Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2023 – 05:01 AM EDT

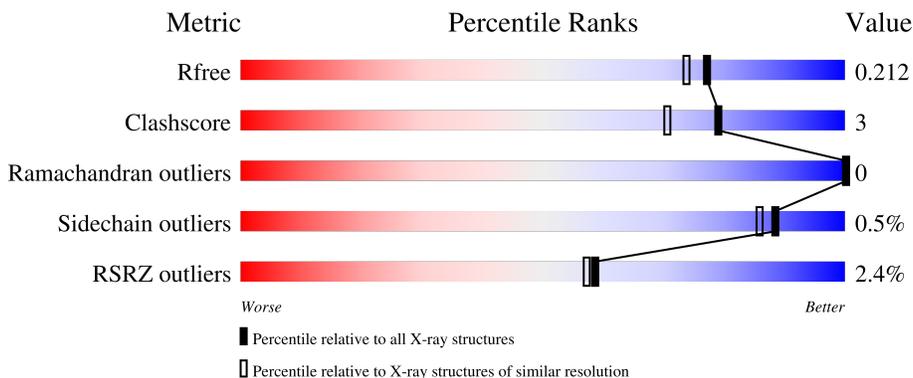|		
PDB ID	:	7JHO
Title	:	Structure of human beta 1,3-N-acetylglucosaminyltransferase 2 with UDP
Authors	:	Hao, Y.; Huang, X.
Deposited on	:	2020-07-21
Resolution	:	1.85 Å(reported)

1   Overall quality at a glance ⓘ

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

The reported resolution of this entry is 1.85 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	Percentile Ranks	Value
Rfree		0.212
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		0.5%
RSRZ outliers		2.4%

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	370	2% ... 84% ... 5% 11%
1	B	370	2% ... 84% ... 6% 11%
2	C	2	50% ... 50%
3	D	5	40% ... 60%
3	F	5	20% ... 80%

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	E	6	

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 6574 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 N-acetyllactosaminide beta-1,3-N-acetylglucosaminyltransferase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	331	Total 2793	C 1795	N 479	O 505	S 14	0	9	0
1	B	331	Total 2832	C 1819	N 485	O 513	S 15	0	13	0

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	28	MET	-	initiating methionine	UNP Q9NY97
A	29	HIS	-	expression tag	UNP Q9NY97
A	30	HIS	-	expression tag	UNP Q9NY97
A	31	HIS	-	expression tag	UNP Q9NY97
A	32	HIS	-	expression tag	UNP Q9NY97
A	33	HIS	-	expression tag	UNP Q9NY97
A	34	HIS	-	expression tag	UNP Q9NY97
A	35	HIS	-	expression tag	UNP Q9NY97
A	36	HIS	-	expression tag	UNP Q9NY97
A	37	GLU	-	expression tag	UNP Q9NY97
A	38	ASN	-	expression tag	UNP Q9NY97
A	39	LEU	-	expression tag	UNP Q9NY97
A	40	TYR	-	expression tag	UNP Q9NY97
A	41	PHE	-	expression tag	UNP Q9NY97
A	42	GLN	-	expression tag	UNP Q9NY97
A	43	GLY	-	expression tag	UNP Q9NY97
A	44	SER	-	expression tag	UNP Q9NY97
B	28	MET	-	initiating methionine	UNP Q9NY97
B	29	HIS	-	expression tag	UNP Q9NY97
B	30	HIS	-	expression tag	UNP Q9NY97
B	31	HIS	-	expression tag	UNP Q9NY97
B	32	HIS	-	expression tag	UNP Q9NY97
B	33	HIS	-	expression tag	UNP Q9NY97
B	34	HIS	-	expression tag	UNP Q9NY97

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	35	HIS	-	expression tag	UNP Q9NY97
B	36	HIS	-	expression tag	UNP Q9NY97
B	37	GLU	-	expression tag	UNP Q9NY97
B	38	ASN	-	expression tag	UNP Q9NY97
B	39	LEU	-	expression tag	UNP Q9NY97
B	40	TYR	-	expression tag	UNP Q9NY97
B	41	PHE	-	expression tag	UNP Q9NY97
B	42	GLN	-	expression tag	UNP Q9NY97
B	43	GLY	-	expression tag	UNP Q9NY97
B	44	SER	-	expression tag	UNP Q9NY97

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-a cetamido-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 alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyran ose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-a cetamido-2-deoxy-beta-D-glucopyranose.
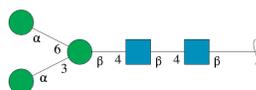
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	D	5	61	34	2	25	0	0	0
3	F	5	61	34	2	25	0	0	0

- Molecule 4 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyran ose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[al pha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.
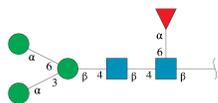
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	E	6	Total 71	C 40	N 2	O 29	0	0	0

- Molecule 5 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: $C_9H_{14}N_2O_{12}P_2$).
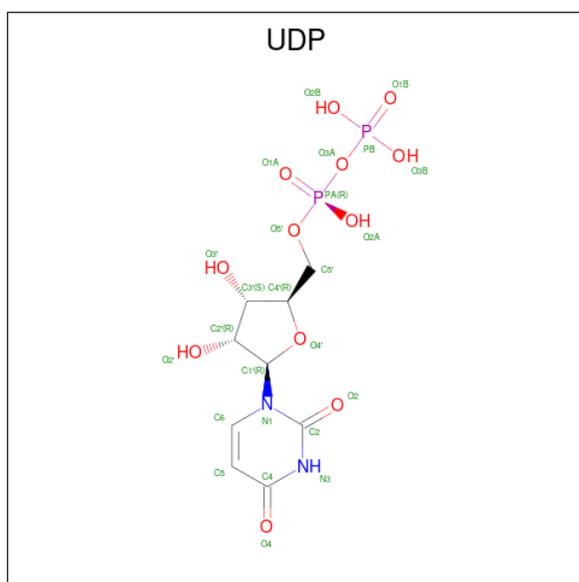
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total 25	C 9	N 2	O 12	P 2	0	0
5	B	1	Total 25	C 9	N 2	O 12	P 2	0	0

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL
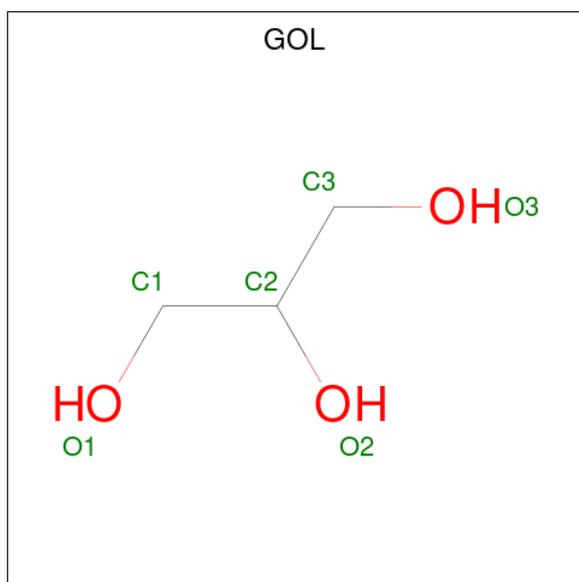
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 6	C 3	O 3	0	0
7	A	1	Total 6	C 3	O 3	0	0
7	B	1	Total 6	C 3	O 3	0	0
7	B	1	Total 6	C 3	O 3	0	0

