Full wwPDB X-ray Structure Validation Report (i)

Sep 6, 2023 – 02:02 AM EDT

PDB ID	:	4EEG
Title	:	Crystal structure of human M340H-beta-1,4-galactosyltransferase-1 (M340H-B4GAL-T1) in complex with GLCNAC-BETA1,6-Gal-Beta
Authors	:	Ramakrishnan, B.; Qasba, P.K.
Deposited on	:	2012-03-28
Resolution	:	2.20 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:

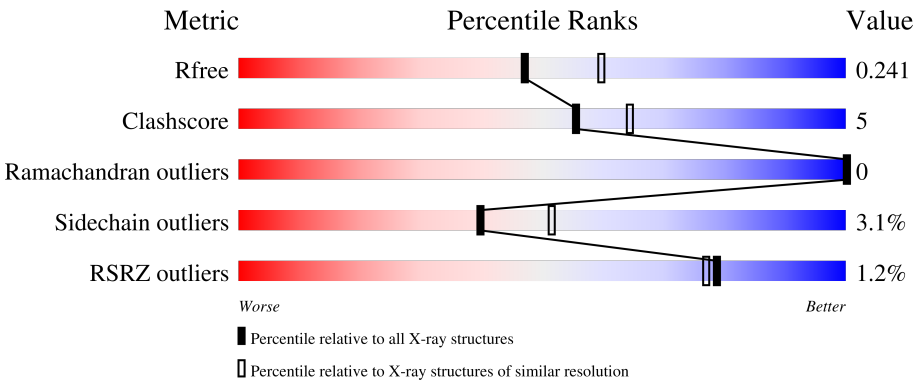
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	287	.% 83% 11% • 5%
1	B	287	85% 9% • 5%
1	C	287	2% 79% 16% 5%
2	D	2	50% 50%
2	E	2	50% 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	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
5	GOL	B	405	-	X	-	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 7417 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 Beta-1,4-galactosyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	273	2229	1427	387	404	11	0	2	0
1	B	273	2223	1423	386	403	11	0	1	0
1	C	273	2218	1420	386	401	11	0	0	0

There are 51 discrepancies between the modelled and reference sequences:

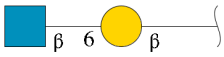
Chain	Residue	Modelled	Actual	Comment	Reference
A	112	ALA	-	expression tag	UNP P15291
A	113	SER	-	expression tag	UNP P15291
A	114	MET	-	expression tag	UNP P15291
A	115	THR	-	expression tag	UNP P15291
A	116	GLY	-	expression tag	UNP P15291
A	117	GLY	-	expression tag	UNP P15291
A	118	GLN	-	expression tag	UNP P15291
A	119	GLN	-	expression tag	UNP P15291
A	120	MET	-	expression tag	UNP P15291
A	121	GLY	-	expression tag	UNP P15291
A	122	ARG	-	expression tag	UNP P15291
A	123	GLY	-	expression tag	UNP P15291
A	124	SER	-	expression tag	UNP P15291
A	125	ALA	-	expression tag	UNP P15291
A	337	THR	ARG	engineered mutation	UNP P15291
A	338	THR	CYS	engineered mutation	UNP P15291
A	340	HIS	MET	engineered mutation	UNP P15291
B	112	ALA	-	expression tag	UNP P15291
B	113	SER	-	expression tag	UNP P15291
B	114	MET	-	expression tag	UNP P15291
B	115	THR	-	expression tag	UNP P15291
B	116	GLY	-	expression tag	UNP P15291
B	117	GLY	-	expression tag	UNP P15291

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	118	GLN	-	expression tag	UNP P15291
B	119	GLN	-	expression tag	UNP P15291
B	120	MET	-	expression tag	UNP P15291
B	121	GLY	-	expression tag	UNP P15291
B	122	ARG	-	expression tag	UNP P15291
B	123	GLY	-	expression tag	UNP P15291
B	124	SER	-	expression tag	UNP P15291
B	125	ALA	-	expression tag	UNP P15291
B	337	THR	ARG	engineered mutation	UNP P15291
B	338	THR	CYS	engineered mutation	UNP P15291
B	340	HIS	MET	engineered mutation	UNP P15291
C	112	ALA	-	expression tag	UNP P15291
C	113	SER	-	expression tag	UNP P15291
C	114	MET	-	expression tag	UNP P15291
C	115	THR	-	expression tag	UNP P15291
C	116	GLY	-	expression tag	UNP P15291
C	117	GLY	-	expression tag	UNP P15291
C	118	GLN	-	expression tag	UNP P15291
C	119	GLN	-	expression tag	UNP P15291
C	120	MET	-	expression tag	UNP P15291
C	121	GLY	-	expression tag	UNP P15291
C	122	ARG	-	expression tag	UNP P15291
C	123	GLY	-	expression tag	UNP P15291
C	124	SER	-	expression tag	UNP P15291
C	125	ALA	-	expression tag	UNP P15291
C	337	THR	ARG	engineered mutation	UNP P15291
C	338	THR	CYS	engineered mutation	UNP P15291
C	340	HIS	MET	engineered mutation	UNP P15291

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-6)-beta-D-galactopyranose.



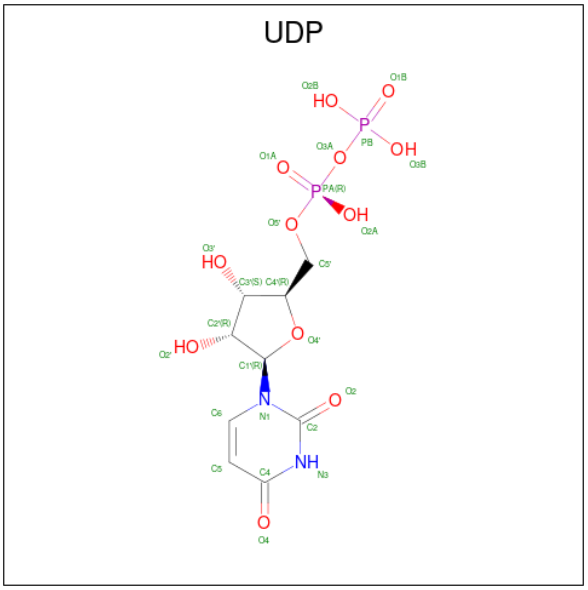
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	D	2	26	14	1	11	0	0	0
			Total	C	N	O			
2	E	2	26	14	1	11	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	F	2	26	14	1	11	0	0	0

- Molecule 3 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: $C_9H_{14}N_2O_{12}P_2$).

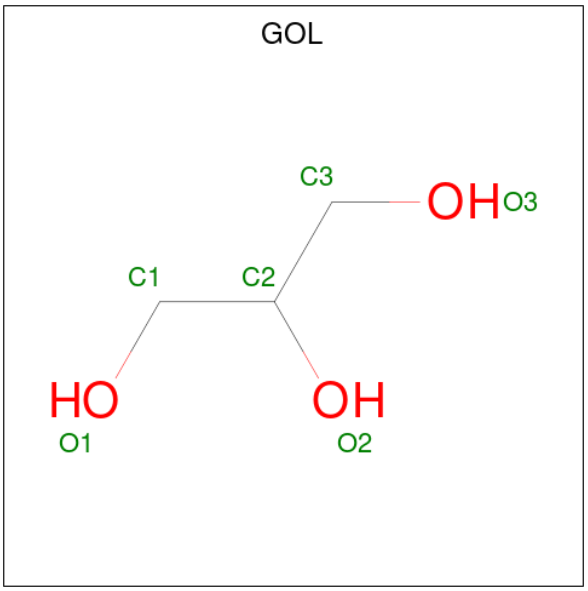


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	25	9	2	12	2	0	0
			Total	C	N	O	P		
3	B	1	25	9	2	12	2	0	0
			Total	C	N	O	P		
3	C	1	25	9	2	12	2	0	0

- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

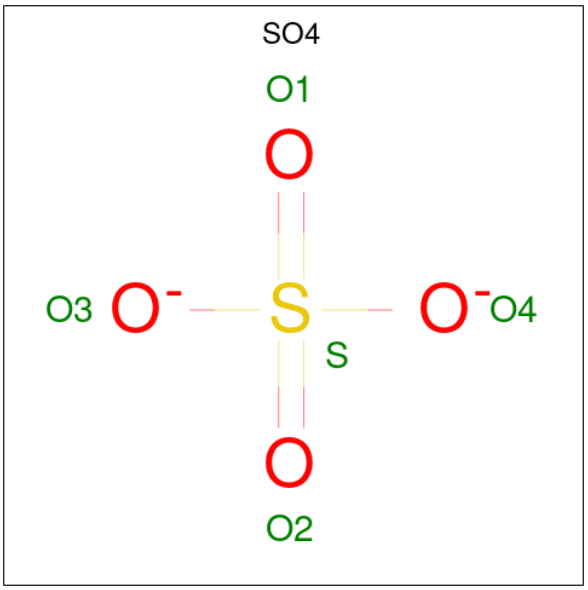
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mn		
4	A	1	1	1	0	0
			Total	Mn		
4	B	1	1	1	0	0
			Total	Mn		
4	C	1	1	1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
6	A	1	5	4	1	0	0
			Total	O	S		
6	A	1	5	4	1	0	0
			Total	O	S		
6	A	1	5	4	1	0	0
			Total	O	S		
6	B	1	5	4	1	0	0
			Total	O	S		
6	B	1	5	4	1	0	0
			Total	O	S		
6	B	1	5	4	1	0	0
			Total	O	S		
6	B	1	5	4	1	0	0
			Total	O	S		
6	B	1	5	4	1	0	0
			Total	O	S		
6	C	1	5	4	1	0	0
			Total	O	S		
6	C	1	5	4	1	0	0

- Molecule 7 is water.

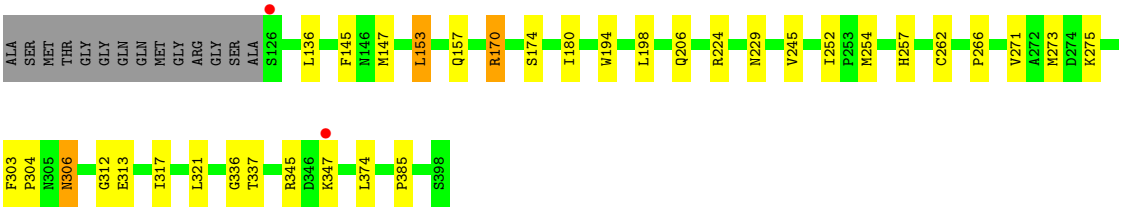
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	189	189	189	0	0
			Total	O		
7	B	202	202	202	0	0
			Total	O		
7	C	126	126	126	0	0

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.

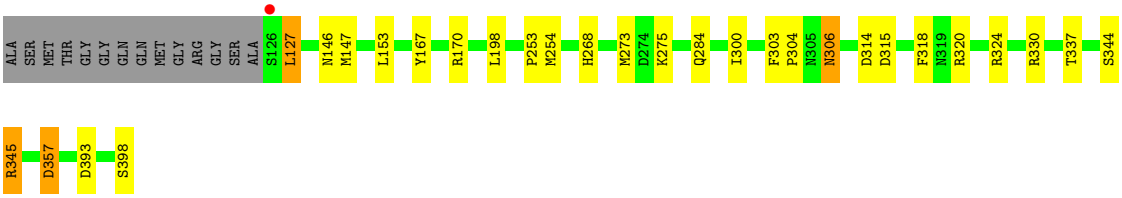
• Molecule 1: Beta-1,4-galactosyltransferase 1

Chain A:



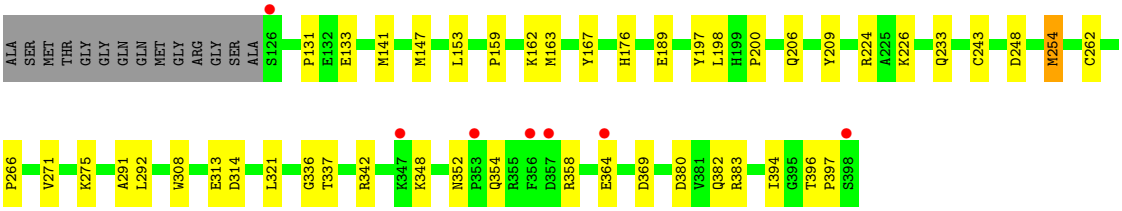
• Molecule 1: Beta-1,4-galactosyltransferase 1

Chain B:



• Molecule 1: Beta-1,4-galactosyltransferase 1

Chain C:



• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-6)-beta-D-galactopyranose

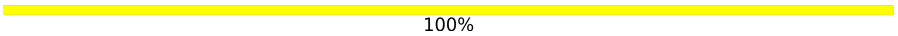
Chain D:

• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-6)-beta-D-galactopyranose

Chain E:

• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-6)-beta-D-galactopyranose

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	107.36Å 195.46Å 143.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.05 – 2.20 47.05 – 2.20	Depositor EDS
% Data completeness (in resolution range)	96.2 (47.05-2.20) 96.2 (47.05-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	3.42 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, R_{free}	0.199 , 0.244 0.198 , 0.241	Depositor DCC
R_{free} test set	3705 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	38.5	Xtriage
Anisotropy	0.020	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 41.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.000 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.014 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7417	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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: UDP, SO4, GAL, MN, NAG, GOL

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	0.83	0/2297	0.84	1/3120 (0.0%)
1	B	0.85	0/2288	0.88	4/3108 (0.1%)
1	C	0.67	0/2280	0.75	0/3097
All	All	0.79	0/6865	0.83	5/9325 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	357	ASP	CB-CG-OD2	-6.47	112.47	118.30
1	A	262	CYS	CA-CB-SG	-5.86	103.46	114.00
1	B	393	ASP	CB-CG-OD2	5.84	123.56	118.30
1	B	315	ASP	CB-CG-OD1	5.50	123.25	118.30
1	B	127	LEU	CA-CB-CG	5.05	126.92	115.30

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	2229	0	2174	24	0
1	B	2223	0	2166	17	0
1	C	2218	0	2162	29	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	26	0	23	1	0
2	E	26	0	24	1	0
2	F	26	0	24	0	0
3	A	25	0	11	0	0
3	B	25	0	11	1	0
3	C	25	0	11	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
5	A	6	0	8	0	0
5	B	12	0	16	2	0
5	C	6	0	8	0	0
6	A	15	0	0	0	0
6	B	25	0	0	2	0
6	C	10	0	0	0	0
7	A	189	0	0	2	0
7	B	202	0	0	4	0
7	C	126	0	0	1	0
All	All	7417	0	6638	69	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 5.

