



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 06:15 PM UTC

PDB ID : 2HVI / pdb_00002hvi
Title : ddCTP:G pair in the polymerase active site (0 position)
Authors : Warren, J.J.; Forsberg, L.J.; Beese, L.S.
Deposited on : 2006-07-28
Resolution : 1.98 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Buster-report : wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

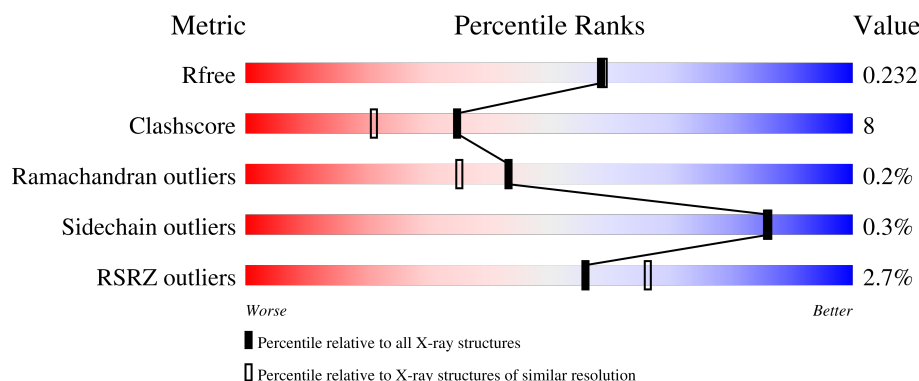
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.98 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.



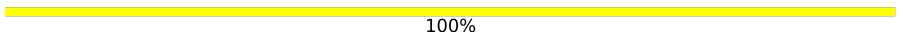
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	1506 (1.98-1.98)
Clashscore	190562	1534 (1.98-1.98)
Ramachandran outliers	187476	1518 (1.98-1.98)
Sidechain outliers	187428	1518 (1.98-1.98)
RSRZ outliers	180081	1506 (1.98-1.98)

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 $\leq 5\%$. 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	B	9	11% 56% 44%
1	E	9	11% 56% 44%
2	C	13	8% 54% 38% 8%
2	F	13	69% 23% 8%
3	A	580	3% 79% 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	580	 2% 83% 17%
4	G	2	 100%
4	H	2	 100%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 10940 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 DNA chain called 5'-D(*CP*CP*TP*GP*AP*CP*TP*CP*(DDG))-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	9	Total	C	N	O	P	0	0	0
			177	86	31	52	8			
1	E	9	Total	C	N	O	P	0	0	0
			177	86	31	52	8			

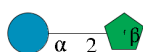
- Molecule 2 is a DNA chain called 5'-D(*C*AP*TP*GP*CP*GP*AP*GP*TP*CP*AP*GP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	12	Total	C	N	O	P	0	0	0
			248	118	50	69	11			
2	F	12	Total	C	N	O	P	0	0	0
			248	118	50	69	11			

- Molecule 3 is a protein called DNA Polymerase I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	580	Total	C	N	O	S	0	0	0
			4652	2958	808	869	17			
3	D	580	Total	C	N	O	S	0	0	0
			4652	2958	808	869	17			

- Molecule 4 is an oligosaccharide called beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose.



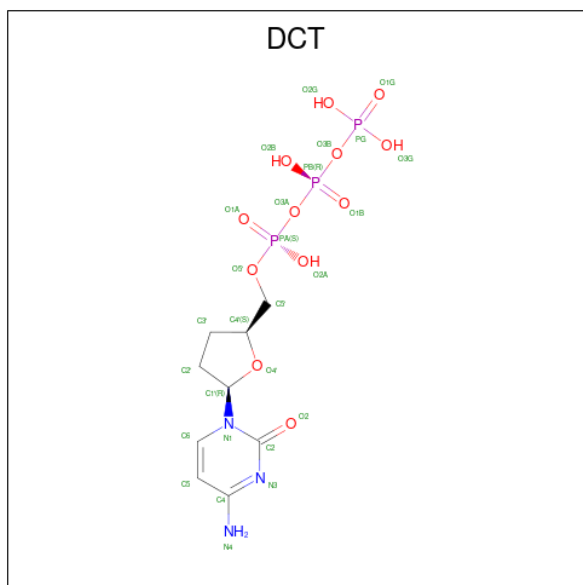
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
4	G	2	Total	C	O	0	0	0
			23	12	11			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
4	H	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 5 is 2',3'-DIDEOXYCYTIDINE 5'-TRIPHOSPHATE (CCD ID: DCT) (formula: $C_9H_{16}N_3O_{12}P_3$).

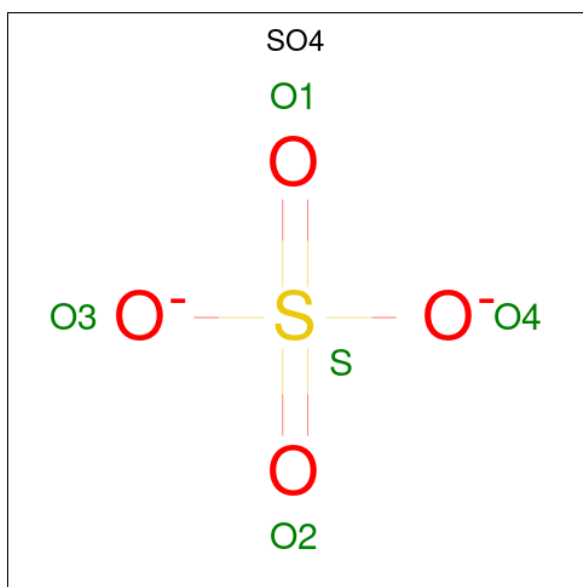


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			27	9	3	12	3		
5	A	1	Total	C	N	O	P	0	0
			27	9	3	12	3		
5	D	1	Total	C	N	O	P	0	0
			27	9	3	12	3		

- Molecule 6 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

- Molecule 7 is SULFATE ION (CCD ID: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	20	Total	O	0	0
			20	20		
8	C	36	Total	O	0	0
			36	36		
8	E	22	Total	O	0	0
			22	22		
8	F	39	Total	O	0	0
			39	39		
8	A	200	Total	O	0	0
			200	200		
8	D	335	Total	O	0	0
			335	335		

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: 5'-D(*CP*CP*TP*GP*AP*CP*TP*CP*(DDG))-3'



- Molecule 1: 5'-D(*CP*CP*TP*GP*AP*CP*TP*CP*(DDG))-3'



