Full wwPDB X-ray Structure Validation Report (i)

Apr 18, 2024 – 04:52 pm BST

PDB ID	:	5IEG
Title	:	Murine endoplasmic reticulum alpha-glucosidase II with N-9'-methoxynonyl-1-deoxynojirimycin.
Authors	:	Caputo, A.T.; Roversi, P.; Alonzi, D.S.; Kiappes, J.L.; Zitzmann, N.
Deposited on	:	2016-02-25
Resolution	:	1.82 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

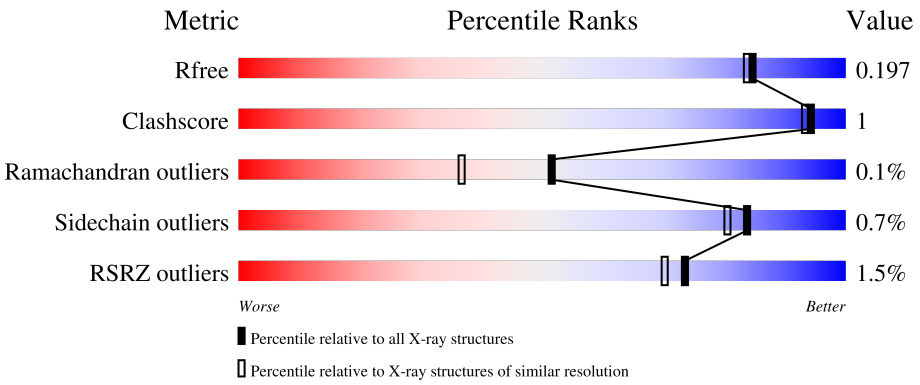
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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 1.82 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	7484 (1.84-1.80)
Clashscore	141614	8401 (1.84-1.80)
Ramachandran outliers	138981	8290 (1.84-1.80)
Sidechain outliers	138945	8290 (1.84-1.80)
RSRZ outliers	127900	7371 (1.84-1.80)

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	914	% 90% • 6%
2	B	88	5% 100%
3	C	2	100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	P6G	A	1015	-	-	-	X

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 8464 atoms, of which 128 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 Neutral alpha-glucosidase AB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	856	6998	4486	1208	1273	31	0	17	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	PHE	deletion	UNP Q8BHN3
A	?	-	SER	deletion	UNP Q8BHN3
A	?	-	ASP	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	VAL	deletion	UNP Q8BHN3
A	?	-	SER	deletion	UNP Q8BHN3
A	?	-	LEU	deletion	UNP Q8BHN3
A	?	-	ALA	deletion	UNP Q8BHN3
A	?	-	LEU	deletion	UNP Q8BHN3
A	?	-	GLY	deletion	UNP Q8BHN3
A	?	-	SER	deletion	UNP Q8BHN3
A	?	-	VAL	deletion	UNP Q8BHN3
A	?	-	TRP	deletion	UNP Q8BHN3
A	?	-	ASP	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	ILE	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	ASN	deletion	UNP Q8BHN3
A	?	-	LEU	deletion	UNP Q8BHN3
A	?	-	PHE	deletion	UNP Q8BHN3
A	?	-	SER	deletion	UNP Q8BHN3
A	?	-	ARG	deletion	UNP Q8BHN3
A	967	ALA	-	expression tag	UNP Q8BHN3
A	968	SER	-	expression tag	UNP Q8BHN3

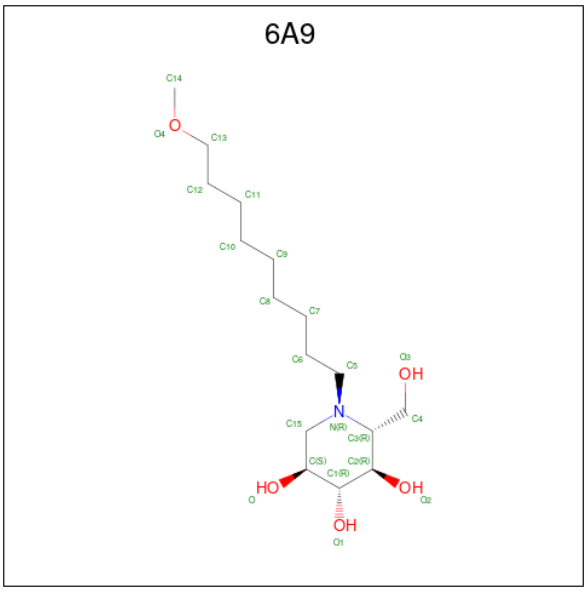
- Molecule 2 is a protein called Glucosidase 2 subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	88	663	399	106	148	10	0	0	0

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

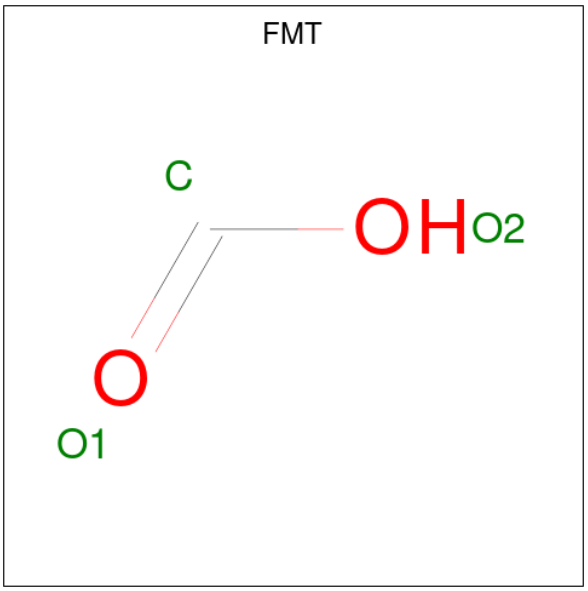
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	2	28	16	2	10	0	0	0

- Molecule 4 is N-9'-methoxynonyl-1-deoxynojirimycin (three-letter code: 6A9) (formula: $C_{16}H_{33}NO_5$).



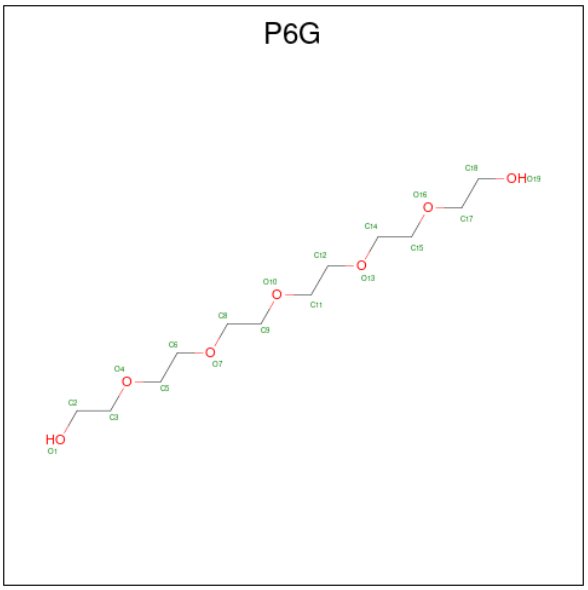
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	44	32	2	10	0	1

- Molecule 5 is FORMIC ACID (three-letter code: FMT) (formula: CH_2O_2).