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

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

- Molecule 9 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

NAG

O6
OH

O1　　　　　O5
HO　　　　　O

C1(R)　　C5(R)　　　　C6

C2(R)　　C4(S)
N2　　　　　C3(R)
HN　　　　　　　OH
O4

C7
OH
C8　　　　O　　　O3
O7
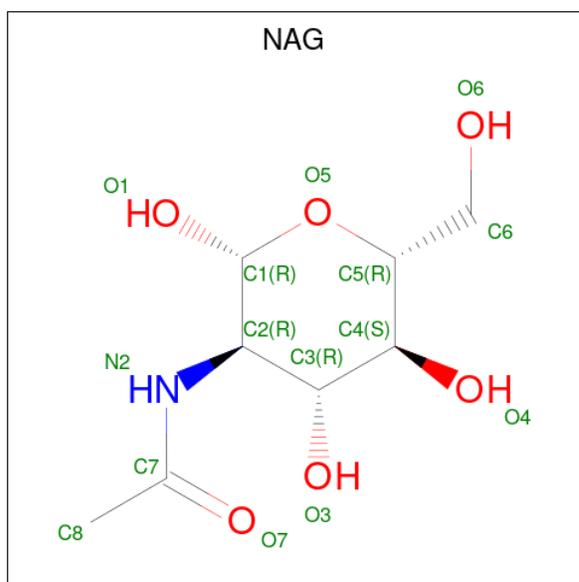
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	335	Total	O	0	0
			335	335		
10	B	301	Total	O	0	0
			301	301		

3 Residue-property plots ⓘ

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: N-acetyllactosaminide beta-1,3-N-acetylglucosaminyltransferase 2
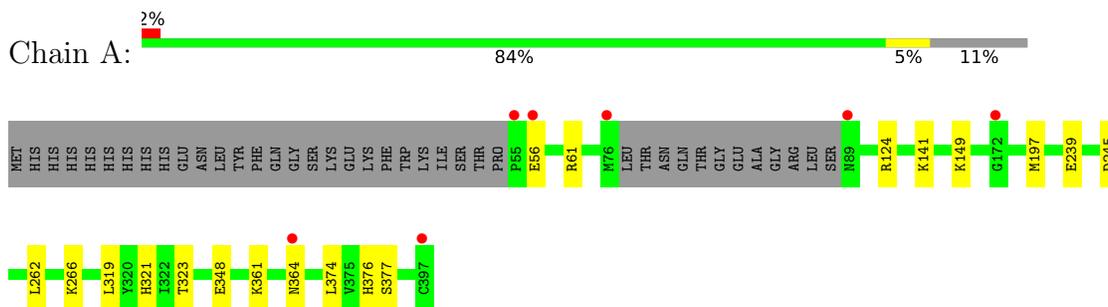
- Molecule 1: N-acetyllactosaminide beta-1,3-N-acetylglucosaminyltransferase 2
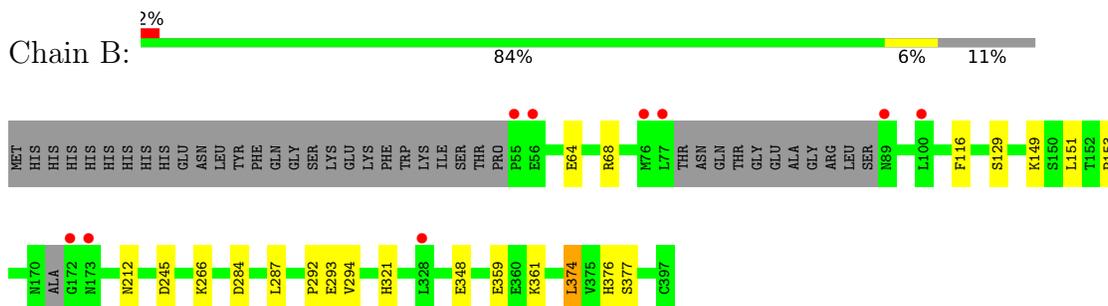
- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

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

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

Chain F:

20%   80%

NAG1 NAG2 BMA3 MAN4 MAN5

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

Chain E:

50%   50%

NAG1 NAG2 BMA3 MAN4 MAN5 FUC6

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	137.65Å 76.51Å 82.70Å 90.00° 109.93° 90.00°	Depositor
Resolution (Å)	29.22 – 1.85 29.22 – 1.85	Depositor EDS
% Data completeness (in resolution range)	93.8 (29.22-1.85) 93.8 (29.22-1.85)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.53 (at 1.85Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.165 , 0.211 0.165 , 0.212	Depositor DCC
R_{free} test set	3223 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	18.2	Xtriage
Anisotropy	0.144	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 56.4	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	6574	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 73.10 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.9534e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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: CL, FUC, UDP, GOL, BMA, NAG, MAN, MG

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	0.36	0/2869	0.54	0/3886
1	B	0.36	0/2907	0.54	1/3936 (0.0%)
All	All	0.36	0/5776	0.54	1/7822 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	374	LEU	CA-CB-CG	5.27	127.42	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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	2793	0	2728	14	0
1	B	2832	0	2766	17	0
2	C	28	0	25	0	0
3	D	61	0	52	0	0
3	F	61	0	52	0	0
4	E	71	0	61	0	0
5	A	25	0	11	1	0
5	B	25	0	11	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	12	0	16	1	0
7	B	12	0	16	3	0
8	A	1	0	0	1	0
8	B	1	0	0	0	0
9	B	14	0	13	0	0
10	A	335	0	0	4	1
10	B	301	0	0	3	1
All	All	6574	0	5751	33	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 3.

All (33) 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:64:GLU:O	1:B:68:ARG:HD2	1.85	0.76
1:A:321[B]:HIS:ND1	7:A:404:GOL:O3	2.20	0.70
1:B:64:GLU:O	1:B:68:ARG:CD	2.44	0.65
8:A:405:CL:CL	10:A:705:HOH:O	2.50	0.65
1:A:141:LYS:HZ1	1:A:239:GLU:HG2	1.62	0.61
1:B:293:GLU:OE1	10:B:501:HOH:O	2.16	0.58
1:A:262:LEU:HD22	1:A:266:LYS:HG2	1.86	0.57
1:A:61:ARG:NH1	10:A:501:HOH:O	2.21	0.57
1:A:141:LYS:NZ	1:A:239:GLU:HG2	2.23	0.54
1:A:266:LYS:NZ	1:A:348:GLU:OE1	2.31	0.52
1:B:321[B]:HIS:HD2	7:B:405:GOL:H32	1.73	0.52
1:A:56:GLU:N	1:A:56:GLU:OE1	2.43	0.51
1:A:245:ASP:CG	5:A:401:UDP:H4'	2.32	0.50
1:A:361:LYS:O	1:A:364:ASN:ND2	2.44	0.50
5:B:402:UDP:H5'2	10:B:685:HOH:O	2.11	0.49
1:B:359:GLU:OE1	1:B:361:LYS:HG2	2.13	0.49
1:B:321[B]:HIS:CD2	7:B:405:GOL:H32	2.50	0.47
1:A:149:LYS:HE3	1:A:245:ASP:OD1	2.16	0.45
1:B:294[B]:VAL:HG23	10:B:501:HOH:O	2.17	0.45
1:B:116:PHE:CZ	1:B:292:PRO:HD3	2.53	0.44
1:B:293:GLU:H	1:B:293:GLU:CD	2.21	0.44
1:A:319:LEU:O	1:A:323:THR:HG23	2.17	0.44
1:B:151:LEU:HB3	1:B:153:PRO:HD2	2.00	0.43
1:B:129:SER:N	1:B:212:ASN:OD1	2.44	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:321[B]:HIS:HD2	7:B:405:GOL:C3	2.32	0.43
1:B:376:HIS:HA	1:B:377:SER:HA	1.76	0.42
1:B:149:LYS:HE3	1:B:245:ASP:OD1	2.19	0.42
1:A:124:ARG:NE	10:A:519:HOH:O	2.52	0.42
1:B:64:GLU:OE2	1:B:68:ARG:NE	2.50	0.42
1:A:376:HIS:HA	1:A:377:SER:HA	1.82	0.41
1:A:321[A]:HIS:HE1	10:A:744:HOH:O	2.03	0.40
1:B:266:LYS:NZ	1:B:348:GLU:OE1	2.47	0.40
1:B:284:ASP:HB3	1:B:287:LEU:HG	2.03	0.40

