1	0
5	C	705	IMD	1	0
3	F	701	PBW	1	0

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.

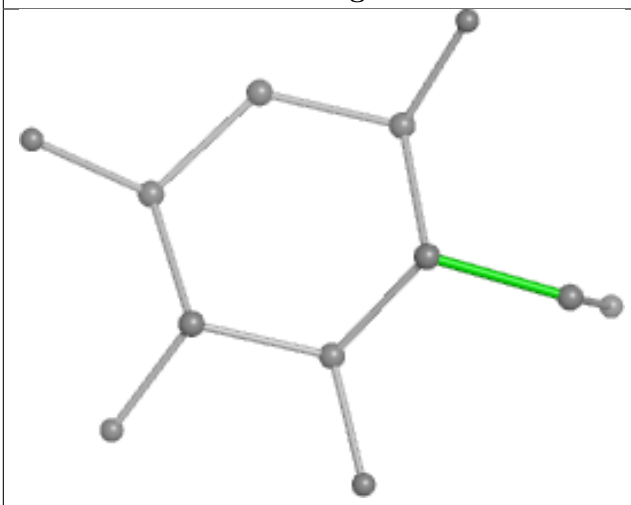
Ligand PBW B 701



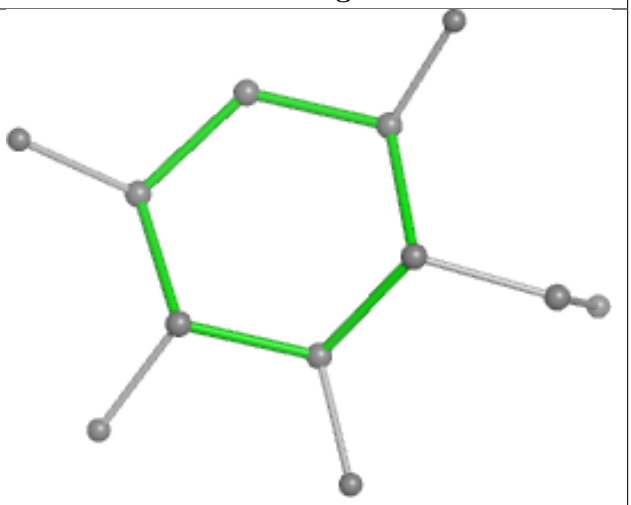
Bond lengths



Bond angles

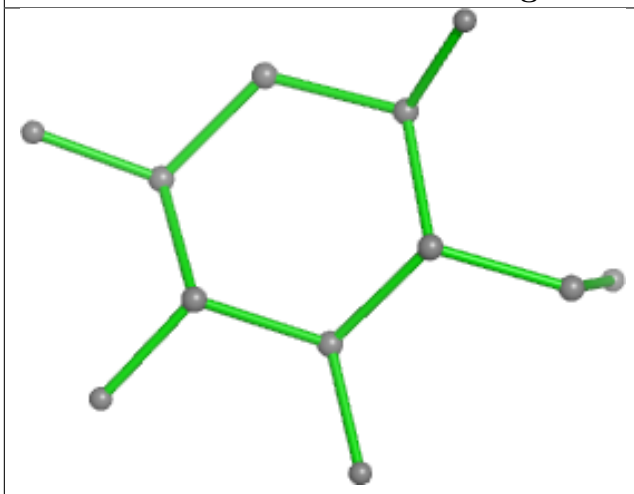


Torsions

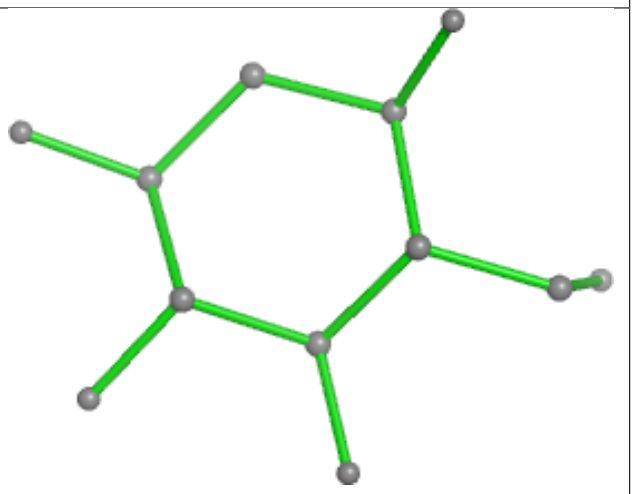


Rings

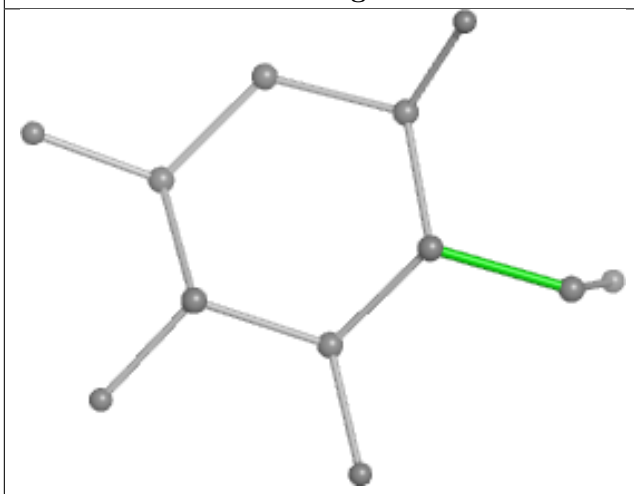
Ligand PBW E 701



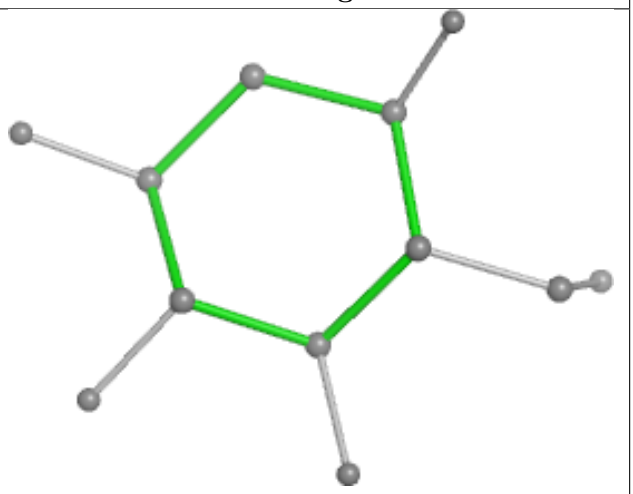
Bond lengths



Bond angles

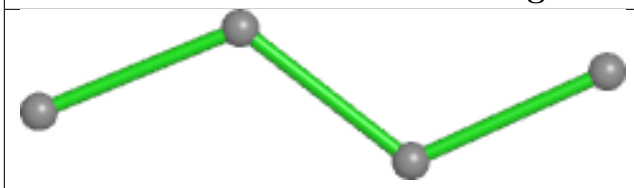


Torsions

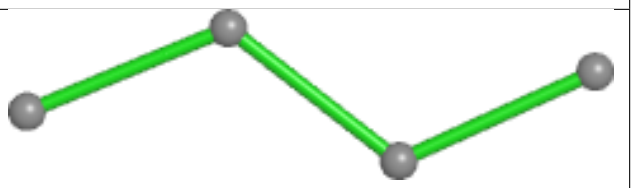


Rings

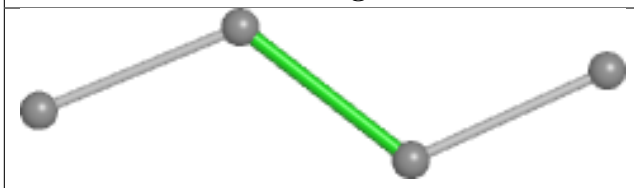
Ligand OC9 B 702



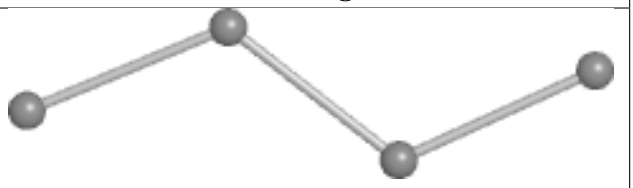
Bond lengths



Bond angles

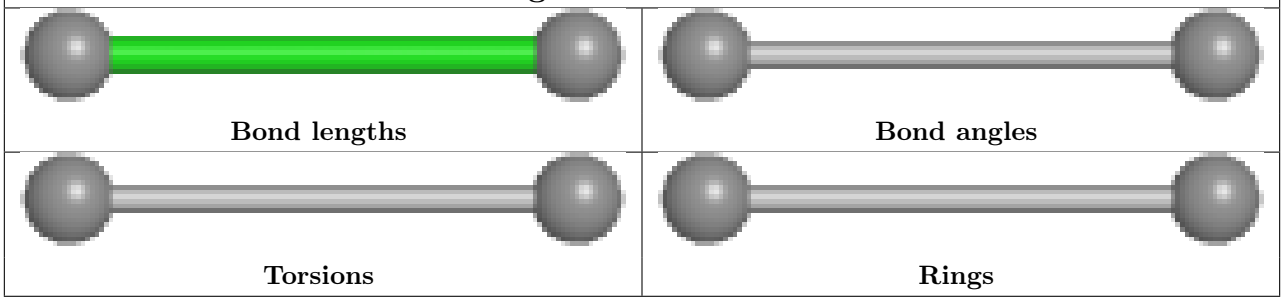


Torsions

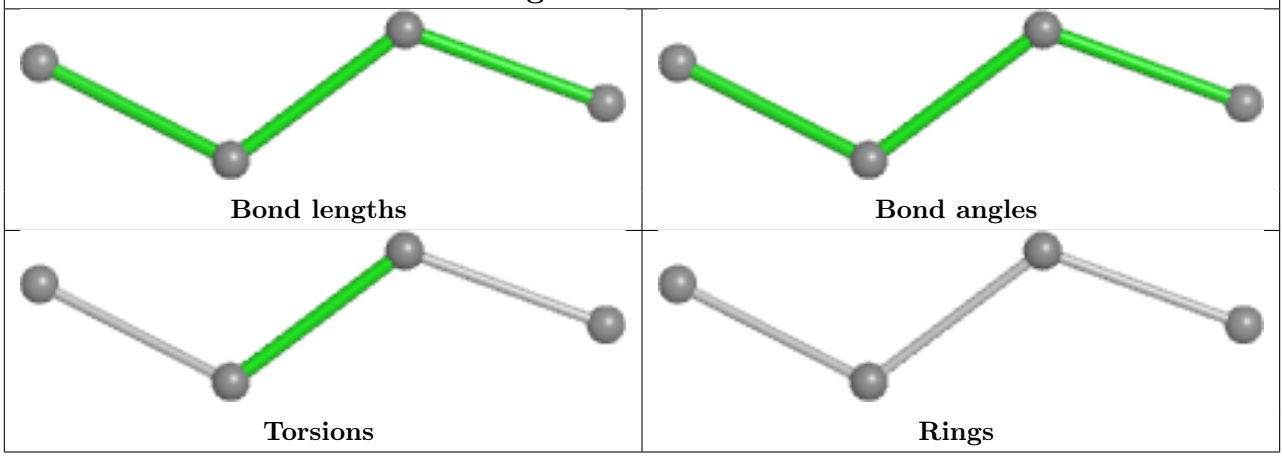


Rings

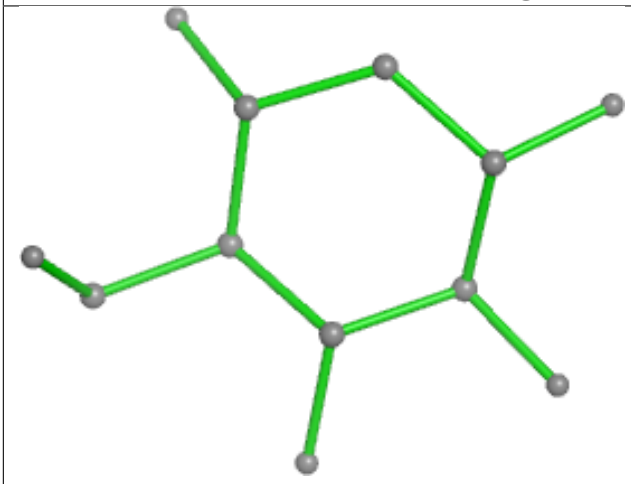
Ligand OC9 E 702