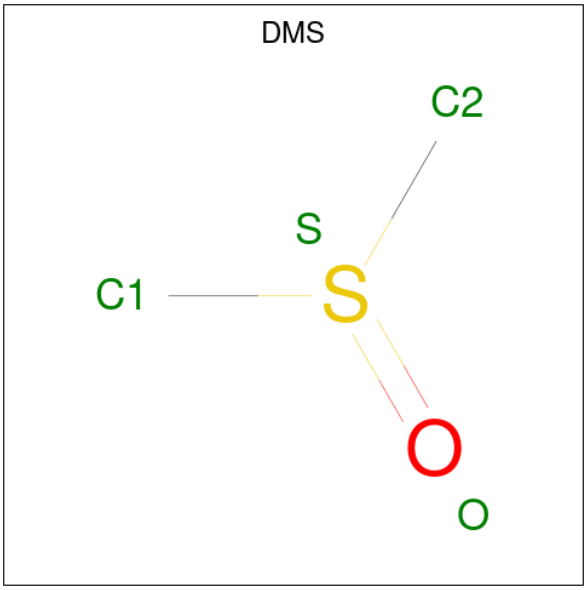
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
5	A	1	5	1	2	2	0	0
			Total	C	H	O		
5	A	1	5	1	2	2	0	0
			Total	C	H	O		
5	A	1	5	1	2	2	0	0
			Total	C	H	O		
5	A	1	5	1	2	2	0	0
			Total	C	H	O		
5	A	1	5	1	2	2	0	0

- Molecule 6 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: $C_{12}H_{26}O_7$).



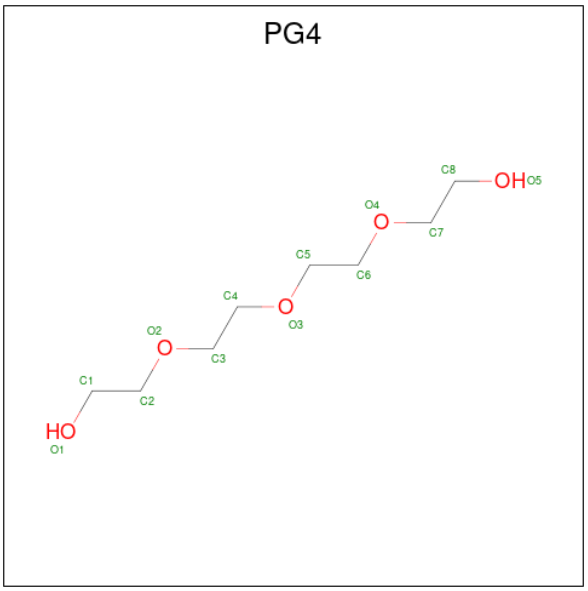
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
6	A	1	30	9	16	5	0	0
			Total	C	H	O		
6	A	1	45	12	26	7	0	0
			Total	C	H	O		
6	A	1	45	12	26	7	0	0
			Total	C	H	O		
6	A	1	30	9	16	5	0	0

- Molecule 7 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	O	S		
7	A	1	10	2	6	1	1	0	0

- Molecule 8 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C H O 17 5 9 3	0	0
8	A	1	Total C H O 20 6 11 3	0	0
8	B	1	Total C H O 17 6 8 3	0	0

- Molecule 9 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	B	2	Total Ca 2 2	0	0

- Molecule 10 is water.

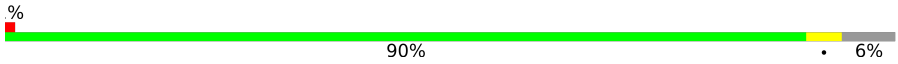
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	445	Total O 445 445	0	0
10	B	45	Total O 45 45	0	0

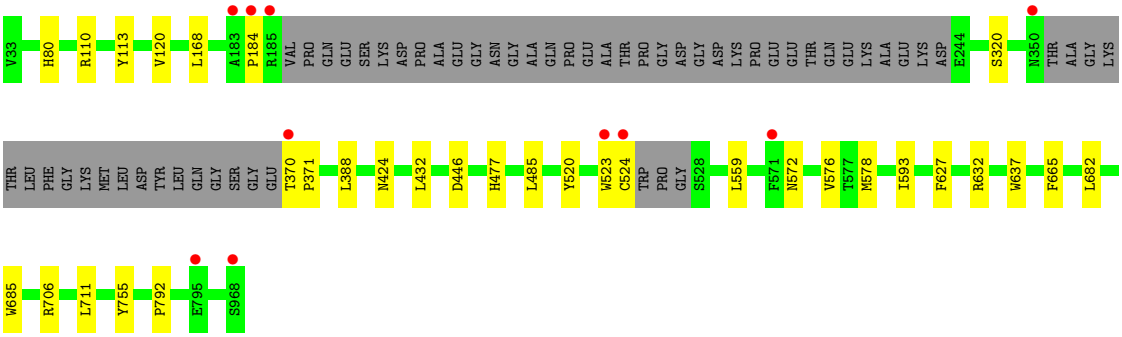
3 Residue-property plots (i)

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: Neutral alpha-glucosidase AB

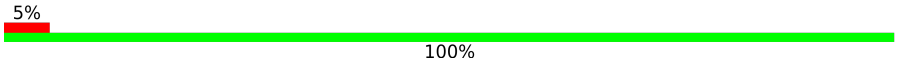
Chain A:

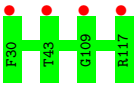




• Molecule 2: Glucosidase 2 subunit beta

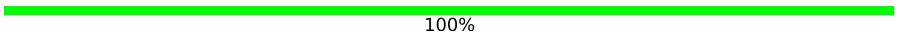
Chain B:





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

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	104.32Å 173.31Å 63.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	89.38 – 1.82 89.38 – 1.82	Depositor EDS
% Data completeness (in resolution range)	99.7 (89.38-1.82) 100.0 (89.38-1.82)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	2.11 (at 1.82Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.175 , 0.195 0.178 , 0.197	Depositor DCC
R_{free} test set	5057 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	21.1	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 44.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8464	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.41% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, PG4, P6G, 6A9, CA, FMT, DMS

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	0.50	0/7247	0.64	0/9858
2	B	0.48	0/677	0.64	0/920
All	All	0.50	0/7924	0.64	0/10778

