Full wwPDB X-ray Structure Validation Report (i)

Dec 18, 2023 – 03:19 am GMT

PDB ID	:	4UFA
Title	:	Crystal structure of the Angiotensin-1 converting enzyme N-domain in complex with Ac-SD
Authors	:	Masuyer, G.; Douglas, R.G.; Sturrock, E.D.; Acharya, K.R.
Deposited on	:	2015-03-16
Resolution	:	1.80 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:

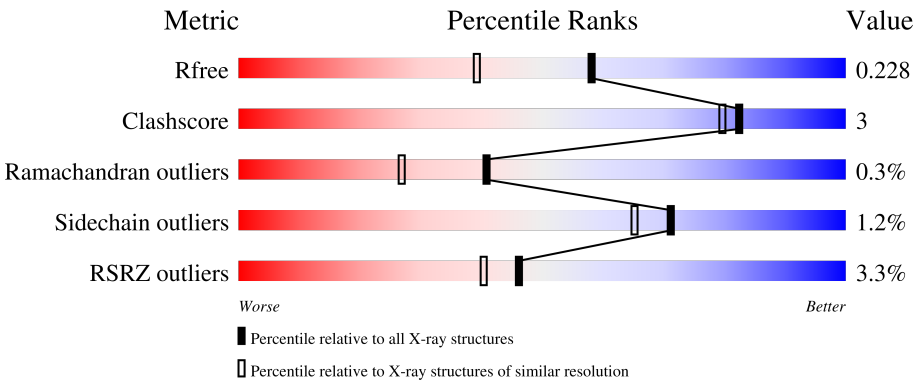
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	629	2%, 91%, 6%, •
1	B	629	4%, 90%, 6%, •
2	C	2	100%
3	D	4	25%, 75%
4	E	2	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	F	3	67% 33%

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
11	SAC	A	1301	-	-	X	-
11	SAC	B	1301	-	-	X	-
12	ASP	A	1302	-	-	X	-
12	ASP	B	1302	-	-	X	-

2 Entry composition (i)

There are 13 unique types of molecules in this entry. The entry contains 10977 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 ANGIOTENSIN-CONVERTING ENZYME.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	607	4967	3190	854	904	19	0	2	0
1	B	607	4941	3175	847	900	19	0	0	0

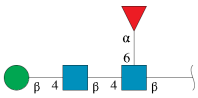
There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	629	LEU	-	expression tag	UNP P12821
A	9	GLN	ASN	engineered mutation	UNP P12821
A	25	GLN	ASN	engineered mutation	UNP P12821
A	82	GLN	ASN	engineered mutation	UNP P12821
A	117	GLN	ASN	engineered mutation	UNP P12821
A	131	GLN	ASN	engineered mutation	UNP P12821
A	289	GLN	ASN	engineered mutation	UNP P12821
A	545	ARG	GLN	engineered mutation	UNP P12821
A	576	LEU	PRO	engineered mutation	UNP P12821
B	629	LEU	-	expression tag	UNP P12821
B	9	GLN	ASN	engineered mutation	UNP P12821
B	25	GLN	ASN	engineered mutation	UNP P12821
B	82	GLN	ASN	engineered mutation	UNP P12821
B	117	GLN	ASN	engineered mutation	UNP P12821
B	131	GLN	ASN	engineered mutation	UNP P12821
B	289	GLN	ASN	engineered mutation	UNP P12821
B	545	ARG	GLN	engineered mutation	UNP P12821
B	576	LEU	PRO	engineered mutation	UNP P12821

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

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

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

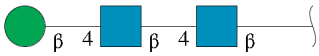


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	D	4	49	28	2	19	0	0	0

- Molecule 4 is an oligosaccharide called alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose.

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

- Molecule 5 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	F	3	39	22	2	15	0	0	0

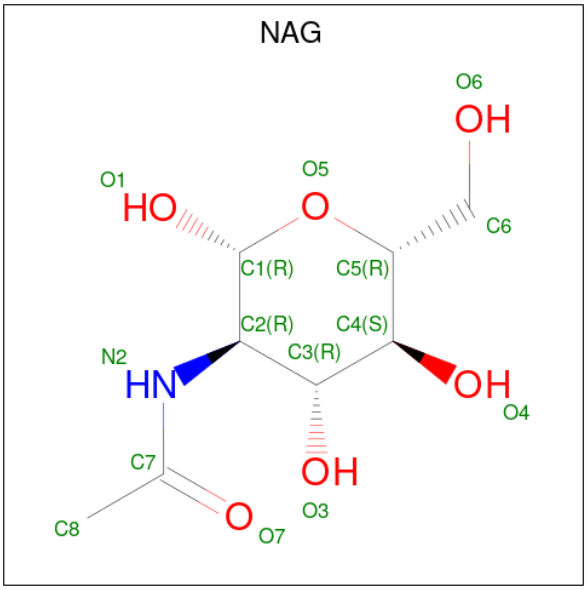
- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Zn 1 1	0	0
6	B	1	Total Zn 1 1	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

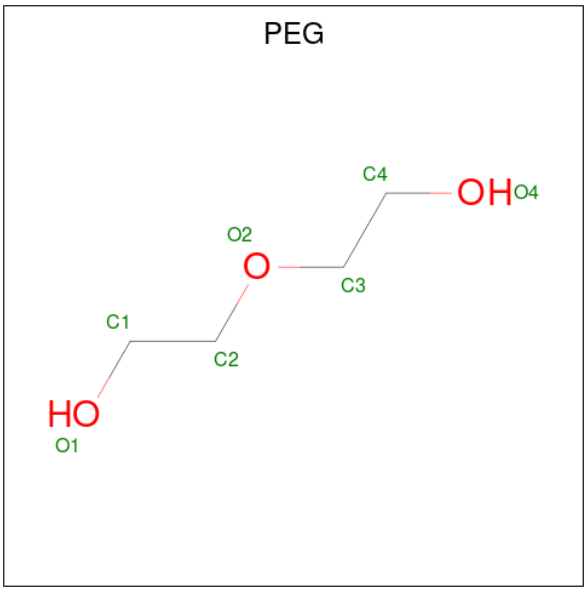
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total Cl 1 1	0	0
7	B	1	Total Cl 1 1	0	0

- Molecule 8 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



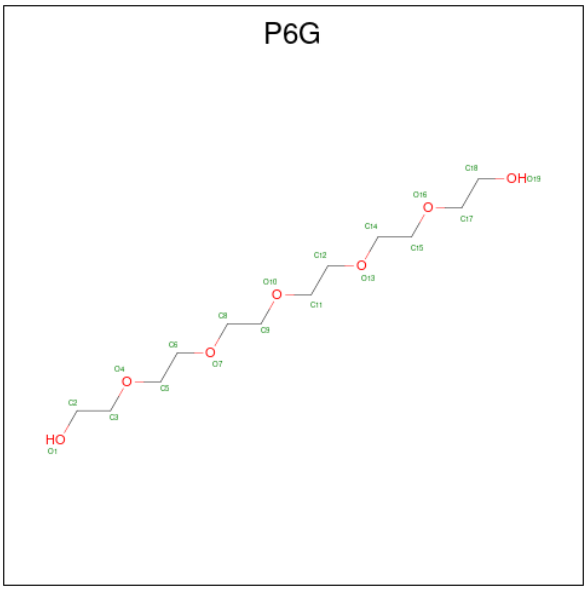
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C N O 14 8 1 5	0	0
8	B	1	Total C N O 14 8 1 5	0	0

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



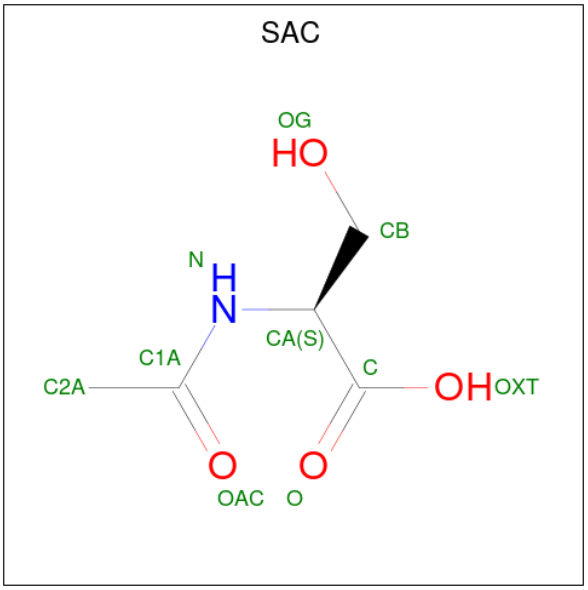
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
9	A	1	7	4	3	0	0
			Total	C	O		
9	A	1	7	4	3	0	0
			Total	C	O		
9	B	1	7	4	3	0	0
			Total	C	O		
9	B	1	7	4	3	0	0

- Molecule 10 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: $C_{12}H_{26}O_7$).