All (69) 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)
5:B:405:GOL:H12	7:B:695:HOH:O	1.83	0.78
1:A:170:ARG:HD2	1:C:342:ARG:HE	1.57	0.69
1:C:266:PRO:HG3	1:C:321:LEU:HD22	1.81	0.63
1:A:312:GLY:HA2	2:D:2:NAG:H83	1.83	0.59
1:B:306:ASN:H	1:B:306:ASN:HD22	1.51	0.58
1:B:146:ASN:HD22	1:C:159:PRO:HG3	1.69	0.57
1:C:271:VAL:HG22	1:C:336:GLY:HA3	1.86	0.57
1:A:306:ASN:H	1:A:306:ASN:HD22	1.58	0.52
1:C:396:THR:HB	1:C:397:PRO:HD2	1.92	0.52
1:A:306:ASN:H	1:A:306:ASN:ND2	2.07	0.52
1:A:266:PRO:HG3	1:A:321:LEU:HD22	1.92	0.51
1:A:174:SER:HB2	1:A:206:GLN:O	2.11	0.51
1:C:254:MET:HG3	1:C:337:THR:HG22	1.94	0.50
1:A:224:ARG:HD3	1:A:313:GLU:OE1	2.11	0.50
1:B:398:SER:HB2	6:B:407:SO4:O1	2.12	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:303:PHE:HB3	1:B:304:PRO:HD2	1.93	0.49
1:A:254:MET:SD	1:A:337:THR:HG22	2.54	0.48
1:B:306:ASN:H	1:B:306:ASN:ND2	2.12	0.48
1:A:271:VAL:HG22	1:A:336:GLY:HA3	1.95	0.48
1:B:254:MET:SD	1:B:337:THR:HG22	2.53	0.48
1:A:306:ASN:HD22	1:A:306:ASN:N	2.12	0.47
1:A:170:ARG:HD3	1:C:342:ARG:HH21	1.79	0.46
1:A:145:PHE:CZ	1:A:252:ILE:HG13	2.51	0.46
1:B:345:ARG:NH1	7:B:508:HOH:O	2.49	0.46
1:B:324:ARG:HA	1:B:324:ARG:HD2	1.73	0.45
7:B:702:HOH:O	2:E:1:GAL:H1	2.16	0.45
1:C:224:ARG:HD3	1:C:313:GLU:OE1	2.16	0.45
1:A:385:PRO:HD2	7:A:594:HOH:O	2.17	0.45
5:B:405:GOL:H31	7:B:503:HOH:O	2.16	0.45
1:C:141:MET:N	1:C:254:MET:O	2.48	0.45
1:B:268:HIS:HB3	1:B:330:ARG:HG2	1.99	0.45
1:B:344:SER:OG	6:B:409:SO4:O1	2.27	0.45
1:C:254:MET:HE3	1:C:254:MET:HB3	1.70	0.45
1:A:273:MET:CE	1:A:275:LYS:HE3	2.47	0.44
1:A:303:PHE:HB3	1:A:304:PRO:HD2	1.98	0.44
1:A:180:ILE:HG13	1:A:245:VAL:HB	1.99	0.44
1:C:243:CYS:SG	1:C:262:CYS:N	2.91	0.44
1:C:248:ASP:HB3	3:C:403:UDP:O3'	2.18	0.44
1:A:306:ASN:ND2	1:A:306:ASN:N	2.66	0.44
1:C:163:MET:HE1	1:C:383:ARG:O	2.18	0.44
1:A:136:LEU:HA	1:C:254:MET:HE1	2.00	0.43
1:C:291:ALA:C	1:C:292:LEU:HD12	2.39	0.43
1:B:167:TYR:CD2	1:B:167:TYR:C	2.92	0.43
1:C:271:VAL:CG2	1:C:336:GLY:HA3	2.48	0.43
1:B:146:ASN:ND2	1:C:159:PRO:HG3	2.32	0.43
1:B:300:ILE:HB	1:B:320:ARG:HB3	2.00	0.43
1:C:314:ASP:N	1:C:314:ASP:OD1	2.52	0.43
1:C:131:PRO:HD2	1:C:206:GLN:HE22	1.84	0.42
1:C:197:TYR:O	1:C:200:PRO:HD2	2.18	0.42
1:C:369:ASP:OD2	1:C:369:ASP:C	2.57	0.42
1:C:167:TYR:HB3	1:C:209:TYR:CE2	2.54	0.42
1:B:284:GLN:HG2	1:B:318:PHE:CZ	2.55	0.42
1:C:198:LEU:HD13	1:C:198:LEU:HA	1.81	0.41
1:B:314:ASP:OD1	1:B:314:ASP:N	2.53	0.41
1:B:198:LEU:HD21	1:B:253:PRO:HD2	2.01	0.41
1:A:229:ASN:HB3	1:A:374:LEU:HD22	2.03	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:313:GLU:O	1:A:317:ILE:HG13	2.20	0.41
1:C:308:TRP:O	1:C:352:ASN:HB2	2.21	0.41
1:C:354:GLN:OE1	1:C:358:ARG:NH2	2.54	0.41
1:A:194:TRP:CZ2	1:A:198:LEU:HG	2.56	0.40
3:B:403:UDP:O2B	3:B:403:UDP:O2A	2.39	0.40
1:C:163:MET:CE	1:C:383:ARG:O	2.69	0.40
1:A:257:HIS:HE1	7:A:601:HOH:O	2.05	0.40
1:C:176:HIS:HA	7:C:562:HOH:O	2.22	0.40
1:A:198:LEU:HD13	1:A:198:LEU:HA	1.97	0.40
1:C:226:LYS:NZ	1:C:394:ILE:O	2.43	0.40
1:A:153:LEU:HG	1:A:157:GLN:HE21	1.87	0.40
1:B:273:MET:CE	1:B:275:LYS:HE3	2.51	0.40
1:C:380:ASP:OD1	1:C:382:GLN:HG2	2.22	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	273/287 (95%)	269 (98%)	4 (2%)	0	100	100
1	B	272/287 (95%)	268 (98%)	4 (2%)	0	100	100
1	C	271/287 (94%)	265 (98%)	6 (2%)	0	100	100
All	All	816/861 (95%)	802 (98%)	14 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	246/252 (98%)	240 (98%)	6 (2%)	49	62
1	B	245/252 (97%)	238 (97%)	7 (3%)	42	54
1	C	244/252 (97%)	234 (96%)	10 (4%)	30	39
All	All	735/756 (97%)	712 (97%)	23 (3%)	40	51

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	147	MET
1	A	153	LEU
1	A	170	ARG
1	A	306	ASN
1	A	345	ARG
1	A	347	LYS
1	B	127	LEU
1	B	147	MET
1	B	153	LEU
1	B	170	ARG
1	B	306	ASN
1	B	345	ARG
1	B	357	ASP
1	C	133	GLU
1	C	147	MET
1	C	153	LEU
1	C	162	LYS
1	C	189	GLU
1	C	233	GLN
1	C	254	MET
1	C	275	LYS
1	C	348	LYS
1	C	364	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	157	GLN
1	A	306	ASN
1	B	186	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	206	GLN
1	B	306	ASN
1	C	157	GLN
1	C	206	GLN
1	C	257	HIS

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)

6 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	GAL	D	1	2	12,12,12	0.51	0	17,17,17	2.34	5 (29%)
2	NAG	D	2	2	14,14,15	0.66	0	17,19,21	1.51	5 (29%)
2	GAL	E	1	2	12,12,12	0.62	0	17,17,17	1.36	1 (5%)
2	NAG	E	2	2	14,14,15	0.73	0	17,19,21	1.38	2 (11%)
2	GAL	F	1	2	12,12,12	0.67	0	17,17,17	1.64	4 (23%)
2	NAG	F	2	2	14,14,15	0.62	0	17,19,21	1.19	2 (11%)

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	GAL	D	1	2	-	1/2/22/22	0/1/1/1
2	NAG	D	2	2	-	2/6/23/26	0/1/1/1
2	GAL	E	1	2	-	0/2/22/22	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1
2	GAL	F	1	2	-	0/2/22/22	0/1/1/1
2	NAG	F	2	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	1	GAL	O5-C5-C6	5.90	121.12	106.44
2	D	1	GAL	C1-O5-C5	-4.75	104.70	113.66
2	F	1	GAL	O6-C6-C5	-4.21	96.84	111.29
2	D	2	NAG	C1-O5-C5	3.27	116.63	112.19
2	D	1	GAL	C6-C5-C4	3.21	120.51	113.00
2	E	1	GAL	O6-C6-C5	-3.12	100.59	111.29
2	D	1	GAL	C3-C4-C5	-2.91	105.04	110.24
2	E	2	NAG	O5-C5-C6	2.90	111.75	107.20
2	F	2	NAG	O7-C7-C8	-2.80	116.85	122.06
2	D	2	NAG	C2-N2-C7	-2.62	119.17	122.90
2	F	1	GAL	O3-C3-C2	2.46	116.03	110.35
2	E	2	NAG	O4-C4-C3	-2.38	104.85	110.35
2	F	2	NAG	O7-C7-N2	2.35	126.27	121.95
2	D	1	GAL	O5-C5-C4	2.30	113.87	109.69
2	D	2	NAG	O5-C5-C6	2.19	110.64	107.20
2	D	2	NAG	O3-C3-C2	-2.14	105.04	109.47
2	F	1	GAL	O1-C1-C2	2.14	115.05	109.03
2	D	2	NAG	C1-C2-N2	2.12	114.12	110.49
2	F	1	GAL	C1-O5-C5	-2.02	109.86	113.66

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	1	GAL	O5-C5-C6-O6
2	D	2	NAG	C4-C5-C6-O6
2	D	2	NAG	O5-C5-C6-O6