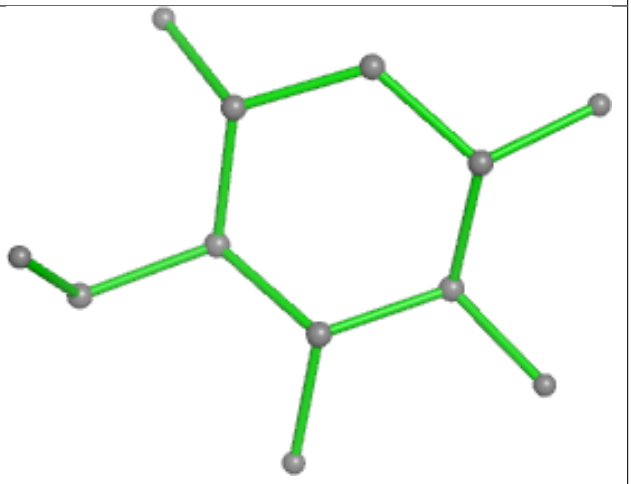
Ligand OC9 A 702



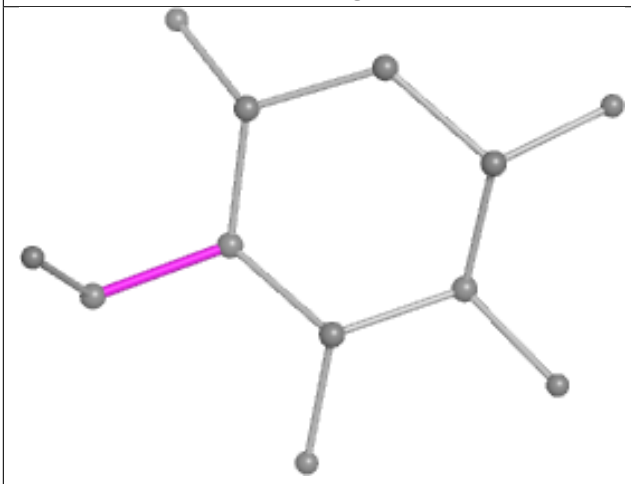
Ligand PBW D 701



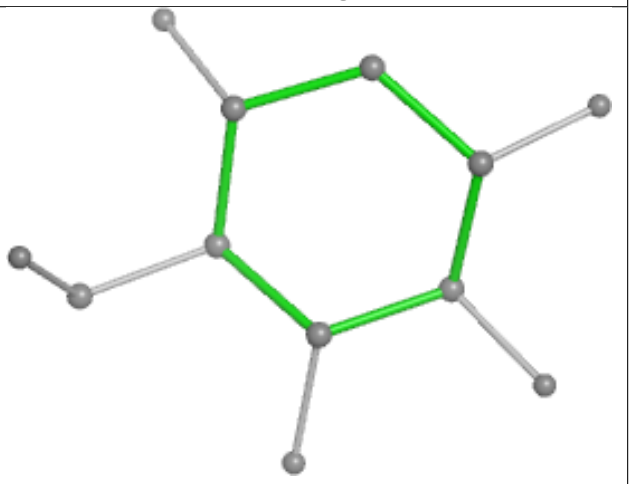
Bond lengths



Bond angles

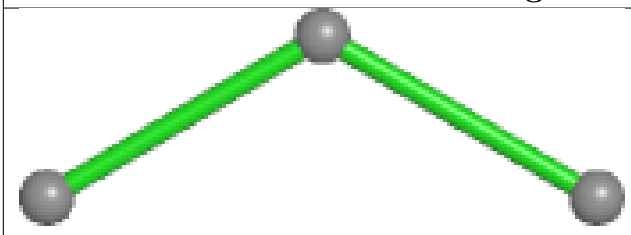


Torsions

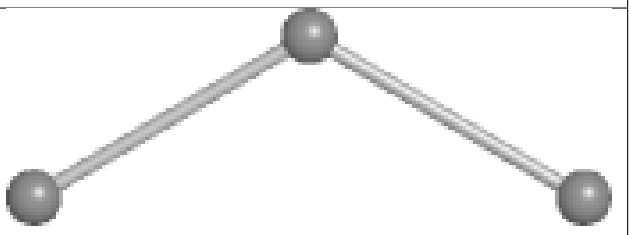


Rings

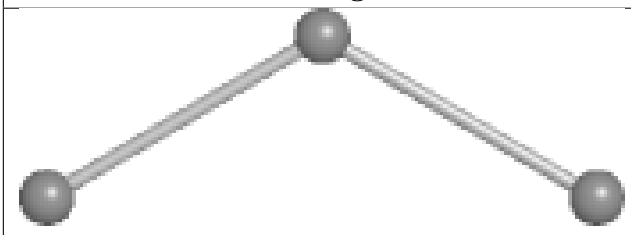
Ligand OC9 D 702



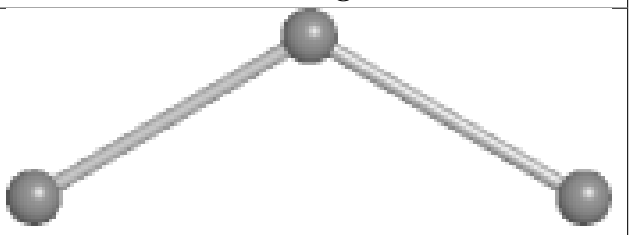
Bond lengths



Bond angles

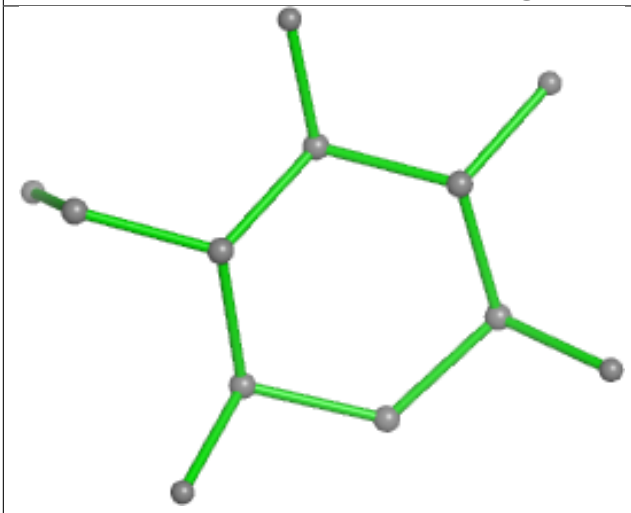


Torsions

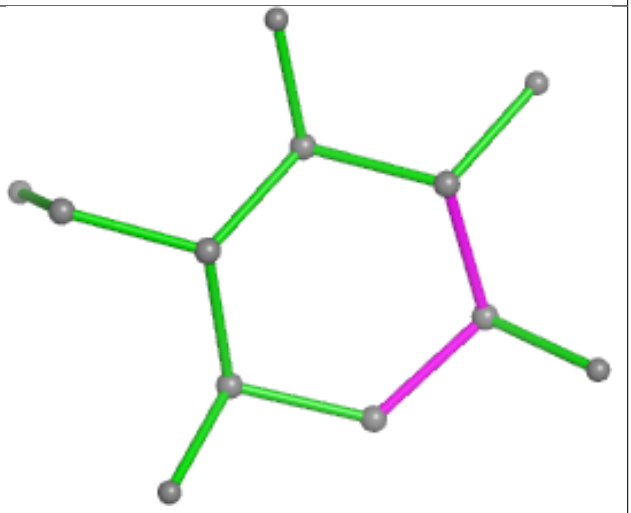


Rings

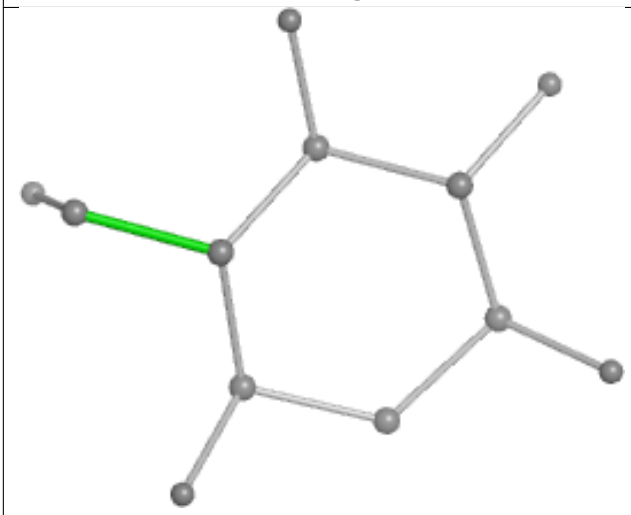
Ligand PBW C 701



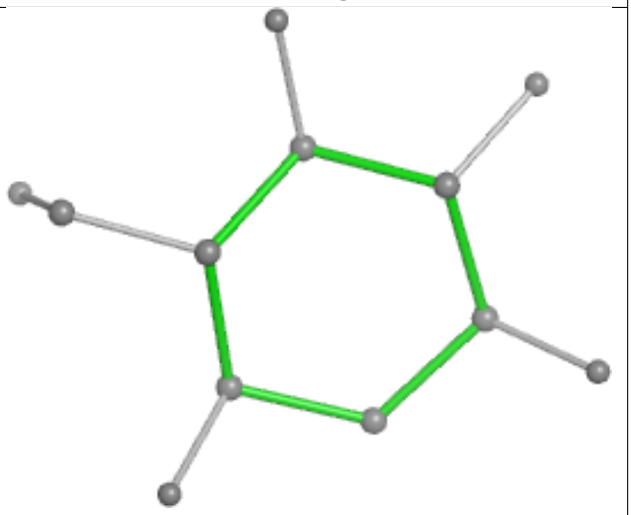
Bond lengths



Bond angles

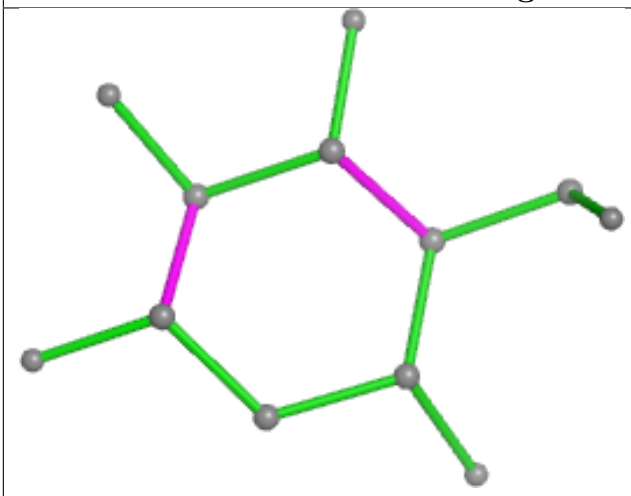


Torsions

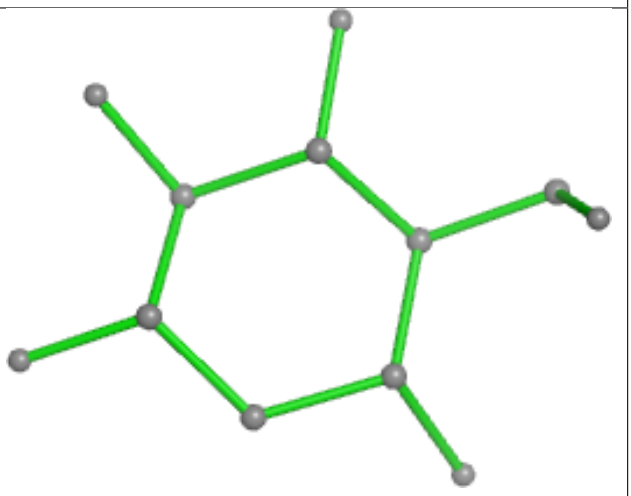


Rings

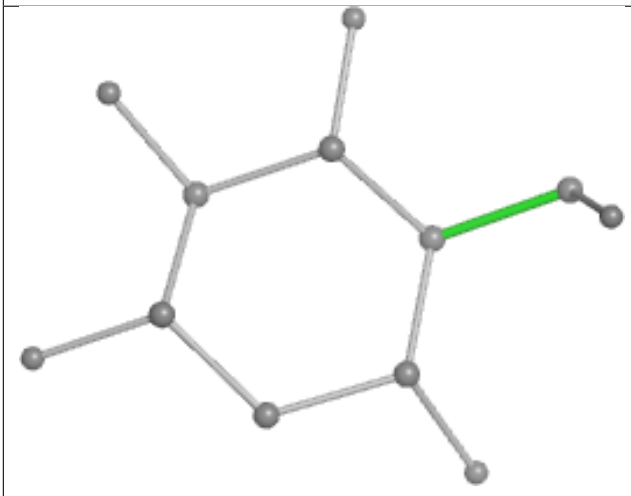
Ligand PBW A 701



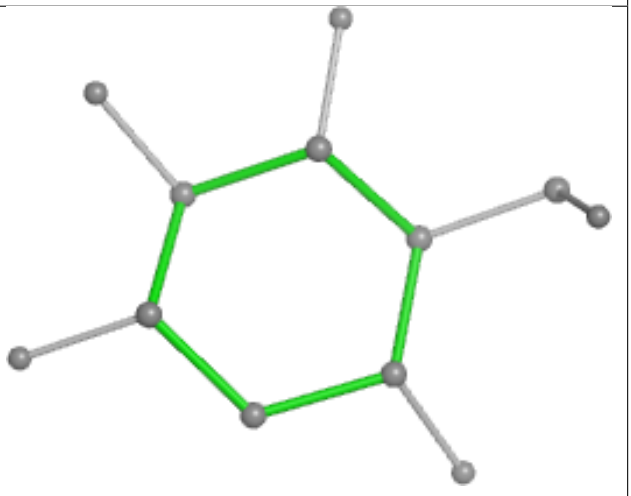
Bond lengths