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6998	0	6778	20	0
2	B	663	0	568	0	0
3	C	28	0	25	0	0
4	A	44	0	0	0	0
5	A	15	10	7	0	0
6	A	66	84	84	5	0
7	A	4	6	6	0	0
8	A	17	20	20	0	0
8	B	9	8	8	0	0
9	B	2	0	0	0	0
10	A	445	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	B	45	0	0	0	0
All	All	8336	128	7496	20	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (20) 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:A:632:ARG:HH11	6:A:1015:P6G:H82	1.38	0.87
1:A:168[B]:LEU:HD11	1:A:388:LEU:HD13	1.70	0.74
1:A:632:ARG:NH1	6:A:1015:P6G:H82	2.11	0.65
1:A:432:LEU:HD22	1:A:477[A]:HIS:CD2	2.34	0.61
1:A:110:ARG:NH2	1:A:184:PRO:HG3	2.22	0.54
1:A:682:LEU:HD23	1:A:711:LEU:HD11	1.93	0.51
1:A:520:TYR:CE2	1:A:578[A]:MET:SD	3.04	0.51
1:A:432:LEU:HD22	1:A:477[A]:HIS:HD2	1.77	0.48
1:A:523:TRP:HD1	1:A:572:ASN:ND2	2.12	0.47
1:A:113:TYR:CZ	1:A:593:ILE:HG22	2.50	0.47
1:A:320:SER:O	1:A:627:PHE:HA	2.15	0.47
1:A:80:HIS:HE2	6:A:1011:P6G:H111	1.80	0.46
1:A:485:LEU:HD23	1:A:559[B]:LEU:CD1	2.46	0.45
1:A:520:TYR:CZ	1:A:578[A]:MET:SD	3.10	0.45
1:A:120:VAL:HG12	6:A:1015:P6G:H141	1.97	0.45
1:A:370:THR:N	1:A:371:PRO:HD2	2.34	0.42
1:A:485:LEU:CD2	1:A:559[B]:LEU:CD1	2.98	0.42
1:A:632:ARG:HE	6:A:1015:P6G:C6	2.34	0.41
1:A:523:TRP:O	1:A:524:CYS:HB2	2.21	0.41
1:A:755:TYR:CE2	1:A:792:PRO:HG2	2.57	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	865/914 (95%)	839 (97%)	25 (3%)	1 (0%)	51	37
2	B	86/88 (98%)	84 (98%)	2 (2%)	0	100	100
All	All	951/1002 (95%)	923 (97%)	27 (3%)	1 (0%)	51	37

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	576	VAL

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	764/791 (97%)	758 (99%)	6 (1%)	81	77
2	B	77/77 (100%)	77 (100%)	0	100	100
All	All	841/868 (97%)	835 (99%)	6 (1%)	84	80

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	424	ASN
1	A	446	ASP
1	A	637	TRP
1	A	665	PHE
1	A	685	TRP
1	A	706	ARG

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)

2 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	$\#\|Z\| > 2$
3	NAG	C	1	3,1	14,14,15	0.36	0	17,19,21	0.68	0
3	NAG	C	2	3	14,14,15	0.33	0	17,19,21	0.51	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	C	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	C	2	3	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

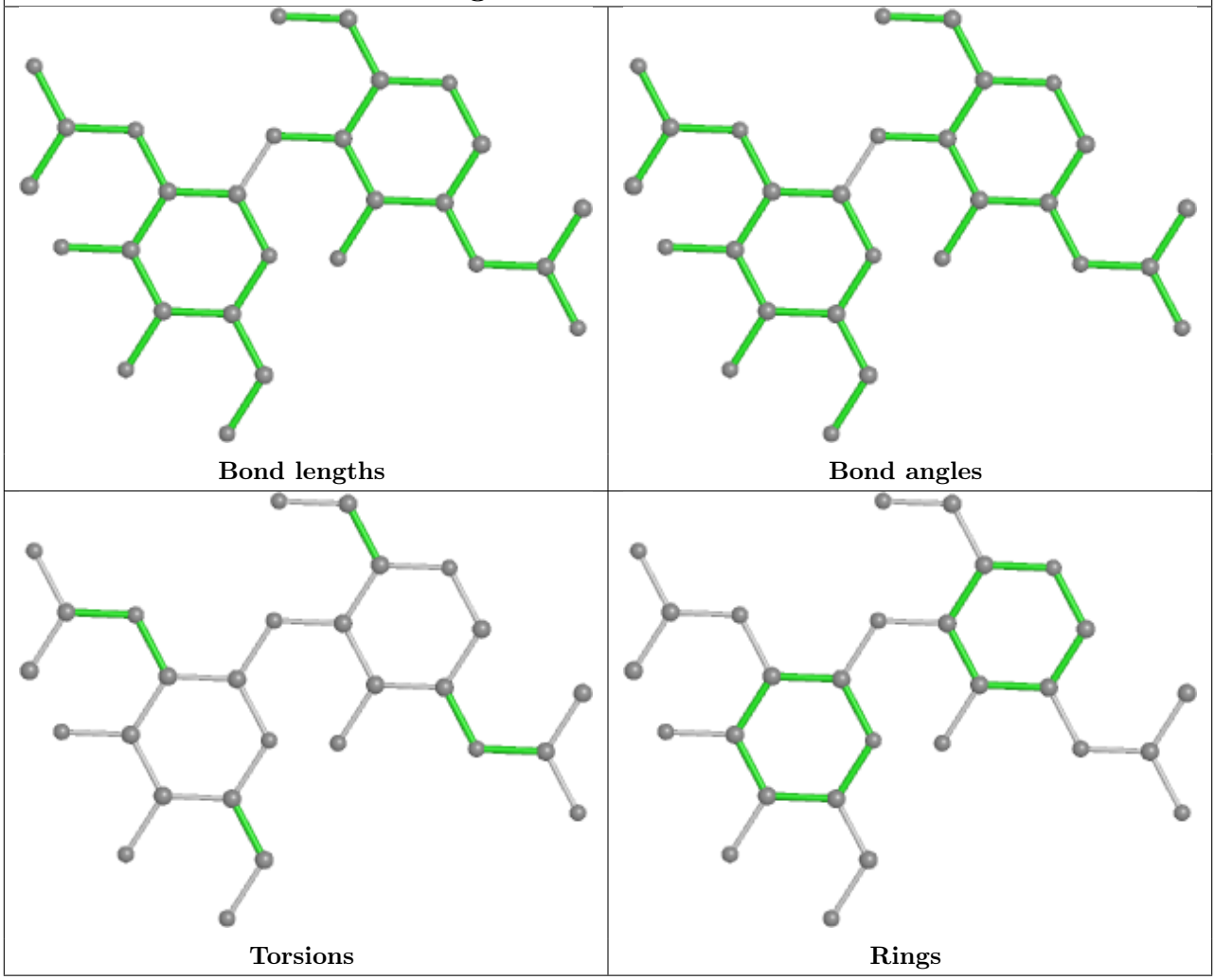
There are no ring outliers.

No monomer is involved in short contacts.