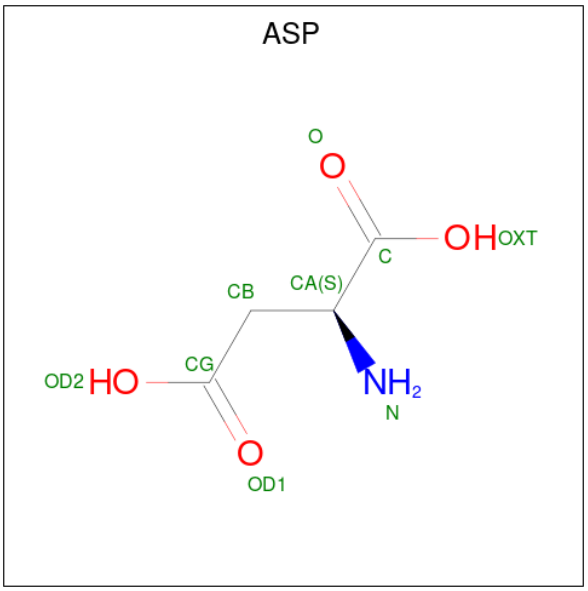
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O			
10	A	1	19	12	7		0	0
			Total	C	O			
10	B	1	19	12	7		0	0
			Total	C	O			
10	B	1	19	12	7		0	0

- Molecule 11 is N-ACETYL-SERINE (three-letter code: SAC) (formula: $C_5H_9NO_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
11	A	1	9	5	1	3	0	0
			Total	C	N	O		
11	B	1	9	5	1	3	0	0

- Molecule 12 is ASPARTIC ACID (three-letter code: ASP) (formula: $C_4H_7NO_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
12	A	1	9	4	1	4	0	0
			Total	C	N	O		
12	B	1	9	4	1	4	0	0

- Molecule 13 is water.

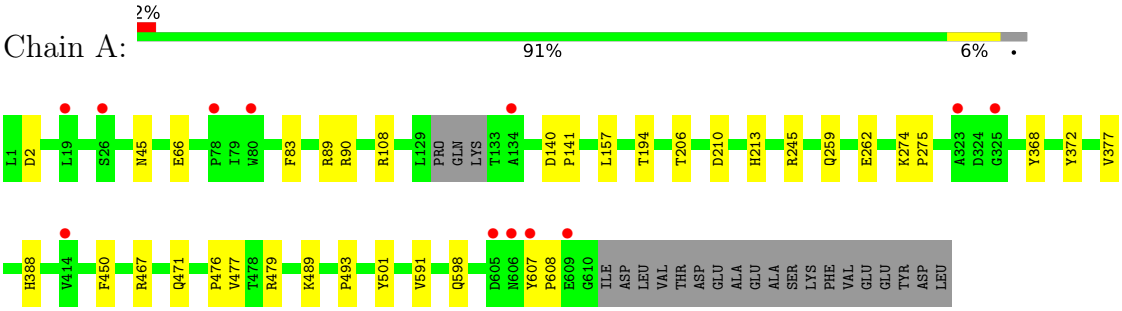
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
13	A	443	443	443	0	0
			Total	O		
13	B	333	333	333	0	0

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.

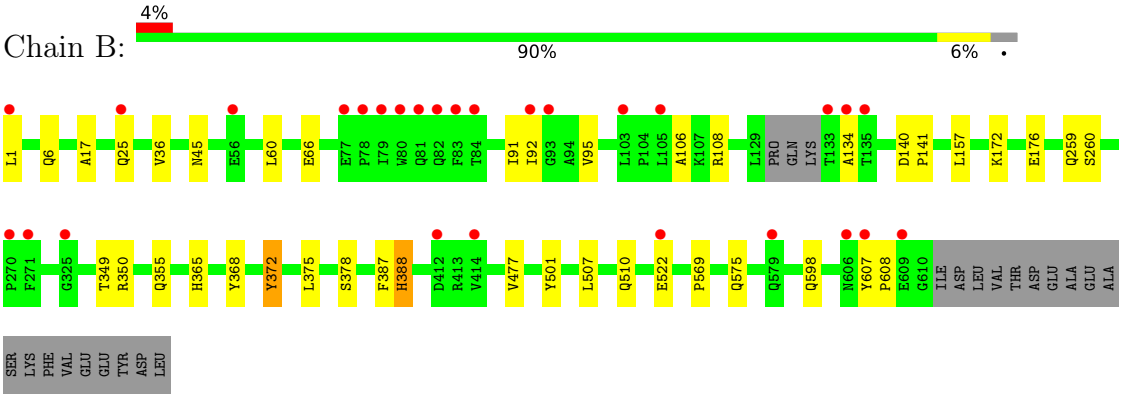
• Molecule 1: ANGIOTENSIN-CONVERTING ENZYME

Chain A:



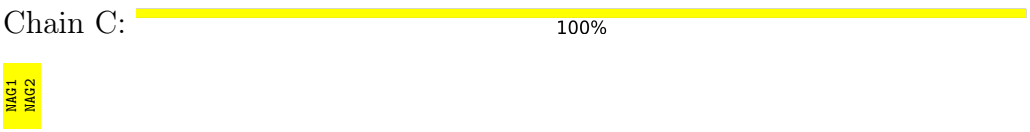
• Molecule 1: ANGIOTENSIN-CONVERTING ENZYME

Chain B:



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

Chain C:



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

Chain D:

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

Chain E: 100%

• Molecule 5: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F: 67% 33%

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	72.87Å 76.72Å 82.80Å 88.65° 64.31° 75.22°	Depositor
Resolution (Å)	74.24 – 1.80 34.79 – 1.80	Depositor EDS
% Data completeness (in resolution range)	90.4 (74.24-1.80) 90.4 (34.79-1.80)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.62 (at 1.81Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.202 , 0.228 0.202 , 0.228	Depositor DCC
R_{free} test set	6455 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	25.0	Xtriage
Anisotropy	0.238	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 38.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	10977	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.97% 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: PEG, BMA, SAC, CL, P6G, FUC, NAG, ZN

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	0.53	0/5126	0.65	0/6981
1	B	0.52	0/5096	0.64	0/6942
All	All	0.52	0/10222	0.65	0/13923

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	4967	0	4737	20	0
1	B	4941	0	4707	25	0
2	C	28	0	25	0	0
3	D	49	0	43	0	0
4	E	24	0	22	0	0
5	F	39	0	34	1	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	1	0	0	0	0
7	B	1	0	0	0	0
8	A	14	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	B	14	0	13	0	0
9	A	14	0	20	0	0
9	B	14	0	20	0	0
10	A	19	0	26	0	0
10	B	38	0	52	1	0
11	A	9	0	8	4	0
11	B	9	0	8	4	0
12	A	9	0	4	4	0
12	B	9	0	4	4	0
13	A	443	0	0	3	0
13	B	333	0	0	2	0
All	All	10977	0	9736	50	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 3.