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)
10:A:730:HOH:O	10:B:722:HOH:O[4_544]	2.09	0.11

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	336/370 (91%)	330 (98%)	6 (2%)	0	100	100
1	B	339/370 (92%)	335 (99%)	4 (1%)	0	100	100
All	All	675/740 (91%)	665 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	313/339 (92%)	311 (99%)	2 (1%)	86	83
1	B	319/339 (94%)	318 (100%)	1 (0%)	92	91
All	All	632/678 (93%)	629 (100%)	3 (0%)	88	86

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	197	MET
1	A	374	LEU
1	B	374	LEU

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

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

18 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	2,1	14,14,15	0.52	0	17,19,21	0.63	0
2	NAG	C	2	2	14,14,15	0.93	1 (7%)	17,19,21	1.81	3 (17%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
3	NAG	D	1	3,1	14,14,15	0.46	0	17,19,21	0.49	0
3	NAG	D	2	3	14,14,15	0.71	1 (7%)	17,19,21	0.59	0
3	BMA	D	3	3	11,11,12	0.89	0	15,15,17	0.72	0
3	MAN	D	4	3	11,11,12	1.00	1 (9%)	15,15,17	0.84	0
3	MAN	D	5	3	11,11,12	0.70	0	15,15,17	1.05	1 (6%)
4	NAG	E	1	4,1	14,14,15	0.42	0	17,19,21	0.57	0
4	NAG	E	2	4	14,14,15	0.40	0	17,19,21	0.49	0
4	BMA	E	3	4	11,11,12	0.58	0	15,15,17	1.10	1 (6%)
4	MAN	E	4	4	11,11,12	1.11	1 (9%)	15,15,17	0.93	1 (6%)
4	MAN	E	5	4	11,11,12	0.90	0	15,15,17	1.03	1 (6%)
4	FUC	E	6	4	10,10,11	0.97	0	14,14,16	0.93	0
3	NAG	F	1	3,1	14,14,15	0.57	0	17,19,21	0.55	0
3	NAG	F	2	3	14,14,15	0.81	1 (7%)	17,19,21	0.53	0
3	BMA	F	3	3	11,11,12	0.81	1 (9%)	15,15,17	0.66	0
3	MAN	F	4	3	11,11,12	0.76	0	15,15,17	1.07	1 (6%)
3	MAN	F	5	3	11,11,12	0.75	0	15,15,17	1.05	1 (6%)

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

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

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MAN	F	4	3	-	2/2/19/22	0/1/1/1
3	MAN	F	5	3	-	0/2/19/22	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	2	NAG	O5-C1	3.16	1.48	1.43
4	E	4	MAN	C2-C3	2.84	1.56	1.52
3	F	2	NAG	O5-C1	-2.56	1.39	1.43
3	D	2	NAG	O5-C1	-2.12	1.40	1.43
3	F	3	BMA	O5-C1	-2.11	1.40	1.43
3	D	4	MAN	O5-C1	-2.10	1.40	1.43

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	2	NAG	C1-O5-C5	5.28	119.34	112.19
2	C	2	NAG	C2-N2-C7	3.26	127.55	122.90
2	C	2	NAG	C1-C2-N2	2.92	115.48	110.49
4	E	5	MAN	C1-O5-C5	2.75	115.92	112.19
3	F	4	MAN	C1-O5-C5	2.68	115.83	112.19
3	F	5	MAN	C1-O5-C5	2.64	115.78	112.19
4	E	4	MAN	C1-O5-C5	2.59	115.70	112.19
3	D	5	MAN	C1-O5-C5	2.57	115.67	112.19
4	E	3	BMA	C1-O5-C5	2.23	115.21	112.19

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	2	NAG	C1-C2-N2-C7
2	C	2	NAG	C4-C5-C6-O6
2	C	2	NAG	O5-C5-C6-O6
3	F	4	MAN	O5-C5-C6-O6
3	F	4	MAN	C4-C5-C6-O6
4	E	5	MAN	C4-C5-C6-O6
4	E	5	MAN	O5-C5-C6-O6
3	F	2	NAG	C8-C7-N2-C2
3	F	2	NAG	O7-C7-N2-C2
3	D	4	MAN	O5-C5-C6-O6
3	D	4	MAN	C4-C5-C6-O6

Continued on next page...

Continued from previous page...

| Mol	Chain	Res	Type	Atoms
3	D	5	MAN	C4-C5-C6-O6
3	D	5	MAN	O5-C5-C6-O6
4	E	1	NAG	C3-C2-N2-C7
4	E	1	NAG	C1-C2-N2-C7