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

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

Oligosaccharide Chain C



5.6 Ligand geometry (i)

Of 17 ligands modelled in this entry, 2 are monoatomic - leaving 15 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	FMT	A	1007	-	2,2,2	1.11	0	1,1,1	1.21	0
8	PG4	A	1014	-	8,8,12	0.23	0	7,7,11	0.30	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
8	PG4	A	1013	-	7,7,12	0.19	0	6,6,11	0.18	0
5	FMT	A	1006	-	2,2,2	1.02	0	1,1,1	1.13	0
6	P6G	A	1010	-	18,18,18	0.26	0	17,17,17	0.37	0
5	FMT	A	1005	-	2,2,2	1.29	0	1,1,1	1.10	0
5	FMT	A	1008	-	2,2,2	1.18	0	1,1,1	1.14	0
6	P6G	A	1009	-	13,13,18	0.15	0	12,12,17	0.27	0
4	6A9	A	1003[A]	-	22,22,22	0.40	0	25,27,27	0.81	1 (4%)
4	6A9	A	1003[B]	1	22,22,22	0.28	0	25,27,27	0.84	1 (4%)
6	P6G	A	1015	-	13,13,18	0.29	0	12,12,17	0.73	0
8	PG4	B	203	-	8,8,12	0.31	0	7,7,11	0.34	0
7	DMS	A	1012	-	3,3,3	0.35	0	3,3,3	0.42	0
5	FMT	A	1004	-	2,2,2	0.67	0	1,1,1	1.29	0
6	P6G	A	1011	-	18,18,18	0.20	0	17,17,17	0.27	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	PG4	A	1014	-	-	2/6/6/10	-
8	PG4	A	1013	-	-	0/5/5/10	-
6	P6G	A	1010	-	-	8/16/16/16	-
6	P6G	A	1009	-	-	2/11/11/16	-
4	6A9	A	1003[A]	-	-	1/13/33/33	0/1/1/1
4	6A9	A	1003[B]	1	-	2/13/33/33	0/1/1/1
6	P6G	A	1015	-	-	2/11/11/16	-
8	PG4	B	203	-	-	5/6/6/10	-
6	P6G	A	1011	-	-	7/16/16/16	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1003[B]	6A9	C4-C3-C2	-2.74	108.70	112.90
4	A	1003[A]	6A9	C4-C3-C2	-2.36	109.28	112.90

There are no chirality outliers.

All (29) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	1015	P6G	C11-C12-O13-C14
6	A	1010	P6G	O4-C5-C6-O7
6	A	1010	P6G	O7-C8-C9-O10
8	B	203	PG4	O3-C5-C6-O4
6	A	1011	P6G	O1-C2-C3-O4
8	B	203	PG4	O2-C3-C4-O3
4	A	1003[A]	6A9	C10-C11-C12-C13
6	A	1015	P6G	O4-C5-C6-O7
6	A	1011	P6G	C8-C9-O10-C11
6	A	1010	P6G	C14-C15-O16-C17
6	A	1011	P6G	O13-C14-C15-O16
6	A	1011	P6G	C12-C11-O10-C9
6	A	1010	P6G	C6-C5-O4-C3
6	A	1010	P6G	O16-C17-C18-O19
6	A	1011	P6G	O10-C11-C12-O13
8	B	203	PG4	C3-C4-O3-C5
6	A	1009	P6G	C15-C14-O13-C12
8	B	203	PG4	C4-C3-O2-C2
6	A	1010	P6G	C2-C3-O4-C5
4	A	1003[B]	6A9	C6-C7-C8-C9
6	A	1009	P6G	C5-C6-O7-C8
4	A	1003[B]	6A9	C7-C8-C9-C10
8	A	1014	PG4	C4-C3-O2-C2
8	B	203	PG4	C6-C5-O3-C4
8	A	1014	PG4	O1-C1-C2-O2
6	A	1010	P6G	C9-C8-O7-C6
6	A	1010	P6G	O10-C11-C12-O13
6	A	1011	P6G	C6-C5-O4-C3
6	A	1011	P6G	O16-C17-C18-O19

There are no ring outliers.

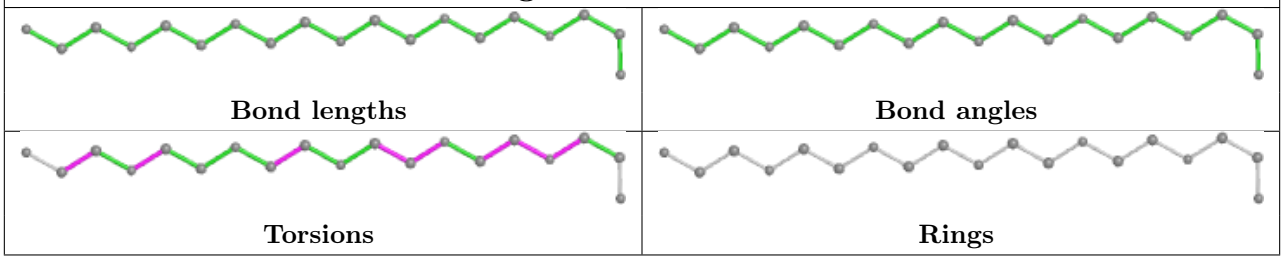
2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1015	P6G	4	0
6	A	1011	P6G	1	0

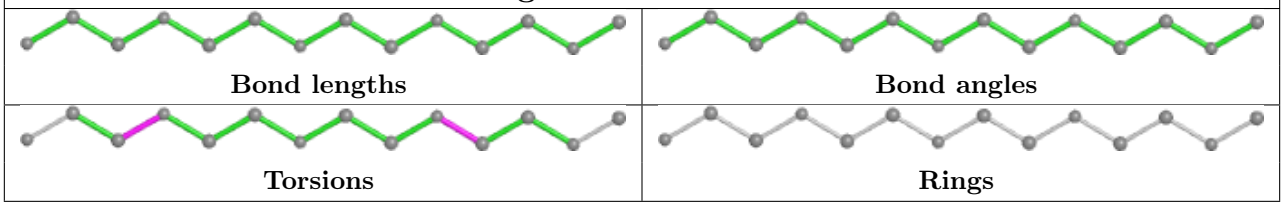
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