All (50) 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:365:HIS:HD1	1:B:388:HIS:HD2	1.18	0.89
1:B:259:GLN:NE2	12:B:1302:ASP:HB2	1.87	0.88
1:B:259:GLN:HE21	12:B:1302:ASP:HB2	1.41	0.85
1:A:467:ARG:HH11	1:A:471:GLN:HE22	1.29	0.80
11:A:1301:SAC:HA	12:A:1302:ASP:OXT	1.81	0.78
1:A:259:GLN:NE2	12:A:1302:ASP:HB2	1.99	0.78
1:B:365:HIS:HD1	1:B:388:HIS:CD2	2.02	0.78
1:B:350:ARG:H	1:B:355:GLN:HE21	1.31	0.77
1:A:259:GLN:HE21	12:A:1302:ASP:HB2	1.50	0.75
1:A:157:LEU:HD11	1:A:477:VAL:HG13	1.69	0.74
11:B:1301:SAC:HA	12:B:1302:ASP:OXT	1.90	0.72
1:B:157:LEU:HD11	1:B:477:VAL:HG13	1.77	0.66
1:B:522:GLU:OE1	5:F:3:BMA:H5	2.02	0.58
1:A:245:ARG:HG2	1:A:591:VAL:HG11	1.85	0.58
1:B:1:LEU:HD23	1:B:6:GLN:HG2	1.85	0.58
1:A:83:PHE:HB2	1:A:89:ARG:HG2	1.86	0.57
11:A:1301:SAC:HB3	13:A:2301:HOH:O	2.03	0.57
11:B:1301:SAC:HB3	13:B:2234:HOH:O	2.04	0.56
11:A:1301:SAC:HA	12:A:1302:ASP:C	2.26	0.55
1:A:213[B]:HIS:HD2	13:A:2196:HOH:O	1.91	0.53
1:B:17:ALA:HB1	1:B:92:ILE:HD11	1.91	0.53
1:A:206:THR:HG23	1:A:210:ASP:OD2	2.09	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:467:ARG:NH1	1:A:471:GLN:HE22	2.04	0.51
1:B:372:TYR:HB2	1:B:375:LEU:HD12	1.93	0.51
1:B:106:ALA:HB1	13:B:2056:HOH:O	2.10	0.51
1:B:91:ILE:O	1:B:95:VAL:HG23	2.12	0.50
1:A:90:ARG:NH2	13:A:2064:HOH:O	2.45	0.50
11:B:1301:SAC:HA	12:B:1302:ASP:C	2.30	0.50
1:B:501:TYR:CE1	11:B:1301:SAC:HB2	2.47	0.49
1:A:501:TYR:CE1	11:A:1301:SAC:HB2	2.48	0.48
1:B:66:GLU:CD	1:B:108:ARG:HH22	2.16	0.48
1:A:489:LYS:O	1:A:493:PRO:HD2	2.16	0.46
1:B:510:GLN:HG2	1:B:569:PRO:HG2	1.97	0.46
1:B:350:ARG:H	1:B:355:GLN:NE2	2.07	0.46
1:A:157:LEU:HD13	1:A:476:PRO:HB2	1.96	0.46
1:B:36:VAL:HG22	10:B:1204:P6G:H31	1.99	0.44
1:A:66:GLU:CD	1:A:108:ARG:HH22	2.22	0.44
1:B:607:TYR:HA	1:B:608:PRO:HA	1.75	0.43
1:A:245:ARG:HG2	1:A:591:VAL:CG1	2.48	0.43
1:A:140:ASP:HA	1:A:141:PRO:HA	1.83	0.43
1:B:25:GLN:HE21	1:B:378:SER:H	1.66	0.43
1:A:479:ARG:H	1:B:598:GLN:NE2	2.17	0.43
1:B:172:LYS:O	1:B:176:GLU:HG3	2.19	0.42
1:A:607:TYR:HA	1:A:608:PRO:HA	1.64	0.42
1:B:387:PHE:CD1	1:B:507:LEU:HD21	2.55	0.42
1:A:274:LYS:HB3	1:A:275:PRO:CD	2.49	0.42
1:B:372:TYR:OH	1:B:388:HIS:HE1	2.03	0.41
1:B:349:THR:HA	1:B:355:GLN:NE2	2.36	0.41
1:A:194:THR:HB	1:A:450:PHE:CD1	2.55	0.41
1:B:140:ASP:HA	1:B:141:PRO:HA	1.90	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	605/629 (96%)	593 (98%)	11 (2%)	1 (0%)	47	33
1	B	603/629 (96%)	592 (98%)	8 (1%)	3 (0%)	29	15
All	All	1208/1258 (96%)	1185 (98%)	19 (2%)	4 (0%)	41	27

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	45	ASN
1	B	575	GLN
1	A	45	ASN
1	B	134	ALA

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	521/541 (96%)	514 (99%)	7 (1%)	69	62
1	B	517/541 (96%)	512 (99%)	5 (1%)	76	71
All	All	1038/1082 (96%)	1026 (99%)	12 (1%)	71	65

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	ASP
1	A	262	GLU
1	A	368	TYR
1	A	372	TYR
1	A	377	VAL
1	A	388	HIS
1	A	598	GLN
1	B	60	LEU
1	B	260	SER
1	B	368	TYR
1	B	372	TYR
1	B	388	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	6	GLN
1	A	9	GLN
1	A	263	ASN
1	A	371	GLN
1	A	471	GLN
1	A	491	HIS
1	B	25	GLN
1	B	87	GLN
1	B	109	GLN
1	B	117	GLN
1	B	355	GLN
1	B	371	GLN
1	B	388	HIS
1	B	491	HIS
1	B	582	GLN
1	B	588	ASN
1	B	598	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

11 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NAG	C	1	1,2	14,14,15	0.57	0	17,19,21	1.29	1 (5%)
2	NAG	C	2	2	14,14,15	0.67	0	17,19,21	1.28	2 (11%)
3	NAG	D	1	3,1	14,14,15	0.45	0	17,19,21	0.96	1 (5%)
3	NAG	D	2	3	14,14,15	0.46	0	17,19,21	1.17	2 (11%)
3	BMA	D	3	3	11,11,12	0.68	0	15,15,17	1.27	2 (13%)
3	FUC	D	4	3	10,10,11	0.77	0	14,14,16	0.62	0
4	NAG	E	1	1,4	14,14,15	0.60	0	17,19,21	1.58	4 (23%)
4	FUC	E	2	4	10,10,11	0.83	0	14,14,16	1.98	4 (28%)
5	NAG	F	1	1,5	14,14,15	0.66	0	17,19,21	1.41	2 (11%)
5	NAG	F	2	5	14,14,15	0.61	0	17,19,21	1.83	3 (17%)
5	BMA	F	3	5	11,11,12	0.73	0	15,15,17	1.22	2 (13%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	C	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	C	2	2	-	2/6/23/26	0/1/1/1
3	NAG	D	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	D	2	3	-	0/6/23/26	0/1/1/1
3	BMA	D	3	3	-	0/2/19/22	0/1/1/1
3	FUC	D	4	3	-	-	0/1/1/1
4	NAG	E	1	1,4	-	2/6/23/26	0/1/1/1
4	FUC	E	2	4	-	-	0/1/1/1
5	NAG	F	1	1,5	-	0/6/23/26	0/1/1/1
5	NAG	F	2	5	-	0/6/23/26	0/1/1/1
5	BMA	F	3	5	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	F	2	NAG	C1-O5-C5	5.27	119.33	112.19
5	F	1	NAG	C1-O5-C5	4.52	118.31	112.19
2	C	1	NAG	C1-O5-C5	4.05	117.68	112.19
4	E	2	FUC	C1-C2-C3	3.68	114.19	109.67
4	E	2	FUC	C1-O5-C5	3.23	120.11	112.78

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	2	FUC	O5-C1-C2	3.17	115.66	110.77
4	E	1	NAG	O5-C5-C6	3.07	112.02	107.20
2	C	2	NAG	C4-C3-C2	3.00	115.42	111.02
5	F	3	BMA	C1-C2-C3	3.00	113.35	109.67
4	E	2	FUC	C3-C4-C5	2.96	114.38	109.77
5	F	2	NAG	C4-C3-C2	2.87	115.23	111.02
4	E	1	NAG	C1-O5-C5	2.76	115.93	112.19
4	E	1	NAG	C4-C3-C2	2.70	114.98	111.02
3	D	3	BMA	C1-O5-C5	2.67	115.81	112.19
3	D	2	NAG	C2-N2-C7	2.58	126.57	122.90
3	D	3	BMA	C1-C2-C3	2.54	112.79	109.67
2	C	2	NAG	O5-C5-C6	2.53	111.17	107.20
3	D	2	NAG	C4-C3-C2	2.52	114.71	111.02
5	F	1	NAG	C2-N2-C7	2.40	126.31	122.90
4	E	1	NAG	O5-C1-C2	2.37	115.03	111.29
5	F	2	NAG	C3-C4-C5	2.37	114.46	110.24
3	D	1	NAG	C2-N2-C7	2.32	126.21	122.90
5	F	3	BMA	C1-O5-C5	2.25	115.24	112.19

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	E	1	NAG	O5-C5-C6-O6
4	E	1	NAG	C4-C5-C6-O6
2	C	2	NAG	C4-C5-C6-O6
2	C	2	NAG	O5-C5-C6-O6

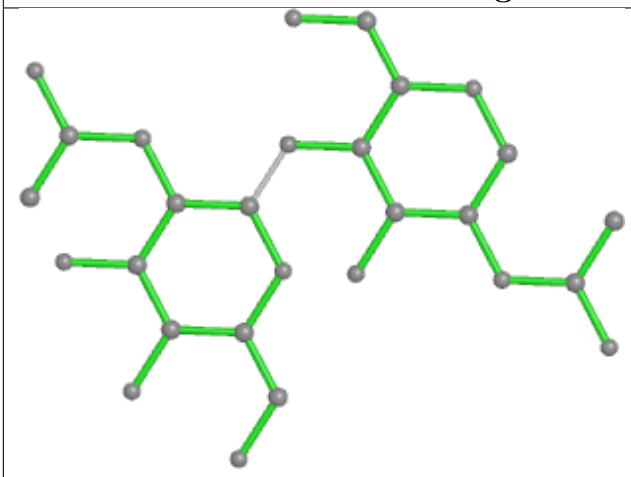
There are no ring outliers.