Bond angles

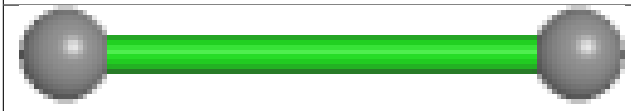


Torsions



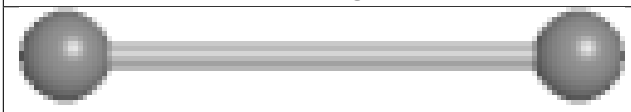
Rings

Ligand OC9 C 702

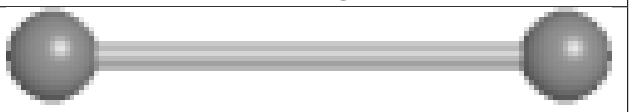


Bond lengths

Bond angles

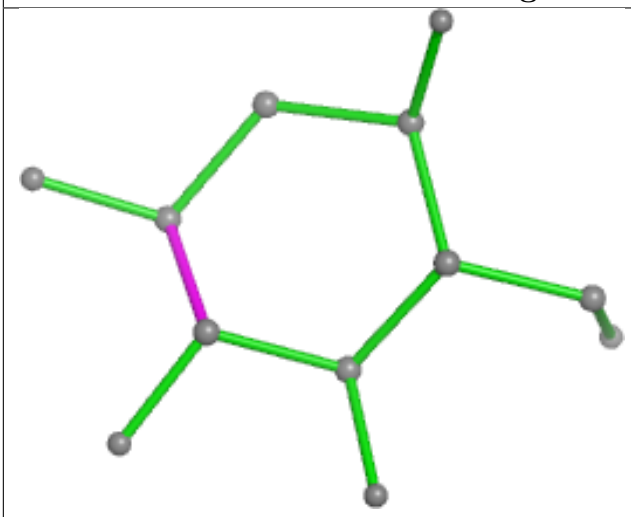


Torsions

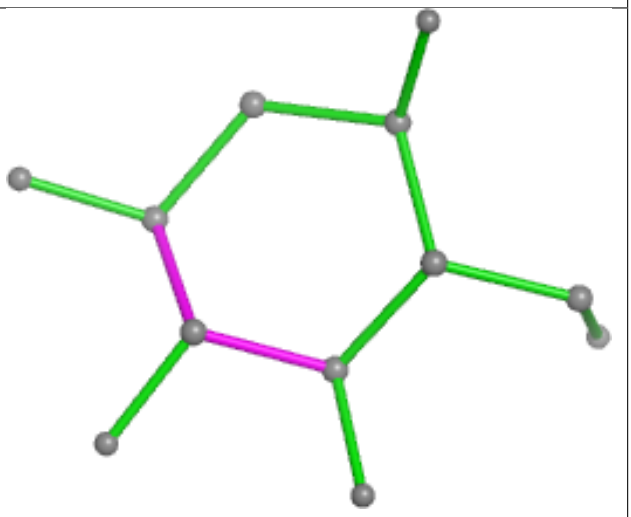


Rings

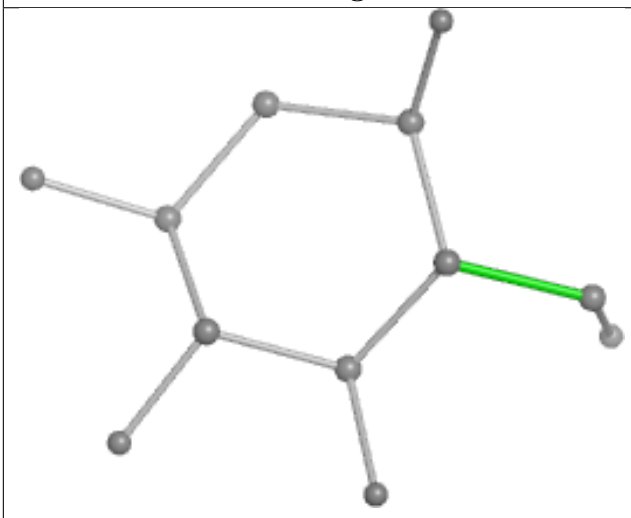
Ligand PBW F 701



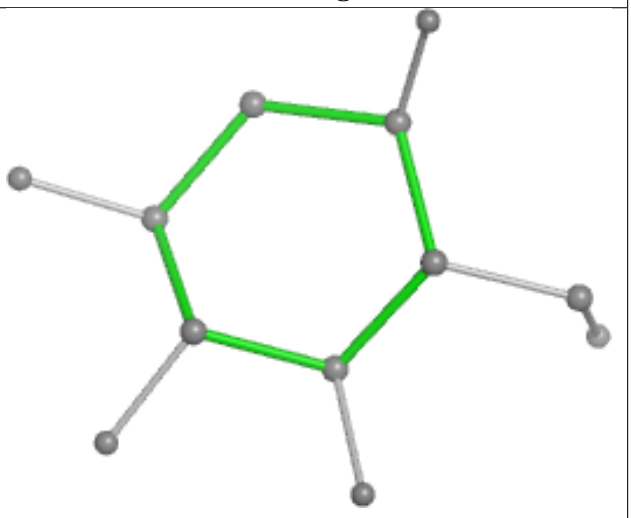
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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	594/610 (97%)	-0.33	0 100 100	26, 43, 69, 98	0
1	B	591/610 (96%)	0.13	22 (3%) 45 45	34, 56, 89, 168	0
1	C	590/610 (96%)	0.53	36 (6%) 28 28	48, 74, 112, 161	0
1	D	594/610 (97%)	0.76	37 (6%) 28 27	50, 83, 120, 151	0
1	E	593/610 (97%)	0.77	35 (5%) 29 28	48, 90, 126, 159	0
1	F	589/610 (96%)	1.07	81 (13%) 8 7	51, 92, 119, 144	0