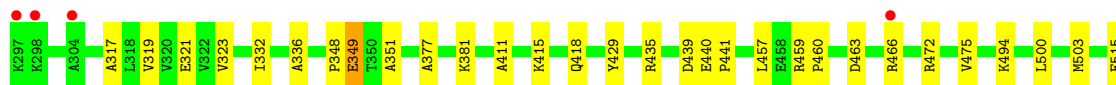
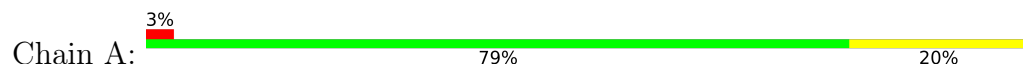
- Molecule 2: 5'-D(*C*AP*TP*GP*CP*GP*AP*GP*TP*CP*AP*GP*G)-3'

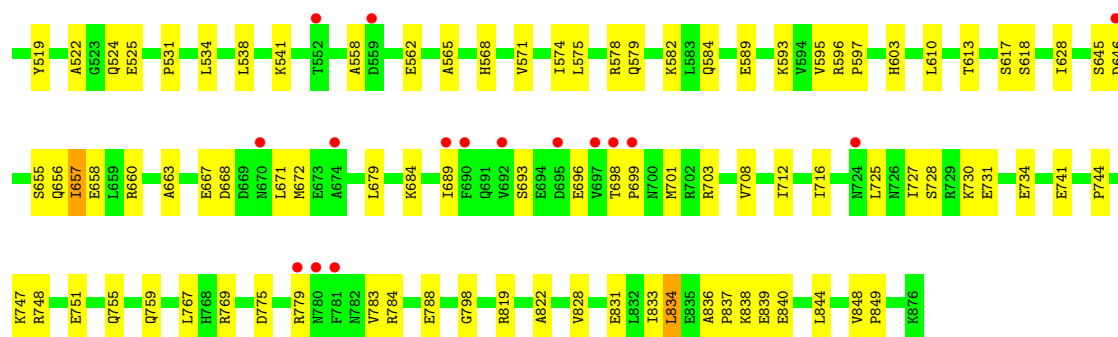


- Molecule 2: 5'-D(*C*AP*TP*GP*CP*GP*AP*GP*TP*CP*AP*GP*G)-3'

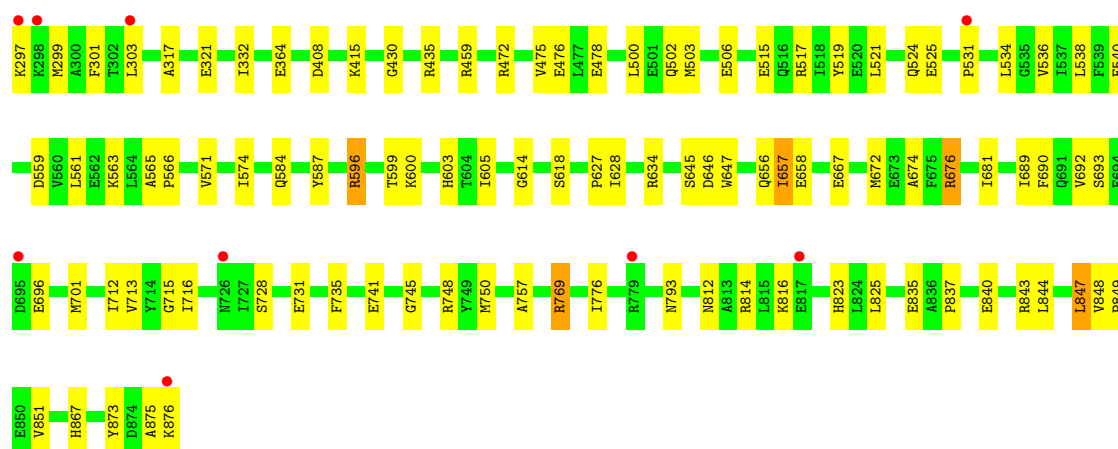
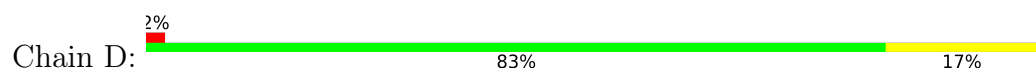


- Molecule 3: DNA Polymerase I





• Molecule 3: DNA Polymerase I



• Molecule 4: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose



GLC1
FRU2

• Molecule 4: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose



GLC1
FRU2

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	93.80Å 108.49Å 149.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.98 50.00 – 1.98	Depositor EDS
% Data completeness (in resolution range)	83.2 (50.00-1.98) 83.2 (50.00-1.98)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.74 (at 1.98Å)	Xtriage
Refinement program	CNS, REFMAC	Depositor
R, R_{free}	0.199 , 0.231 0.199 , 0.232	Depositor DCC
R_{free} test set	4484 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	27.2	Xtriage
Anisotropy	0.160	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 44.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	10940	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.37% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: DDG, DCT, SO4, FRU, MG, GLC

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	B	0.30	0/173	0.79	0/264
1	E	0.29	0/173	0.79	0/264
2	C	0.29	0/279	0.61	0/430
2	F	0.29	0/279	0.66	0/430
3	A	0.35	0/4736	0.86	10/6400 (0.2%)
3	D	0.38	0/4736	0.88	15/6400 (0.2%)
All	All	0.36	0/10376	0.85	25/14188 (0.2%)

There are no bond length outliers.

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	657	ILE	N-CA-C	10.28	121.16	110.36
3	D	321	GLU	N-CA-C	7.30	120.42	109.25
3	A	321	GLU	N-CA-C	6.91	119.82	109.25
3	D	769	ARG	N-CA-C	-6.85	100.84	110.50
3	D	657	ILE	N-CA-C	6.62	117.11	110.23
3	D	475	VAL	N-CA-C	6.51	119.23	111.09
3	D	847	LEU	N-CA-C	6.32	117.97	111.14
3	D	715	GLY	N-CA-C	6.18	124.67	115.64
3	A	351	ALA	N-CA-C	6.15	117.65	111.07
3	D	538	LEU	N-CA-C	6.03	117.65	111.14
3	A	769	ARG	N-CA-C	-5.97	102.08	110.50
3	A	317	ALA	N-CA-C	-5.71	99.94	109.24
3	A	538	LEU	N-CA-C	5.40	116.97	111.14
3	D	317	ALA	N-CA-C	-5.35	100.51	109.24
3	A	617	SER	N-CA-C	-5.31	101.29	109.52
3	D	478	GLU	N-CA-C	5.31	116.75	111.07
3	A	741	GLU	N-CA-C	-5.24	105.64	111.36
3	D	332	ILE	N-CA-C	-5.22	99.74	107.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	475	VAL	N-CA-C	5.21	117.57	111.05
3	D	741	GLU	N-CA-C	-5.20	105.67	112.23
3	D	596	ARG	CA-C-N	5.20	124.65	119.24
3	D	596	ARG	C-N-CA	5.20	124.65	119.24
3	A	323	VAL	N-CA-C	5.07	115.84	110.72
3	D	676	ARG	N-CA-C	-5.03	106.40	112.54
3	D	793	ASN	N-CA-C	5.02	119.42	112.45

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	B	177	0	103	4	0
1	E	177	0	103	2	0
2	C	248	0	136	3	0
2	F	248	0	136	2	0
3	A	4652	0	4707	80	0
3	D	4652	0	4707	66	0
4	G	23	0	21	0	0
4	H	23	0	21	2	0
5	A	54	0	24	6	0
5	D	27	0	12	2	0
6	A	1	0	0	0	0
6	D	1	0	0	0	0
7	D	5	0	0	0	0
8	A	200	0	0	5	0
8	B	20	0	0	0	0
8	C	36	0	0	0	0
8	D	335	0	0	6	0
8	E	22	0	0	0	0
8	F	39	0	0	0	0
All	All	10940	0	9970	152	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 8.