1 monomer is involved in 1 short contact:

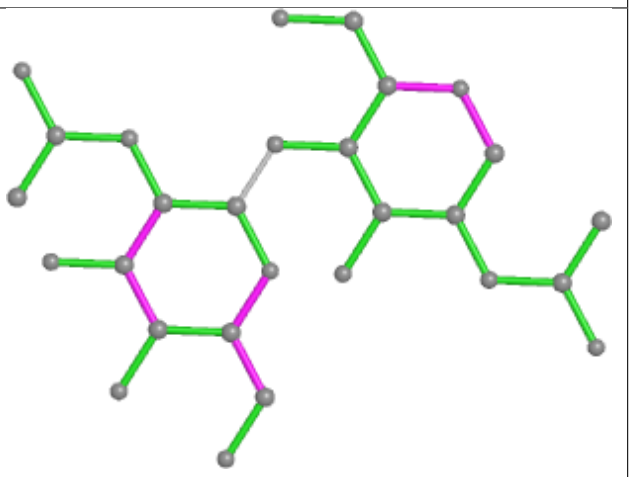
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	F	3	BMA	1	0

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

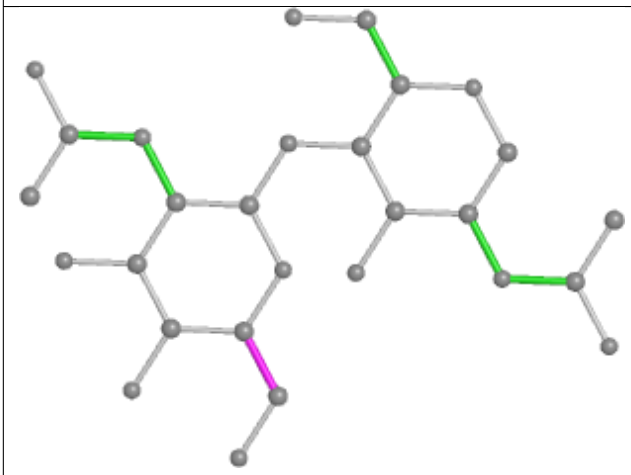
Oligosaccharide Chain C



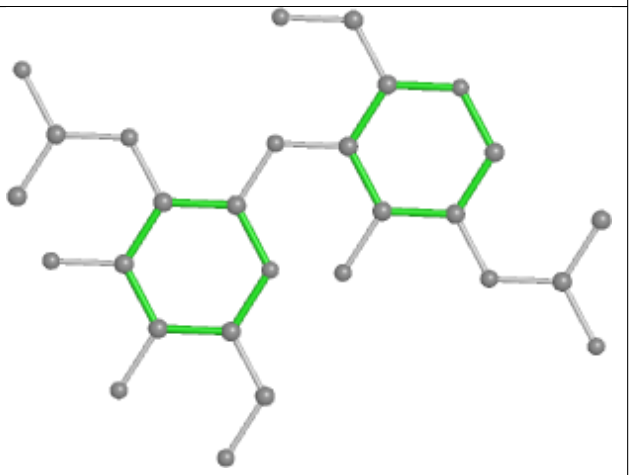
Bond lengths



Bond angles

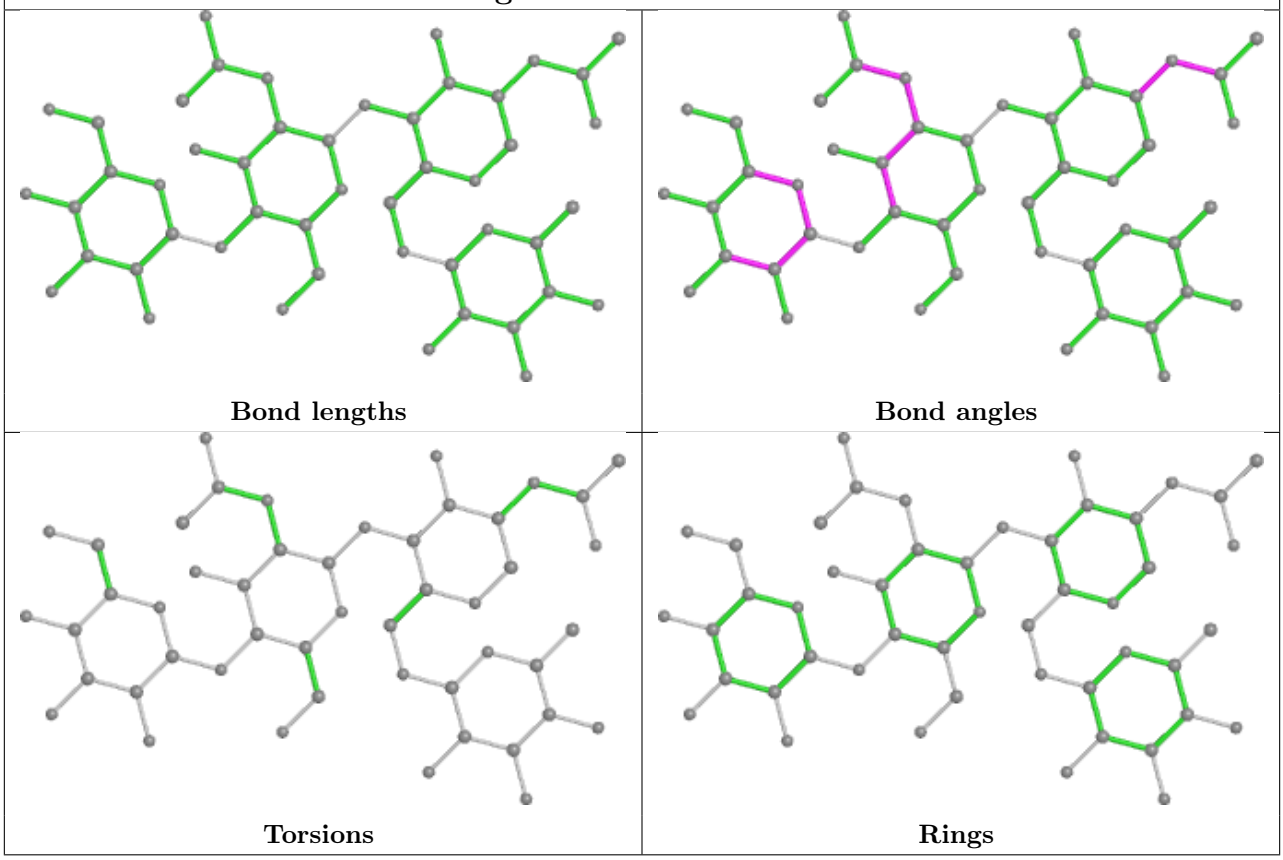


Torsions

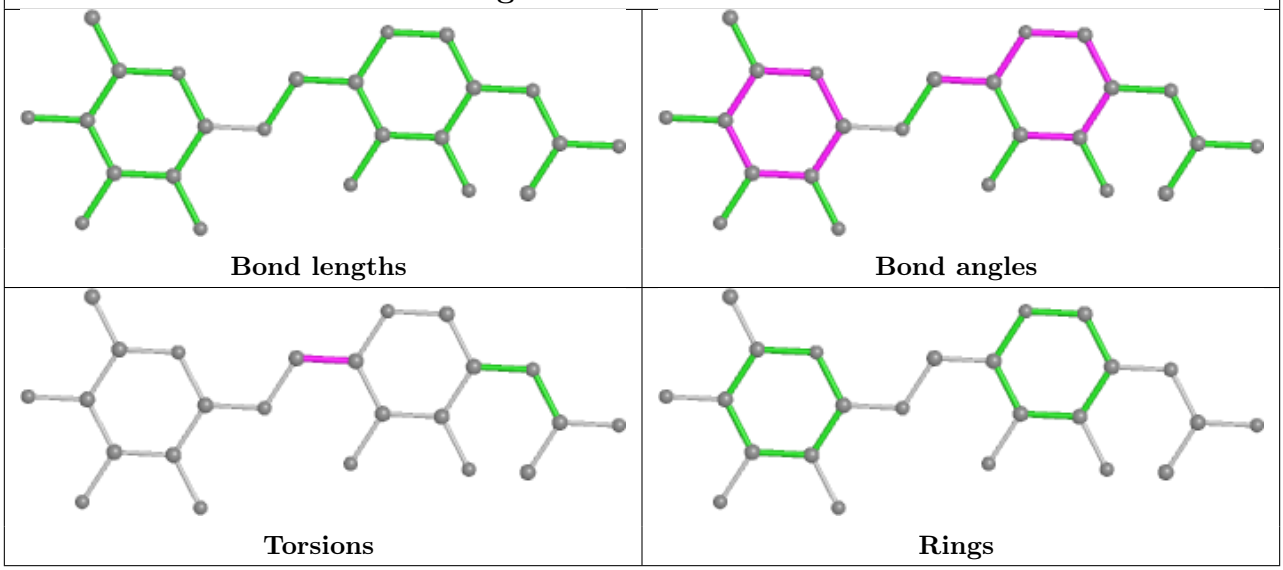


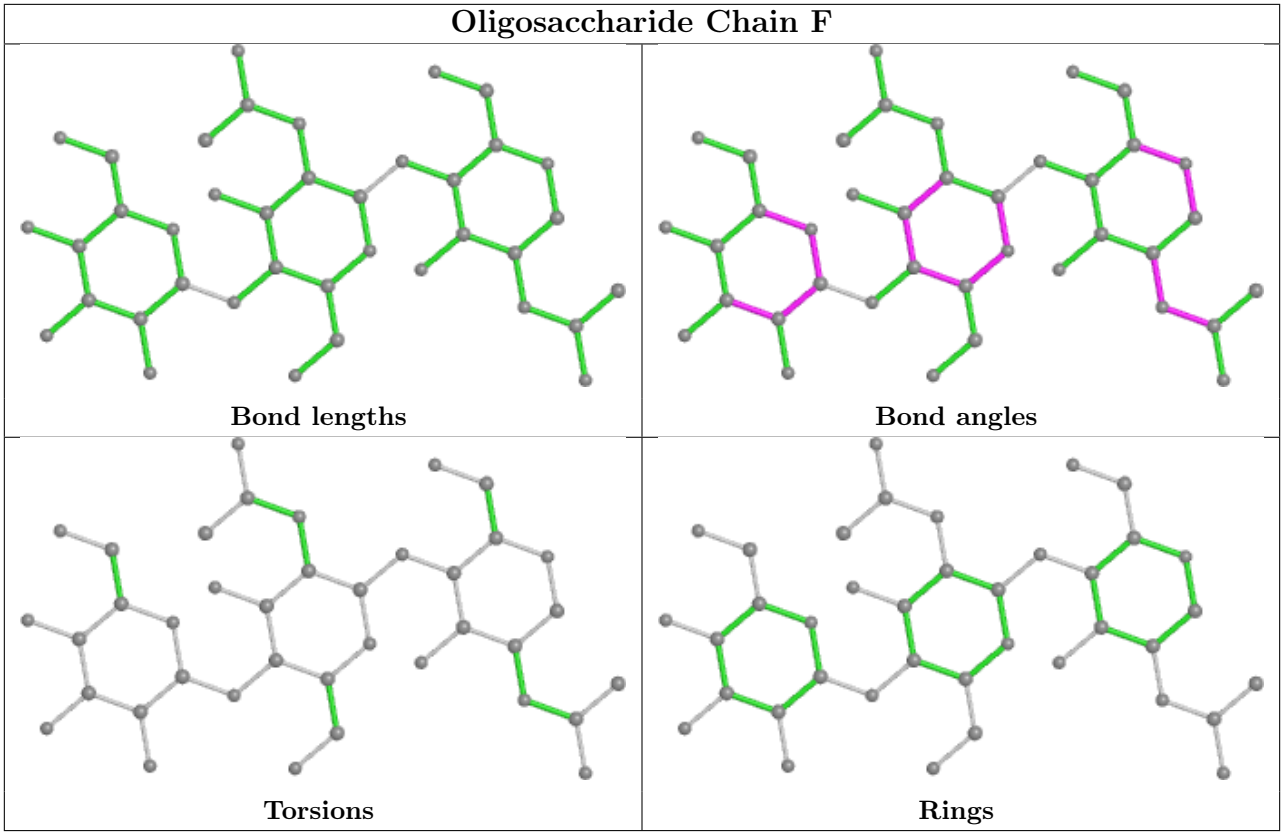
Rings

Oligosaccharide Chain D



Oligosaccharide Chain E





5.6 Ligand geometry (i)

Of 17 ligands modelled in this entry, 4 are monoatomic - leaving 13 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	$\#\|Z\| > 2$
12	ASP	B	1302	11	6,8,8	1.43	1 (16%)	8,10,10	1.95	2 (25%)
10	P6G	B	1203	-	18,18,18	0.58	0	17,17,17	0.26	0
10	P6G	B	1204	-	18,18,18	0.49	0	17,17,17	0.26	0
8	NAG	B	1102	1	14,14,15	0.63	0	17,19,21	1.10	1 (5%)
9	PEG	A	1201	-	6,6,6	0.42	0	5,5,5	0.34	0
9	PEG	B	1201	-	6,6,6	0.50	0	5,5,5	0.24	0
8	NAG	A	1100	1	14,14,15	0.64	0	17,19,21	1.57	4 (23%)
10	P6G	A	1202	-	18,18,18	0.49	0	17,17,17	0.26	0
11	SAC	A	1301	12	7,8,9	1.21	1 (14%)	8,9,11	1.56	2 (25%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
11	SAC	B	1301	12	7,8,9	1.26	1 (14%)	8,9,11	1.69	1 (12%)
12	ASP	A	1302	11	6,8,8	1.48	1 (16%)	8,10,10	2.01	2 (25%)
9	PEG	B	1202	-	6,6,6	0.40	0	5,5,5	0.36	0
9	PEG	A	1200	-	6,6,6	0.40	0	5,5,5	0.32	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
12	ASP	B	1302	11	-	3/8/8/8	-
10	P6G	B	1203	-	-	11/16/16/16	-
10	P6G	B	1204	-	-	7/16/16/16	-
8	NAG	B	1102	1	-	2/6/23/26	0/1/1/1
9	PEG	A	1201	-	-	2/4/4/4	-
9	PEG	B	1201	-	-	2/4/4/4	-
8	NAG	A	1100	1	-	2/6/23/26	0/1/1/1
10	P6G	A	1202	-	-	13/16/16/16	-
11	SAC	A	1301	12	-	4/7/8/10	-
11	SAC	B	1301	12	-	5/7/8/10	-
12	ASP	A	1302	11	-	2/8/8/8	-
9	PEG	B	1202	-	-	1/4/4/4	-
9	PEG	A	1200	-	-	2/4/4/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	B	1301	SAC	CA-N	3.13	1.50	1.46
11	A	1301	SAC	CA-N	3.04	1.50	1.46
12	A	1302	ASP	OXT-C	-2.18	1.23	1.30
12	B	1302	ASP	OXT-C	-2.13	1.23	1.30

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
12	A	1302	ASP	OXT-C-O	-4.83	113.13	124.09
12	B	1302	ASP	OXT-C-O	-4.80	113.20	124.09
