Full wwPDB X-ray Structure Validation Report

Aug 7, 2020 – 04:44 PM BST

PDB ID	:	5U6N
Title	:	Crystal structure of UDP-glucosyltransferase, UGT74F2 (T15S), with UDP and salicylic acid
Authors	:	George Thompson, A.M.; Iancu, C.V.; Dean, J.V.; Choe, J.
Deposited on	:	2016-12-08
Resolution	:	2.00 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 symbol.

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

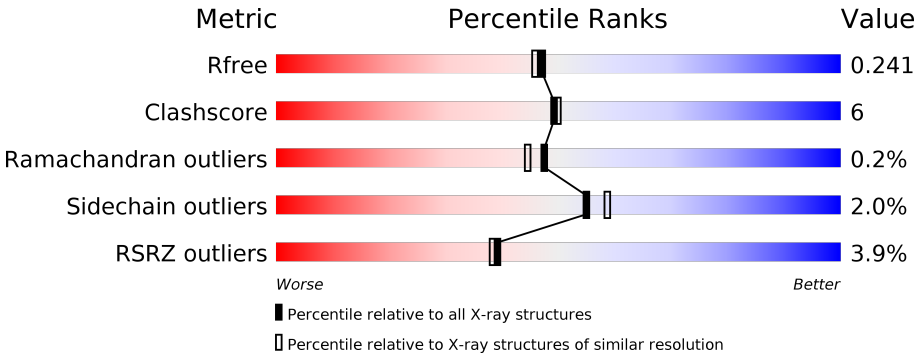
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 on 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	449	4% 86% 12%
1	B	449	4% 85% 14%
2	C	2	50% 50%
2	D	2	100%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7299 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.

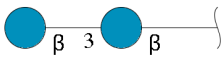
- Molecule 1 is a protein called UDP-glycosyltransferase 74F2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	442	Total	C	N	O	S	0	0	0
			3518	2265	568	669	16			
1	B	448	Total	C	N	O	S	0	0	0
			3569	2295	578	679	17			

There are 2 discrepancies between the modelled and reference sequences:

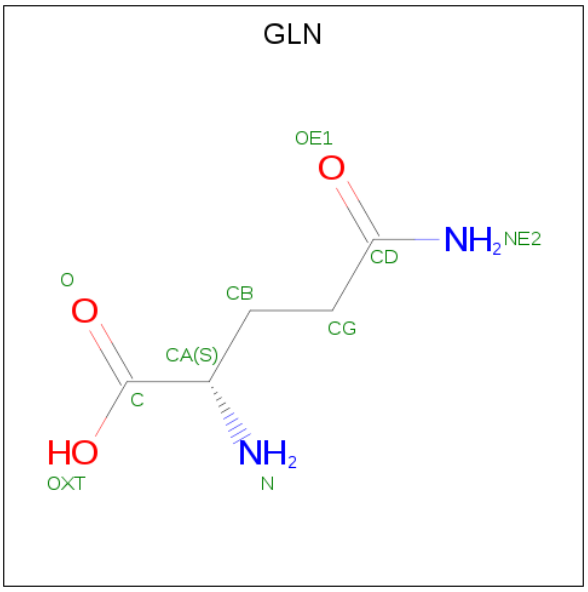
Chain	Residue	Modelled	Actual	Comment	Reference
A	15	SER	THR	conflict	UNP O22822
B	15	SER	THR	conflict	UNP O22822

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



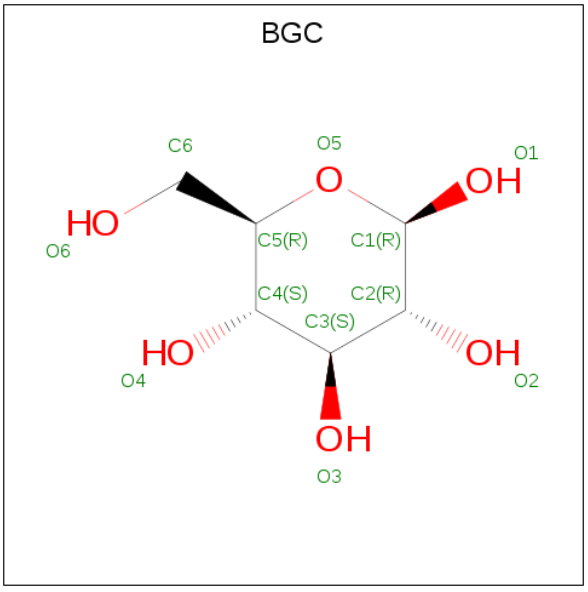
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	C	2	Total	C	O	0	0	0
			22	12	10			
2	D	2	Total	C	O	0	0	0
			22	12	10			

- Molecule 3 is GLUTAMINE (three-letter code: GLN) (formula: $C_5H_{10}N_2O_3$).



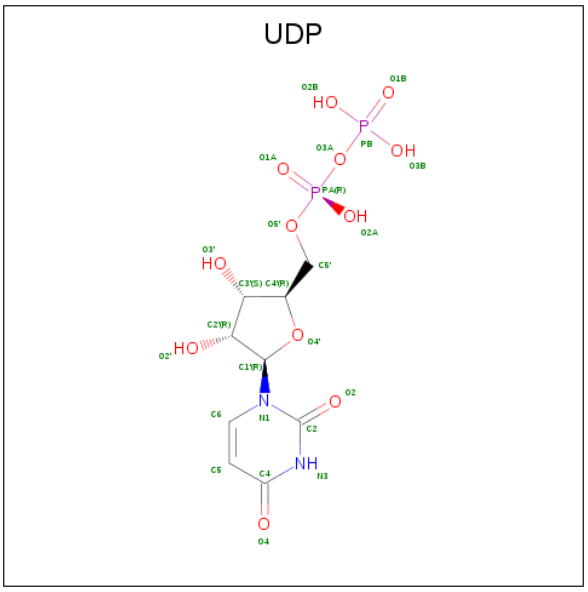
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			9	5	2	2		

- Molecule 4 is beta-D-glucopyranose (three-letter code: BGC) (formula: $C_6H_{12}O_6$).



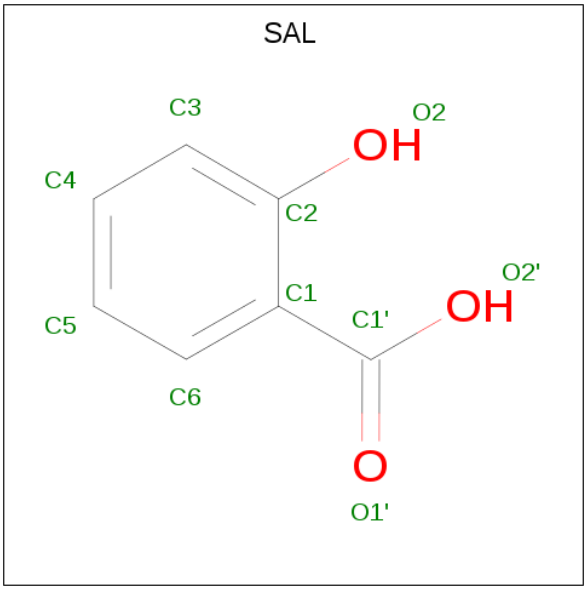
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	B	1	Total	C	O	0	0
			11	6	5		

- 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	C	N	O	P	0	0
			25	9	2	12	2		
5	B	1	Total	C	N	O	P	0	0
			25	9	2	12	2		

- Molecule 6 is 2-HYDROXYBENZOIC ACID (three-letter code: SAL) (formula: $C_7H_6O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	O		0	0
			10	7	3			
6	B	1	Total	C	O		0	0
			10	7	3			

- Molecule 7 is water.