All (152) 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)
3:D:634:ARG:NH2	3:D:876:LYS:HA	1.96	0.81
3:A:828:VAL:HB	3:A:831:GLU:HG3	1.65	0.79
3:D:658:GLU:CD	5:D:202:DCT:H2''	2.11	0.75
3:A:679:LEU:HD13	3:A:684:LYS:NZ	2.05	0.71
3:D:534:LEU:HD11	3:D:574:ILE:HD13	1.74	0.70
3:A:656:GLN:HB2	3:A:660:ARG:NH1	2.07	0.69
3:A:522:ALA:O	3:A:541:LYS:HE2	1.93	0.69
3:A:575:LEU:O	3:A:579:GLN:HG3	1.92	0.68
3:A:565:ALA:HA	3:A:571:VAL:HG21	1.77	0.67
3:A:747:LYS:HE3	3:A:751:GLU:OE1	1.96	0.66
3:D:408:ASP:HB2	4:H:2:FRU:H11	1.77	0.65
3:A:838:LYS:HG3	3:A:839:GLU:OE2	1.99	0.62
3:A:658:GLU:CD	5:A:201:DCT:H2''	2.24	0.62
3:A:349:GLU:H	3:A:349:GLU:CD	2.07	0.61
3:A:730:LYS:O	3:A:734:GLU:HG3	1.99	0.61
3:A:667:GLU:C	3:A:672:MET:HE2	2.26	0.60
3:D:713:VAL:O	3:D:750:MET:HE3	2.01	0.60
3:A:848:VAL:HB	3:A:849:PRO:HD3	1.83	0.59
3:A:833:ILE:N	3:A:833:ILE:HD12	2.18	0.59
3:A:524:GLN:HG2	3:A:525:GLU:N	2.18	0.59
3:A:584:GLN:HG3	8:A:4059:HOH:O	2.02	0.58
3:A:332:ILE:O	3:A:348:PRO:HG2	2.04	0.58
3:A:728:SER:HB3	3:A:731:GLU:HG3	1.85	0.58
3:D:848:VAL:HB	3:D:849:PRO:HD3	1.86	0.57
3:D:812:ASN:O	3:D:816:LYS:HG2	2.04	0.57
3:D:584:GLN:HG3	8:D:3721:HOH:O	2.05	0.57
3:D:472:ARG:HD2	3:D:476:GLU:OE1	2.05	0.56
3:D:500:LEU:HA	3:D:503:MET:HE3	1.87	0.56
1:B:28:DC:H2'	1:B:29:DDG:H8	1.87	0.56
3:D:605:ILE:HG21	4:H:1:GLC:H62	1.87	0.56
3:D:825:LEU:HD11	3:D:835:GLU:HB3	1.88	0.56
3:D:531:PRO:HD2	8:D:3472:HOH:O	2.06	0.55
3:D:517:ARG:NH2	3:D:521:LEU:HD21	2.21	0.55
3:D:840:GLU:O	3:D:844:LEU:HD23	2.06	0.55
3:A:679:LEU:HD13	3:A:684:LYS:HZ1	1.71	0.54
3:D:634:ARG:HH21	3:D:876:LYS:HA	1.72	0.54
3:D:843:ARG:HD2	8:D:3074:HOH:O	2.06	0.54
3:A:534:LEU:HD11	3:A:574:ILE:HD13	1.90	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:656:GLN:HA	5:D:202:DCT:O1B	2.08	0.53
3:D:634:ARG:HH22	3:D:876:LYS:HA	1.73	0.53
3:A:565:ALA:HA	3:A:571:VAL:CG2	2.37	0.53
3:D:634:ARG:HG2	3:D:634:ARG:HH11	1.73	0.53
3:A:459:ARG:HB2	3:A:460:PRO:HD3	1.91	0.52
3:D:430:GLY:O	3:D:435:ARG:HD2	2.10	0.52
3:D:728:SER:OG	3:D:731:GLU:HG3	2.10	0.52
3:A:656:GLN:HA	5:A:201:DCT:O1B	2.10	0.52
3:A:463:ASP:HA	3:A:466:ARG:NH1	2.25	0.52
3:A:500:LEU:HA	3:A:503:MET:HE3	1.92	0.52
3:A:699:PRO:O	3:A:703:ARG:HG3	2.10	0.52
3:A:568:HIS:O	3:A:571:VAL:HG22	2.09	0.51
3:D:757:ALA:HB3	3:D:776:ILE:HD13	1.92	0.51
3:D:561:LEU:O	3:D:571:VAL:HG11	2.10	0.51
3:D:587:TYR:CE2	3:D:627:PRO:HD3	2.45	0.51
1:B:26:DC:H2''	3:A:582:LYS:HG2	1.93	0.51
1:E:25:DA:H2'	1:E:26:DC:C6	2.46	0.51
3:A:779:ARG:HG3	3:D:676:ARG:NH2	2.26	0.51
3:A:698:THR:OG1	3:A:701:MET:HG3	2.11	0.50
3:A:463:ASP:HB2	8:A:4056:HOH:O	2.11	0.50
3:A:767:LEU:HA	5:A:203:DCT:H3'2	1.94	0.50
3:A:515:GLU:HG2	3:A:519:TYR:CE2	2.47	0.50
3:A:660:ARG:HH11	3:A:660:ARG:HG3	1.77	0.50
3:A:837:PRO:HG2	3:A:840:GLU:HG3	1.94	0.50
3:D:301:PHE:HE1	3:D:303:LEU:HD13	1.77	0.49
3:A:755:GLN:O	3:A:759:GLN:HG3	2.12	0.49
3:A:725:LEU:O	3:A:727:ILE:HG23	2.13	0.49
3:D:519:TYR:CD2	3:D:525:GLU:HG2	2.47	0.49
3:D:515:GLU:HG2	3:D:519:TYR:CE2	2.47	0.49
3:A:663:ALA:HB2	3:A:671:LEU:HG	1.95	0.48
3:D:657:ILE:HG23	3:D:658:GLU:N	2.28	0.48
3:A:411:ALA:O	3:A:415:LYS:HG2	2.14	0.48
3:D:415:LYS:HE3	8:D:3608:HOH:O	2.14	0.48
1:B:25:DA:H2'	1:B:26:DC:C6	2.49	0.48
3:A:610:LEU:HD23	3:A:610:LEU:C	2.39	0.47
3:A:819:ARG:HH11	3:A:819:ARG:HG3	1.79	0.47
3:A:657:ILE:HG23	3:A:658:GLU:N	2.29	0.47