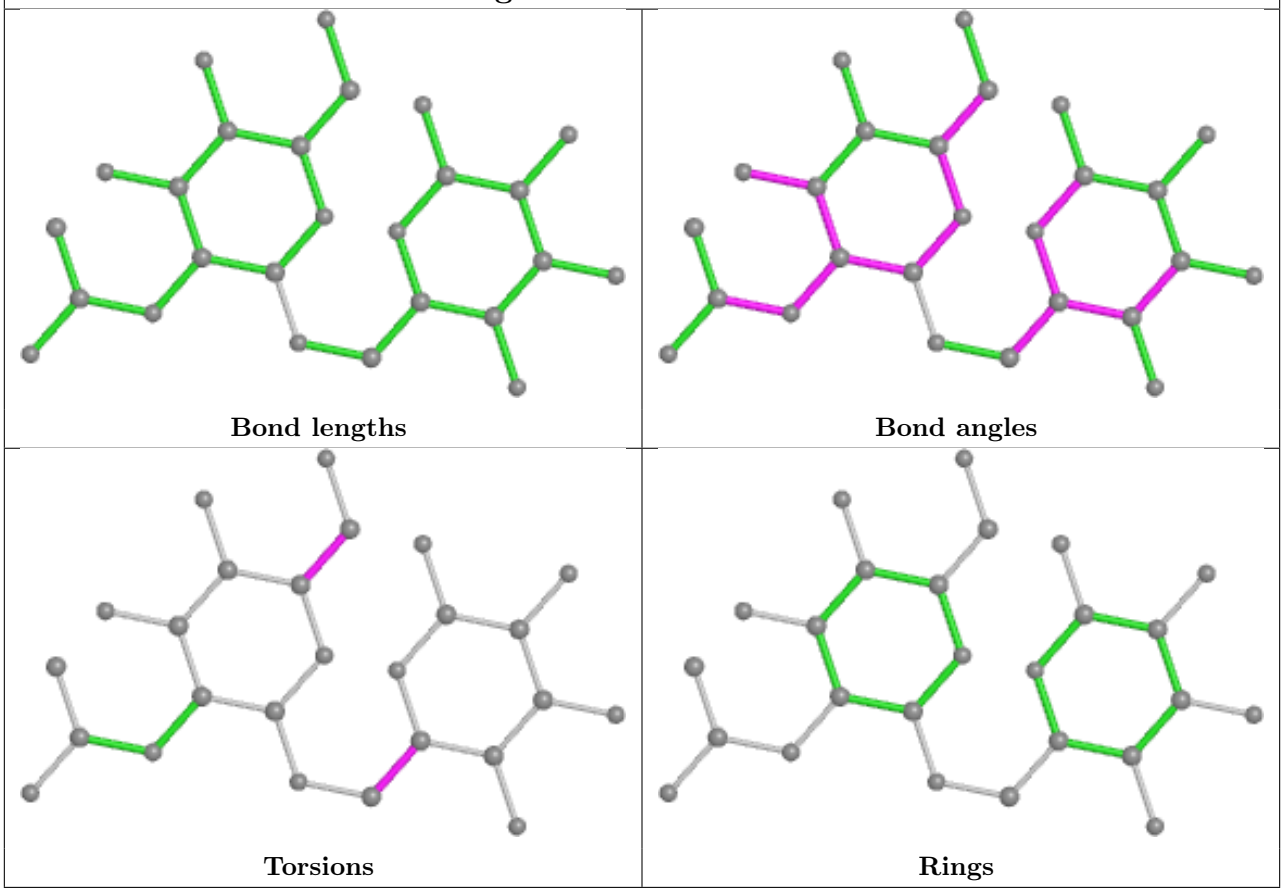
There are no ring outliers.

2 monomers are involved in 2 short contacts:

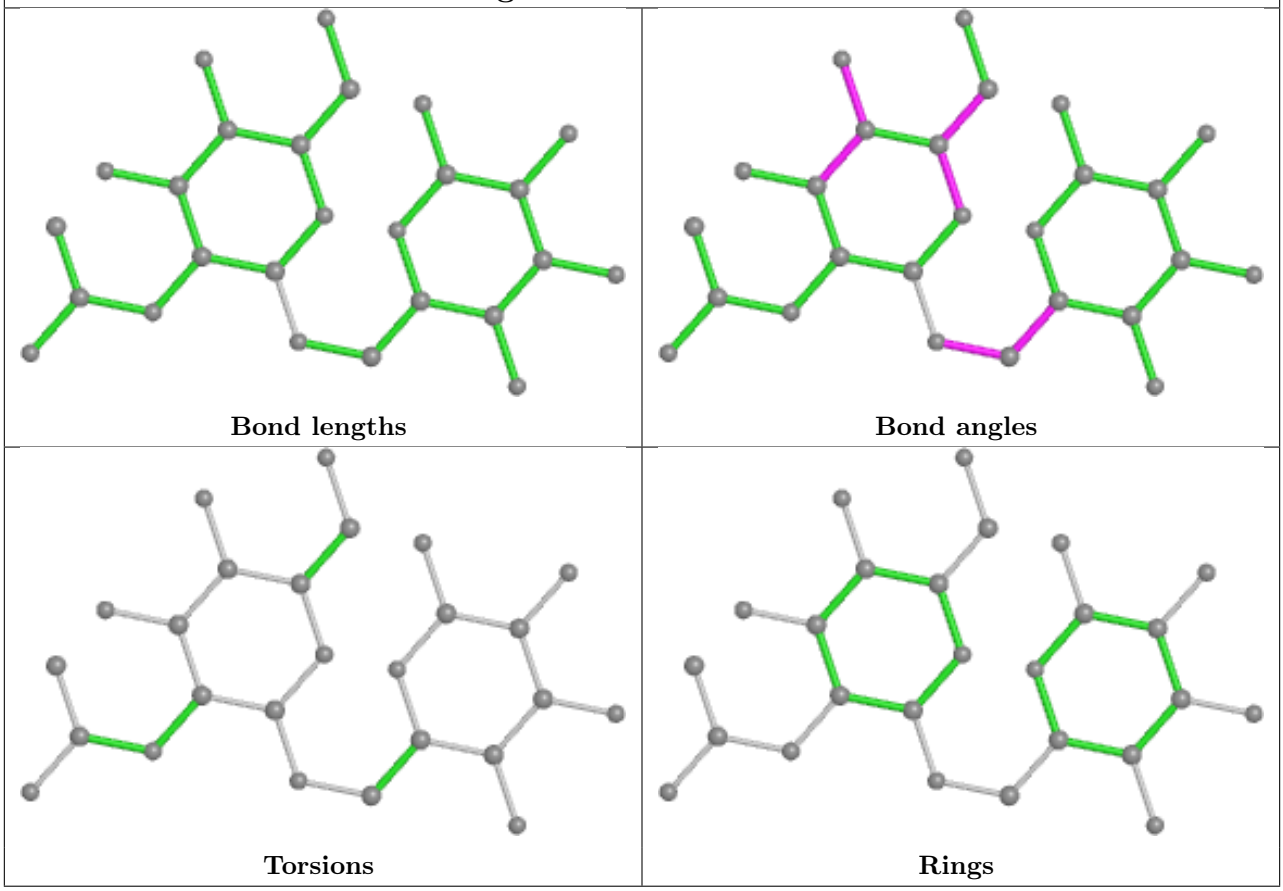
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	1	GAL	1	0
2	D	2	NAG	1	0