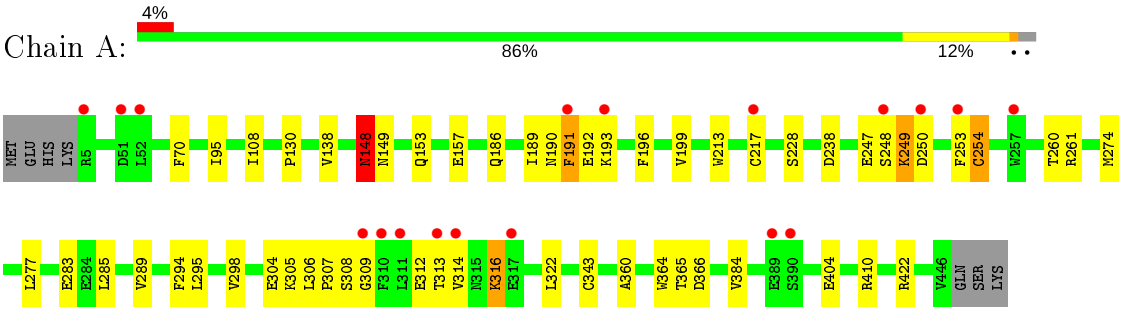
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	25	Total O 25 25	0	0
7	B	42	Total O 42 42	0	0

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.

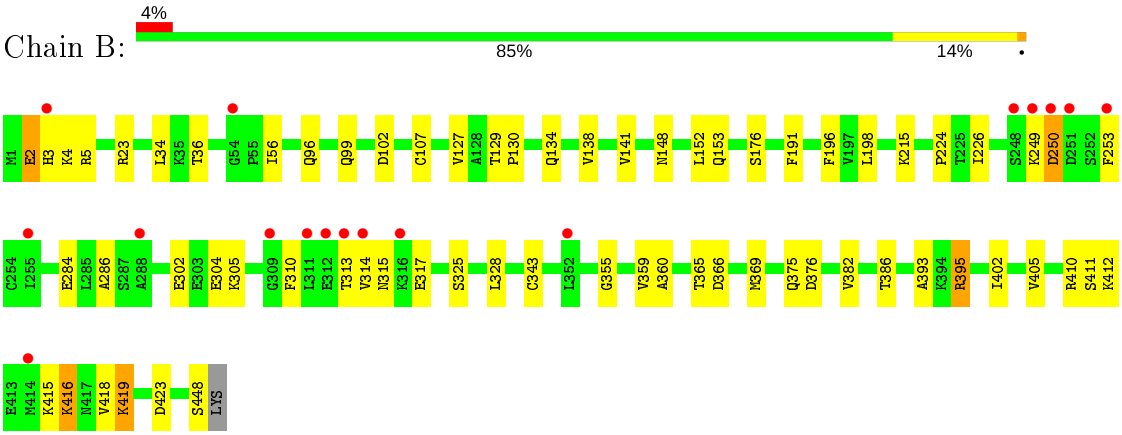
- Molecule 1: UDP-glycosyltransferase 74F2

Chain A:



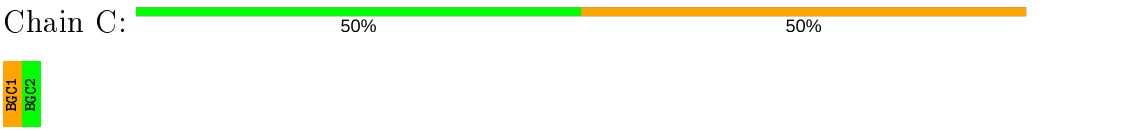
- Molecule 1: UDP-glycosyltransferase 74F2

Chain B:



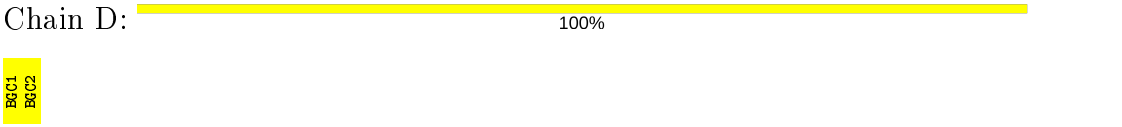
- Molecule 2: beta-D-glucopyranose-(1-3)-beta-D-glucopyranose

Chain C:



- Molecule 2: beta-D-glucopyranose-(1-3)-beta-D-glucopyranose

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	65.18Å 87.56Å 163.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.91 – 2.00 50.91 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.91-2.00) 99.7 (50.91-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.57 (at 2.00Å)	Xtriage
Refinement program	PHENIX (dev_2481: ???)	Depositor
R, R_{free}	0.194 , 0.241 0.194 , 0.241	Depositor DCC
R_{free} test set	3084 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	51.0	Xtriage
Anisotropy	0.186	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 48.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	7299	wwPDB-VP
Average B, all atoms (Å^2)	68.0	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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: UDP, BGC, SAL

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	0.45	0/3607	0.59	0/4898
1	B	0.50	1/3659 (0.0%)	0.61	0/4966
All	All	0.48	1/7266 (0.0%)	0.60	0/9864

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	325	SER	C-N	-8.50	1.18	1.34

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	148	ASN	Peptide
1	A	250	ASP	Peptide

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	3518	0	3449	45	0
1	B	3569	0	3500	41	0
2	C	22	0	19	1	0
2	D	22	0	19	0	0
3	A	9	0	7	0	0
4	A	11	0	9	0	0
4	B	11	0	10	0	0
5	A	25	0	11	0	0
5	B	25	0	11	0	0
6	A	10	0	5	1	0
6	B	10	0	5	2	0
7	A	25	0	0	0	0
7	B	42	0	0	1	0
All	All	7299	0	7045	84	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 6.