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

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain D

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain F

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings
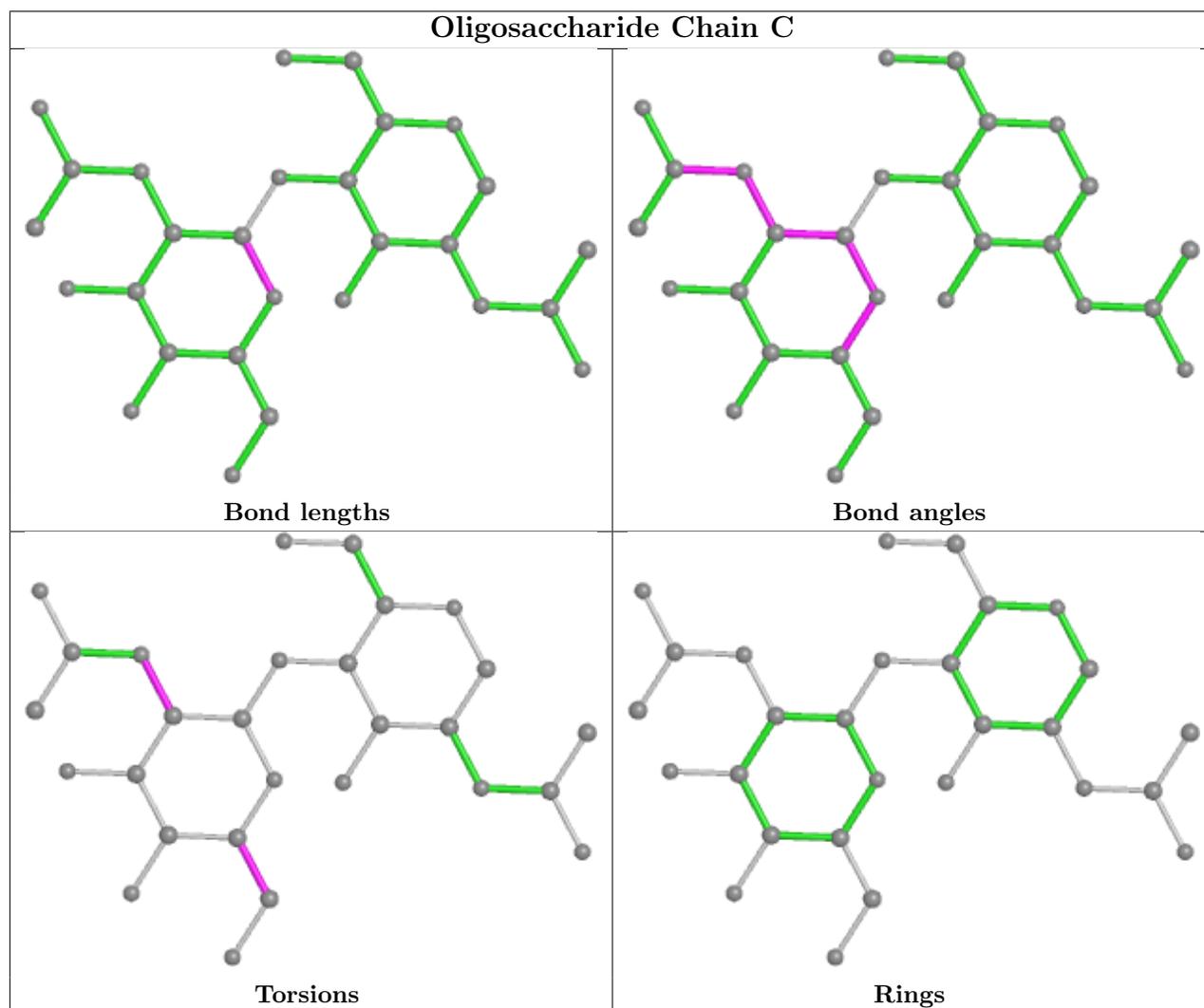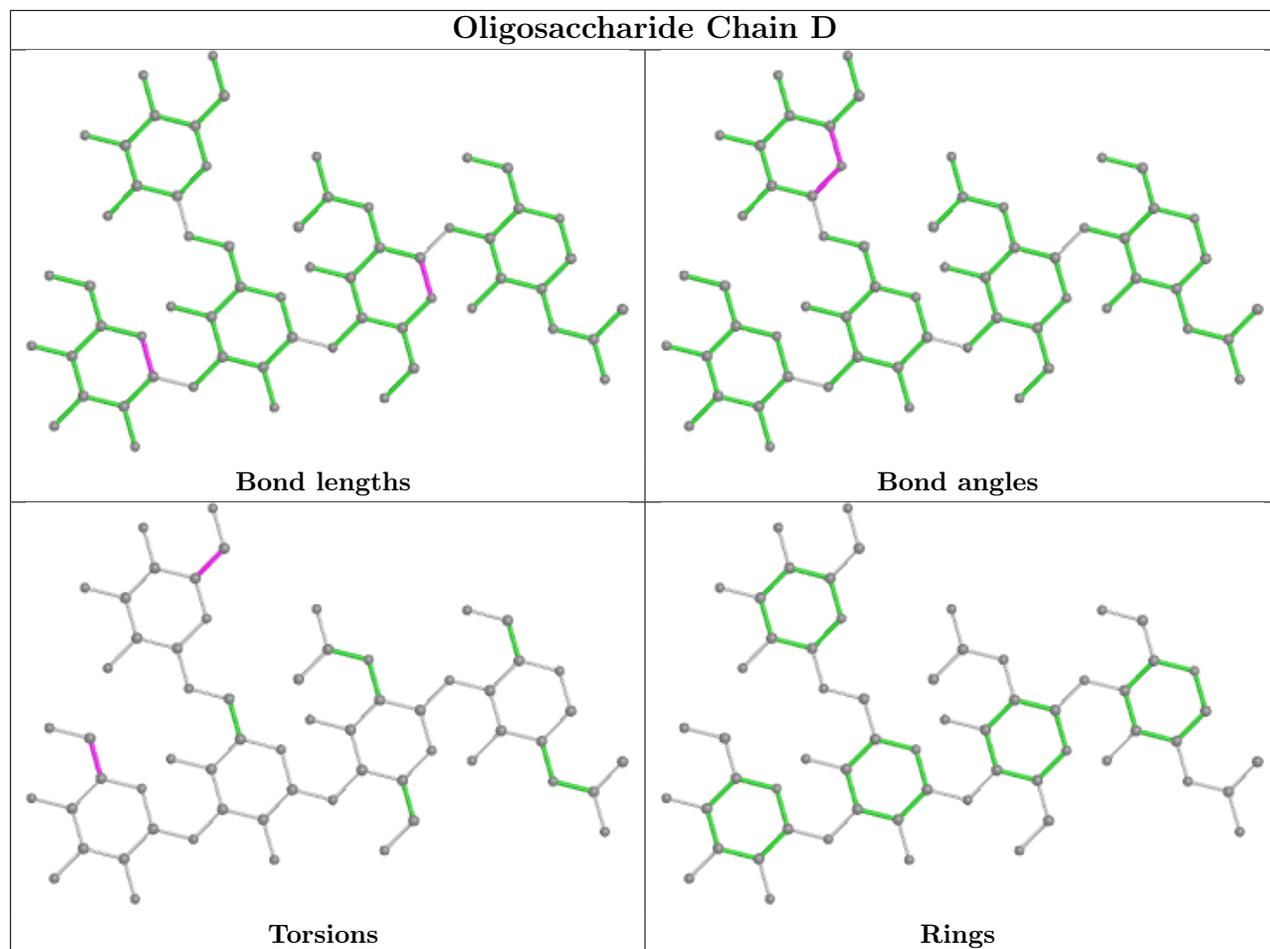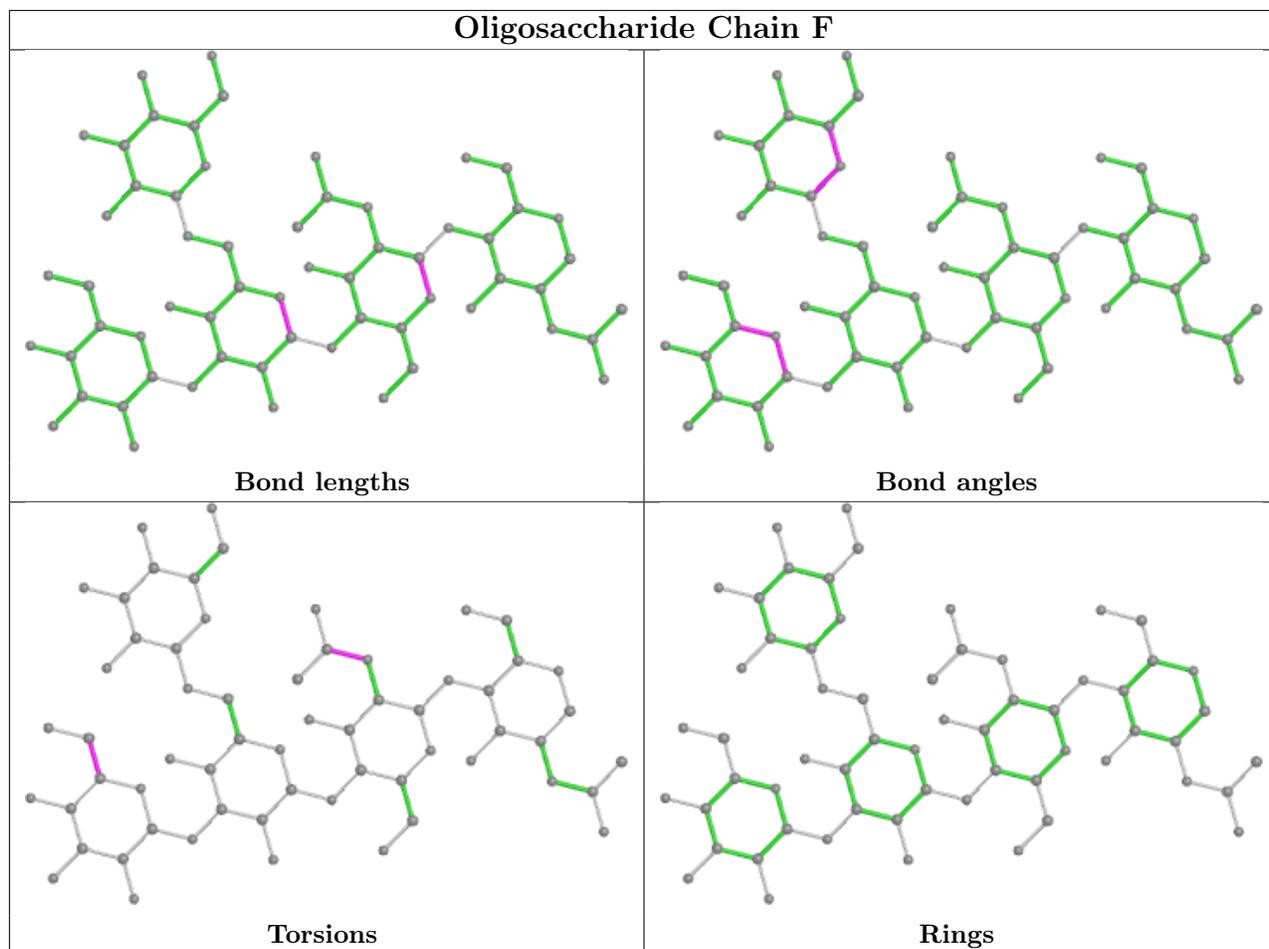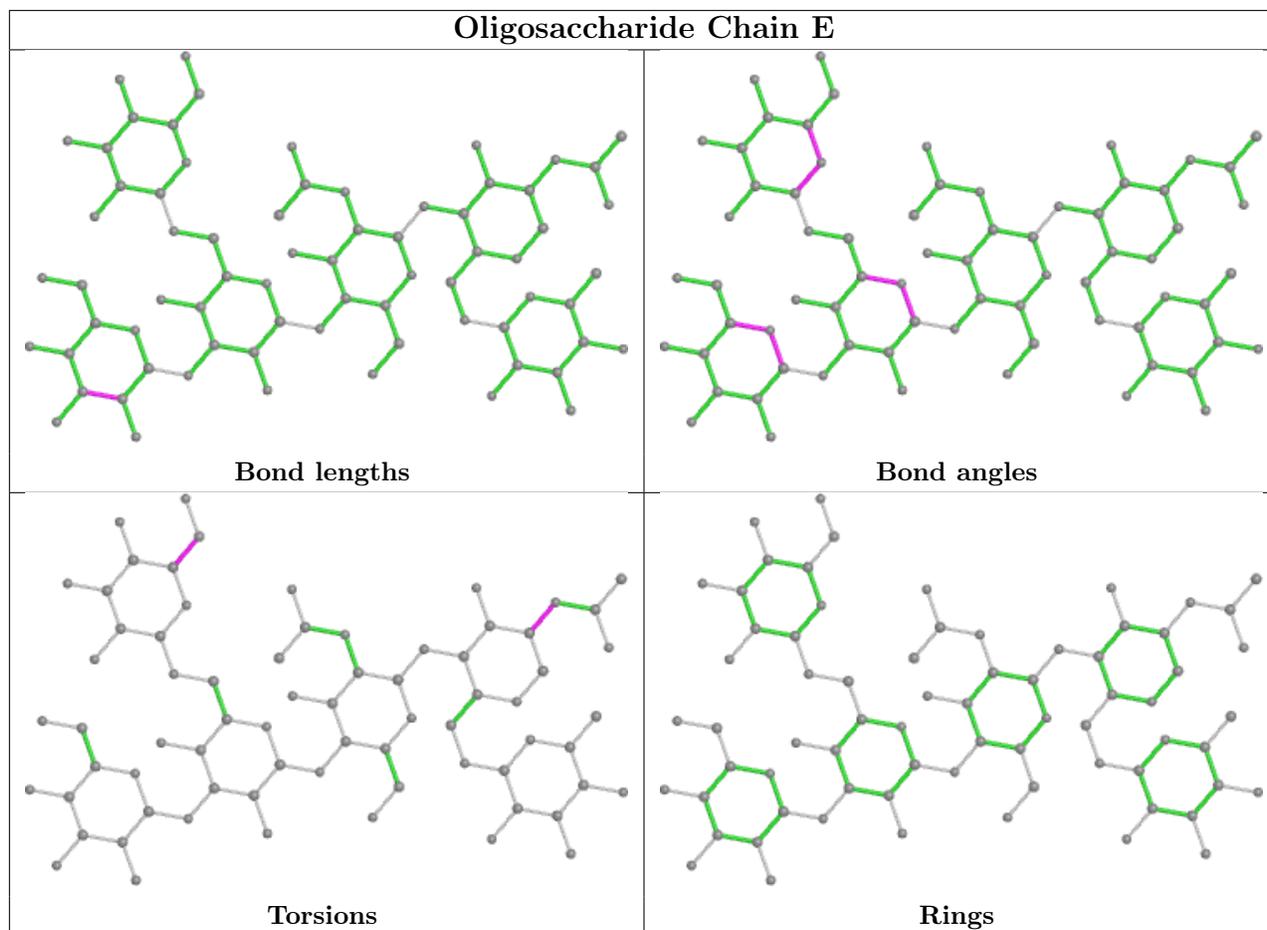
5.6 Ligand geometry ⓘ