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

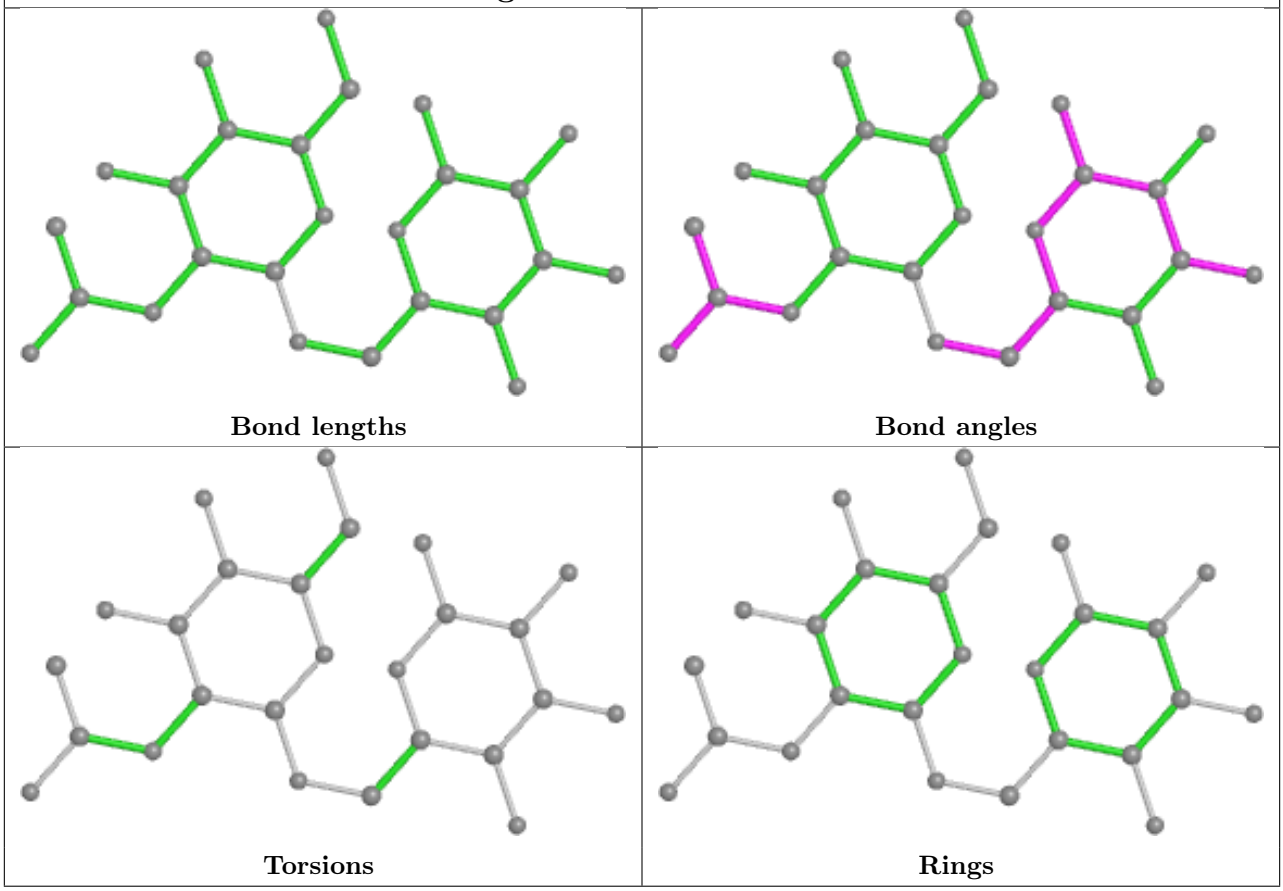
Oligosaccharide Chain D



Oligosaccharide Chain E



Oligosaccharide Chain F



5.6 Ligand geometry (i)

Of 20 ligands modelled in this entry, 3 are monoatomic - leaving 17 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
6	SO4	B	409	-	4,4,4	0.13	0	6,6,6	0.19	0
5	GOL	B	405	-	5,5,5	0.23	0	5,5,5	2.17	3 (60%)
6	SO4	C	406	-	4,4,4	0.30	0	6,6,6	0.63	0
6	SO4	C	407	-	4,4,4	0.13	0	6,6,6	0.46	0
5	GOL	B	406	-	5,5,5	0.32	0	5,5,5	1.40	1 (20%)
6	SO4	B	411	-	4,4,4	0.30	0	6,6,6	0.66	0
6	SO4	B	408	-	4,4,4	0.17	0	6,6,6	0.40	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	GOL	C	405	-	5,5,5	0.43	0	5,5,5	0.81	0
6	SO4	A	407	-	4,4,4	0.27	0	6,6,6	0.67	0
6	SO4	B	407	-	4,4,4	0.16	0	6,6,6	0.47	0
3	UDP	B	403	4	24,26,26	1.18	2 (8%)	37,40,40	2.31	10 (27%)
3	UDP	A	403	4	24,26,26	1.16	1 (4%)	37,40,40	1.92	8 (21%)
3	UDP	C	403	4	24,26,26	0.95	2 (8%)	37,40,40	1.91	9 (24%)
6	SO4	B	410	-	4,4,4	0.23	0	6,6,6	0.58	0
6	SO4	A	408	-	4,4,4	0.22	0	6,6,6	0.60	0
6	SO4	A	406	-	4,4,4	0.21	0	6,6,6	0.51	0
5	GOL	A	405	-	5,5,5	0.44	0	5,5,5	0.97	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	GOL	B	405	-	-	3/4/4/4	-
5	GOL	B	406	-	-	2/4/4/4	-
5	GOL	C	405	-	-	4/4/4/4	-
3	UDP	B	403	4	-	5/16/32/32	0/2/2/2
3	UDP	A	403	4	-	3/16/32/32	0/2/2/2
3	UDP	C	403	4	-	1/16/32/32	0/2/2/2
5	GOL	A	405	-	-	2/4/4/4	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	403	UDP	C2-N1	3.14	1.43	1.38
3	A	403	UDP	C6-C5	2.83	1.41	1.35
3	B	403	UDP	C6-C5	2.59	1.41	1.35
3	C	403	UDP	C5-C4	-2.05	1.39	1.43
3	C	403	UDP	C6-C5	2.02	1.39	1.35

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	403	UDP	C4-N3-C2	-6.99	117.36	126.58
3	A	403	UDP	C5-C4-N3	5.74	123.43	114.84
3	C	403	UDP	C4-N3-C2	-5.67	119.10	126.58

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	403	UDP	N3-C2-N1	5.46	122.14	114.89
3	A	403	UDP	C4-N3-C2	-5.13	119.81	126.58
3	C	403	UDP	N3-C2-N1	4.98	121.50	114.89
3	B	403	UDP	C5-C4-N3	4.89	122.15	114.84
3	B	403	UDP	O4'-C1'-N1	4.17	117.89	108.36
3	A	403	UDP	PA-O3A-PB	-4.03	118.98	132.83
3	B	403	UDP	PA-O3A-PB	-3.89	119.47	132.83
3	A	403	UDP	C6-C5-C4	-3.75	114.39	119.52
3	B	403	UDP	O2-C2-N3	-3.65	114.71	121.50
3	C	403	UDP	C5-C4-N3	3.37	119.89	114.84
3	A	403	UDP	O4-C4-C5	-3.25	119.44	125.16
3	C	403	UDP	PA-O3A-PB	-3.14	122.06	132.83
3	B	403	UDP	O4-C4-C5	-3.11	119.70	125.16
3	C	403	UDP	C6-N1-C2	-3.08	117.05	120.99
3	C	403	UDP	O4-C4-C5	-3.03	119.83	125.16
5	B	405	GOL	O2-C2-C1	2.94	122.09	109.12
3	C	403	UDP	O2-C2-N1	-2.65	119.26	122.79
3	A	403	UDP	O5'-C5'-C4'	2.55	117.76	108.99
5	B	405	GOL	O3-C3-C2	-2.39	98.75	110.20
5	B	405	GOL	C3-C2-C1	-2.38	102.43	111.70
3	B	403	UDP	O3B-PB-O3A	2.36	112.55	104.64
3	A	403	UDP	N3-C2-N1	2.32	117.97	114.89
3	A	403	UDP	O2B-PB-O3A	2.20	112.02	104.64
3	C	403	UDP	O3B-PB-O3A	2.14	111.80	104.64
3	C	403	UDP	O4'-C1'-N1	2.13	113.24	108.36
3	B	403	UDP	O2B-PB-O3A	2.13	111.76	104.64
5	B	406	GOL	O2-C2-C3	2.03	118.06	109.12
3	B	403	UDP	O5'-PA-O1A	2.00	116.89	109.07

There are no chirality outliers.

All (20) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	406	GOL	C1-C2-C3-O3
5	A	405	GOL	C1-C2-C3-O3
5	B	405	GOL	O1-C1-C2-C3
5	C	405	GOL	O1-C1-C2-C3
5	C	405	GOL	C1-C2-C3-O3
5	C	405	GOL	O1-C1-C2-O2
5	B	406	GOL	O2-C2-C3-O3
3	A	403	UDP	PA-O3A-PB-O1B
3	B	403	UDP	PA-O3A-PB-O1B

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	B	405	GOL	O1-C1-C2-O2
5	C	405	GOL	O2-C2-C3-O3
3	C	403	UDP	C4'-C5'-O5'-PA
3	A	403	UDP	C4'-C5'-O5'-PA
3	B	403	UDP	C4'-C5'-O5'-PA
3	B	403	UDP	C3'-C4'-C5'-O5'
3	B	403	UDP	O4'-C4'-C5'-O5'
5	A	405	GOL	O2-C2-C3-O3
5	B	405	GOL	O2-C2-C3-O3
3	A	403	UDP	C5'-O5'-PA-O1A
3	B	403	UDP	C5'-O5'-PA-O1A

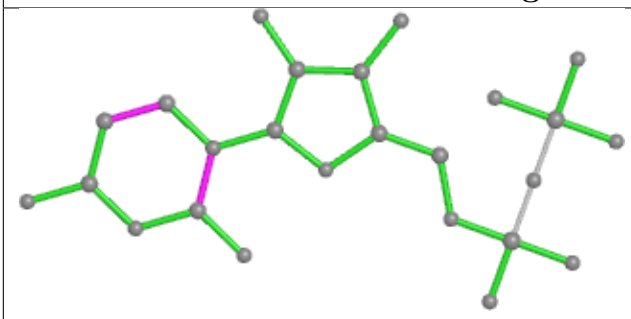
There are no ring outliers.