All (84) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:191:PHE:HE2	1:A:217:CYS:HG	0.99	0.94
1:B:4:LYS:HD2	1:B:5:ARG:H	1.40	0.84
1:B:4:LYS:CD	1:B:5:ARG:H	1.96	0.77
1:A:289:VAL:HG12	1:A:294:PHE:CD2	2.23	0.73
1:A:191:PHE:HE2	1:A:217:CYS:SG	2.09	0.73
1:A:153:GLN:N	1:A:153:GLN:OE1	2.19	0.70
1:A:314:VAL:HB	1:A:316:LYS:HE3	1.72	0.69
1:A:285:LEU:O	1:A:289:VAL:HG23	1.93	0.68
1:B:419:LYS:O	1:B:419:LYS:HD2	1.94	0.67
1:B:415:LYS:O	1:B:416:LYS:HD2	1.95	0.67
1:B:412:LYS:HE2	1:B:415:LYS:HE3	1.81	0.62
1:B:2:GLU:O	1:B:3:HIS:HD2	1.82	0.62
1:A:309:GLY:O	1:A:313:THR:HG23	1.99	0.61
1:A:199:VAL:HG11	1:A:213:TRP:HZ3	1.68	0.59
1:A:190:ASN:HB2	1:A:193:LYS:HD2	1.85	0.59
1:A:95:ILE:HD11	1:A:108:ILE:HD11	1.84	0.59
1:A:199:VAL:HG11	1:A:213:TRP:CZ3	2.38	0.58
1:A:130:PRO:HD2	1:A:196:PHE:O	2.05	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:284:GLU:OE1	1:B:395:ARG:HD2	2.06	0.56
1:B:382:VAL:HG22	1:B:405:VAL:HG23	1.88	0.55
1:A:316:LYS:HD2	1:A:316:LYS:N	2.21	0.55
1:A:247:GLU:OE1	1:B:102:ASP:HB2	2.07	0.54
1:A:312:GLU:OE2	1:A:312:GLU:N	2.40	0.54
1:B:134:GLN:HE22	6:B:504:SAL:H5	1.73	0.54
1:B:36:THR:OG1	1:B:56:ILE:HD12	2.07	0.53
1:B:304:GLU:HG3	1:B:305:LYS:HG3	1.92	0.52
1:B:148:ASN:ND2	1:B:153:GLN:OE1	2.43	0.52
1:A:283:GLU:HG2	1:A:307:PRO:HG3	1.92	0.52
1:B:141:VAL:HG13	1:B:152:LEU:HD11	1.91	0.52
1:B:386:THR:HG22	1:B:393:ALA:HB2	1.92	0.52
1:A:295:LEU:HD11	1:A:322:LEU:HD12	1.93	0.51
1:A:404:GLU:OE1	1:A:410:ARG:HD3	2.10	0.51
1:B:419:LYS:HE3	1:B:423:ASP:OD1	2.10	0.51
1:A:247:GLU:OE1	1:B:102:ASP:N	2.40	0.51
1:B:328:LEU:HB2	7:B:640:HOH:O	2.10	0.51
1:A:298:VAL:HG11	1:A:306:LEU:HD11	1.93	0.50
1:A:186:GLN:O	1:A:189:ILE:HG12	2.12	0.50
1:A:248:SER:O	1:A:249:LYS:HB3	2.12	0.50
1:B:365:THR:HG23	1:B:366:ASP:H	1.77	0.49
1:A:422:ARG:HH11	1:A:422:ARG:HG2	1.76	0.49
1:A:138:VAL:HG21	1:A:366:ASP:HB2	1.95	0.49
1:B:302:GLU:OE2	1:B:305:LYS:NZ	2.33	0.48
1:B:310:PHE:O	1:B:314:VAL:HG13	2.12	0.48
1:B:355:GLY:HA2	1:B:418:VAL:HG13	1.94	0.48
1:A:249:LYS:HZ3	1:A:253:PHE:HB3	1.78	0.48
1:B:249:LYS:HG3	1:B:250:ASP:OD1	2.14	0.48
1:B:286:ALA:HB1	1:B:310:PHE:CE1	2.49	0.48
1:B:134:GLN:HE22	6:B:504:SAL:C5	2.26	0.47
1:A:254:CYS:HB2	1:A:322:LEU:HD11	1.97	0.47
1:A:191:PHE:CD1	1:A:192:GLU:N	2.83	0.47
1:B:382:VAL:HG22	1:B:405:VAL:CG2	2.45	0.47
1:A:422:ARG:NH1	1:A:422:ARG:HG2	2.30	0.46
1:A:314:VAL:HB	1:A:316:LYS:CE	2.43	0.46
1:B:375:GLN:HG2	1:B:376:ASP:OD1	2.16	0.46
1:B:129:THR:CG2	1:B:198:LEU:HG	2.47	0.45
1:A:365:THR:OG1	6:A:505:SAL:H5	2.17	0.45
1:A:148:ASN:O	1:A:149:ASN:HB3	2.17	0.45
1:A:277:LEU:O	1:A:305:LYS:HD2	2.17	0.44
1:B:359:VAL:HG21	1:B:402:ILE:HA	1.99	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:224:PRO:HB2	1:B:226:ILE:HD12	1.99	0.44
1:B:411:SER:OG	1:B:412:LYS:HE2	2.18	0.44
1:A:157:GLU:CD	1:A:157:GLU:H	2.21	0.44
1:A:70:PHE:CE2	2:C:1:BGC:H4	2.53	0.44
1:A:360:ALA:O	1:A:384:VAL:HG22	2.18	0.43
1:A:260:THR:OG1	1:A:261:ARG:HD2	2.19	0.43
1:B:315:ASN:OD1	1:B:317:GLU:HB3	2.18	0.43
1:A:149:ASN:O	1:A:149:ASN:OD1	2.37	0.43
1:A:343:CYS:SG	1:A:360:ALA:HB1	2.59	0.43
1:B:138:VAL:HG22	1:B:369:MET:HG2	2.01	0.43
1:B:107:CYS:HB2	1:B:127:VAL:HB	2.01	0.42
1:A:289:VAL:HG12	1:A:294:PHE:CG	2.54	0.42
1:B:2:GLU:O	1:B:3:HIS:CD2	2.69	0.42
1:B:96:GLN:O	1:B:99:GLN:HB2	2.19	0.42
1:B:310:PHE:HA	1:B:313:THR:HG22	2.02	0.41
1:B:130:PRO:HD2	1:B:196:PHE:O	2.21	0.41
1:B:34:LEU:HA	1:B:34:LEU:HD23	1.80	0.41
1:A:148:ASN:HB3	1:A:149:ASN:H	1.73	0.41
1:B:343:CYS:SG	1:B:360:ALA:HB1	2.60	0.41
1:A:304:GLU:OE2	1:A:305:LYS:HG2	2.21	0.41
1:B:365:THR:HG23	1:B:366:ASP:N	2.36	0.41
1:A:249:LYS:HB3	1:A:249:LYS:HE3	1.81	0.41
1:A:228:SER:HB3	1:A:238:ASP:HB3	2.03	0.40
1:A:274:MET:HE2	1:A:364:TRP:HZ2	1.86	0.40
1:A:304:GLU:OE2	1:A:305:LYS:CG	2.70	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	440/449 (98%)	420 (96%)	19 (4%)	1 (0%)	47 44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	446/449 (99%)	434 (97%)	11 (2%)	1 (0%)	47 44
All	All	886/898 (99%)	854 (96%)	30 (3%)	2 (0%)	47 44