Of 11 ligands modelled in this entry, 4 are monoatomic - leaving 7 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	UDP	A	401	6	24,26,26	0.98	1 (4%)	37,40,40	1.87	7 (18%)
7	GOL	B	404	-	5,5,5	0.37	0	5,5,5	0.21	0
7	GOL	A	403	-	5,5,5	0.36	0	5,5,5	0.21	0
5	UDP	B	402	6	24,26,26	0.94	0	37,40,40	1.79	6 (16%)
9	NAG	B	401	1	14,14,15	0.48	0	17,19,21	0.44	0
7	GOL	A	404	-	5,5,5	0.35	0	5,5,5	0.57	0
7	GOL	B	405	-	5,5,5	0.36	0	5,5,5	0.29	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
5	UDP	A	401	6	-	1/16/32/32	0/2/2/2
7	GOL	B	404	-	-	4/4/4/4	-
7	GOL	A	403	-	-	2/4/4/4	-
5	UDP	B	402	6	-	1/16/32/32	0/2/2/2
9	NAG	B	401	1	-	2/6/23/26	0/1/1/1
7	GOL	A	404	-	-	0/4/4/4	-
7	GOL	B	405	-	-	4/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	401	UDP	C6-C5	2.11	1.39	1.35

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	401	UDP	C4-N3-C2	-5.74	119.00	126.58
5	B	402	UDP	C4-N3-C2	-5.72	119.04	126.58
5	A	401	UDP	O2-C2-N1	-4.54	116.75	122.79
5	B	402	UDP	C5-C4-N3	4.03	120.87	114.84
5	B	402	UDP	O2-C2-N1	-3.94	117.55	122.79
5	A	401	UDP	N3-C2-N1	3.86	120.01	114.89
5	A	401	UDP	C5-C4-N3	3.85	120.60	114.84
5	B	402	UDP	N3-C2-N1	3.73	119.84	114.89
5	A	401	UDP	PA-O3A-PB	-2.92	122.81	132.83
5	B	402	UDP	PA-O3A-PB	-2.68	123.62	132.83
5	B	402	UDP	O4-C4-C5	-2.52	120.72	125.16
5	A	401	UDP	C5-C6-N1	-2.41	117.78	121.81
5	A	401	UDP	O4-C4-C5	-2.13	121.41	125.16

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	403	GOL	C1-C2-C3-O3
7	B	404	GOL	C1-C2-C3-O3
7	B	404	GOL	O2-C2-C3-O3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
7	B	405	GOL	O1-C1-C2-C3
7	B	405	GOL	C1-C2-C3-O3
7	A	403	GOL	O2-C2-C3-O3
7	B	405	GOL	O1-C1-C2-O2
7	B	404	GOL	O1-C1-C2-C3
5	A	401	UDP	O4'-C4'-C5'-O5'
7	B	404	GOL	O1-C1-C2-O2
9	B	401	NAG	C4-C5-C6-O6
7	B	405	GOL	O2-C2-C3-O3
9	B	401	NAG	O5-C5-C6-O6
5	B	402	UDP	C5'-O5'-PA-O3A

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	401	UDP	1	0
5	B	402	UDP	1	0
7	A	404	GOL	1	0
7	B	405	GOL	3	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 UDP A 401