3:A:656:GLN:HB2	3:A:660:ARG:HH12	1.77	0.47
3:A:775:ASP:HB3	3:A:783:VAL:HG12	1.97	0.47
3:D:364:GLU:HG2	8:D:3332:HOH:O	2.14	0.47
3:D:814:ARG:NH2	3:D:847:LEU:HD13	2.29	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:828:VAL:HB	3:A:831:GLU:CG	2.39	0.46
3:A:698:THR:HB	3:A:699:PRO:HD2	1.98	0.46
3:A:472:ARG:HB2	5:A:203:DCT:C2	2.46	0.46
3:A:767:LEU:HD22	5:A:203:DCT:H3'1	1.98	0.46
3:A:834:LEU:HD12	3:A:834:LEU:N	2.30	0.46
3:D:596:ARG:HD3	3:D:603:HIS:CD2	2.51	0.46
3:D:689:ILE:HD11	3:D:735:PHE:CZ	2.51	0.46
3:A:457:LEU:C	3:A:460:PRO:HD2	2.41	0.46
3:A:767:LEU:HD22	5:A:203:DCT:C3'	2.46	0.46
3:D:745:GLY:HA2	3:D:748:ARG:HE	1.81	0.46
3:A:466:ARG:NH1	3:A:466:ARG:HB2	2.31	0.46
1:B:26:DC:H5'	3:A:578:ARG:HB3	1.99	0.45
2:F:7:DG:H4'	3:D:618:SER:O	2.17	0.45
3:D:823:HIS:CE1	3:D:835:GLU:HG2	2.52	0.45
2:C:1:DA:H2''	2:C:2:DT:O5'	2.17	0.45
3:A:531:PRO:HD2	8:A:3574:HOH:O	2.16	0.45
3:A:784:ARG:O	3:A:788:GLU:HG3	2.16	0.45
3:D:634:ARG:NH1	3:D:873:TYR:CD1	2.85	0.45
3:D:674:ALA:HB1	3:D:681:ILE:HD11	1.98	0.45
3:A:319:VAL:HB	3:A:336:ALA:HB3	1.99	0.45
3:A:822:ALA:CB	3:A:836:ALA:HB2	2.47	0.45
3:A:439:ASP:OD2	3:A:441:PRO:HD2	2.17	0.45
3:A:440:GLU:HB3	3:A:441:PRO:HD3	1.99	0.45
3:D:524:GLN:HG2	3:D:525:GLU:N	2.31	0.45
2:C:9:DC:H2'	2:C:10:DA:C8	2.52	0.45
3:A:494:LYS:HE2	8:A:3540:HOH:O	2.17	0.45
3:A:377:ALA:O	3:A:381:LYS:HG3	2.17	0.44
3:D:816:LYS:NZ	8:D:3912:HOH:O	2.49	0.44
3:A:429:TYR:O	3:A:435:ARG:HA	2.18	0.44
3:A:728:SER:HB3	3:A:731:GLU:CG	2.47	0.44
3:D:867:HIS:HB2	3:D:875:ALA:O	2.18	0.44
3:D:297:LYS:C	3:D:299:MET:H	2.25	0.44
3:D:536:VAL:O	3:D:540:GLU:HB2	2.17	0.44
2:F:9:DC:H2'	2:F:10:DA:C8	2.53	0.44
3:A:655:SER:O	3:A:656:GLN:C	2.59	0.44
3:D:667:GLU:C	3:D:672:MET:HE2	2.43	0.44
3:D:645:SER:O	3:D:646:ASP:HB2	2.17	0.44
3:A:348:PRO:HD2	3:A:349:GLU:OE2	2.18	0.44
3:A:595:VAL:O	3:A:597:PRO:HD3	2.16	0.43
3:D:712:ILE:HA	3:D:716:ILE:HG22	2.00	0.43
3:D:517:ARG:HA	3:D:517:ARG:HD2	1.81	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:867:HIS:CG	3:D:876:LYS:HD3	2.53	0.43
1:E:28:DC:H2'	1:E:29:DDG:H8	2.00	0.43
3:D:502:GLN:O	3:D:506:GLU:HG3	2.18	0.43
3:D:647:TRP:CE2	3:D:837:PRO:HG3	2.53	0.43
3:D:565:ALA:N	3:D:566:PRO:CD	2.81	0.43
3:A:689:ILE:HD11	3:A:708:VAL:HG11	2.01	0.43
3:A:744:PRO:O	3:A:748:ARG:HG3	2.19	0.42
3:D:867:HIS:CB	3:D:876:LYS:HD3	2.48	0.42
3:D:614:GLY:HA3	3:D:769:ARG:HD3	2.01	0.42
3:D:692:VAL:HG21	3:D:701:MET:HE1	2.01	0.42
3:A:668:ASP:O	3:A:672:MET:HG3	2.19	0.42
3:D:517:ARG:HG2	3:D:517:ARG:HH11	1.84	0.42
3:A:693:SER:OG	3:A:696:GLU:HG3	2.18	0.42
3:D:823:HIS:HE1	3:D:835:GLU:OE2	2.02	0.42
3:D:435:ARG:HH11	3:D:435:ARG:HG2	1.85	0.42
3:A:589:GLU:O	3:A:593:LYS:HG3	2.19	0.42
3:A:558:ALA:O	3:A:562:GLU:HG3	2.19	0.42
3:A:596:ARG:HD3	3:A:603:HIS:CD2	2.55	0.42
3:D:690:PHE:O	3:D:692:VAL:HG13	2.20	0.42
3:A:466:ARG:HD2	8:A:3694:HOH:O	2.18	0.42
3:A:613:THR:O	3:A:798:GLY:HA2	2.20	0.41
3:D:559:ASP:O	3:D:563:LYS:HG3	2.20	0.41
2:C:7:DG:H4'	3:A:618:SER:O	2.20	0.41
3:D:459:ARG:O	3:D:459:ARG:HD3	2.20	0.41
3:A:679:LEU:HD13	3:A:684:LYS:HZ2	1.81	0.41
3:A:712:ILE:HA	3:A:716:ILE:HG22	2.01	0.41
3:D:599:THR:O	3:D:600:LYS:HB2	2.20	0.41
3:A:645:SER:O	3:A:646:ASP:HB2	2.21	0.40
3:D:847:LEU:O	3:D:851:VAL:HG23	2.21	0.40
3:D:614:GLY:HA3	3:D:769:ARG:CD	2.50	0.40
3:D:693:SER:HB3	3:D:696:GLU:HG3	2.02	0.40
3:A:418:GLN:N	3:A:418:GLN:CD	2.80	0.40