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	148	ASN
1	B	250	ASP

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	394/401 (98%)	389 (99%)	5 (1%)	69 74
1	B	400/401 (100%)	389 (97%)	11 (3%)	43 44
All	All	794/802 (99%)	778 (98%)	16 (2%)	55 58

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	191	PHE
1	A	249	LYS
1	A	254	CYS
1	A	308	SER
1	A	316	LYS
1	B	2	GLU
1	B	23	ARG
1	B	176	SER
1	B	191	PHE
1	B	215	LYS
1	B	253	PHE
1	B	395	ARG
1	B	410	ARG
1	B	416	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	419	LYS
1	B	448	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	3	HIS
1	B	134	GLN
1	B	148	ASN
1	B	153	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)

4 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	BGC	C	1	1,2	11,11,12	0.95	0	15,15,17	0.81	1 (6%)
2	BGC	C	2	2	11,11,12	0.39	0	15,15,17	0.80	0
2	BGC	D	1	1,2	11,11,12	1.25	1 (9%)	15,15,17	1.16	1 (6%)
2	BGC	D	2	2	11,11,12	0.60	0	15,15,17	1.36	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	BGC	C	1	1,2	-	0/2/19/22	0/1/1/1
2	BGC	C	2	2	-	0/2/19/22	0/1/1/1
2	BGC	D	1	1,2	-	0/2/19/22	0/1/1/1
2	BGC	D	2	2	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1	BGC	C1-C2	2.63	1.58	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	2	BGC	C1-O5-C5	3.50	116.94	112.19
2	D	1	BGC	C1-O5-C5	-2.22	109.19	112.19
2	C	1	BGC	O5-C5-C6	2.03	110.39	107.20

There are no chirality outliers.

There are no torsion outliers.

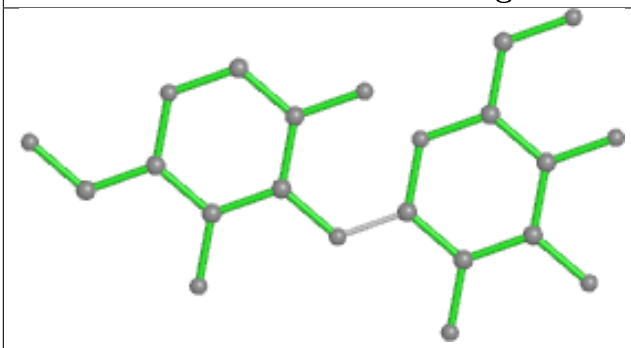
There are no ring outliers.

1 monomer is involved in 1 short contact:

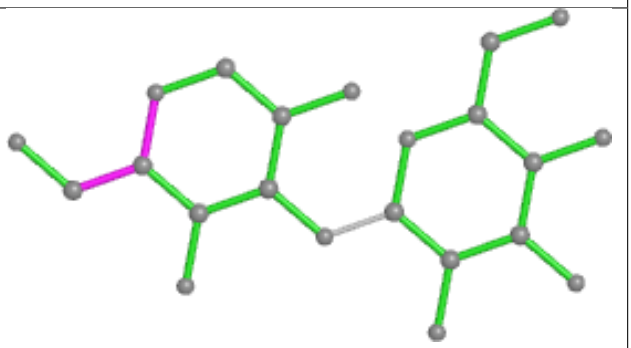
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	BGC	1	0