11	B	1301	SAC	C-CA-N	4.23	117.36	109.73

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	A	1100	NAG	C1-O5-C5	3.94	117.54	112.19
11	A	1301	SAC	C-CA-N	3.72	116.44	109.73
8	A	1100	NAG	O5-C5-C6	2.91	111.76	107.20
12	A	1302	ASP	OXT-C-CA	2.79	122.89	113.38
8	A	1100	NAG	C4-C3-C2	2.69	114.95	111.02
8	B	1102	NAG	O5-C1-C2	-2.56	107.25	111.29
12	B	1302	ASP	OXT-C-CA	2.47	121.79	113.38
8	A	1100	NAG	O5-C1-C2	2.25	114.84	111.29
11	A	1301	SAC	O-C-CA	-2.06	119.38	124.78

There are no chirality outliers.

All (56) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
11	A	1301	SAC	C2A-C1A-N-CA
11	A	1301	SAC	OAC-C1A-N-CA
11	A	1301	SAC	C-CA-CB-OG
11	B	1301	SAC	C2A-C1A-N-CA
11	B	1301	SAC	OAC-C1A-N-CA
11	B	1301	SAC	C-CA-N-C1A
11	B	1301	SAC	C-CA-CB-OG
12	A	1302	ASP	O-C-CA-N
12	B	1302	ASP	O-C-CA-N
8	A	1100	NAG	O5-C5-C6-O6
10	A	1202	P6G	O4-C5-C6-O7
10	B	1203	P6G	O13-C14-C15-O16
10	B	1203	P6G	O4-C5-C6-O7
10	B	1203	P6G	O7-C8-C9-O10
10	A	1202	P6G	O10-C11-C12-O13
10	B	1203	P6G	O10-C11-C12-O13
12	A	1302	ASP	OXT-C-CA-N
8	A	1100	NAG	C4-C5-C6-O6
10	A	1202	P6G	O13-C14-C15-O16
12	B	1302	ASP	OXT-C-CA-N
10	B	1203	P6G	O1-C2-C3-O4
8	B	1102	NAG	C4-C5-C6-O6
10	B	1203	P6G	O16-C17-C18-O19
10	A	1202	P6G	O1-C2-C3-O4
9	A	1200	PEG	C4-C3-O2-C2
9	A	1200	PEG	C1-C2-O2-C3
10	A	1202	P6G	C12-C11-O10-C9
10	A	1202	P6G	C18-C17-O16-C15

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
10	A	1202	P6G	C2-C3-O4-C5
10	B	1204	P6G	C8-C9-O10-C11
10	B	1204	P6G	C18-C17-O16-C15
10	B	1203	P6G	C15-C14-O13-C12
8	B	1102	NAG	O5-C5-C6-O6
9	B	1201	PEG	C4-C3-O2-C2
11	B	1301	SAC	CB-CA-N-C1A
10	B	1204	P6G	C14-C15-O16-C17
10	A	1202	P6G	C5-C6-O7-C8
9	B	1202	PEG	C1-C2-O2-C3
10	B	1203	P6G	C8-C9-O10-C11
10	B	1203	P6G	C9-C8-O7-C6
10	B	1204	P6G	C12-C11-O10-C9
12	B	1302	ASP	C-CA-CB-CG
10	A	1202	P6G	C15-C14-O13-C12
10	A	1202	P6G	C11-C12-O13-C14
10	B	1204	P6G	C6-C5-O4-C3
10	B	1203	P6G	C2-C3-O4-C5
9	A	1201	PEG	O2-C3-C4-O4
9	A	1201	PEG	C4-C3-O2-C2
9	B	1201	PEG	O1-C1-C2-O2
10	B	1204	P6G	C15-C14-O13-C12
10	A	1202	P6G	C6-C5-O4-C3
10	B	1203	P6G	C6-C5-O4-C3
11	A	1301	SAC	C-CA-N-C1A
10	B	1204	P6G	C9-C8-O7-C6
10	A	1202	P6G	O7-C8-C9-O10
10	A	1202	P6G	C14-C15-O16-C17

There are no ring outliers.