All	All	3551/3660 (97%)	0.49	211 (5%) 29 28	26, 74, 116, 168	0

The worst 5 of 211 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	333	HIS	5.7
1	C	24	SER	5.3
1	F	93	ASP	5.2
1	C	333	HIS	5.0
1	B	267	VAL	5.0

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

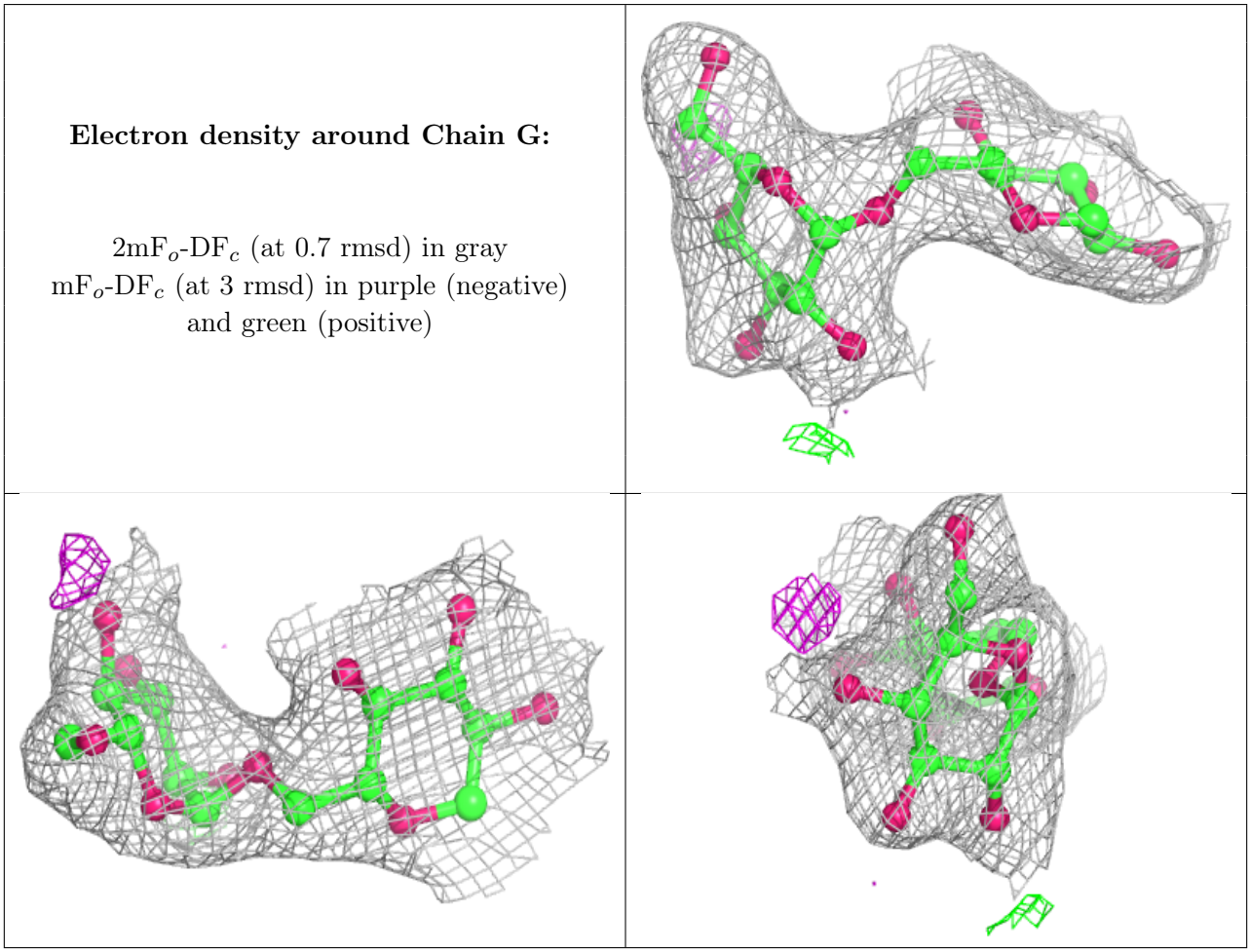
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
2	GLC	J	2	11/12	0.84	0.13	89,94,103,104	0
2	GLC	L	2	11/12	0.85	0.10	84,96,106,110	0
2	GLC	I	2	11/12	0.86	0.11	64,79,99,99	0
2	GLC	K	2	11/12	0.88	0.10	77,89,95,95	0
2	GLC	I	1	11/12	0.91	0.09	69,77,82,84	0
2	GLC	J	1	11/12	0.92	0.10	65,80,85,111	0
2	GLC	G	2	11/12	0.93	0.07	39,45,50,59	0
2	GLC	L	1	11/12	0.94	0.07	73,79,87,92	0
2	GLC	H	1	11/12	0.94	0.07	47,59,72,73	0
2	GLC	K	1	11/12	0.95	0.07	69,81,88,93	0
2	GLC	H	2	11/12	0.95	0.07	62,69,77,93	0
2	GLC	G	1	11/12	0.98	0.05	30,32,37,42	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.