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

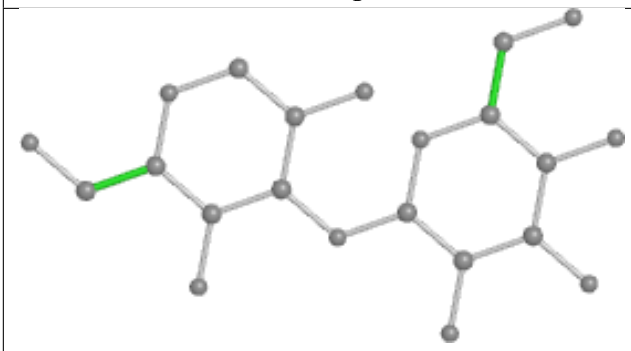
Oligosaccharide Chain C



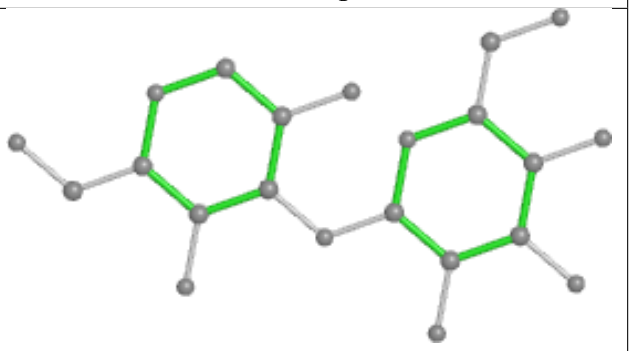
Bond lengths



Bond angles

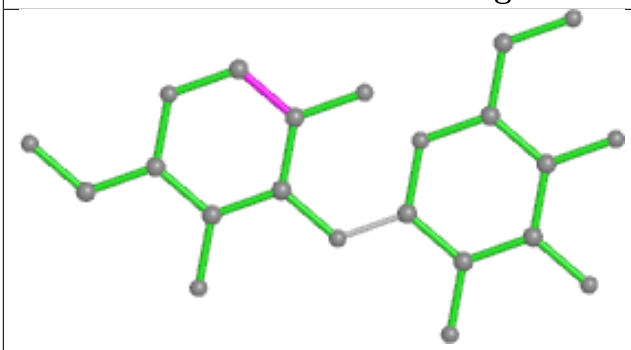


Torsions

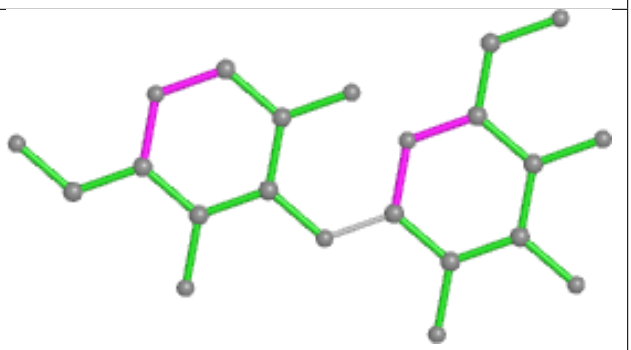


Rings

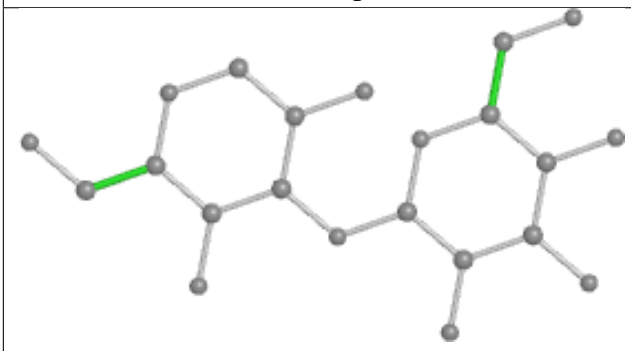
Oligosaccharide Chain D



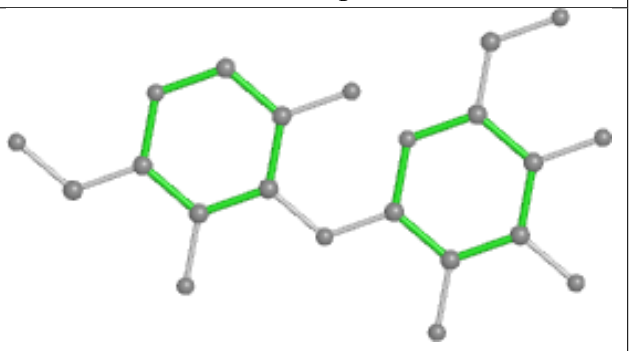
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

7 ligands 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	$\#\|Z\| > 2$
6	SAL	B	504	-	8,10,10	4.39	4 (50%)	9,13,13	0.96	1 (11%)
5	UDP	A	504	-	20,26,26	3.61	8 (40%)	25,40,40	1.27	1 (4%)
4	BGC	A	502	1	11,11,12	1.21	0	15,15,17	1.44	3 (20%)
6	SAL	A	505	-	8,10,10	4.17	4 (50%)	9,13,13	1.06	1 (11%)
4	BGC	B	501	1	11,11,12	1.31	1 (9%)	15,15,17	1.21	3 (20%)
5	UDP	B	503	-	20,26,26	3.76	9 (45%)	25,40,40	1.20	1 (4%)

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
6	SAL	B	504	-	-	0/0/4/4	0/1/1/1
5	UDP	A	504	-	-	2/14/32/32	0/2/2/2
4	BGC	A	502	1	-	0/2/19/22	0/1/1/1
6	SAL	A	505	-	-	0/0/4/4	0/1/1/1
4	BGC	B	501	1	-	0/2/19/22	0/1/1/1
5	UDP	B	503	-	-	5/14/32/32	0/2/2/2

All (26) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	504	SAL	C1-C1'	11.09	1.58	1.47
6	A	505	SAL	C1-C1'	10.43	1.57	1.47
5	B	503	UDP	C6-N1	8.68	1.46	1.35
5	A	504	UDP	C6-N1	7.30	1.44	1.35
5	A	504	UDP	C4-N3	6.98	1.45	1.33
5	B	503	UDP	C4-N3	6.72	1.44	1.33

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	504	UDP	O4'-C4'	6.25	1.59	1.45
5	B	503	UDP	O4'-C4'	5.90	1.58	1.45
5	B	503	UDP	O4'-C1'	-5.87	1.32	1.41
5	B	503	UDP	C3'-C4'	-5.87	1.38	1.53
5	A	504	UDP	C3'-C4'	-5.63	1.38	1.53
5	A	504	UDP	C6-C5	5.16	1.49	1.38
5	A	504	UDP	O4'-C1'	-5.16	1.33	1.41
5	B	503	UDP	C6-C5	5.11	1.49	1.38
5	A	504	UDP	C2-N3	3.97	1.46	1.38
6	A	505	SAL	C6-C1	-3.54	1.34	1.40
5	B	503	UDP	C2-N3	3.45	1.45	1.38
6	B	504	SAL	C6-C1	-3.06	1.35	1.40
6	B	504	SAL	C5-C6	3.00	1.45	1.38
5	A	504	UDP	O3'-C3'	2.74	1.49	1.43
6	B	504	SAL	O2-C2	2.72	1.41	1.36
6	A	505	SAL	O2-C2	2.66	1.41	1.36
6	A	505	SAL	C5-C6	2.45	1.44	1.38
5	B	503	UDP	O3'-C3'	2.44	1.48	1.43
5	B	503	UDP	O4-C4	-2.18	1.19	1.24
4	B	501	BGC	O5-C1	-2.12	1.40	1.43

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	504	UDP	C2'-C3'-C4'	3.79	110.00	102.64
4	A	502	BGC	C1-C2-C3	3.75	114.27	109.67
5	B	503	UDP	C2'-C3'-C4'	3.17	108.79	102.64
6	A	505	SAL	C6-C1-C1'	-2.48	116.50	120.20
6	B	504	SAL	C6-C1-C1'	-2.32	116.75	120.20
4	B	501	BGC	O5-C5-C6	-2.15	103.83	107.20
4	A	502	BGC	O3-C3-C2	2.08	113.97	109.99
4	A	502	BGC	O2-C2-C3	-2.08	105.97	110.14
4	B	501	BGC	C1-O5-C5	2.06	114.98	112.19
4	B	501	BGC	C1-C2-C3	2.03	112.16	109.67

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	503	UDP	C2'-C1'-N1-C6
5	B	503	UDP	O4'-C1'-N1-C6
5	B	503	UDP	O4'-C4'-C5'-O5'

Continued on next page...

Continued from previous page...