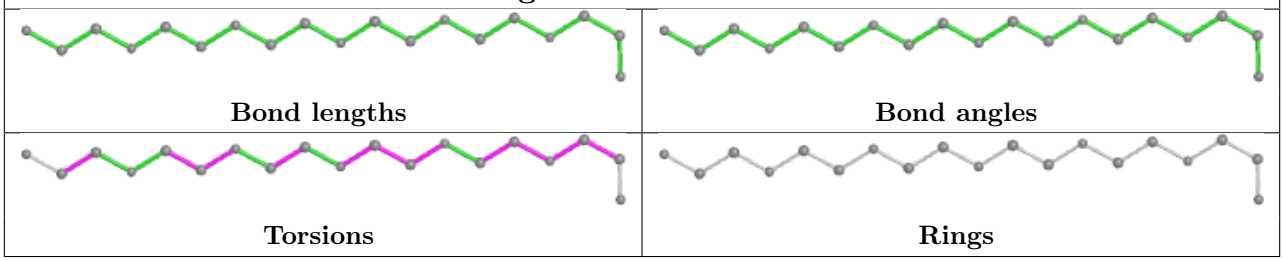
5 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
12	B	1302	ASP	4	0
10	B	1204	P6G	1	0
11	A	1301	SAC	4	0
11	B	1301	SAC	4	0
12	A	1302	ASP	4	0

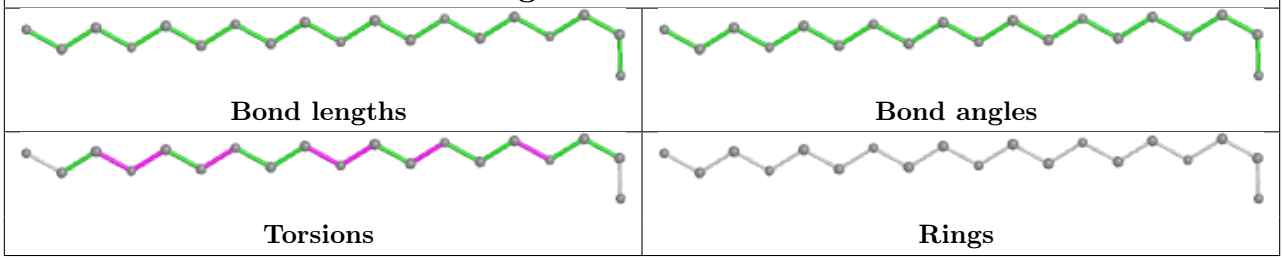
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

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.

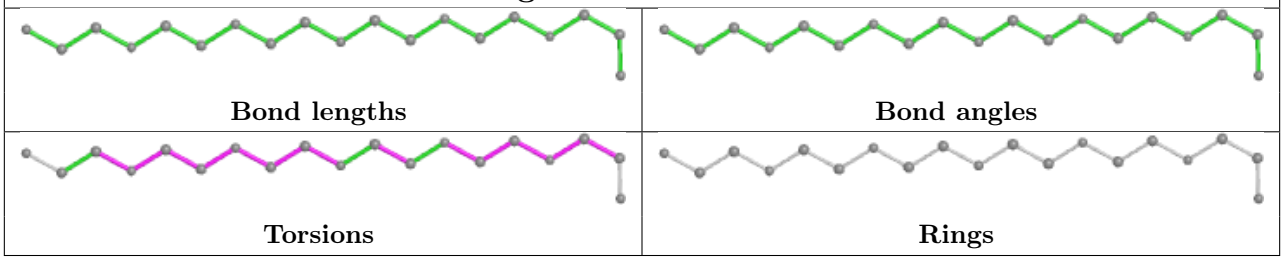
Ligand P6G B 1203



Ligand P6G B 1204



Ligand P6G A 1202



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	607/629 (96%)	-0.10	12 (1%)	65	61	13, 23, 43, 69	0
1	B	607/629 (96%)	0.20	28 (4%)	32	26	14, 28, 51, 73	0
All	All	1214/1258 (96%)	0.05	40 (3%)	46	40	13, 25, 48, 73	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	325	GLY	6.0
1	A	325	GLY	5.4
1	A	323	ALA	5.0
1	B	134	ALA	4.8
1	B	81	GLN	4.5
1	B	78	PRO	4.2
1	A	607	TYR	3.8
1	B	105	LEU	3.8
1	B	606	ASN	3.7
1	A	78	PRO	3.5
1	B	135	THR	3.5
1	A	606	ASN	3.1
1	B	84	THR	3.0
1	A	134	ALA	3.0
1	A	414	VAL	3.0
1	B	133	THR	2.9
1	B	82	GLN	2.9
1	B	1	LEU	2.9
1	A	80	TRP	2.9
1	B	83	PHE	2.9
1	B	270	PRO	2.7
1	A	609	GLU	2.6
1	A	19	LEU	2.6
1	B	79	ILE	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	92	ILE	2.4
1	A	26	SER	2.4
1	B	25	GLN	2.3
1	A	605	ASP	2.3
1	B	607	TYR	2.3
1	B	609	GLU	2.3
1	B	103	LEU	2.3
1	B	93	GLY	2.2
1	B	579	GLN	2.2
1	B	80	TRP	2.2
1	B	522	GLU	2.1
1	B	414	VAL	2.1
1	B	77	GLU	2.1
1	B	56	GLU	2.1
1	B	271	PHE	2.1
1	B	412	ASP	2.0

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

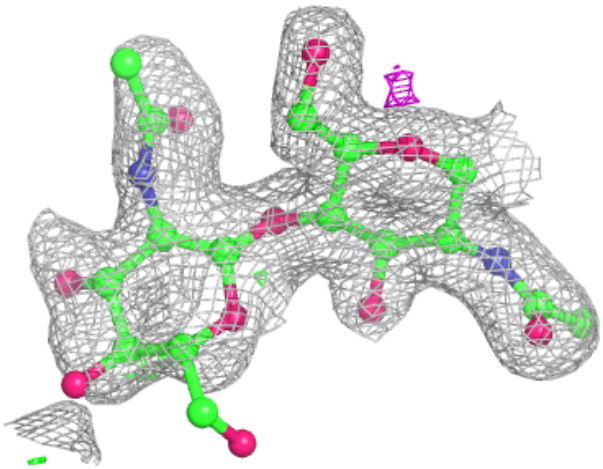
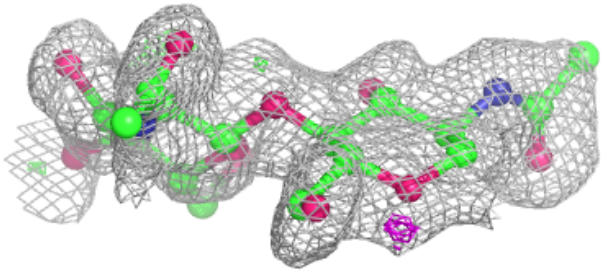
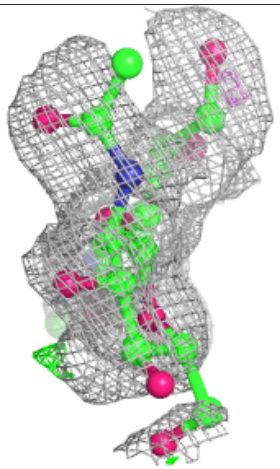
There are no non-standard protein/DNA/RNA residues in this entry.

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.