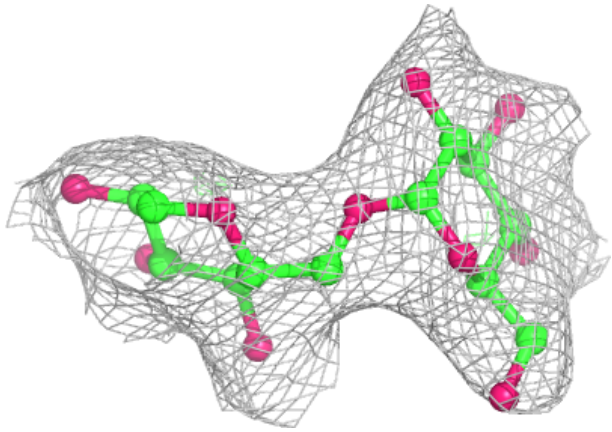
Electron density around Chain G:

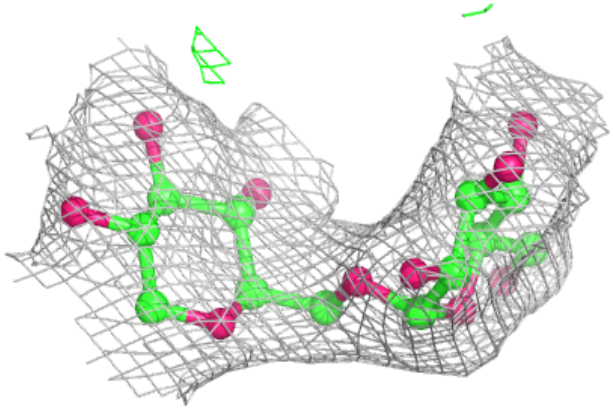
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

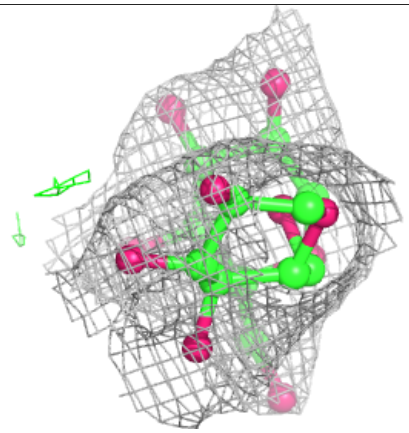


Electron density around Chain H:

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

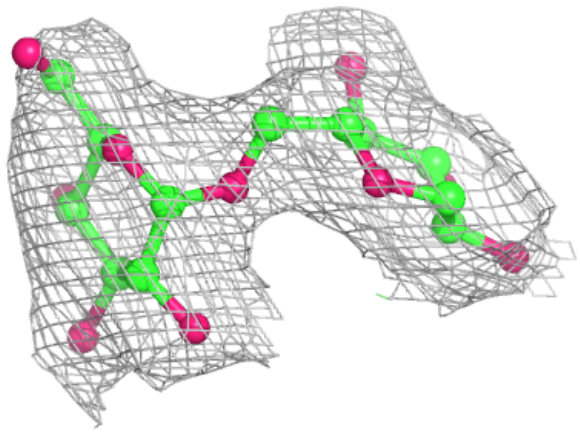


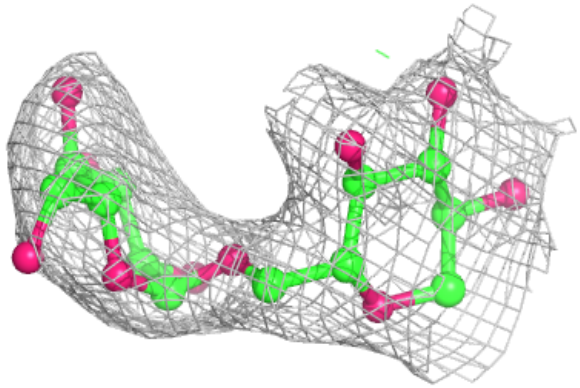


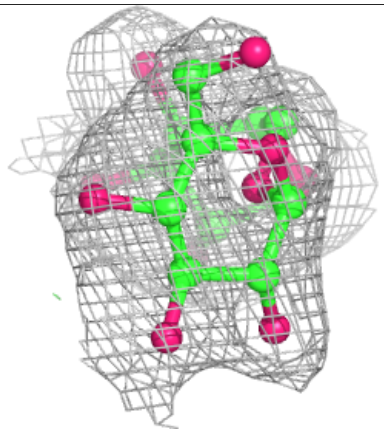


Electron density around Chain I:

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

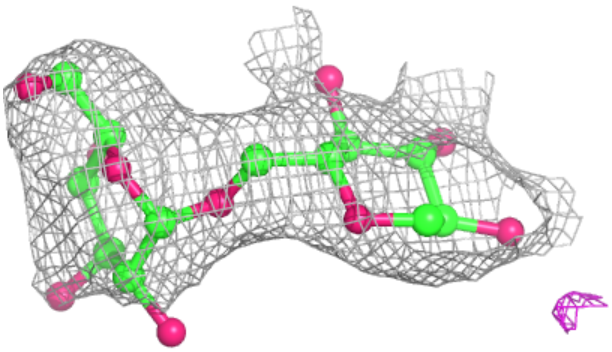


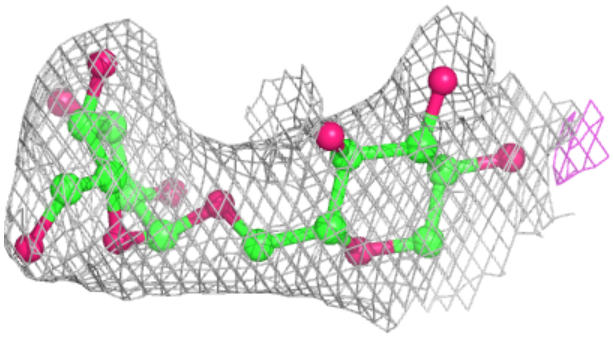


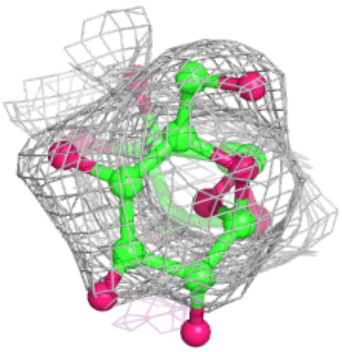


Electron density around Chain J:

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

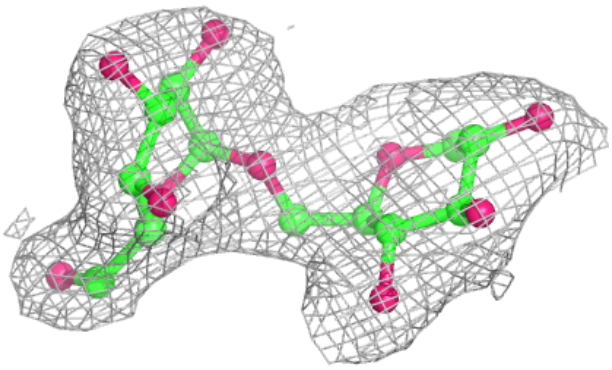


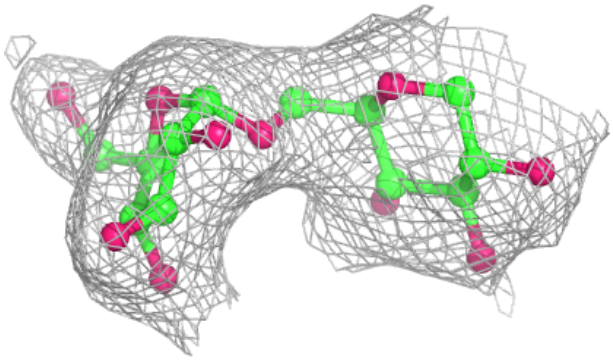


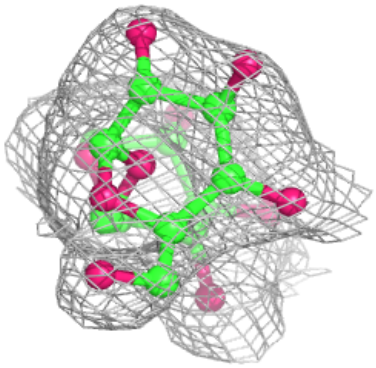


Electron density around Chain K:

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

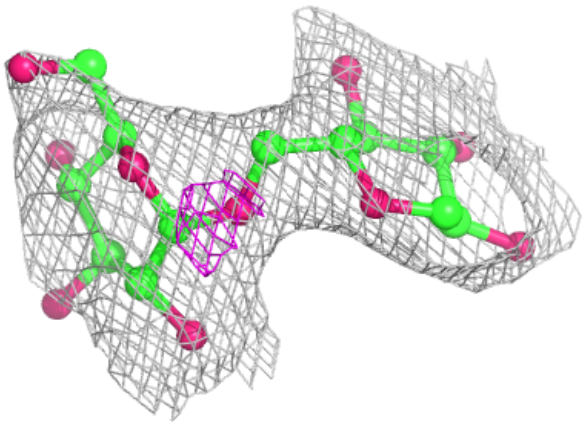


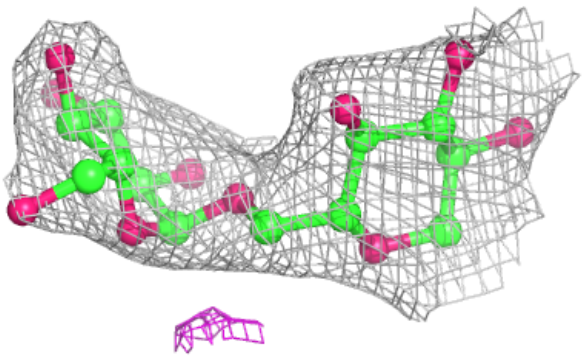


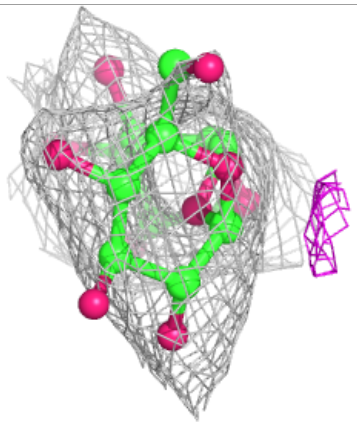


Electron density around Chain L:

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







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, 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.