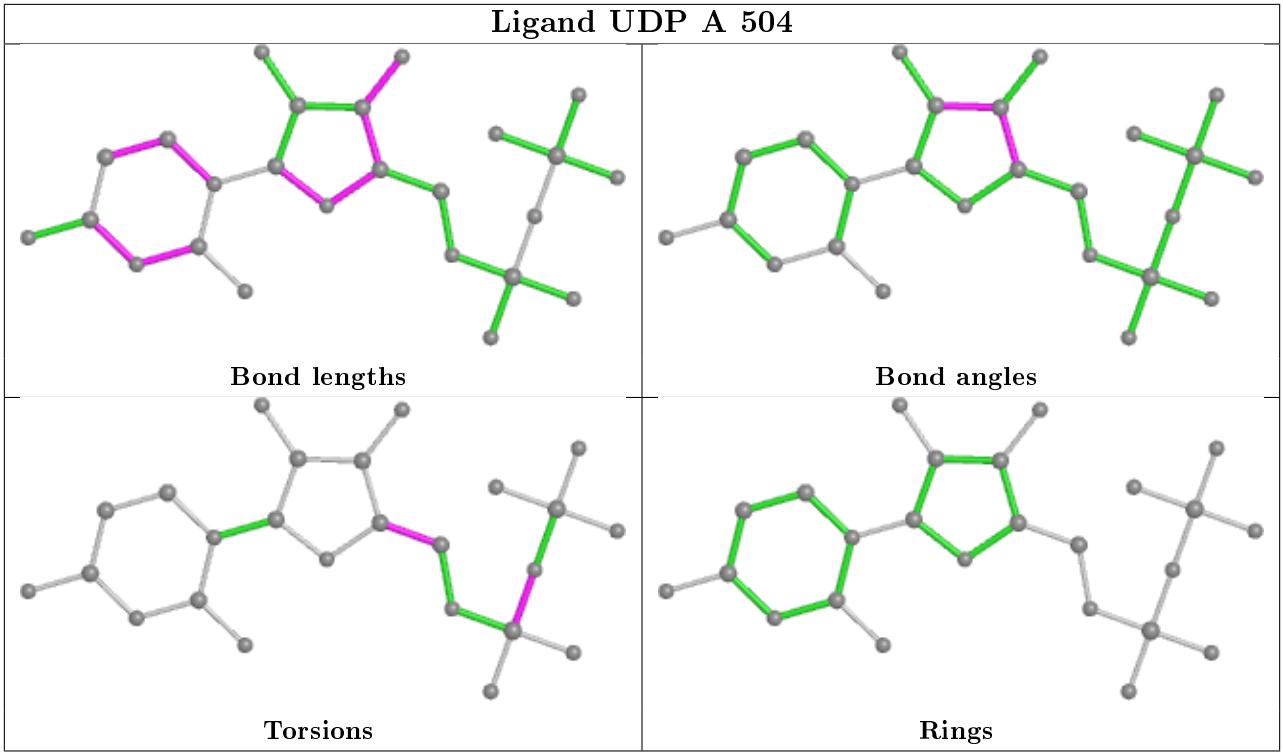
Mol	Chain	Res	Type	Atoms
5	B	503	UDP	C3'-C4'-C5'-O5'
5	A	504	UDP	PB-O3A-PA-O5'
5	B	503	UDP	PB-O3A-PA-O5'
5	A	504	UDP	O4'-C4'-C5'-O5'

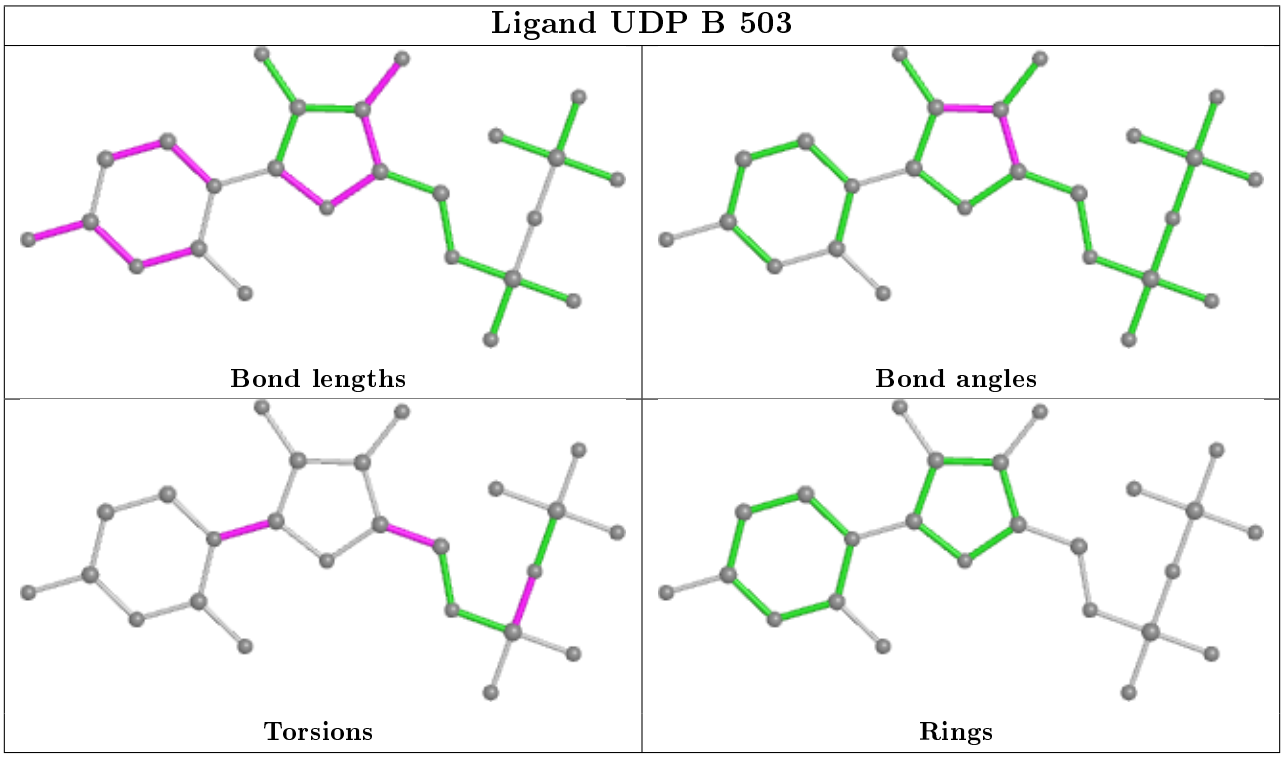
There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	504	SAL	2	0
6	A	505	SAL	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.





5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	325:SER	C	326:PRO	N	1.18

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	442/449 (98%)	0.19	18 (4%)	37	36	42, 64, 107, 152	0
1	B	448/449 (99%)	0.16	17 (3%)	40	39	37, 64, 111, 144	0
All	All	890/898 (99%)	0.17	35 (3%)	39	38	37, 64, 109, 152	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	253	PHE	5.9
1	A	5	ARG	5.2
1	A	250	ASP	5.1
1	B	414	MET	4.4
1	A	52	LEU	3.6
1	B	314	VAL	3.5
1	B	312	GLU	3.3
1	B	255	ILE	3.3
1	B	249	LYS	3.3
1	A	389	GLU	3.2
1	B	309	GLY	3.0
1	B	250	ASP	2.9
1	A	257	TRP	2.9
1	A	191	PHE	2.7
1	A	309	GLY	2.6
1	A	313	THR	2.5
1	A	253	PHE	2.5
1	A	314	VAL	2.5
1	A	310	PHE	2.5
1	B	248	SER	2.4
1	B	251	ASP	2.3
1	B	3	HIS	2.3
1	B	288	ALA	2.2
1	B	311	LEU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	248	SER	2.2
1	A	311	LEU	2.2
1	A	193	LYS	2.2
1	B	54	GLY	2.2
1	A	217	CYS	2.2
1	A	51	ASP	2.1
1	B	316	LYS	2.1
1	A	317	GLU	2.1
1	B	313	THR	2.1
1	A	390	SER	2.0
1	B	352	LEU	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)