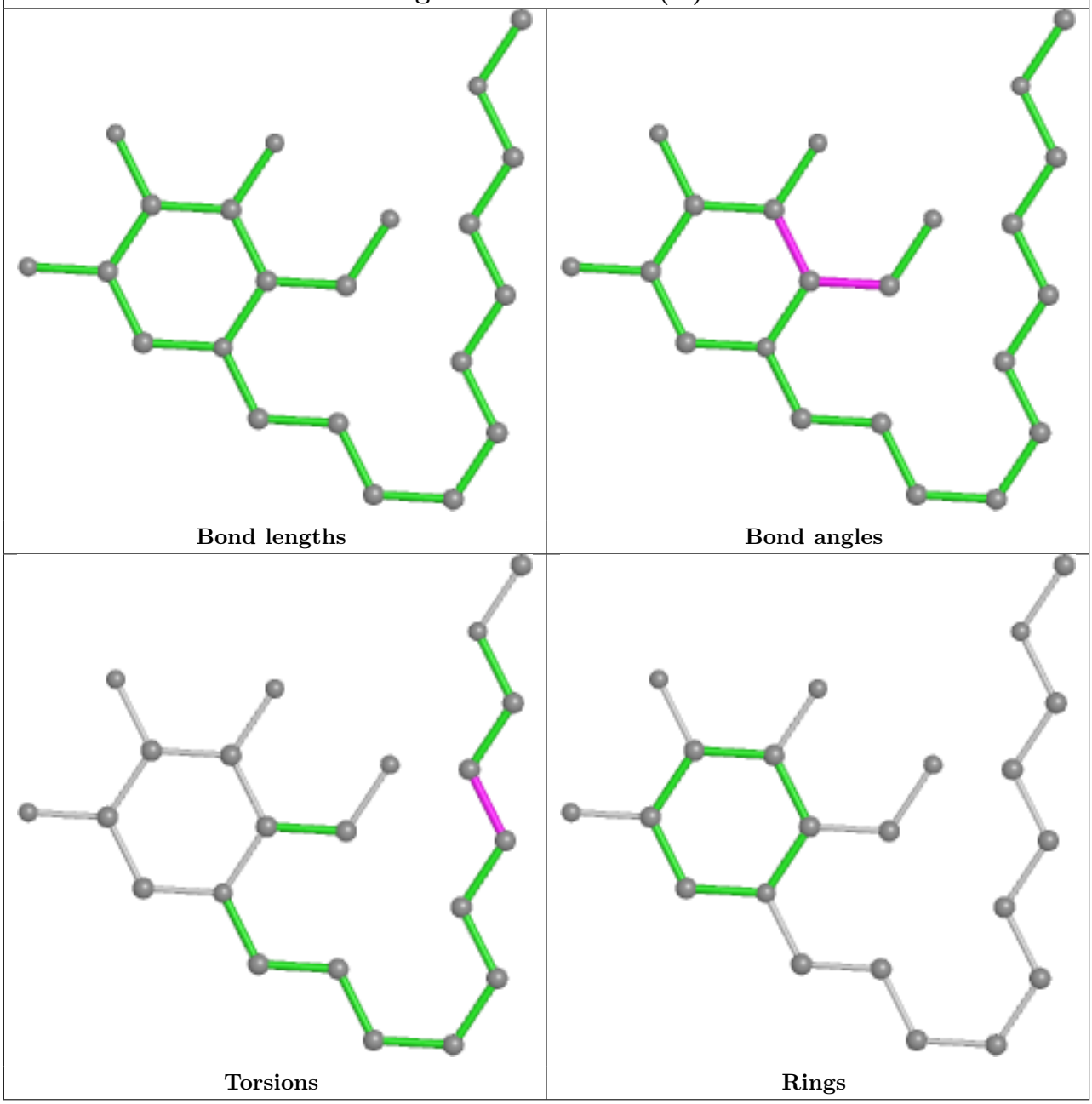
Ligand P6G A 1010



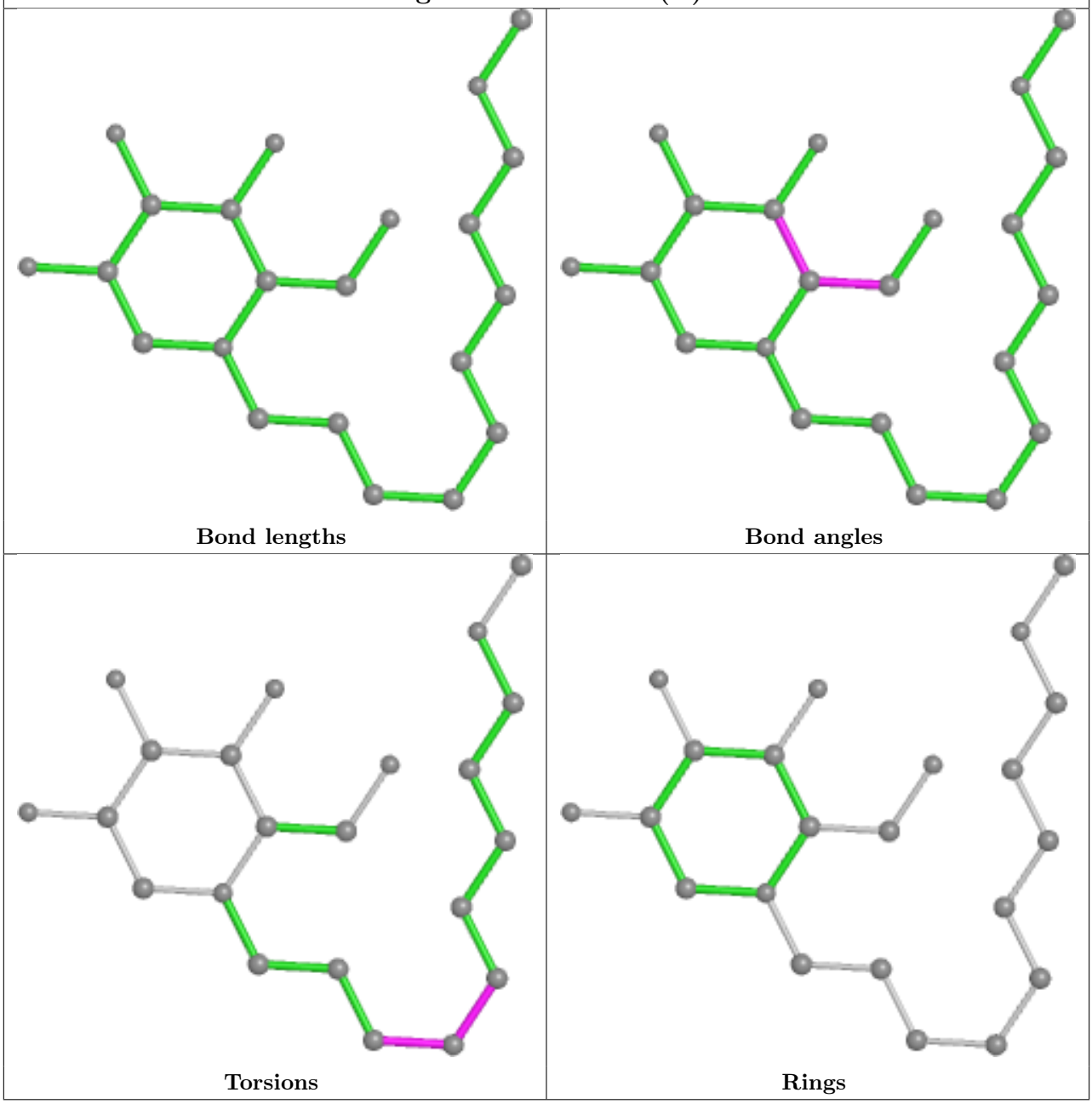
Ligand P6G A 1009



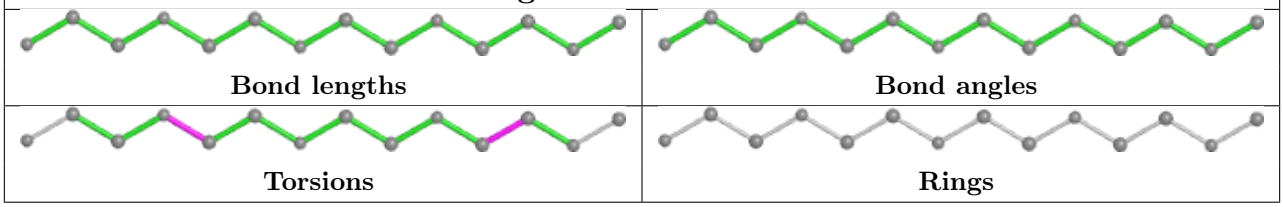
Ligand 6A9 A 1003 (A)