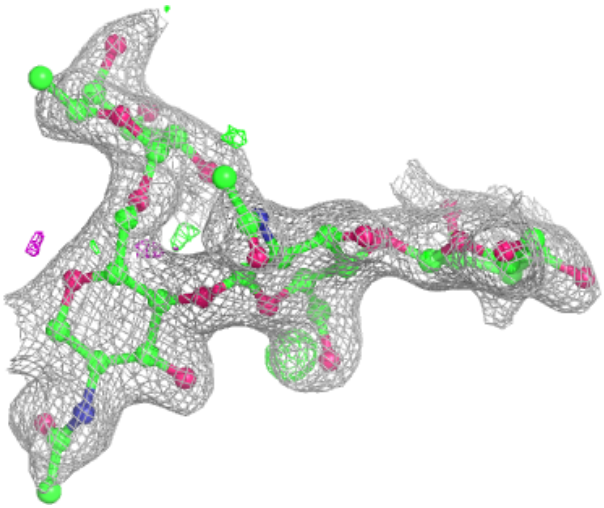
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	FUC	E	2	10/11	0.65	0.38	61,66,70,72	0
5	BMA	F	3	11/12	0.70	0.29	82,85,85,87	0
2	NAG	C	2	14/15	0.76	0.33	57,66,73,74	0
3	BMA	D	3	11/12	0.78	0.22	63,65,67,68	0
2	NAG	C	1	14/15	0.83	0.21	38,45,51,57	0
4	NAG	E	1	14/15	0.84	0.14	39,45,52,53	0
5	NAG	F	1	14/15	0.85	0.25	58,63,68,68	0
3	FUC	D	4	10/11	0.86	0.24	44,46,48,50	0
5	NAG	F	2	14/15	0.87	0.25	71,73,75,79	0
3	NAG	D	2	14/15	0.89	0.19	45,51,58,58	0
3	NAG	D	1	14/15	0.94	0.14	36,39,44,45	0

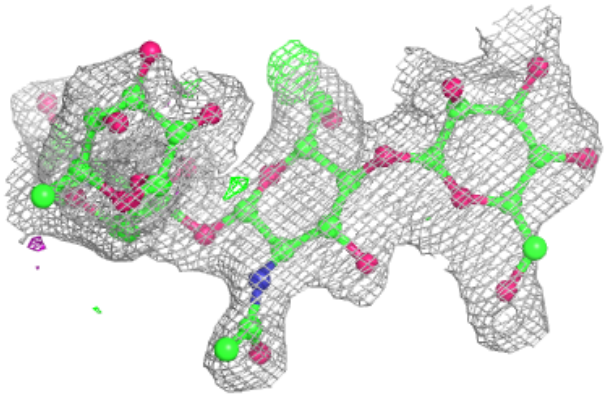
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

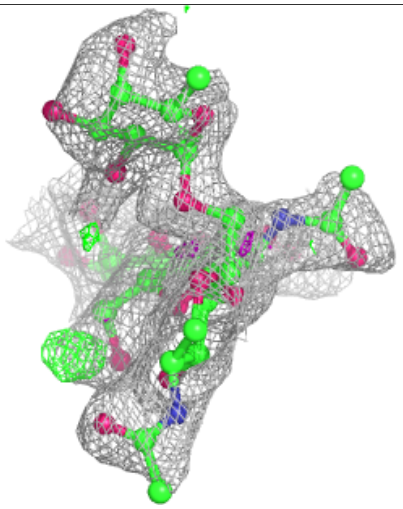
Electron density around Chain C: $2mF_o$-DF_c (at 0.7 rmsd) in gray mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)	
	

Electron density around Chain D:

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

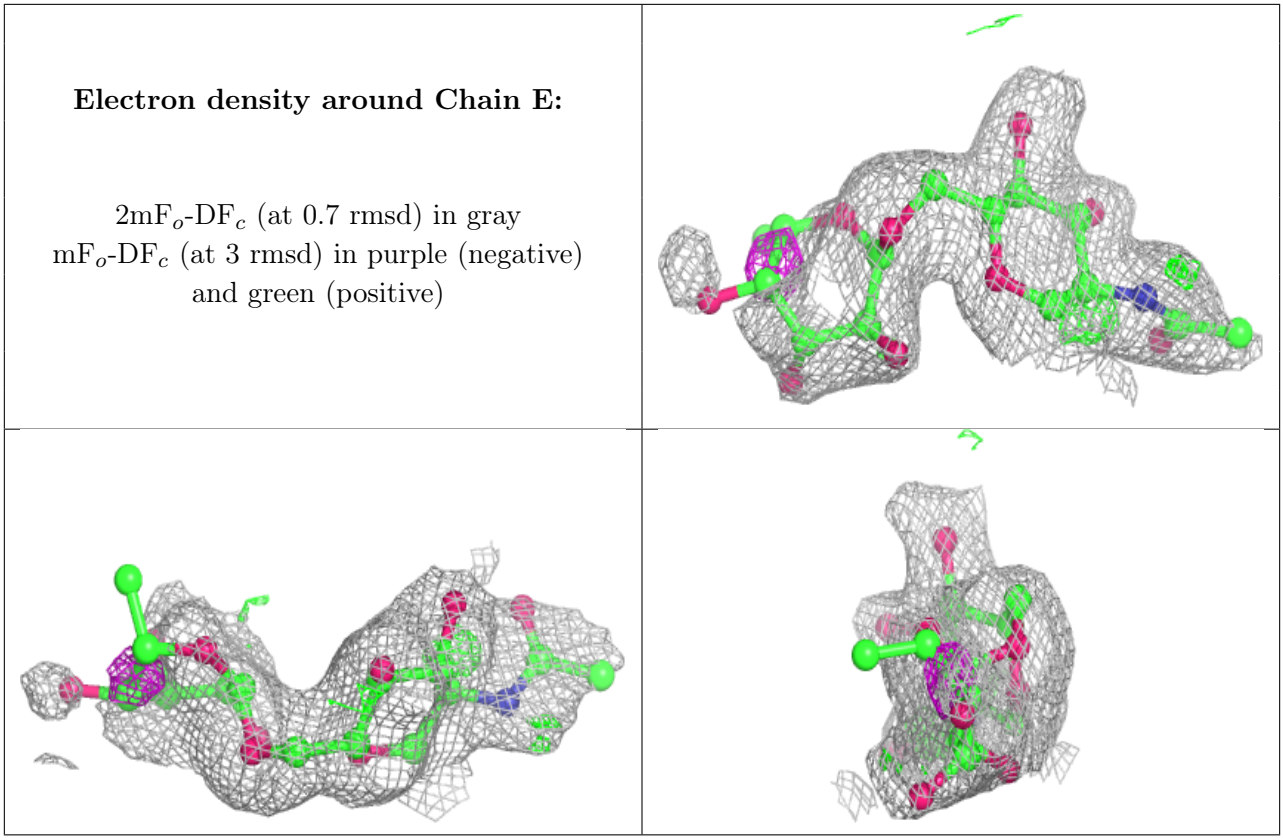






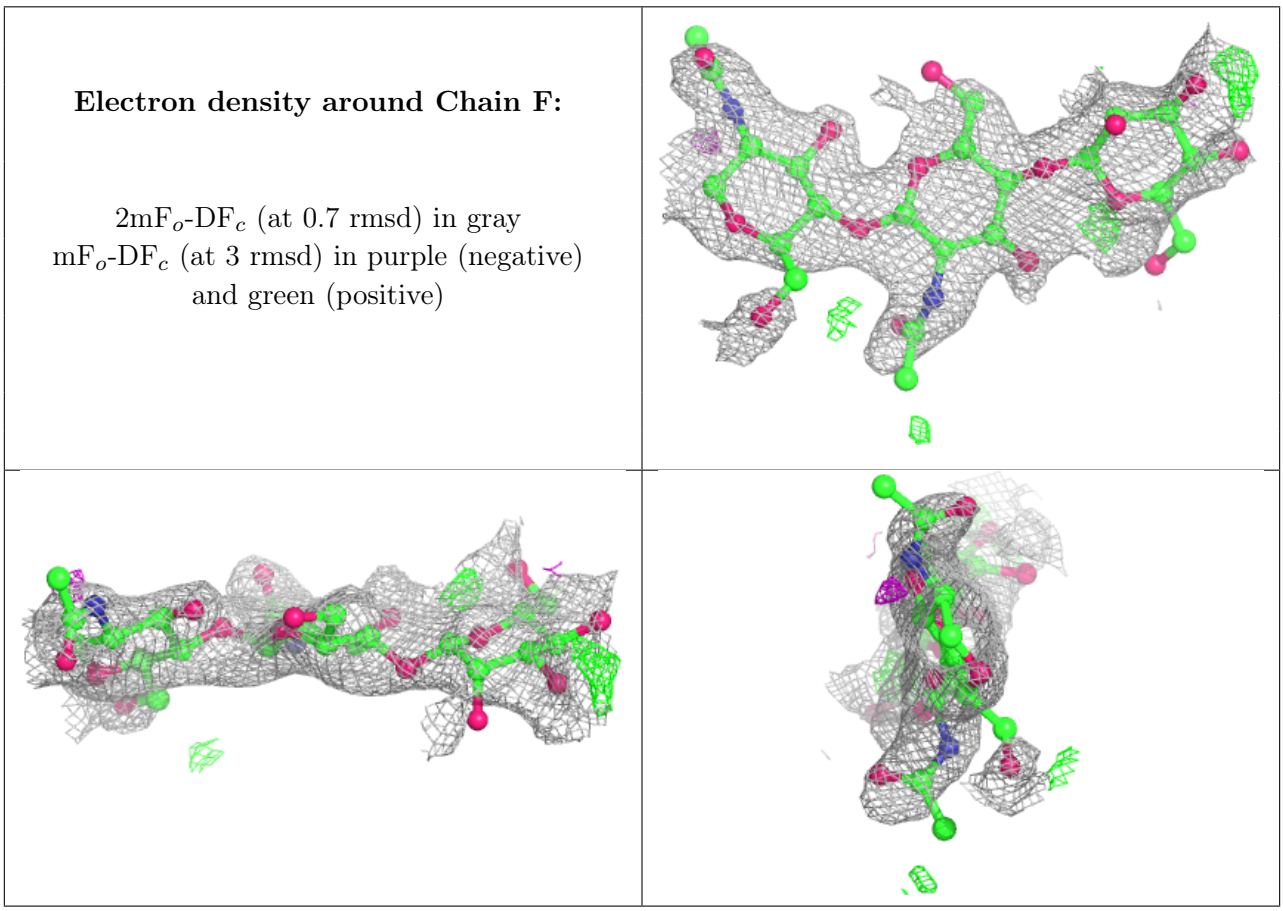
Electron density around Chain E:

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



Electron density around Chain F:

$2mF_o$-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)

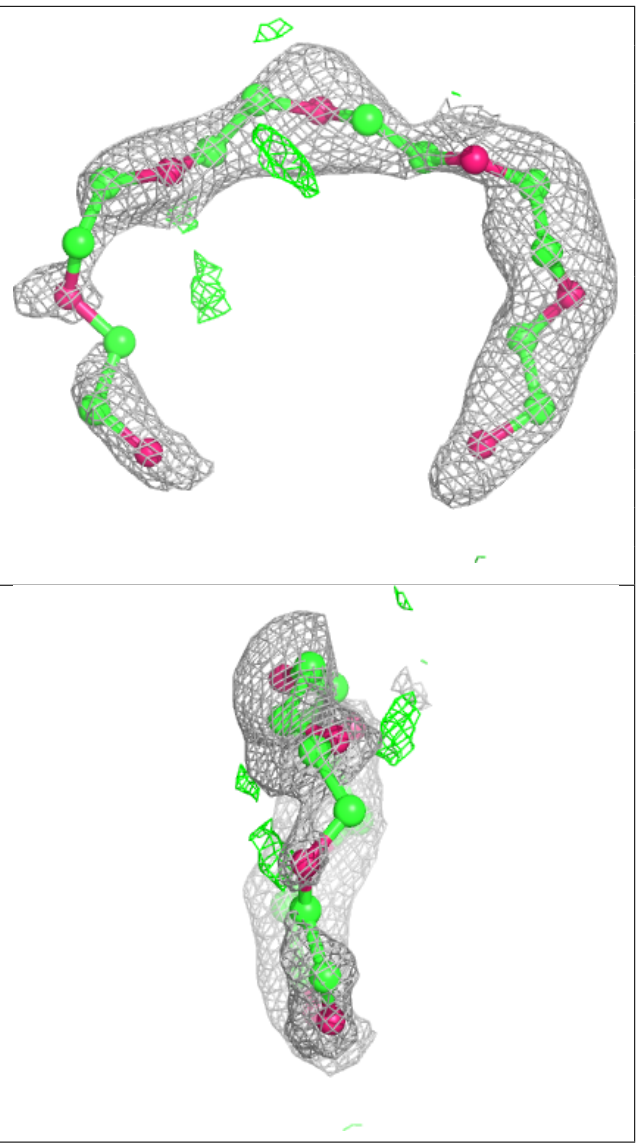
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
9	PEG	A	1201	7/7	0.68	0.23	53,56,61,62	0
10	P6G	A	1202	19/19	0.72	0.25	50,65,72,73	0
10	P6G	B	1204	19/19	0.76	0.26	60,64,74,76	0
10	P6G	B	1203	19/19	0.79	0.17	41,43,52,53	0
11	SAC	B	1301	9/10	0.79	0.23	41,45,49,50	0
12	ASP	A	1302	9/9	0.79	0.24	23,35,41,42	0
12	ASP	B	1302	9/9	0.79	0.23	25,31,38,41	0
8	NAG	B	1102	14/15	0.80	0.29	43,47,48,51	0
8	NAG	A	1100	14/15	0.82	0.16	30,35,40,44	0
11	SAC	A	1301	9/10	0.85	0.18	47,49,50,51	0
9	PEG	B	1201	7/7	0.87	0.21	39,41,43,46	0
9	PEG	A	1200	7/7	0.89	0.17	43,45,48,50	0
9	PEG	B	1202	7/7	0.93	0.15	35,36,41,41	0
7	CL	B	1002	1/1	0.99	0.11	21,21,21,21	0
7	CL	A	1002	1/1	1.00	0.09	17,17,17,17	0
6	ZN	A	1001	1/1	1.00	0.08	17,17,17,17	0
6	ZN	B	1001	1/1	1.00	0.07	17,17,17,17	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.

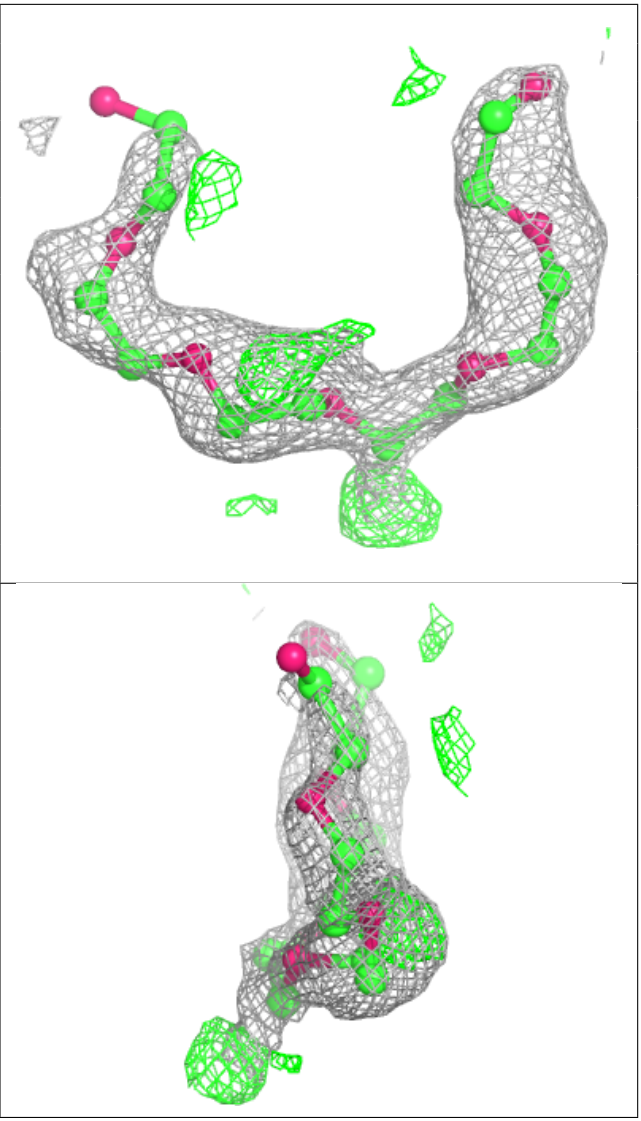
Electron density around P6G A 1202:

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



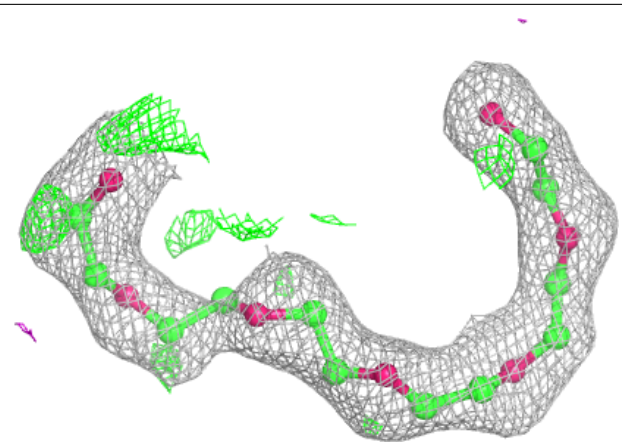
Electron density around P6G B 1204:

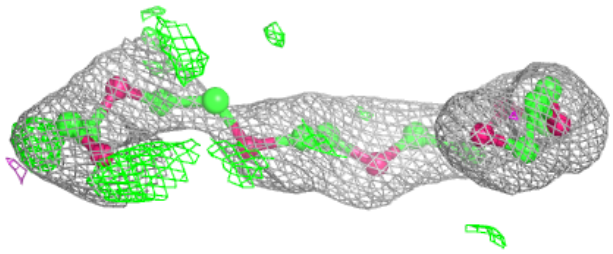
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

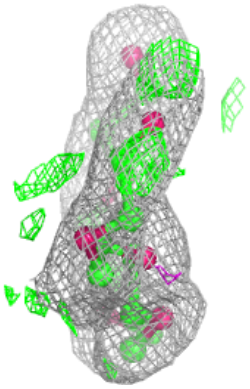


Electron density around P6G B 1203:

$2mF_o$-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.