There are no symmetry-related clashes.

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	
3	A	578/580 (100%)	561 (97%)	16 (3%)	1 (0%)	43	35
3	D	578/580 (100%)	561 (97%)	16 (3%)	1 (0%)	43	35
All	All	1156/1160 (100%)	1122 (97%)	32 (3%)	2 (0%)	43	35

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	628	ILE
3	D	628	ILE

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	
3	A	495/496 (100%)	492 (99%)	3 (1%)	78	76
3	D	495/496 (100%)	495 (100%)	0	100	100
All	All	990/992 (100%)	987 (100%)	3 (0%)	86	86

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

Mol	Chain	Res	Type
3	A	349	GLU
3	A	834	LEU
3	A	844	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	529	ASN
3	A	579	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	608	GLN
3	A	723	GLN
3	A	782	ASN
3	D	502	GLN
3	D	608	GLN
3	D	624	GLN
3	D	625	ASN
3	D	782	ASN
3	D	823	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	$\# Z > 2$
1	DDG	E	29	1,2	20,23,24	0.37	0	27,33,36	0.34	0
1	DDG	B	29	1,2	20,23,24	0.34	0	27,33,36	0.31	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
1	DDG	E	29	1,2	-	0/7/18/19	0/3/3/3
1	DDG	B	29	1,2	-	0/7/18/19	0/3/3/3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	29	DDG	1	0
1	B	29	DDG	1	0

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	$\# Z > 2$
4	GLC	G	1	4	11,11,12	3.37	4 (36%)	15,15,17	1.52	2 (13%)
4	FRU	G	2	4	11,12,12	1.60	2 (18%)	10,18,18	0.83	0
4	GLC	H	1	4	11,11,12	3.32	4 (36%)	15,15,17	1.52	3 (20%)
4	FRU	H	2	4	11,12,12	1.52	1 (9%)	10,18,18	0.82	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
4	GLC	G	1	4	-	0/2/19/22	0/1/1/1
4	FRU	G	2	4	-	2/5/24/24	0/1/1/1
4	GLC	H	1	4	-	0/2/19/22	0/1/1/1
4	FRU	H	2	4	-	0/5/24/24	0/1/1/1

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	1	GLC	C2-C3	9.75	1.67	1.52
4	H	1	GLC	C2-C3	9.62	1.67	1.52
4	G	2	FRU	O2-C2	4.43	1.48	1.40
4	H	2	FRU	O2-C2	4.31	1.48	1.40
4	G	1	GLC	O5-C1	2.89	1.48	1.43
4	G	1	GLC	O5-C5	2.80	1.48	1.43
4	H	1	GLC	O5-C5	2.80	1.48	1.43
4	H	1	GLC	O5-C1	2.78	1.48	1.43
4	G	1	GLC	C4-C5	2.43	1.58	1.53
4	H	1	GLC	C4-C5	2.26	1.57	1.53
4	G	2	FRU	C1-C2	2.01	1.57	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	1	GLC	C1-O5-C5	4.32	117.97	112.19
4	G	1	GLC	C1-O5-C5	4.25	117.88	112.19
4	G	1	GLC	C1-C2-C3	-2.52	105.97	109.64
4	H	1	GLC	C1-C2-C3	-2.44	106.09	109.64
4	H	1	GLC	O3-C3-C2	-2.12	105.74	110.05

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	G	2	FRU	O1-C1-C2-C3
4	G	2	FRU	O1-C1-C2-O5

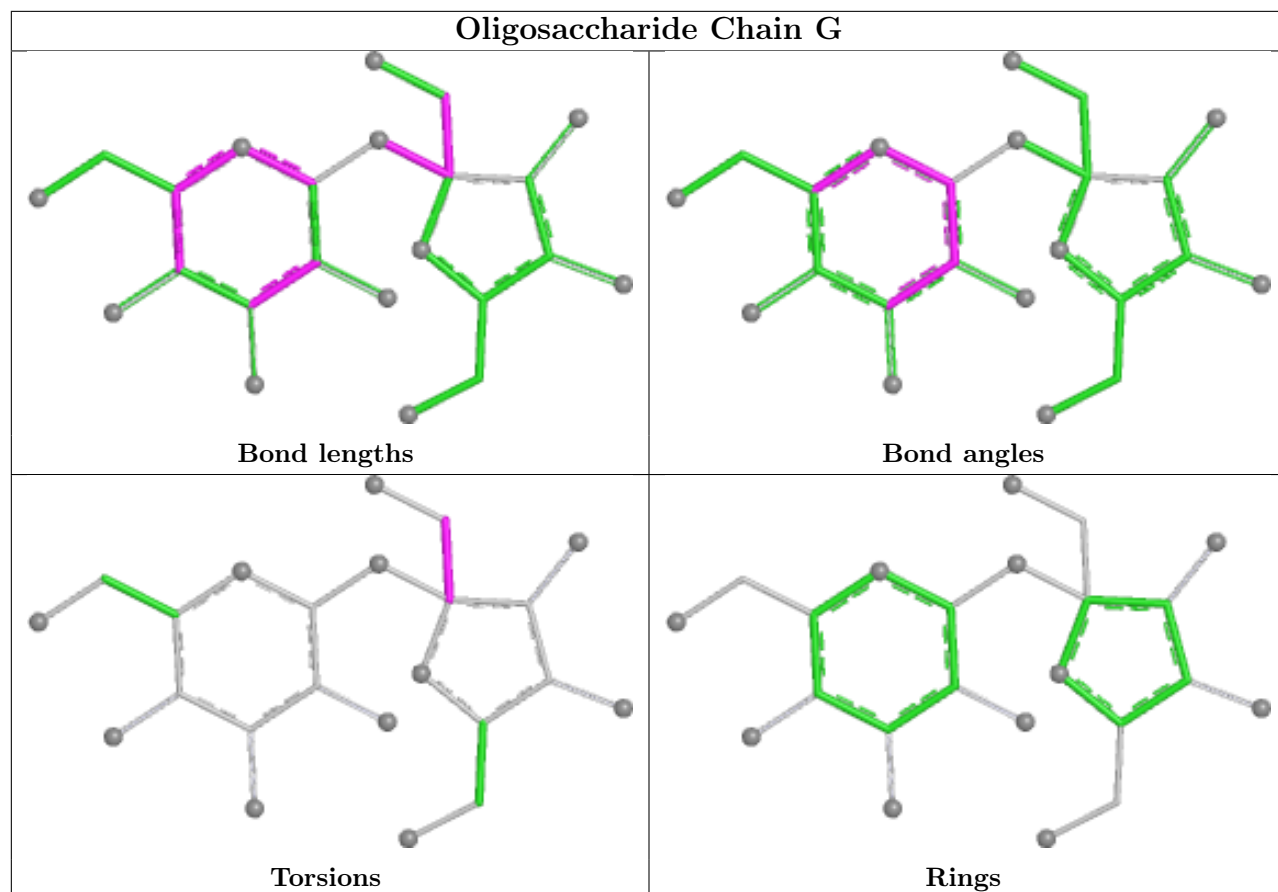
There are no ring outliers.