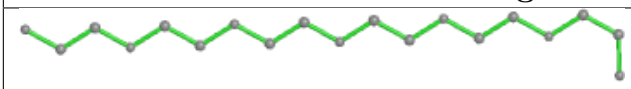
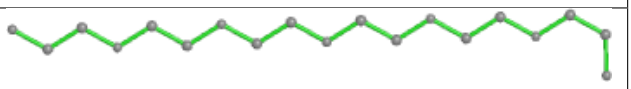
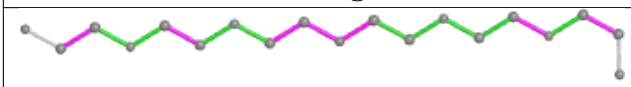
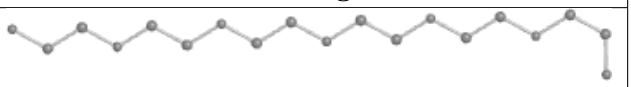


Ligand 6A9 A 1003 (B)



Ligand P6G A 1015



Ligand P6G A 1011	
 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 (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	856/914 (93%)	-0.27	10 (1%)	79	76	13, 21, 43, 78	0
2	B	88/88 (100%)	0.11	4 (4%)	33	27	17, 29, 55, 80	0
All	All	944/1002 (94%)	-0.23	14 (1%)	73	70	13, 21, 44, 80	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	184	PRO	5.5
1	A	524	CYS	5.5
2	B	30	PHE	4.5
1	A	185	ARG	4.1
1	A	968	SER	3.5
2	B	43	THR	3.5
1	A	523	TRP	3.4
1	A	350	ASN	3.2
2	B	117	ARG	2.9
1	A	183	ALA	2.7
1	A	370	THR	2.5
1	A	571[A]	PHE	2.4
1	A	795	GLU	2.3
2	B	109	GLY	2.0

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

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

6.3 Carbohydrates (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

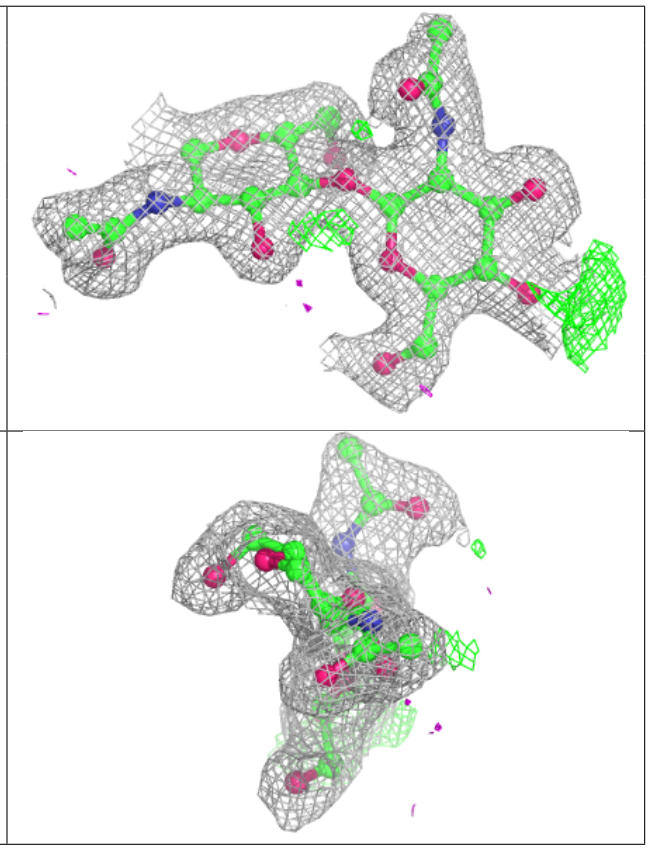
median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NAG	C	2	14/15	0.90	0.12	32,35,37,39	0
3	NAG	C	1	14/15	0.97	0.07	21,25,28,29	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)



6.4 Ligands (i)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	P6G	A	1010	19/19	0.55	0.39	60,70,76,77	0
6	P6G	A	1011	19/19	0.56	0.36	63,66,78,79	0
6	P6G	A	1015	14/19	0.62	0.43	53,58,68,70	0
8	PG4	A	1014	9/13	0.65	0.30	42,54,60,60	0

Continued on next page...

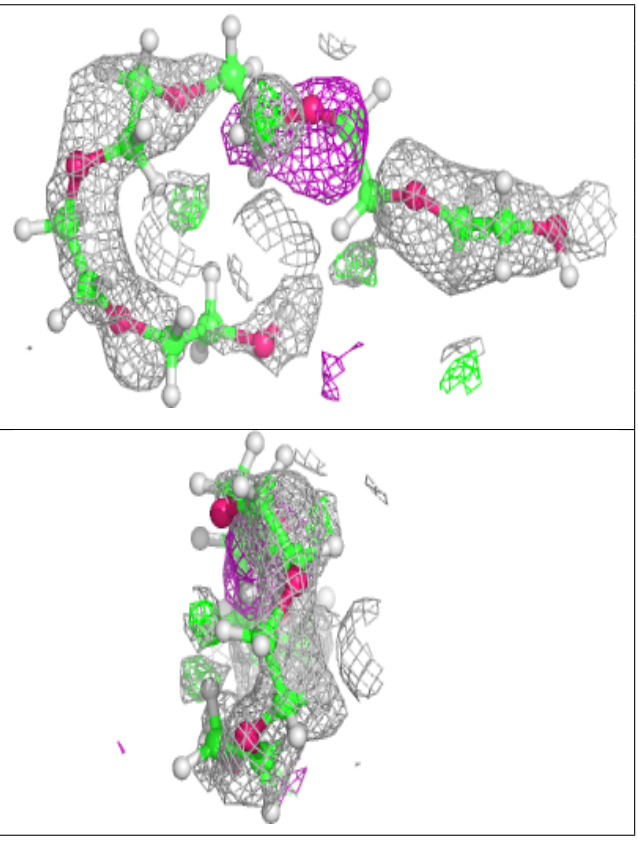
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
8	PG4	B	203	9/13	0.75	0.19	50,54,61,62	0
6	P6G	A	1009	14/19	0.85	0.15	33,42,56,58	0
7	DMS	A	1012	4/4	0.87	0.26	59,60,61,61	0
5	FMT	A	1007	3/3	0.91	0.16	55,56,57,57	0
8	PG4	A	1013	8/13	0.91	0.14	47,47,52,54	0
5	FMT	A	1006	3/3	0.92	0.14	46,46,46,48	0
5	FMT	A	1008	3/3	0.92	0.12	35,36,39,39	0
5	FMT	A	1005	3/3	0.94	0.12	36,36,37,37	0
4	6A9	A	1003[A]	22/22	0.95	0.14	9,12,35,36	22
4	6A9	A	1003[B]	22/22	0.95	0.14	29,30,35,37	22
5	FMT	A	1004	3/3	0.95	0.17	35,35,43,48	0
9	CA	B	201	1/1	0.98	0.10	21,21,21,21	0
9	CA	B	202	1/1	1.00	0.12	18,18,18,18	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.