Bond lengths

Bond angles

Torsions

Rings

Ligand UDP B 402

Bond lengths

Bond angles

Torsions

Rings
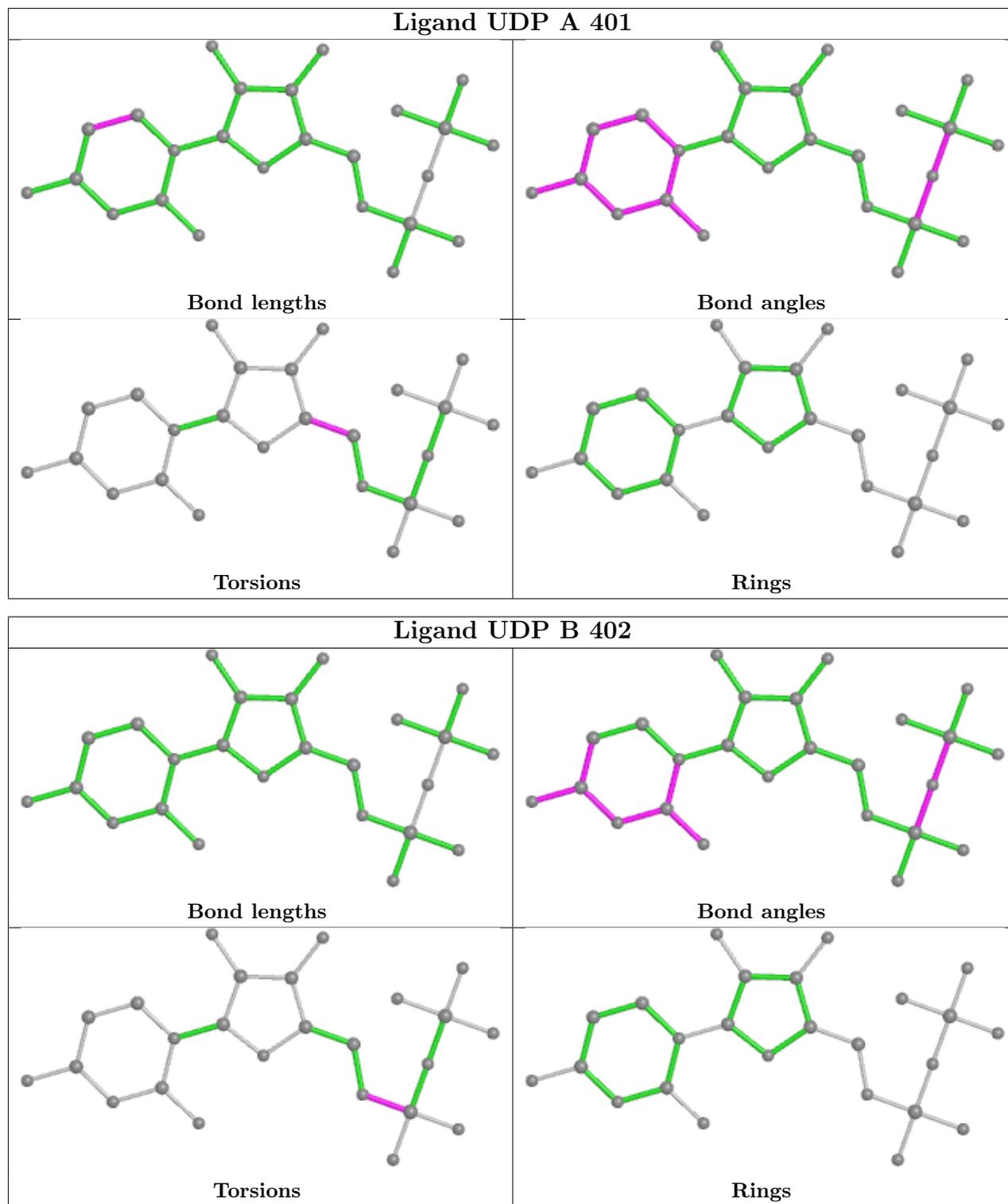
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 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	331/370 (89%)	-0.36	7 (2%)	63	63	10, 20, 40, 73	0
1	B	331/370 (89%)	-0.27	9 (2%)	54	53	10, 21, 45, 73	0
All	All	662/740 (89%)	-0.31	16 (2%)	59	57	10, 21, 42, 73	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	77	LEU	6.3
1	B	89	ASN	4.2
1	B	100	LEU	4.0
1	B	172	GLY	4.0
1	A	55	PRO	3.7
1	A	397	CYS	3.1
1	B	76	MET	2.9
1	B	173	ASN	2.8
1	A	76	MET	2.6
1	B	328	LEU	2.6
1	B	55	PRO	2.4
1	A	172	GLY	2.4
1	B	56	GLU	2.3
1	A	364	ASN	2.1
1	A	56	GLU	2.1
1	A	89	ASN	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.
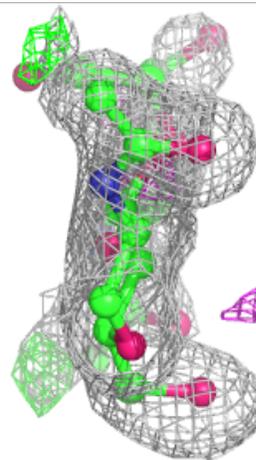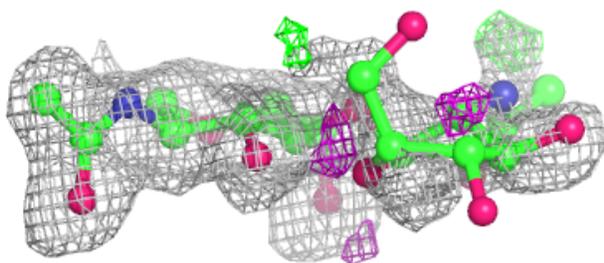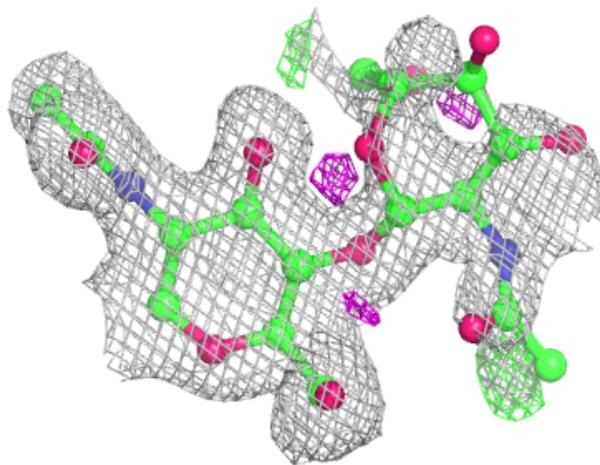
6.3 Carbohydrates (i)