2 monomers are involved in 2 short contacts:

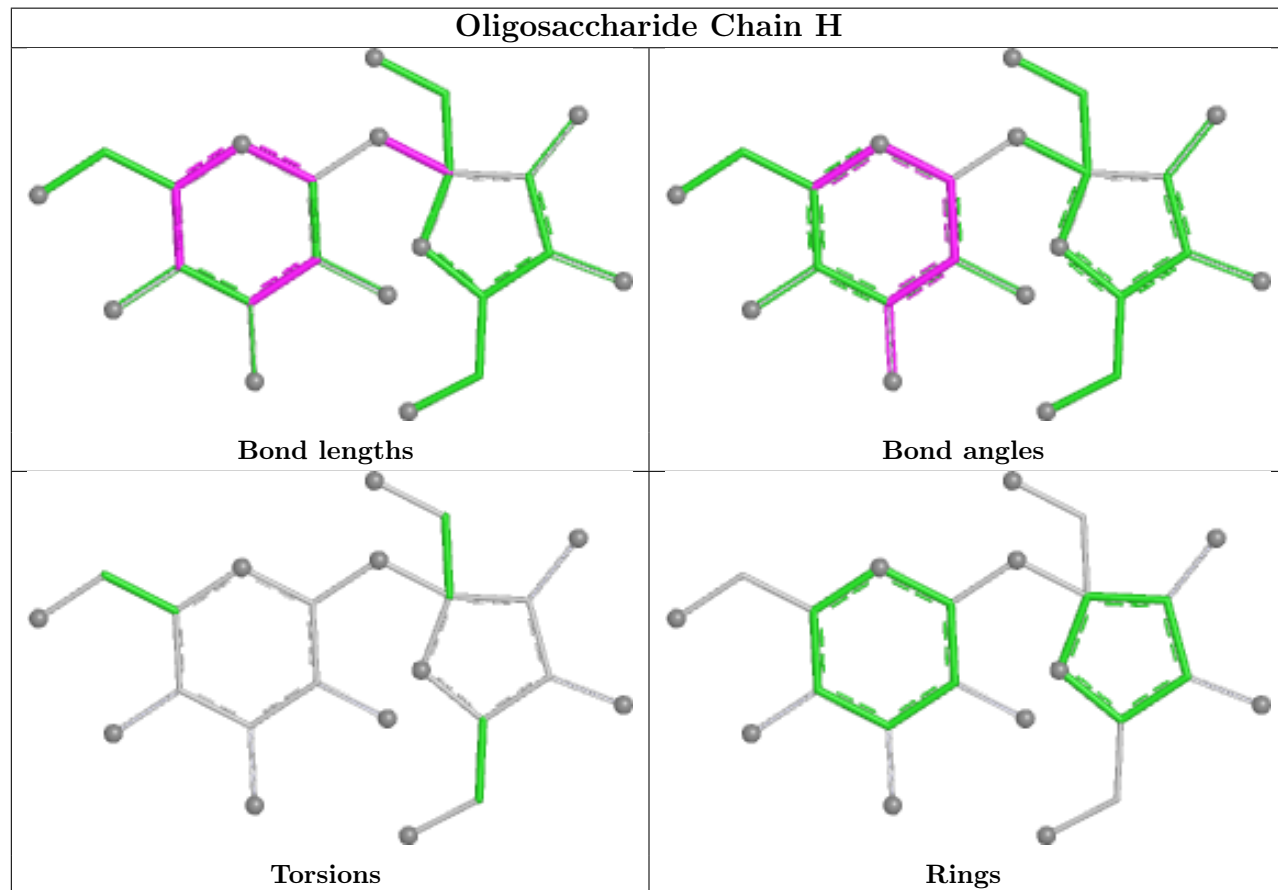
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	1	GLC	1	0
4	H	2	FRU	1	0

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 G



Oligosaccharide Chain H



5.6 Ligand geometry

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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	DCT	A	203	-	27,28,28	1.30	2 (7%)	37,43,43	1.08	2 (5%)
5	DCT	A	201	6	27,28,28	1.31	3 (11%)	37,43,43	1.14	3 (8%)
7	SO4	D	879	-	4,4,4	0.32	0	6,6,6	0.06	0
5	DCT	D	202	6	27,28,28	1.31	2 (7%)	37,43,43	1.16	3 (8%)

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	DCT	A	203	-	-	6/22/31/31	0/2/2/2
5	DCT	A	201	6	-	4/22/31/31	0/2/2/2
5	DCT	D	202	6	-	5/22/31/31	0/2/2/2

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	202	DCT	C4-N3	3.05	1.40	1.34
5	A	203	DCT	C4-N3	3.03	1.40	1.34
5	A	201	DCT	C4-N3	2.96	1.40	1.34
5	A	203	DCT	C2-N3	2.10	1.40	1.36
5	D	202	DCT	C2-N3	2.07	1.40	1.36
5	A	201	DCT	C2-N3	2.07	1.40	1.36
5	A	201	DCT	PB-O3B	2.02	1.61	1.59

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	203	DCT	C3'-C2'-C1'	-3.66	98.66	102.87

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	202	DCT	C2'-C3'-C4'	-3.35	96.53	102.75
5	A	201	DCT	C2'-C3'-C4'	-3.21	96.78	102.75
5	A	201	DCT	C3'-C2'-C1'	-2.70	99.76	102.87
5	D	202	DCT	C4'-O4'-C1'	-2.60	107.35	109.81
5	A	201	DCT	C4'-O4'-C1'	-2.47	107.47	109.81
5	A	203	DCT	O4'-C1'-N1	2.36	112.05	107.86
5	D	202	DCT	C3'-C2'-C1'	-2.33	100.19	102.87

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	203	DCT	C5'-O5'-PA-O2A
5	A	203	DCT	C5'-O5'-PA-O3A
5	D	202	DCT	PB-O3B-PG-O1G
5	D	202	DCT	PB-O3B-PG-O2G
5	A	201	DCT	PG-O3B-PB-O2B
5	D	202	DCT	PG-O3B-PB-O2B
5	A	201	DCT	PB-O3B-PG-O1G
5	A	201	DCT	PB-O3A-PA-O2A
5	D	202	DCT	PB-O3A-PA-O2A
5	A	203	DCT	PA-O3A-PB-O1B
5	A	203	DCT	PA-O3A-PB-O2B
5	A	203	DCT	PG-O3B-PB-O2B
5	D	202	DCT	PG-O3B-PB-O1B
5	A	201	DCT	PG-O3B-PB-O1B
5	A	203	DCT	PG-O3B-PB-O1B