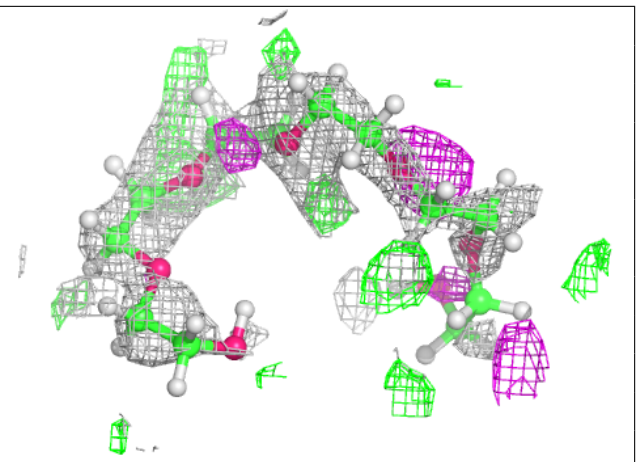
Electron density around P6G A 1010:

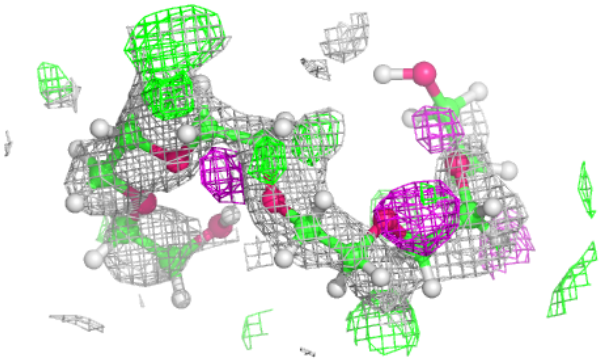
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)

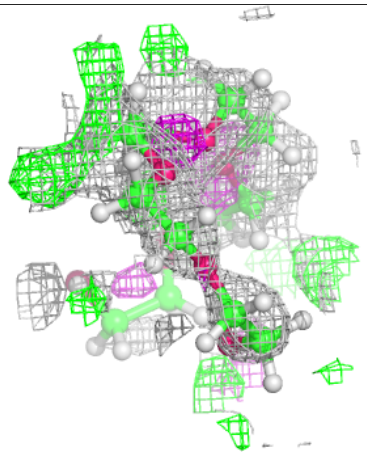


Electron density around P6G A 1011:

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

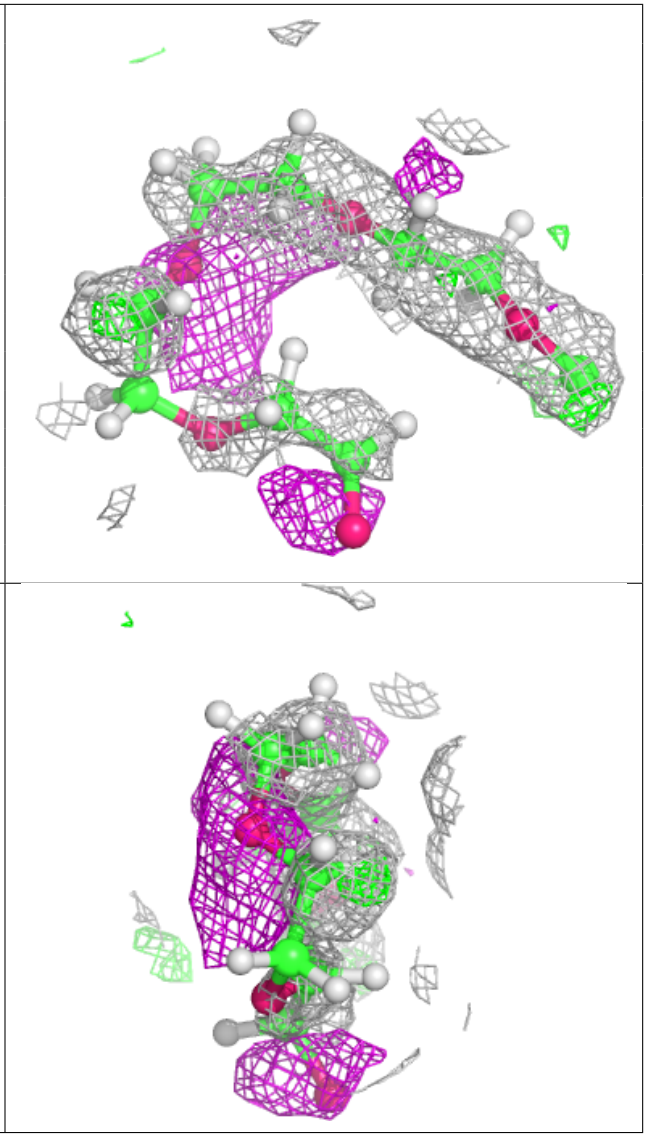






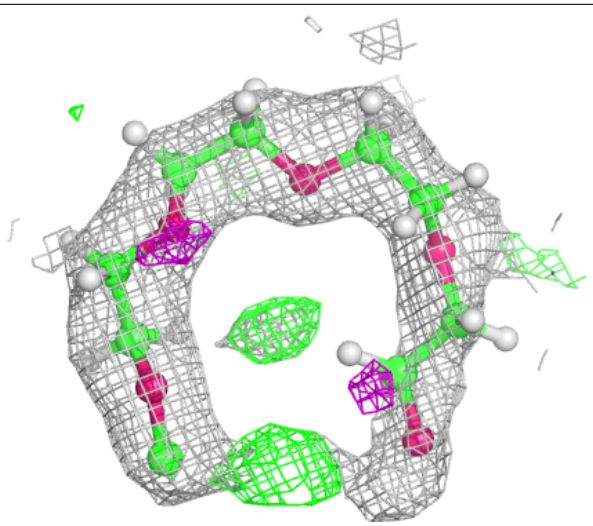
Electron density around P6G A 1015:

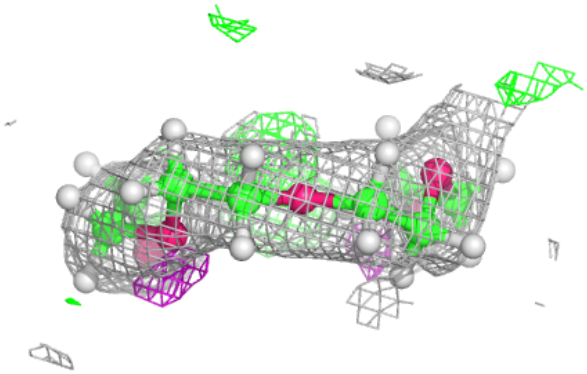
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

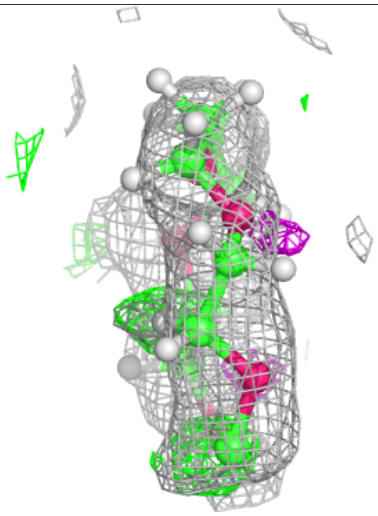


Electron density around P6G A 1009:

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

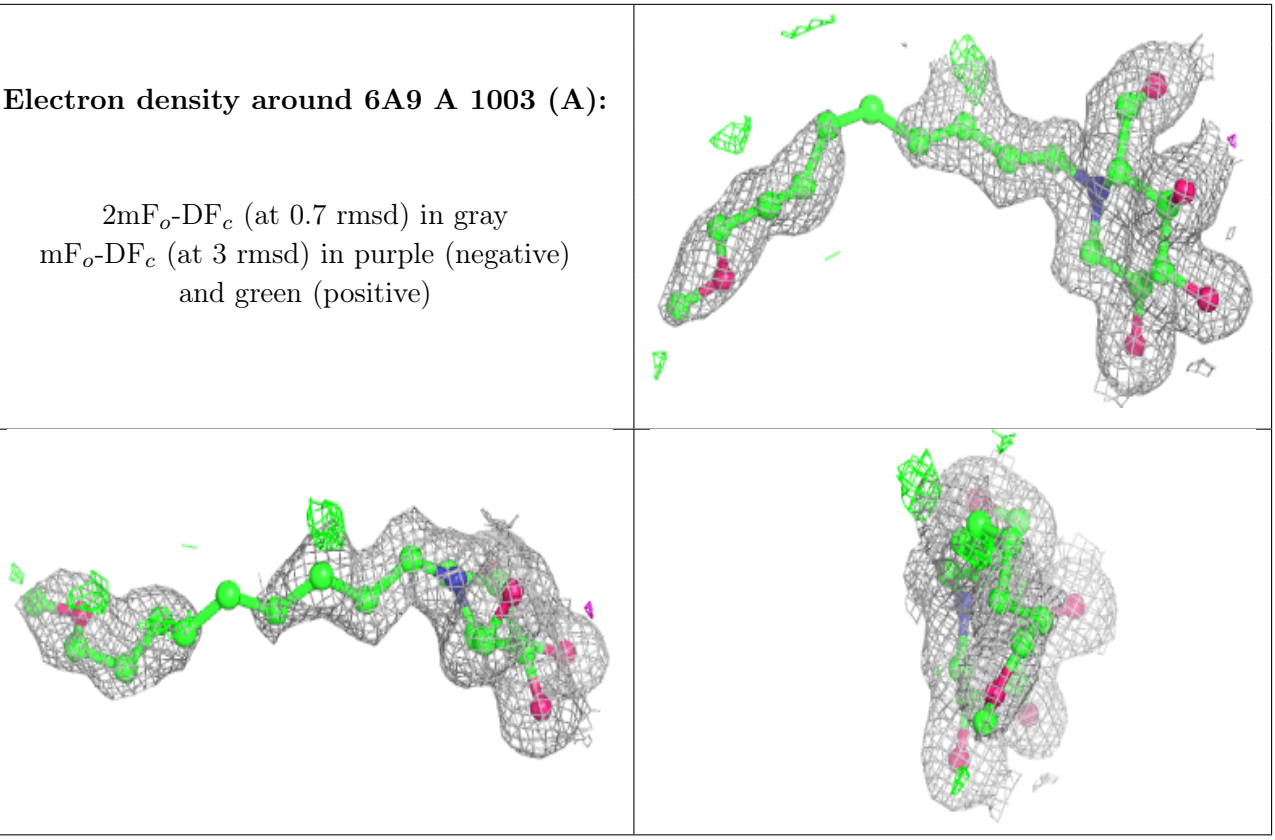






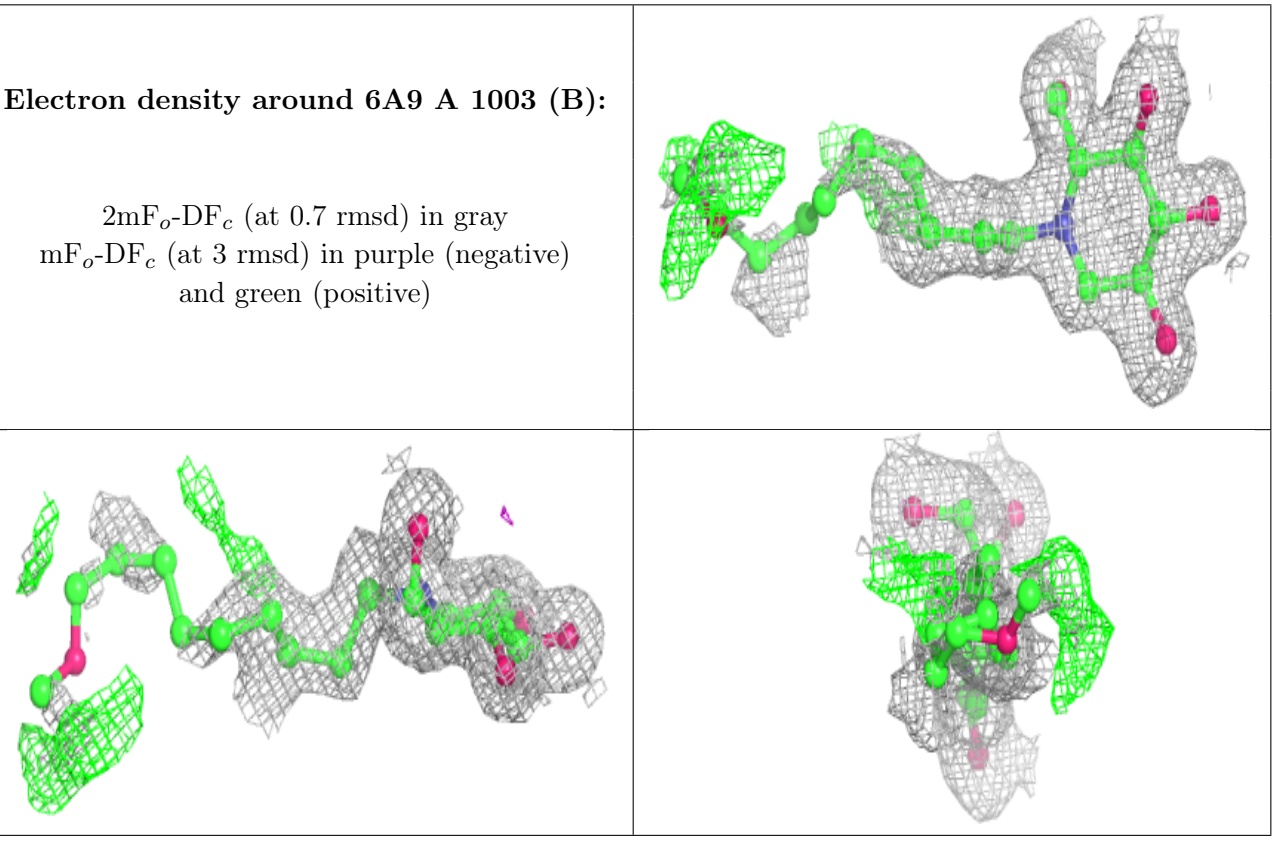
Electron density around 6A9 A 1003 (A):

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



Electron density around 6A9 A 1003 (B):

$2mF_o$-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.