5 monomers are involved in 6 short contacts:

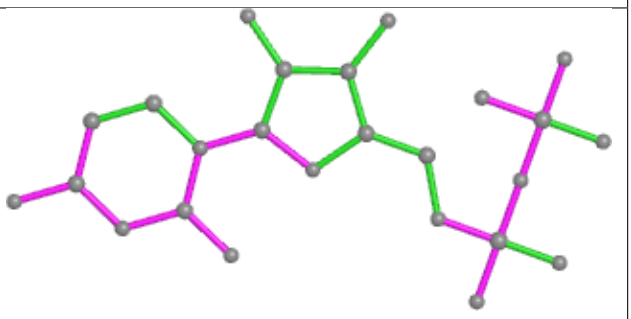
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	409	SO4	1	0
5	B	405	GOL	2	0
6	B	407	SO4	1	0
3	B	403	UDP	1	0
3	C	403	UDP	1	0

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.

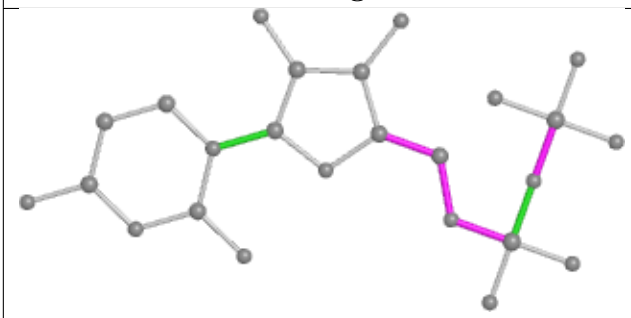
Ligand UDP B 403



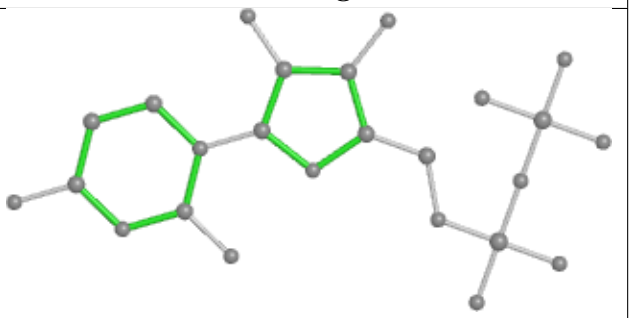
Bond lengths



Bond angles

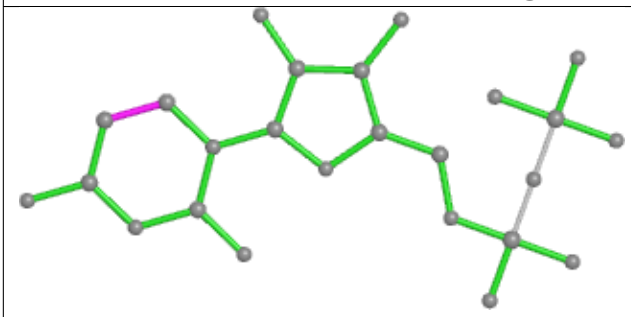


Torsions

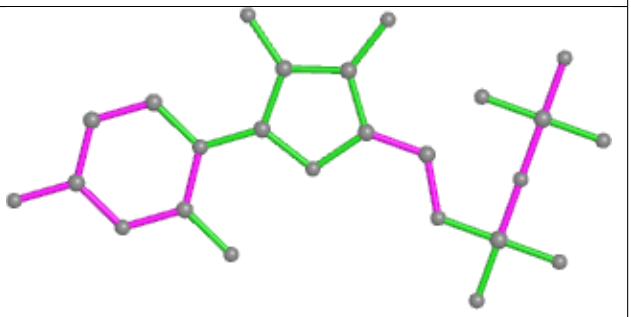


Rings

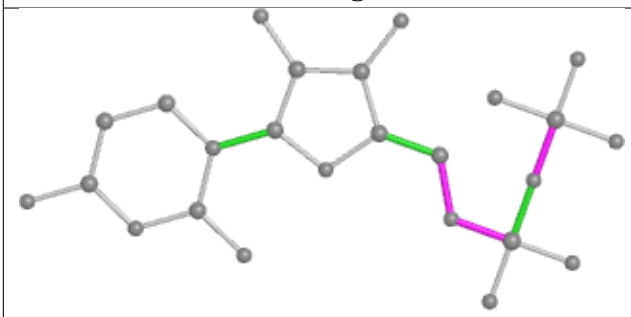
Ligand UDP A 403



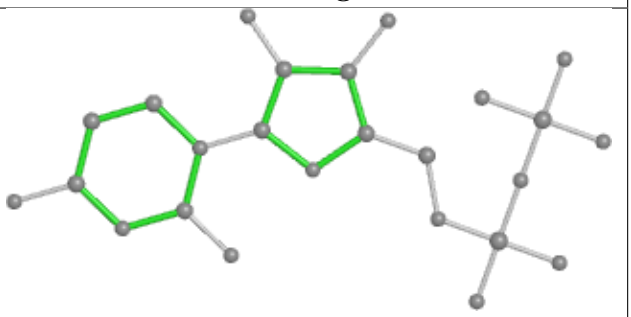
Bond lengths



Bond angles

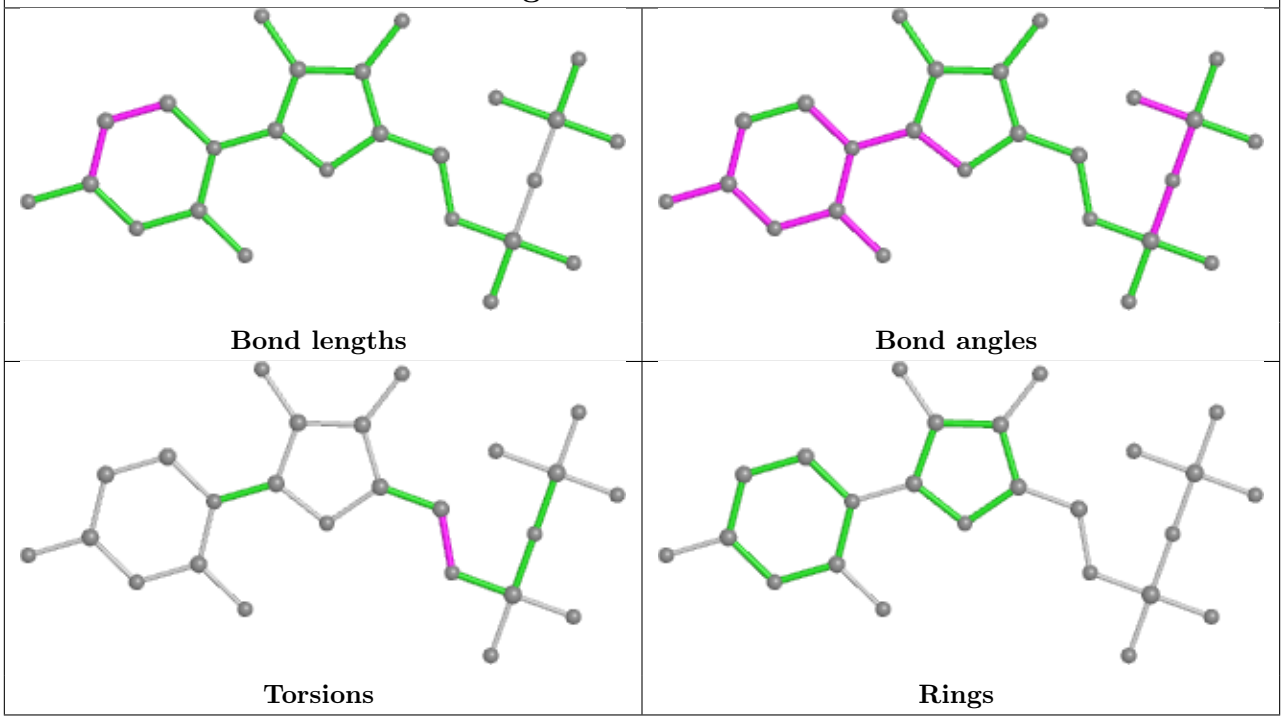


Torsions



Rings

Ligand UDP C 403



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	273/287 (95%)	-0.45	2 (0%)	87	86	24, 33, 53, 68	0
1	B	273/287 (95%)	-0.46	1 (0%)	92	91	22, 30, 46, 63	0
1	C	273/287 (95%)	-0.12	7 (2%)	56	53	30, 42, 68, 79	1 (0%)
All	All	819/861 (95%)	-0.34	10 (1%)	79	77	22, 35, 57, 79	1 (0%)