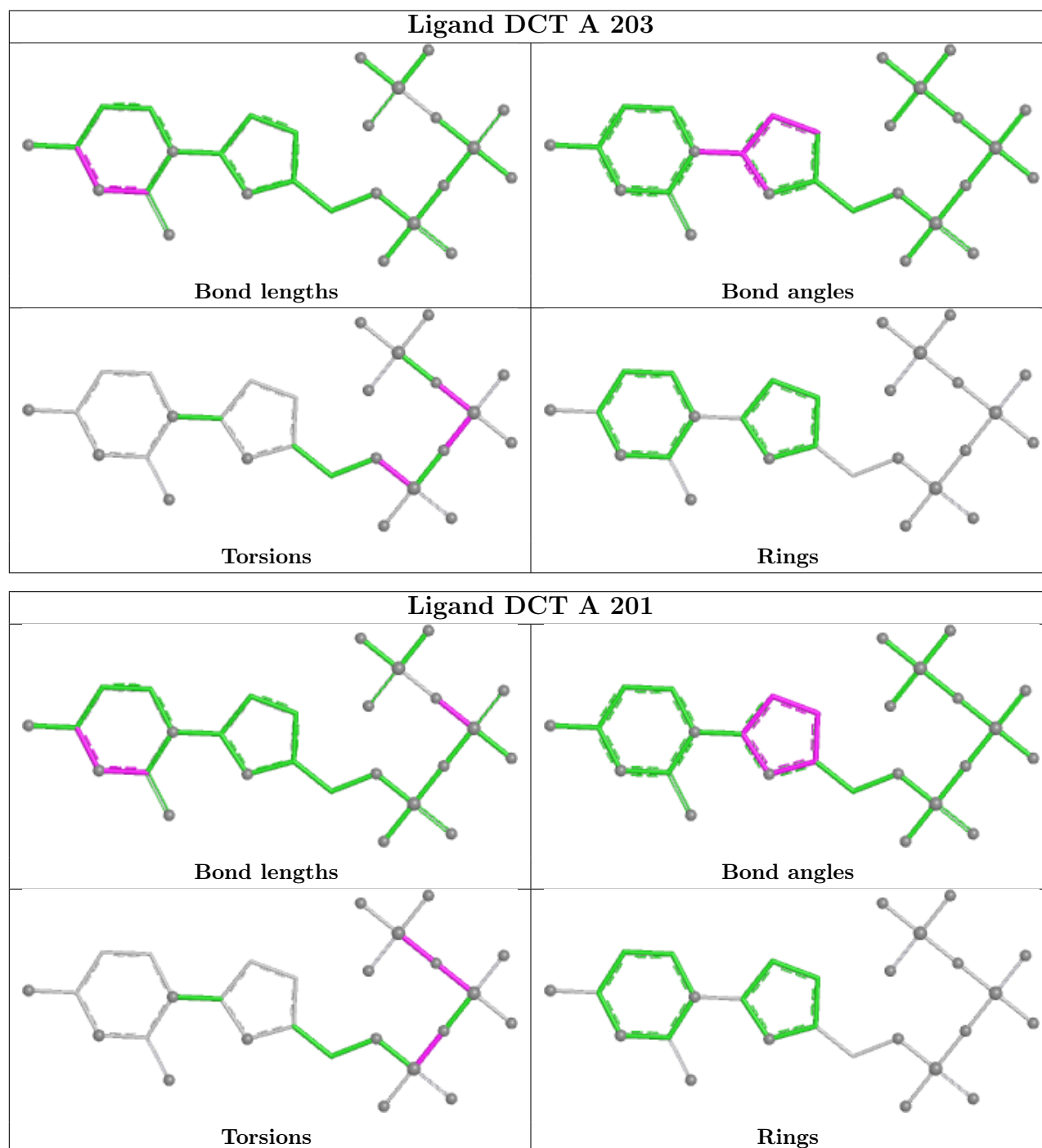
There are no ring outliers.

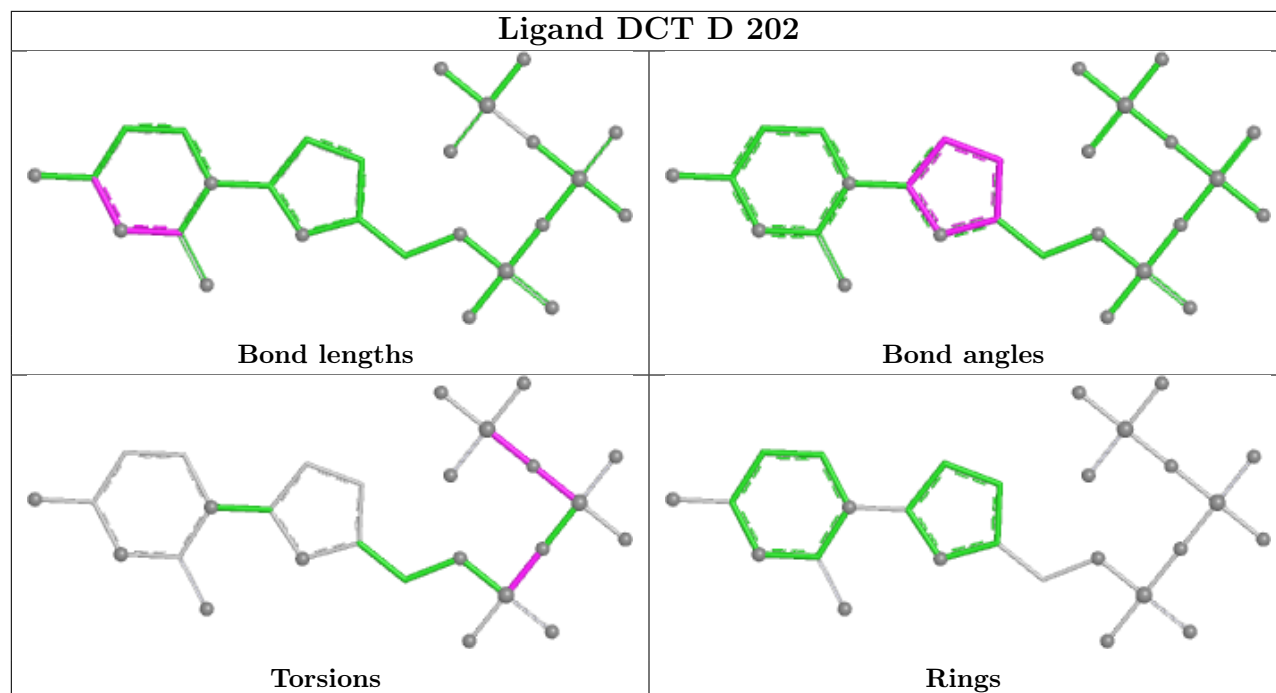
3 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	203	DCT	4	0
5	A	201	DCT	2	0
5	D	202	DCT	2	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 than 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 [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