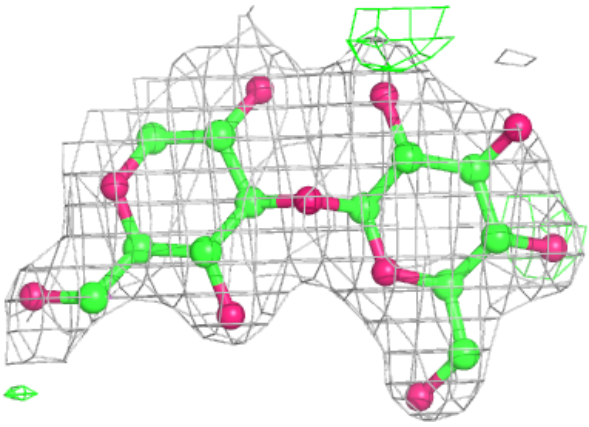
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

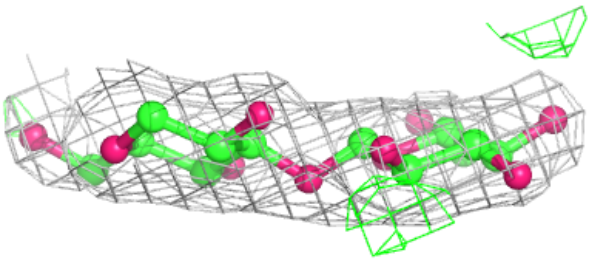
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	BGC	C	2	11/12	0.95	0.12	55,63,73,76	0
2	BGC	D	2	11/12	0.96	0.11	41,46,54,56	0
2	BGC	D	1	11/12	0.96	0.10	41,49,52,63	0
2	BGC	C	1	11/12	0.96	0.09	55,59,64,67	0

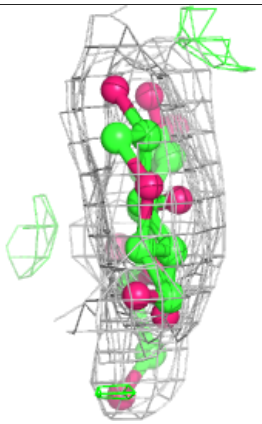
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

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)

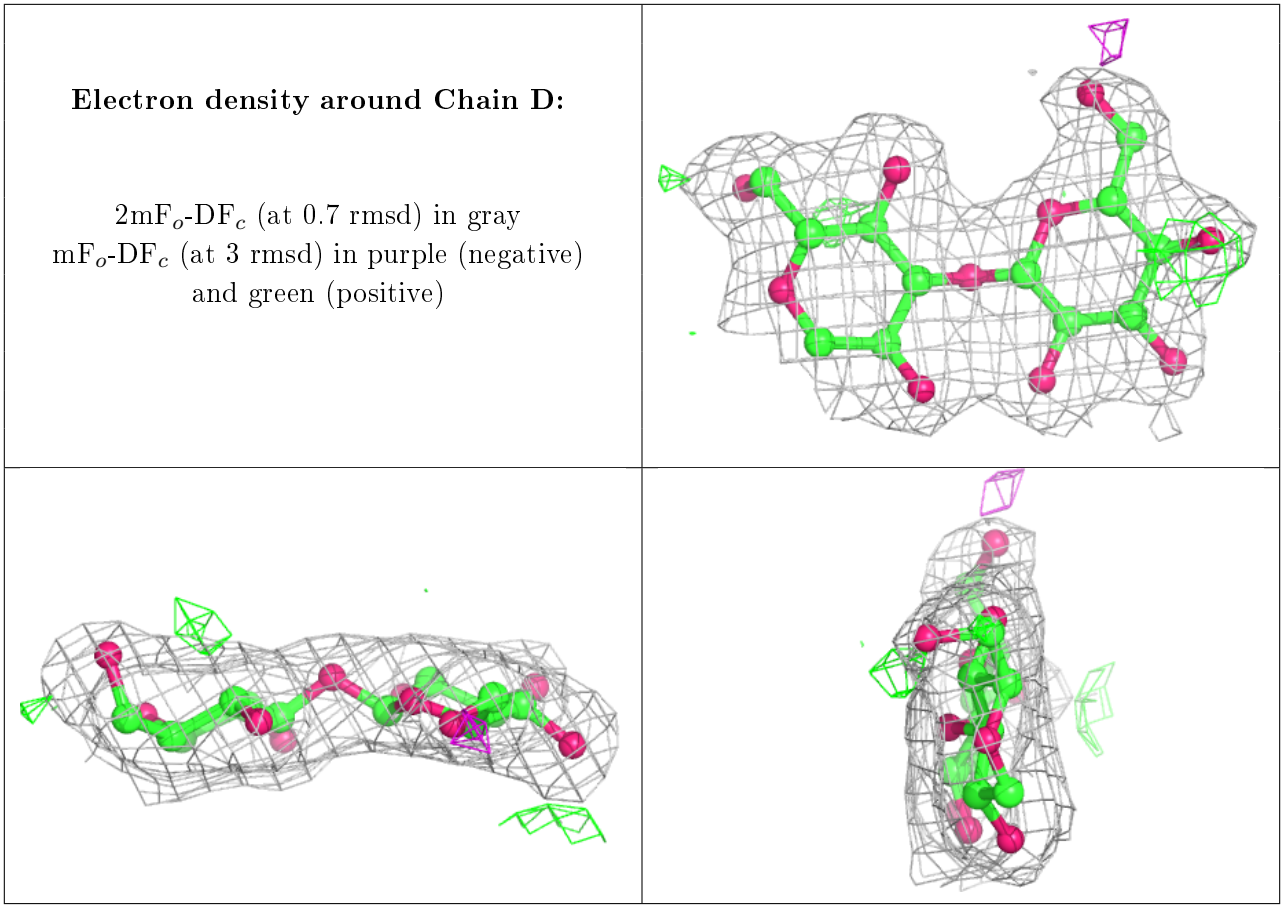






Electron density around Chain D:

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



6.4 Ligands

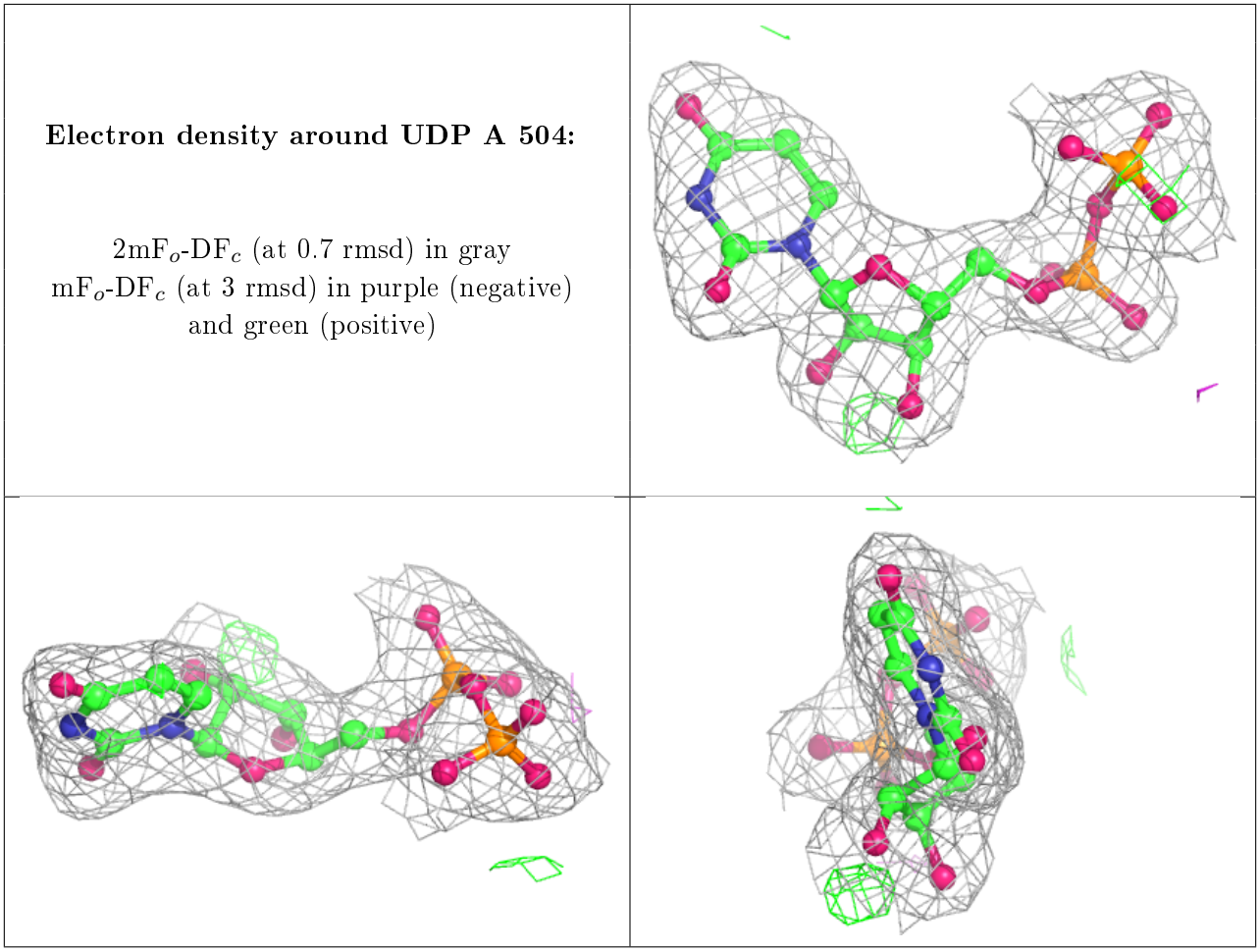
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	SAL	B	504	10/10	0.73	0.19	58,79,86,92	0
6	SAL	A	505	10/10	0.74	0.21	77,89,98,100	0
3	GLN	A	501	9/10	0.89	0.16	77,89,95,96	0
4	BGC	B	501	11/12	0.94	0.11	55,65,71,79	0
4	BGC	A	502	11/12	0.94	0.16	60,64,70,81	0
5	UDP	A	504	25/25	0.98	0.13	40,46,55,57	0
5	UDP	B	503	25/25	0.98	0.13	46,55,62,63	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.

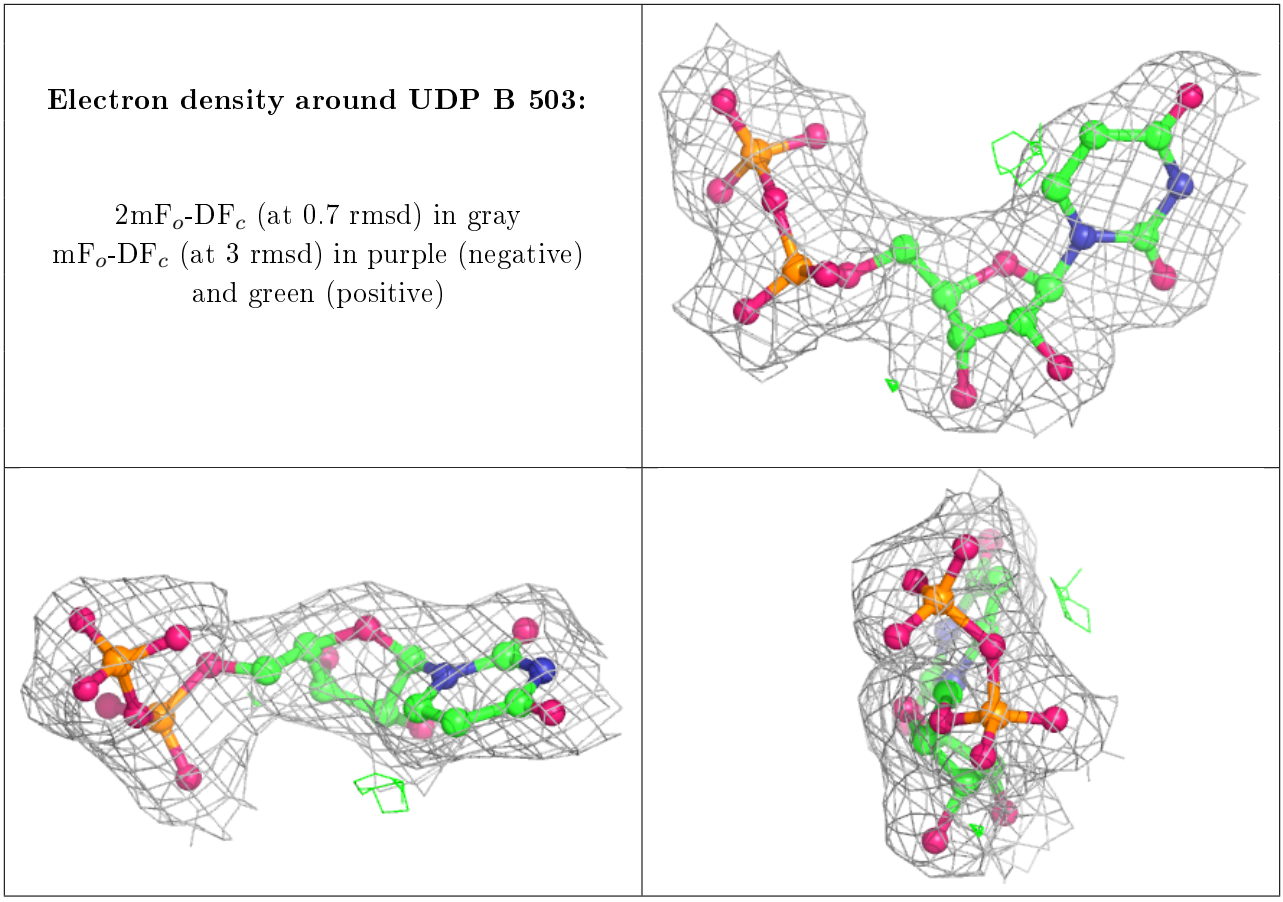
Electron density around UDP A 504:

$2mF_o$-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 503:

$2mF_o$-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

There are no such residues in this entry.