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	MAN	E	4	11/12	0.76	0.35	43,54,77,78	0
4	MAN	E	5	11/12	0.78	0.37	36,63,73,77	0
3	MAN	D	5	11/12	0.79	0.26	50,60,71,72	0
3	MAN	F	4	11/12	0.80	0.26	46,50,69,79	0
2	NAG	C	2	14/15	0.81	0.44	49,69,78,84	0
4	FUC	E	6	10/11	0.82	0.27	36,60,68,71	0
3	MAN	D	4	11/12	0.83	0.31	44,65,72,84	0
3	MAN	F	5	11/12	0.88	0.31	57,59,65,75	0
4	NAG	E	2	14/15	0.90	0.28	28,40,54,60	0
4	NAG	E	1	14/15	0.90	0.27	39,46,60,67	0
4	BMA	E	3	11/12	0.91	0.30	28,57,62,63	0
3	BMA	D	3	11/12	0.93	0.16	28,42,48,49	0
3	BMA	F	3	11/12	0.93	0.20	31,42,56,59	0
3	NAG	F	2	14/15	0.94	0.08	16,27,36,36	0
2	NAG	C	1	14/15	0.94	0.14	25,34,41,51	0
3	NAG	D	2	14/15	0.95	0.07	15,20,32,33	0
3	NAG	D	1	14/15	0.97	0.08	10,15,18,21	0
3	NAG	F	1	14/15	0.97	0.08	13,18,20,21	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.
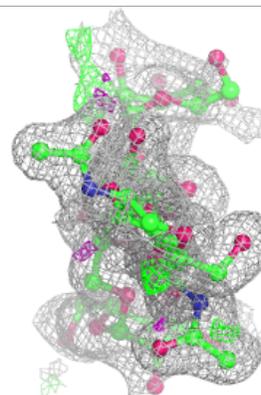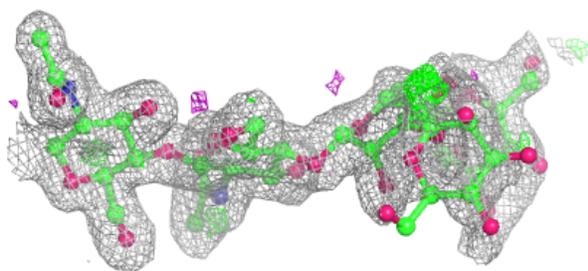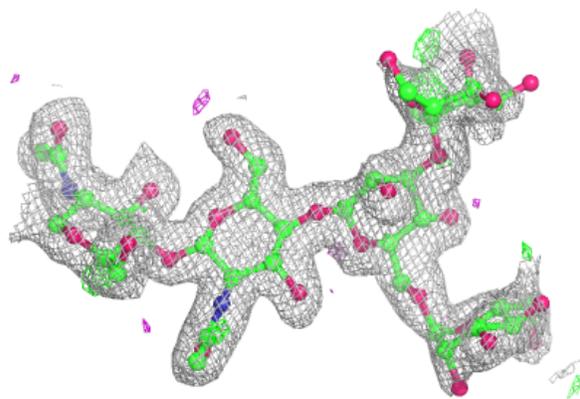
Electron density around Chain C:

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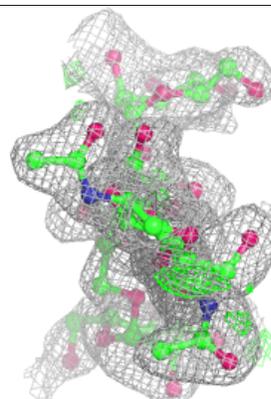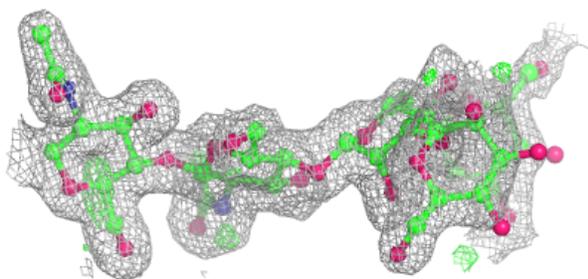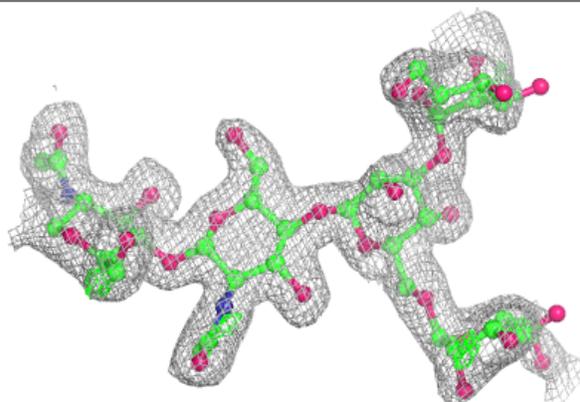and green (positive)

Electron density around Chain D:

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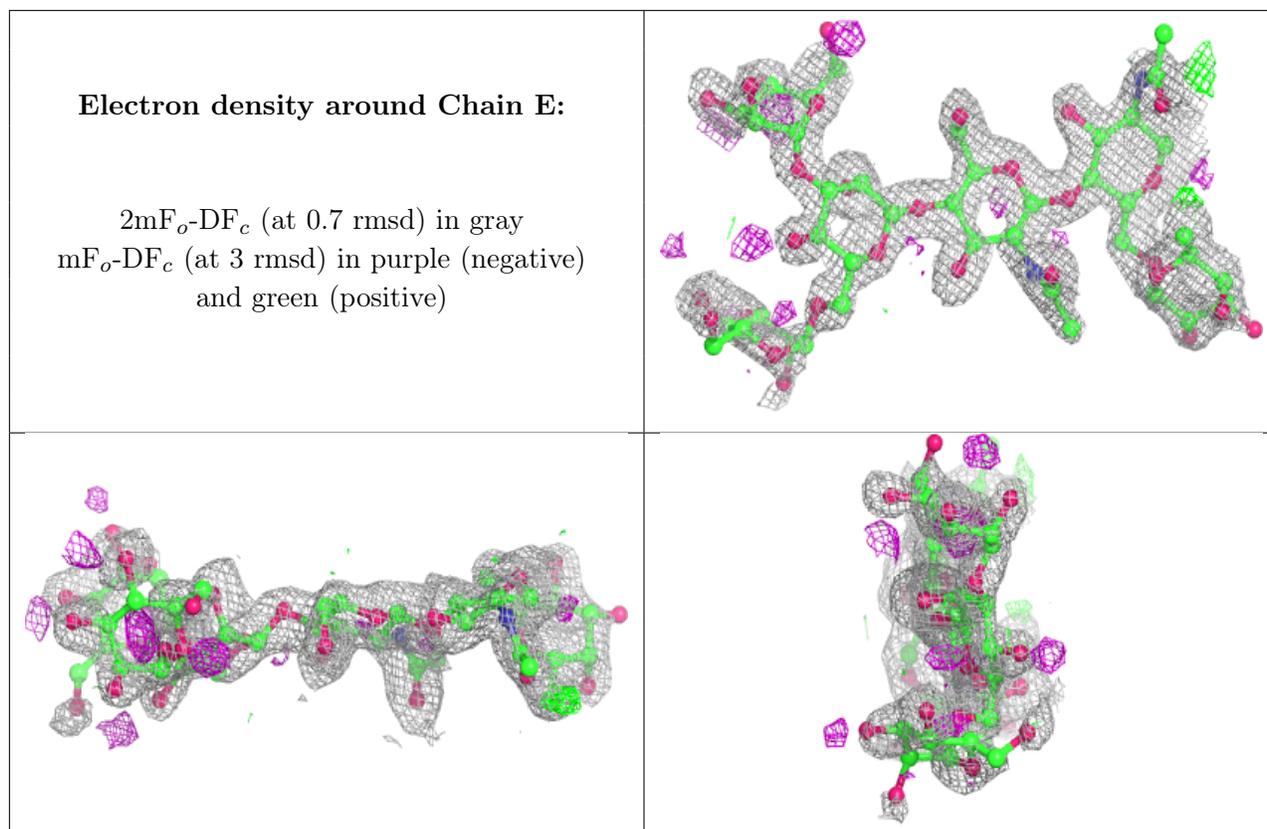and green (positive)

Electron density around Chain F:

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

Electron density around Chain E:

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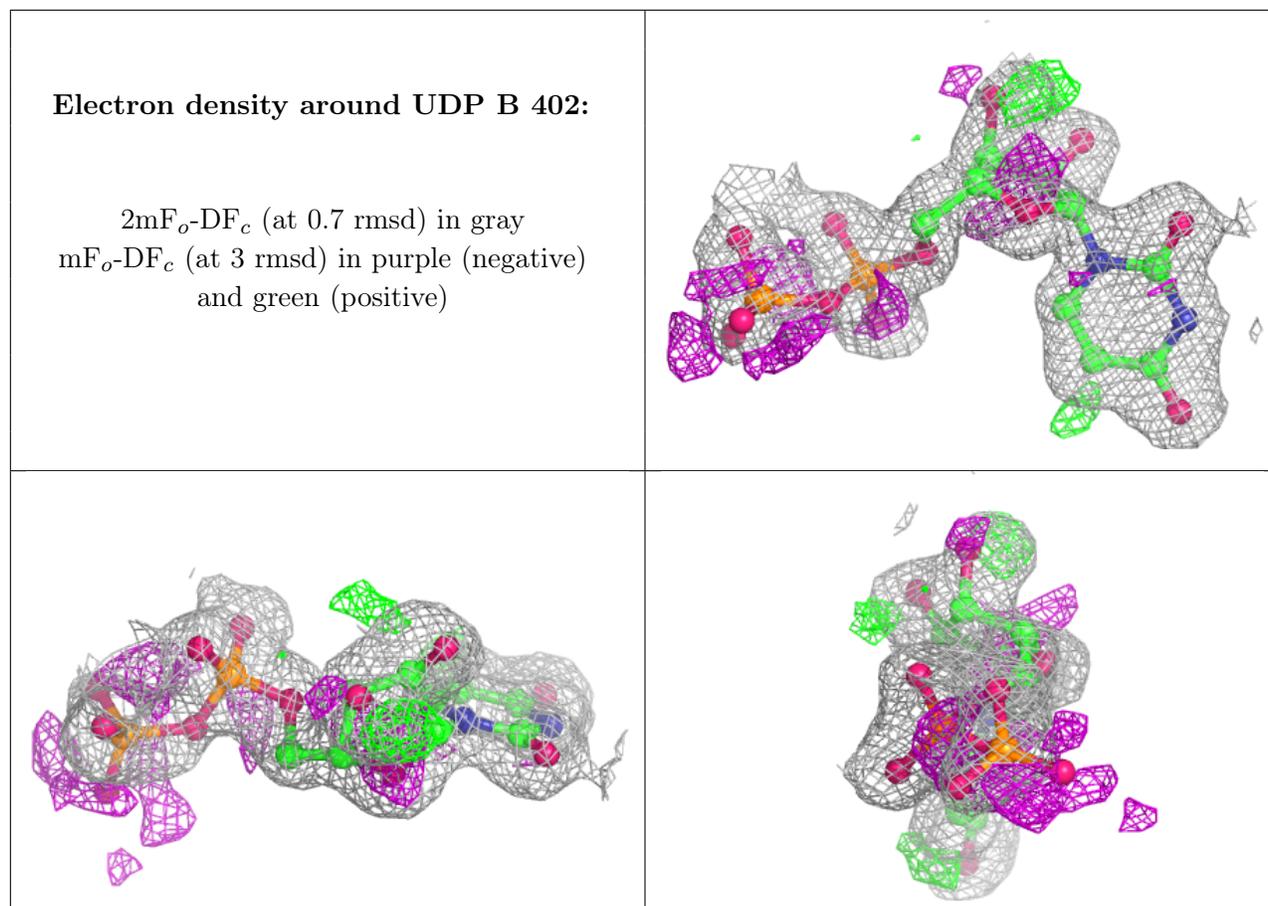mF_o-DF_c (at 3 rmsd) in purple (negative)
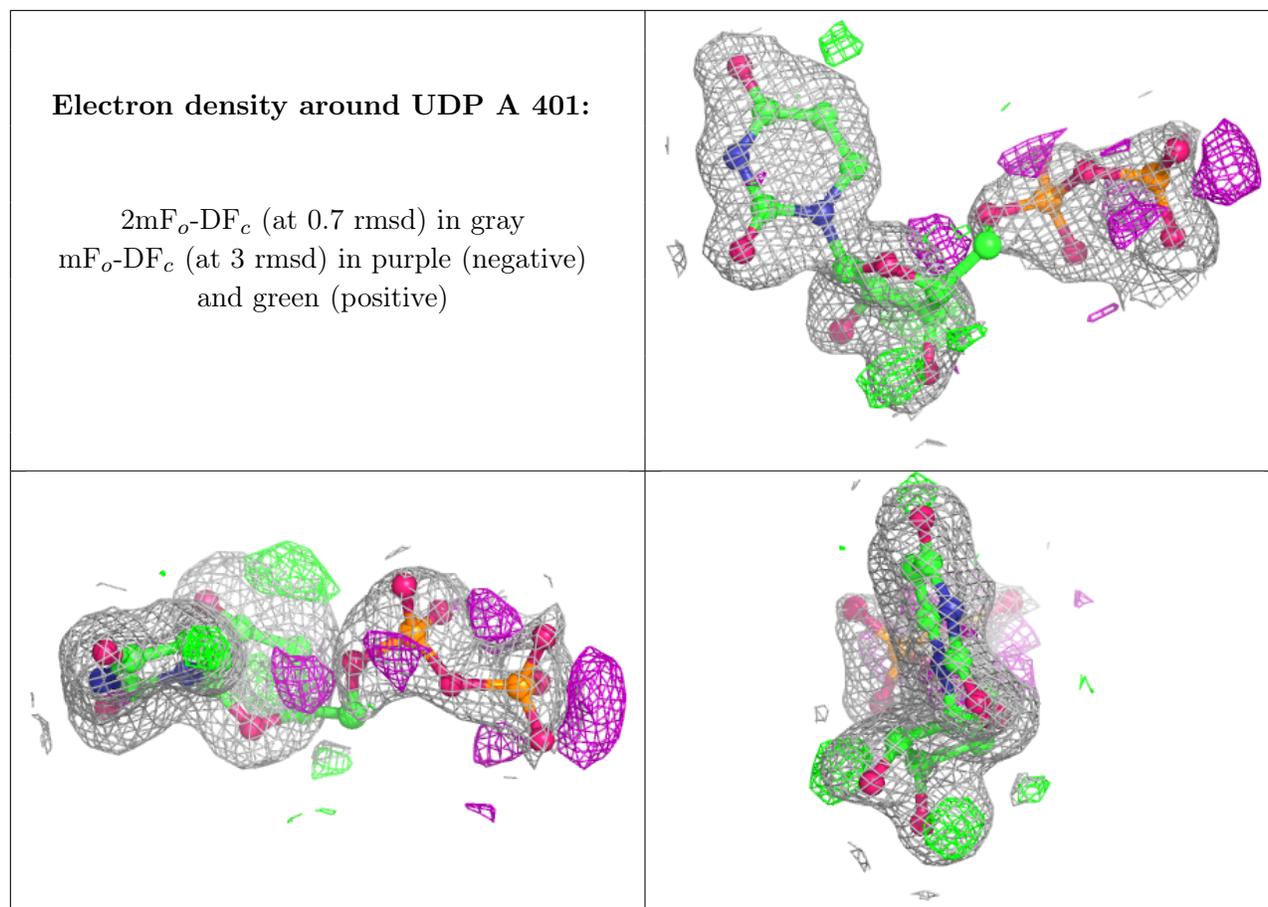and green (positive)

6.4 Ligands ⓘ

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
5	UDP	B	402	25/25	0.86	0.16	21,44,75,87	0
7	GOL	A	403	6/6	0.86	0.14	33,48,54,57	0
9	NAG	B	401	14/15	0.87	0.15	34,38,53,57	0
5	UDP	A	401	25/25	0.89	0.14	18,46,80,118	0
7	GOL	B	405	6/6	0.90	0.19	26,36,55,60	0
7	GOL	B	404	6/6	0.90	0.17	30,52,53,57	0
6	MG	A	402	1/1	0.93	0.23	43,43,43,43	0
7	GOL	A	404	6/6	0.94	0.20	27,33,44,59	0
6	MG	B	403	1/1	0.95	0.15	36,36,36,36	0
8	CL	A	405	1/1	0.98	0.04	30,30,30,30	0
8	CL	B	406	1/1	0.99	0.05	28,28,28,28	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 UDP B 402:

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

Electron density around UDP A 401:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.