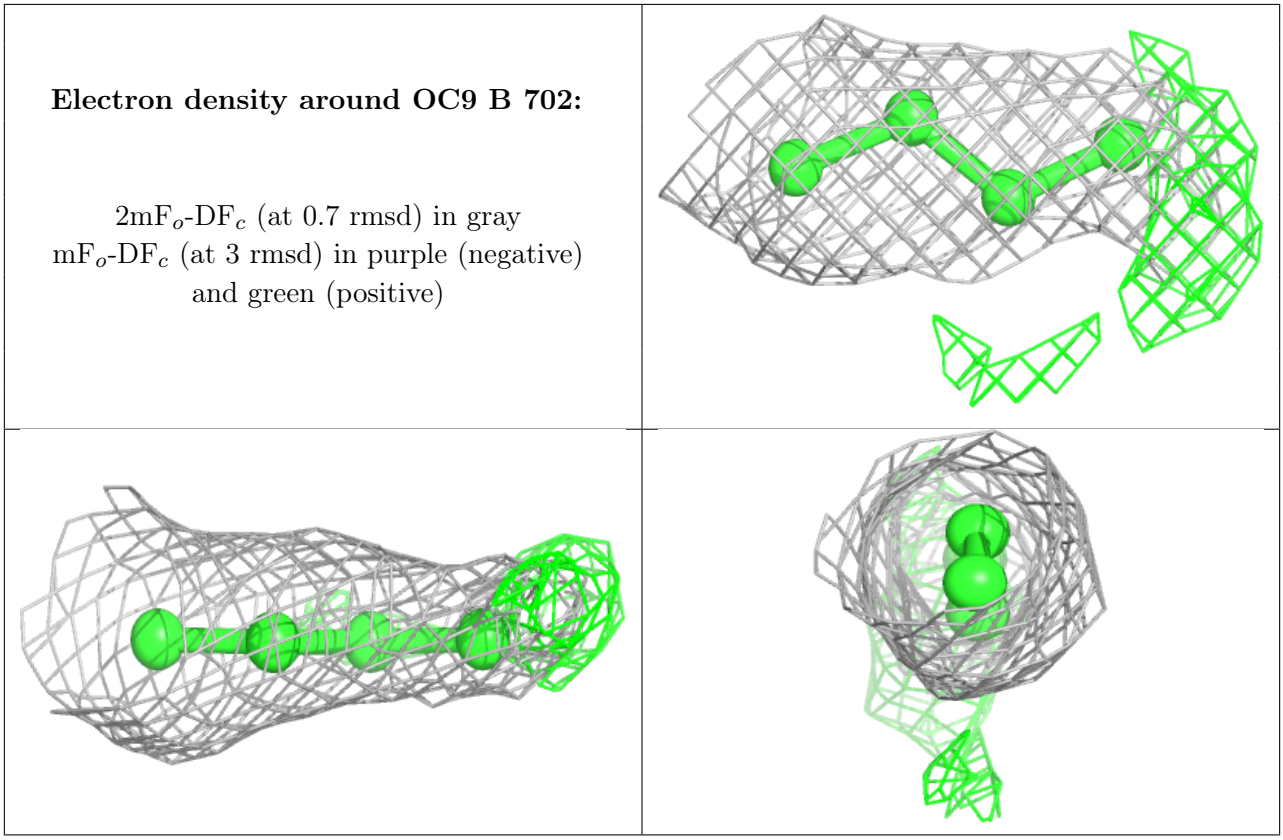
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	OC9	B	702	4/9	0.63	0.36	74,75,83,88	0
5	IMD	E	703	5/5	0.72	0.20	88,96,102,103	0
4	OC9	C	702	2/9	0.74	0.36	84,84,84,88	0
5	IMD	C	704	5/5	0.78	0.20	75,84,91,95	0
5	IMD	D	703	5/5	0.79	0.23	83,93,96,99	0
4	OC9	D	702	3/9	0.80	0.23	78,78,86,99	0
4	OC9	E	702	2/9	0.80	0.38	96,96,96,105	0
5	IMD	C	703	5/5	0.80	0.14	96,96,97,100	0
6	CA	C	707	1/1	0.80	0.12	135,135,135,135	0
5	IMD	A	704	5/5	0.85	0.17	57,67,76,79	0
5	IMD	D	705	5/5	0.85	0.13	65,68,71,71	0
6	CA	B	705	1/1	0.89	0.11	93,93,93,93	0
5	IMD	C	705	5/5	0.89	0.13	56,62,72,74	0
5	IMD	D	704	5/5	0.90	0.13	88,92,92,93	0
5	IMD	F	702	5/5	0.90	0.12	78,80,82,85	0
3	PBW	C	701	12/13	0.91	0.09	68,79,89,90	0
5	IMD	B	703	5/5	0.91	0.15	58,58,67,75	0
3	PBW	F	701	12/13	0.91	0.08	77,90,97,98	0
6	CA	F	704	1/1	0.92	0.09	128,128,128,128	0
4	OC9	A	702	4/9	0.93	0.12	42,45,51,57	0
3	PBW	E	701	12/13	0.93	0.08	64,74,84,85	0
5	IMD	A	705	5/5	0.93	0.15	58,62,65,73	0
3	PBW	B	701	12/13	0.94	0.07	51,61,72,84	0
6	CA	E	704	1/1	0.94	0.06	77,77,77,77	0
5	IMD	A	703	5/5	0.94	0.10	35,43,44,49	0
6	CA	C	706	1/1	0.95	0.06	102,102,102,102	0
3	PBW	D	701	12/13	0.95	0.08	63,68,72,74	0
6	CA	F	703	1/1	0.96	0.06	99,99,99,99	0
3	PBW	A	701	12/13	0.97	0.06	29,33,48,48	0
6	CA	E	705	1/1	0.97	0.06	131,131,131,131	0
6	CA	A	706	1/1	0.98	0.10	58,58,58,58	0
6	CA	D	706	1/1	0.98	0.07	105,105,105,105	0
6	CA	B	704	1/1	0.99	0.10	73,73,73,73	0
6	CA	D	707	1/1	0.99	0.03	75,75,75,75	0
6	CA	A	707	1/1	0.99	0.04	43,43,43,43	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.

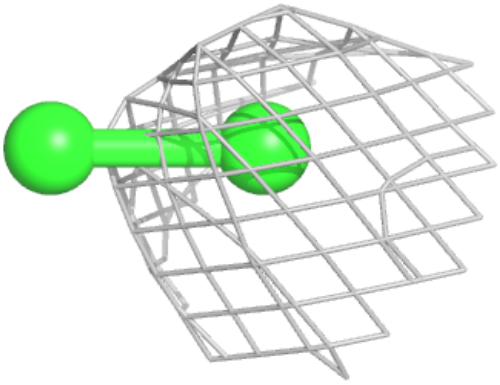
Electron density around OC9 B 702:

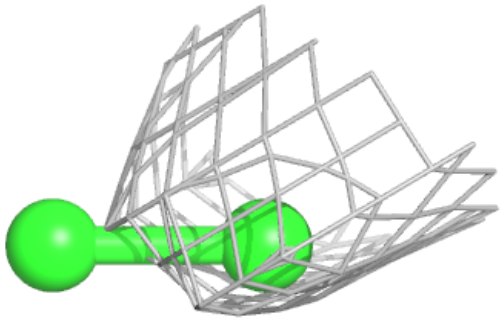
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

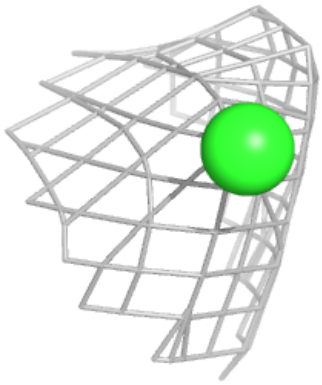


Electron density around OC9 C 702:

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

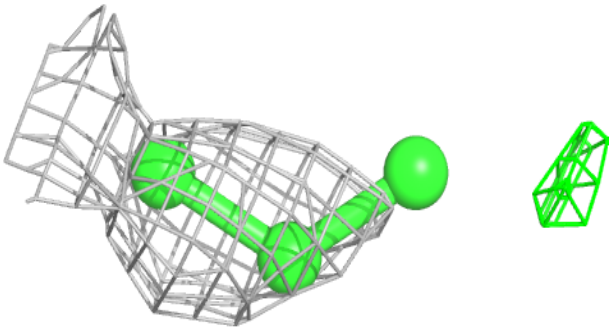


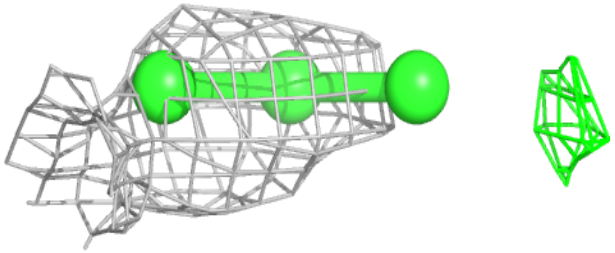


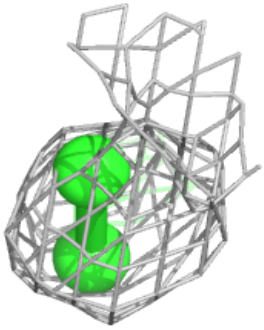


Electron density around OC9 D 702:

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

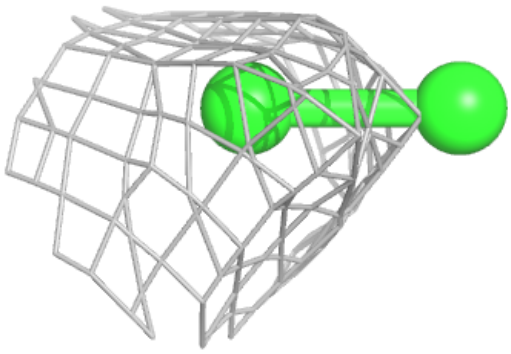


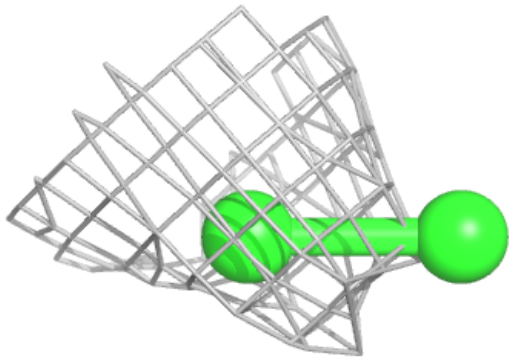


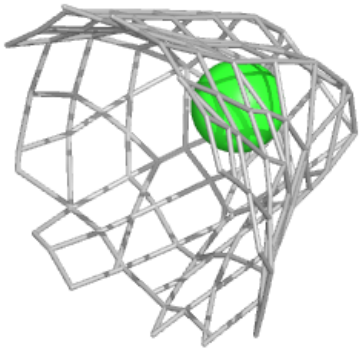


Electron density around OC9 E 702:

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

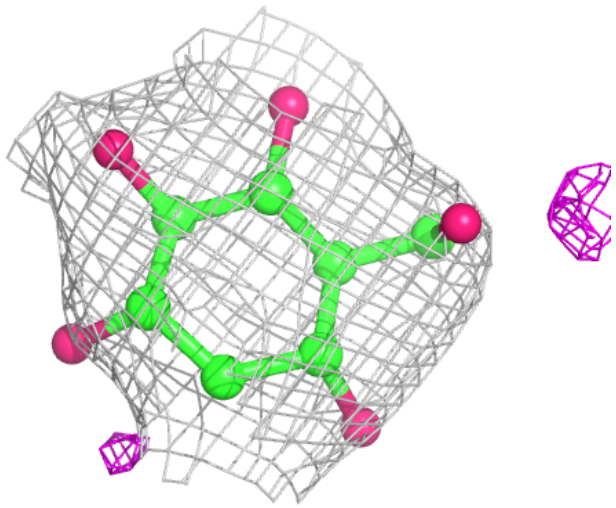


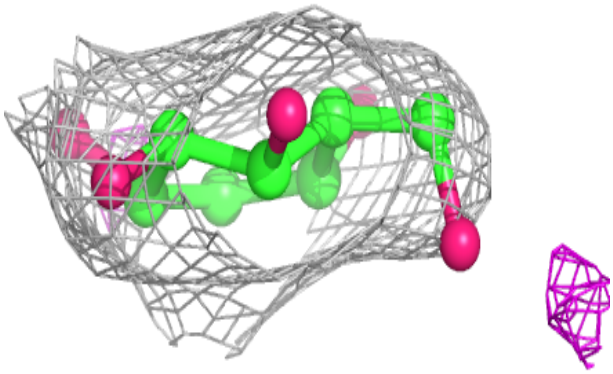


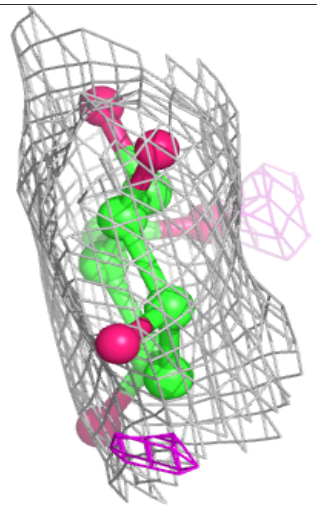


Electron density around PBW C 701:

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

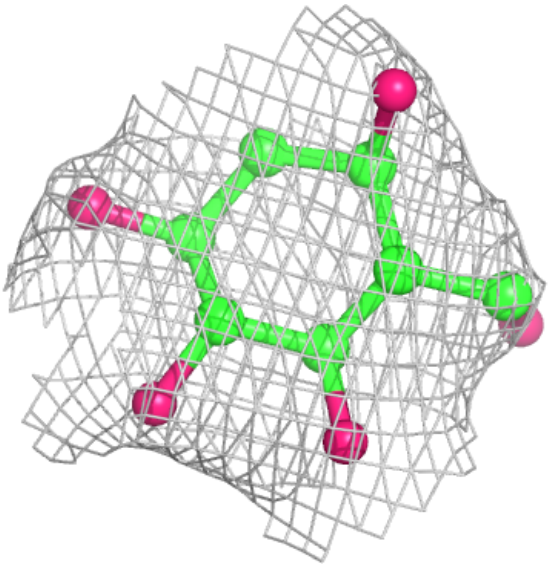


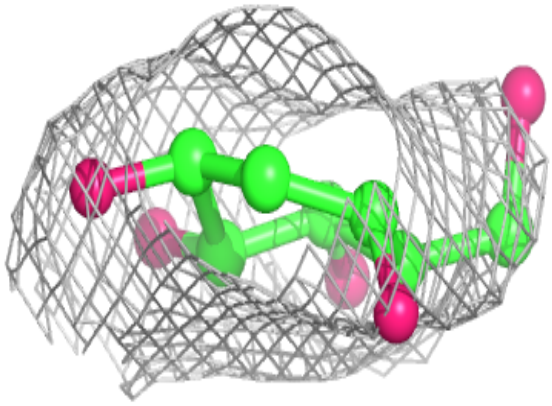


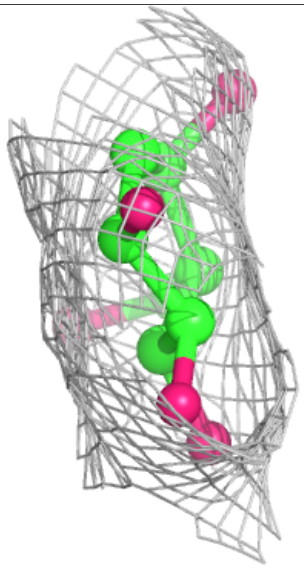


Electron density around PBW F 701:

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

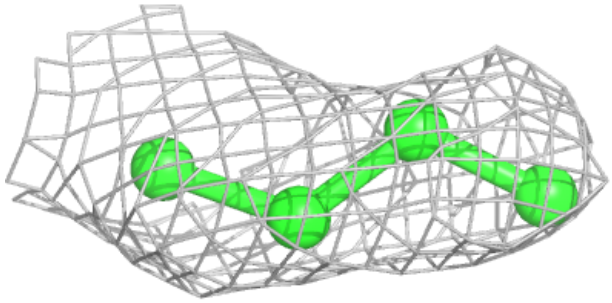


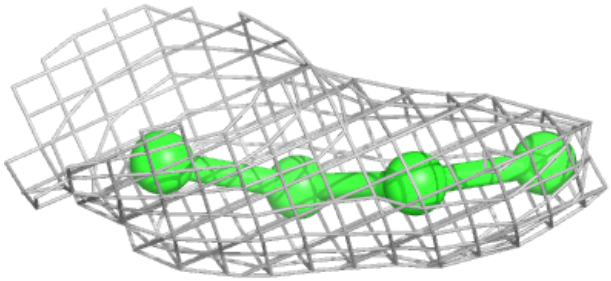


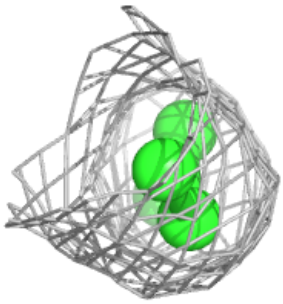


Electron density around OC9 A 702:

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

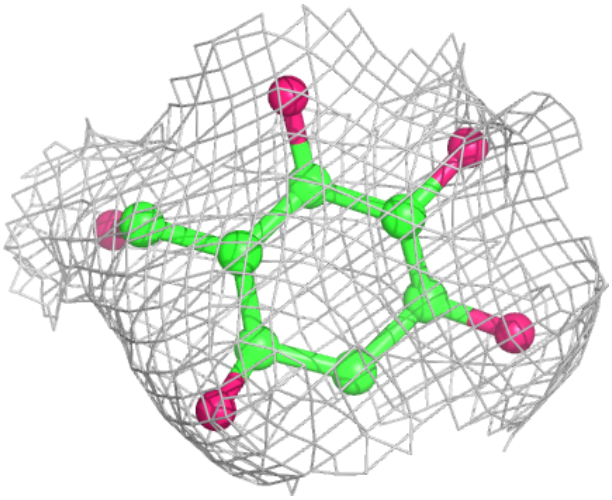


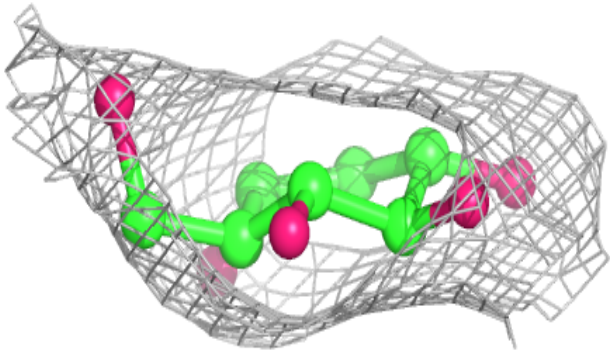


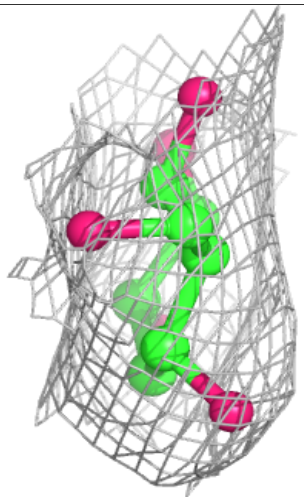


Electron density around PBW E 701:

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

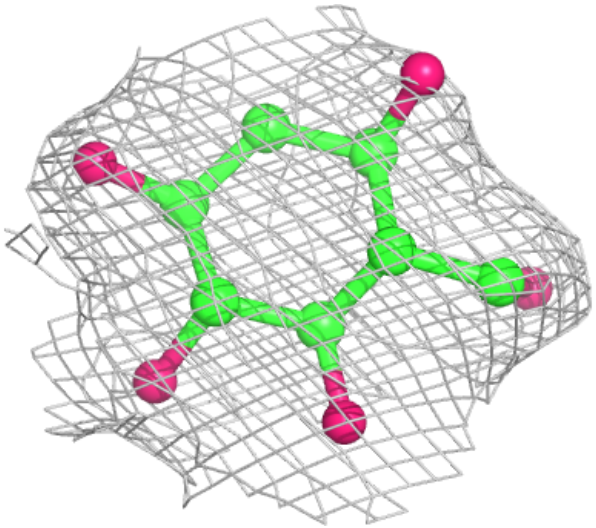


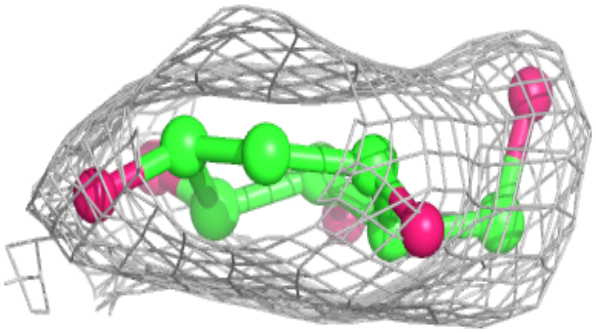


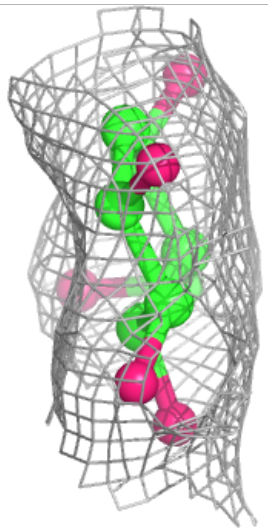


Electron density around PBW B 701:

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

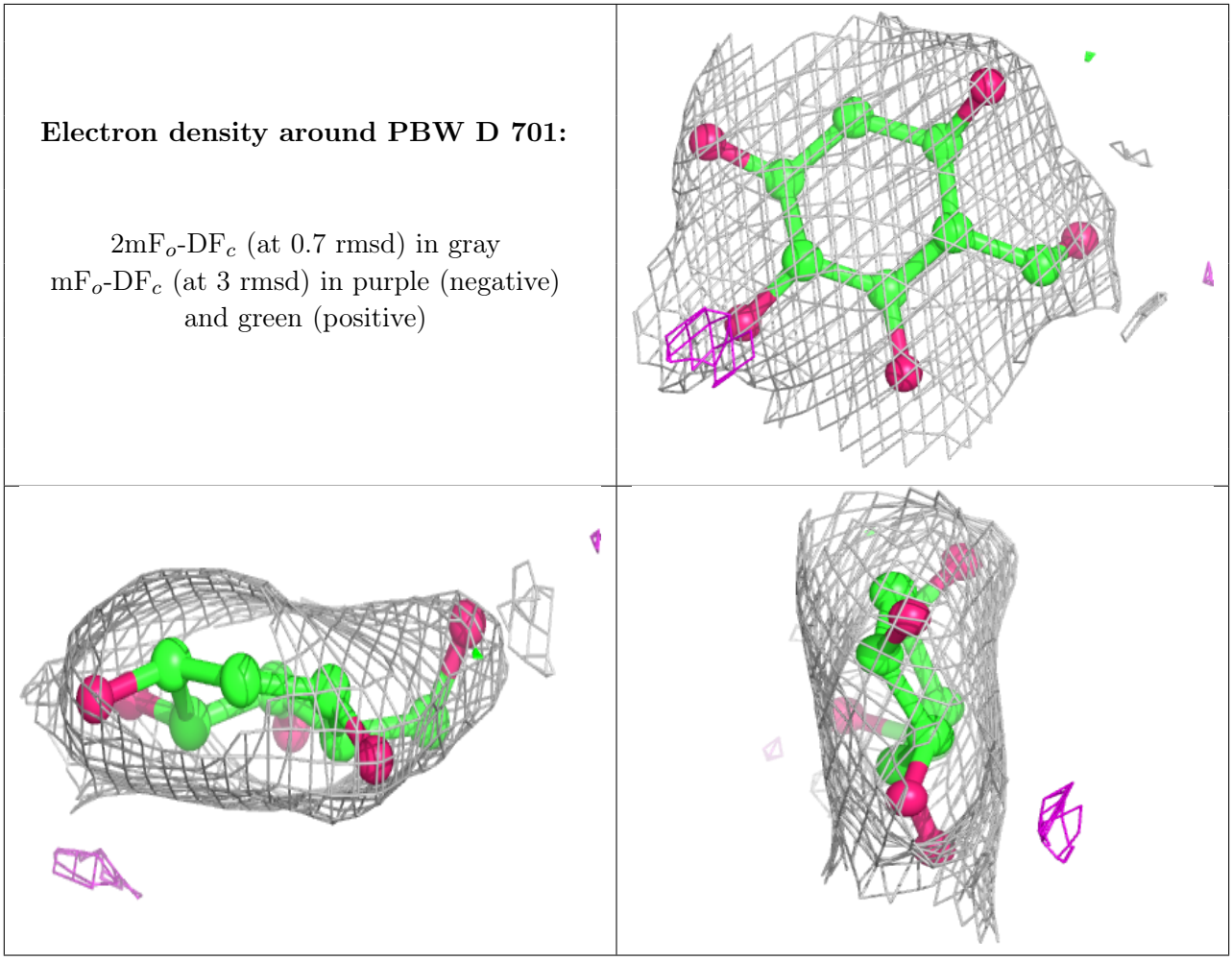






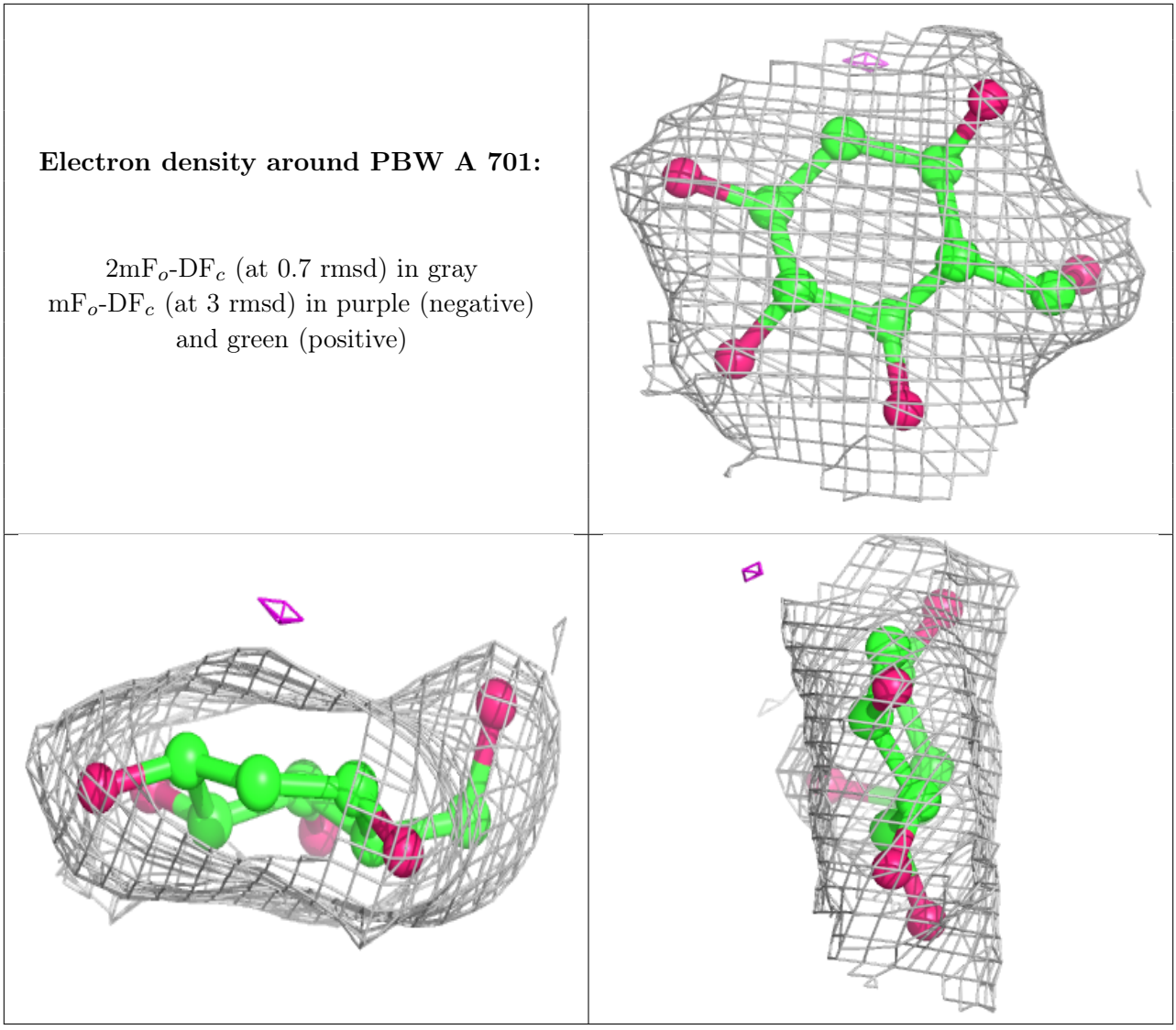
Electron density around PBW D 701:

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



Electron density around PBW A 701:

$2mF_o$-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.