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	126	SER	3.9
1	C	347	LYS	3.7
1	B	126	SER	3.2
1	C	364	GLU	3.1
1	C	357	ASP	3.0
1	A	347	LYS	2.8
1	C	356	PHE	2.6
1	A	126	SER	2.4
1	C	353	PRO	2.4
1	C	398	SER	2.3

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)

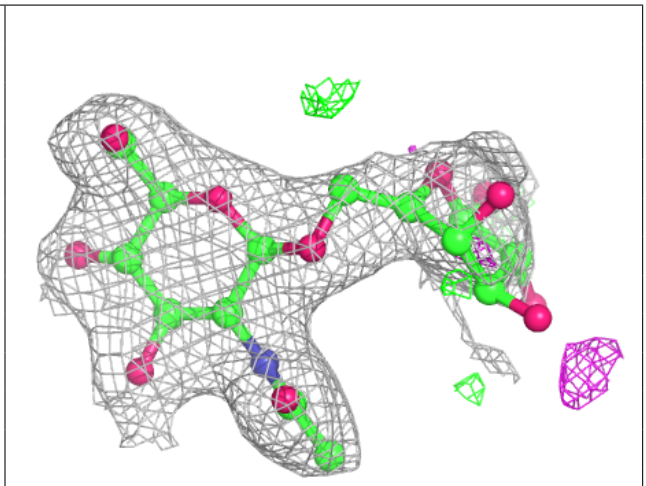
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

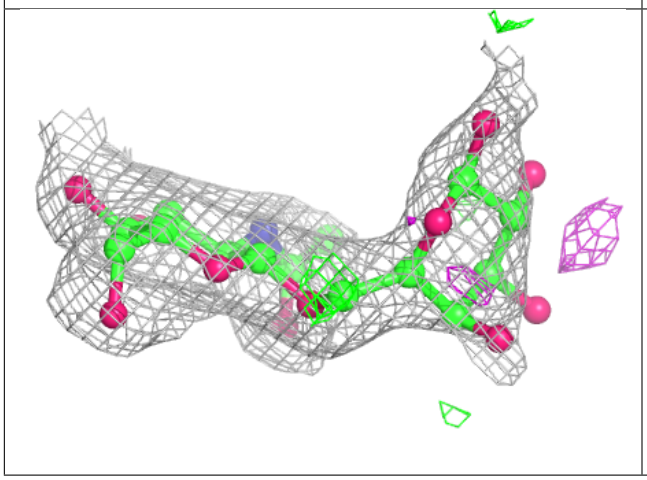
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GAL	F	1	12/12	0.55	0.31	69,91,92,93	0
2	GAL	D	1	12/12	0.75	0.36	56,87,90,91	0
2	GAL	E	1	12/12	0.79	0.38	49,81,88,89	0
2	NAG	F	2	14/15	0.95	0.09	52,56,59,60	0
2	NAG	D	2	14/15	0.97	0.10	37,41,44,47	0
2	NAG	E	2	14/15	0.97	0.09	28,32,35,37	0

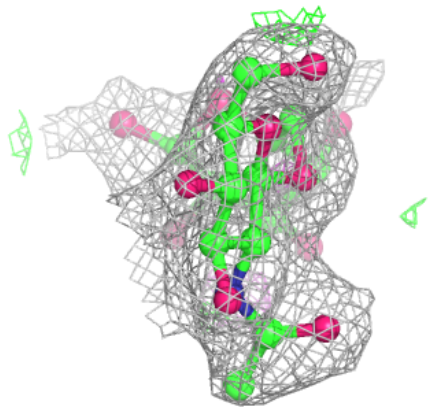
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain D:

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

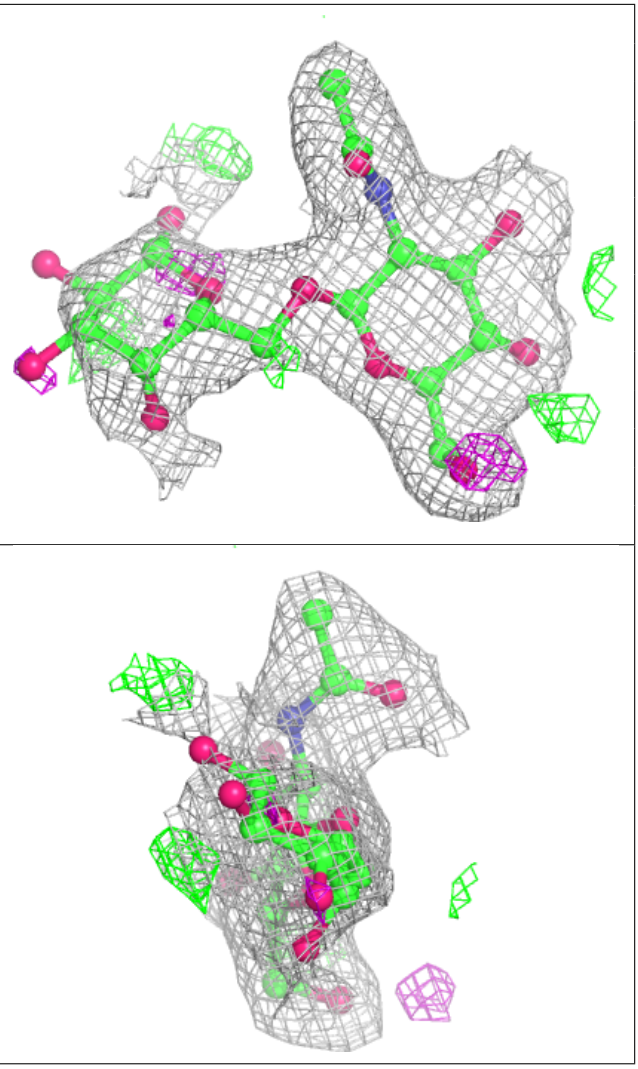






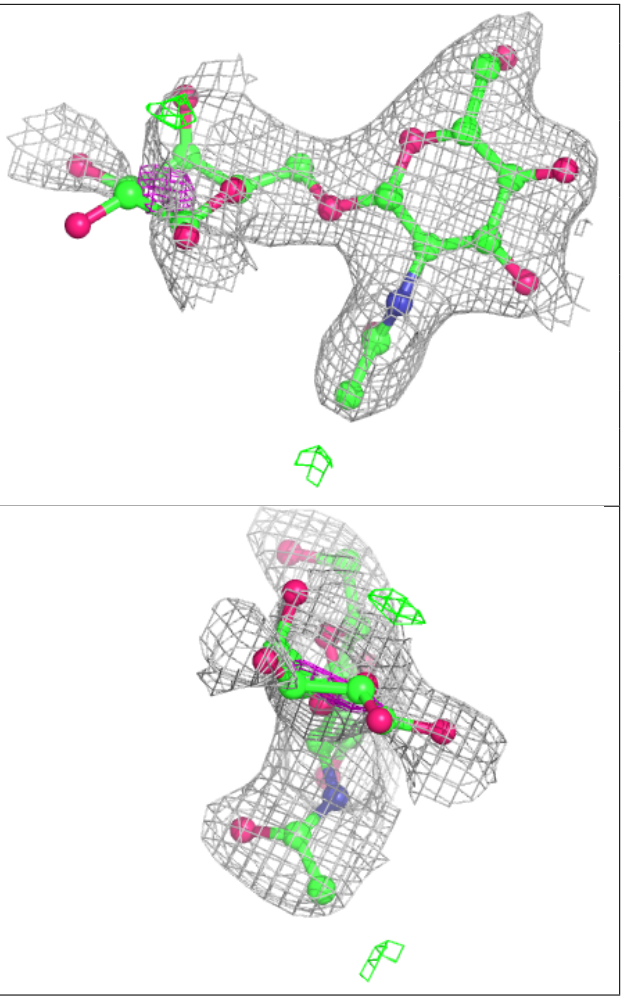
Electron density around Chain E:

$2mF_o$-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
6	SO4	B	407	5/5	0.81	0.21	80,83,85,86	0
3	UDP	C	403	25/25	0.86	0.17	41,49,98,99	0
6	SO4	A	408	5/5	0.88	0.28	79,82,83,84	0
3	UDP	B	403	25/25	0.88	0.16	23,30,86,89	0
6	SO4	B	411	5/5	0.89	0.17	76,77,79,80	0
5	GOL	C	405	6/6	0.91	0.21	68,70,71,71	0
3	UDP	A	403	25/25	0.91	0.15	27,34,87,88	0
5	GOL	B	405	6/6	0.92	0.13	31,41,47,51	0
5	GOL	B	406	6/6	0.92	0.20	34,36,38,43	0
6	SO4	A	406	5/5	0.93	0.17	70,70,75,78	0

Continued on next page...

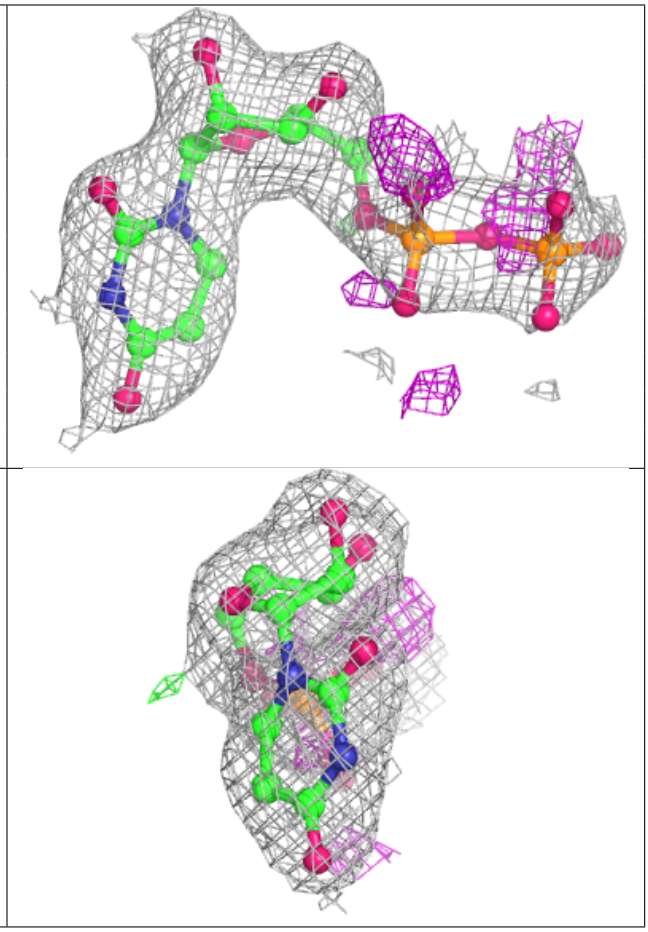
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	GOL	A	405	6/6	0.93	0.14	41,44,52,58	0
6	SO4	C	407	5/5	0.93	0.19	79,80,82,82	0
6	SO4	A	407	5/5	0.95	0.09	69,70,72,73	0
6	SO4	B	409	5/5	0.97	0.15	85,85,86,86	0
6	SO4	C	406	5/5	0.97	0.17	54,56,60,61	0
6	SO4	B	410	5/5	0.97	0.18	58,63,64,66	0
6	SO4	B	408	5/5	0.98	0.17	68,68,71,71	0
4	MN	B	404	1/1	0.99	0.03	40,40,40,40	0
4	MN	C	404	1/1	0.99	0.05	61,61,61,61	0
4	MN	A	404	1/1	1.00	0.03	50,50,50,50	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.

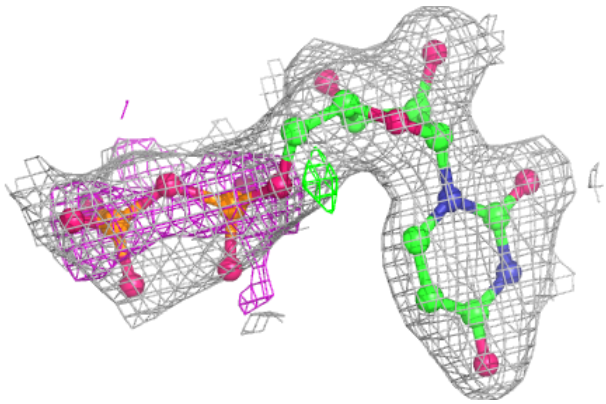
Electron density around UDP C 403:

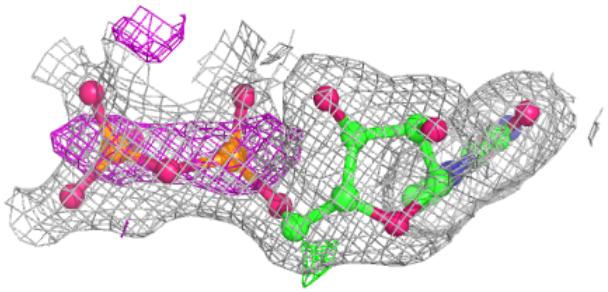
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)

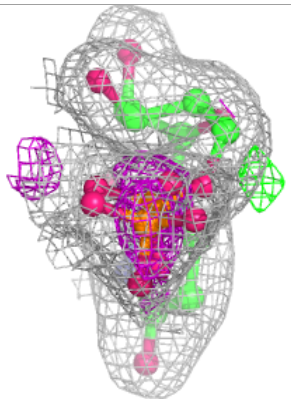


Electron density around UDP B 403:

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

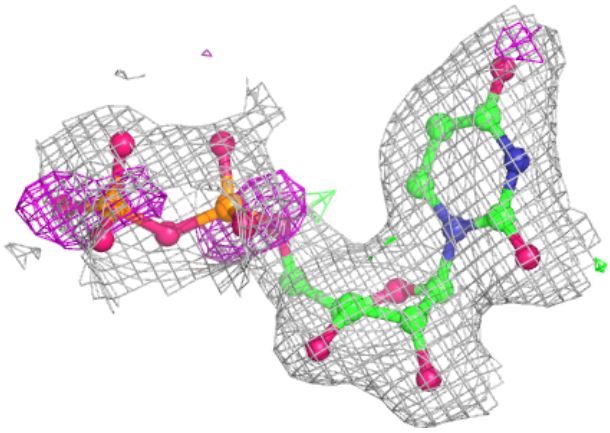


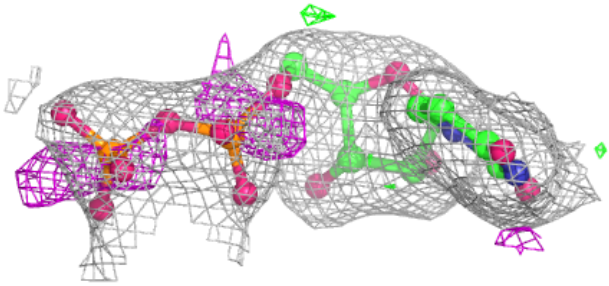


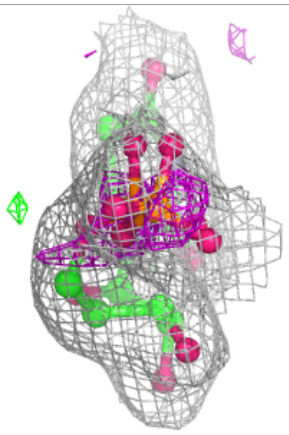


Electron density around UDP A 403:

$2mF_o$-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.