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	B	8/9 (88%)	0.29	1 (12%) 8 11	22, 31, 57, 72	0
1	E	8/9 (88%)	0.09	1 (12%) 8 11	18, 27, 47, 69	0
2	C	12/13 (92%)	0.02	1 (8%) 17 24	21, 28, 63, 87	0
2	F	12/13 (92%)	-0.18	0 100 100	18, 24, 58, 80	0
3	A	580/580 (100%)	0.40	20 (3%) 48 58	19, 32, 53, 71	0
3	D	580/580 (100%)	0.00	9 (1%) 70 79	14, 24, 42, 55	0
All	All	1200/1204 (99%)	0.20	32 (2%) 56 66	14, 29, 48, 87	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	297	LYS	5.9
3	A	297	LYS	5.0
3	A	298	LYS	3.7
3	A	724	ASN	3.5
3	D	531	PRO	3.5
1	E	21	DC	3.2
3	D	876	LYS	3.1
2	C	1	DA	3.1
3	A	689	ILE	3.0
3	A	697	VAL	2.8
1	B	21	DC	2.8
3	D	695	ASP	2.7
3	A	779	ARG	2.6
3	A	695	ASP	2.6
3	D	298	LYS	2.6
3	A	304	ALA	2.5
3	A	781	PHE	2.4
3	A	674	ALA	2.4
3	A	466	ARG	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	D	303	LEU	2.3
3	A	690	PHE	2.2
3	A	698	THR	2.2
3	A	699	PRO	2.2
3	D	817	GLU	2.2
3	D	779	ARG	2.2
3	A	559	ASP	2.1
3	A	646	ASP	2.1
3	A	692	VAL	2.1
3	A	552	THR	2.1
3	A	670	ASN	2.0
3	A	780	ASN	2.0
3	D	726	ASN	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	DDG	B	29	21/22	0.97	0.06	20,23,24,25	0
1	DDG	E	29	21/22	0.98	0.05	14,16,19,21	0

6.3 Carbohydrates [i](#)

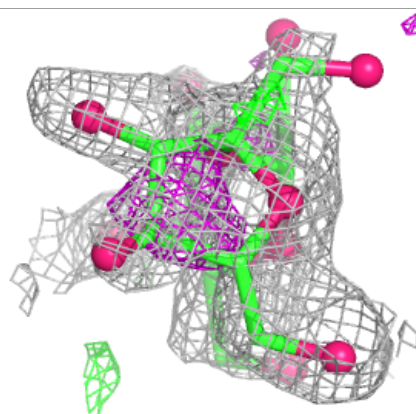
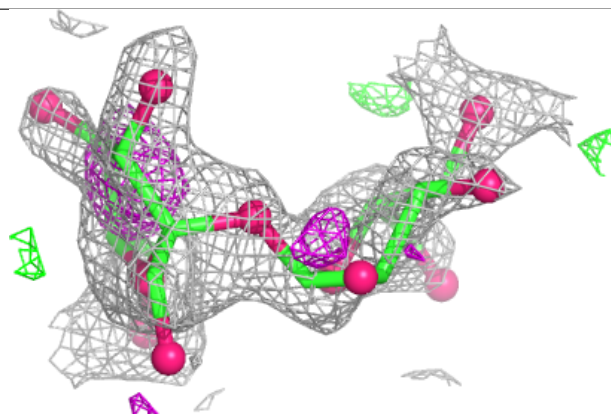
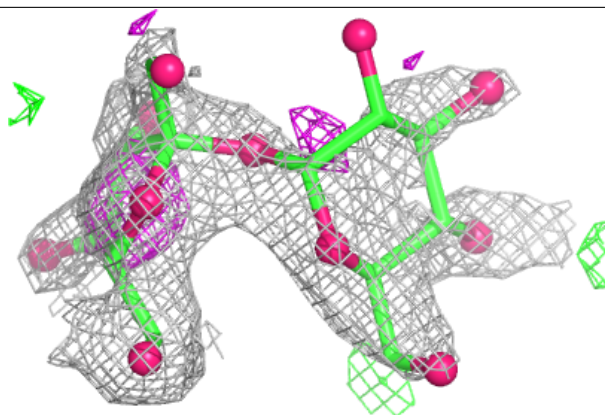
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	GLC	G	1	11/12	0.75	0.15	58,59,60,60	0
4	FRU	G	2	12/12	0.82	0.14	48,55,57,59	0
4	GLC	H	1	11/12	0.85	0.13	50,52,53,56	0
4	FRU	H	2	12/12	0.88	0.11	42,48,49,50	0

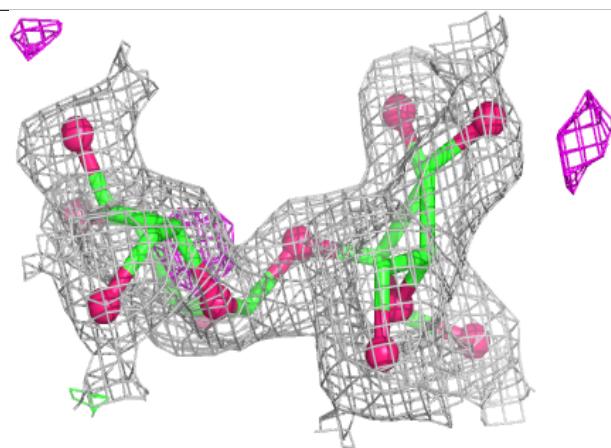
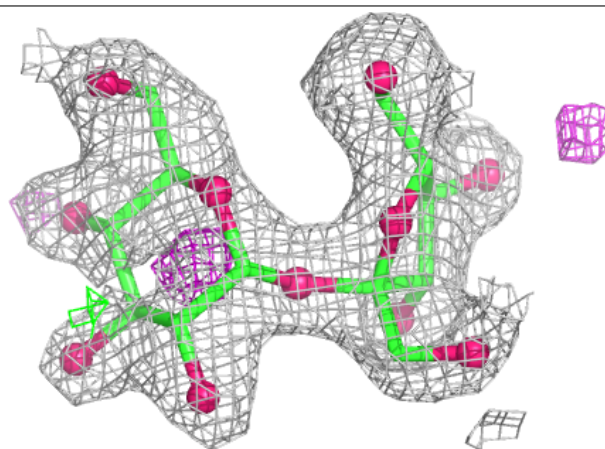
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain G:

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

**Electron density around Chain H:**

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

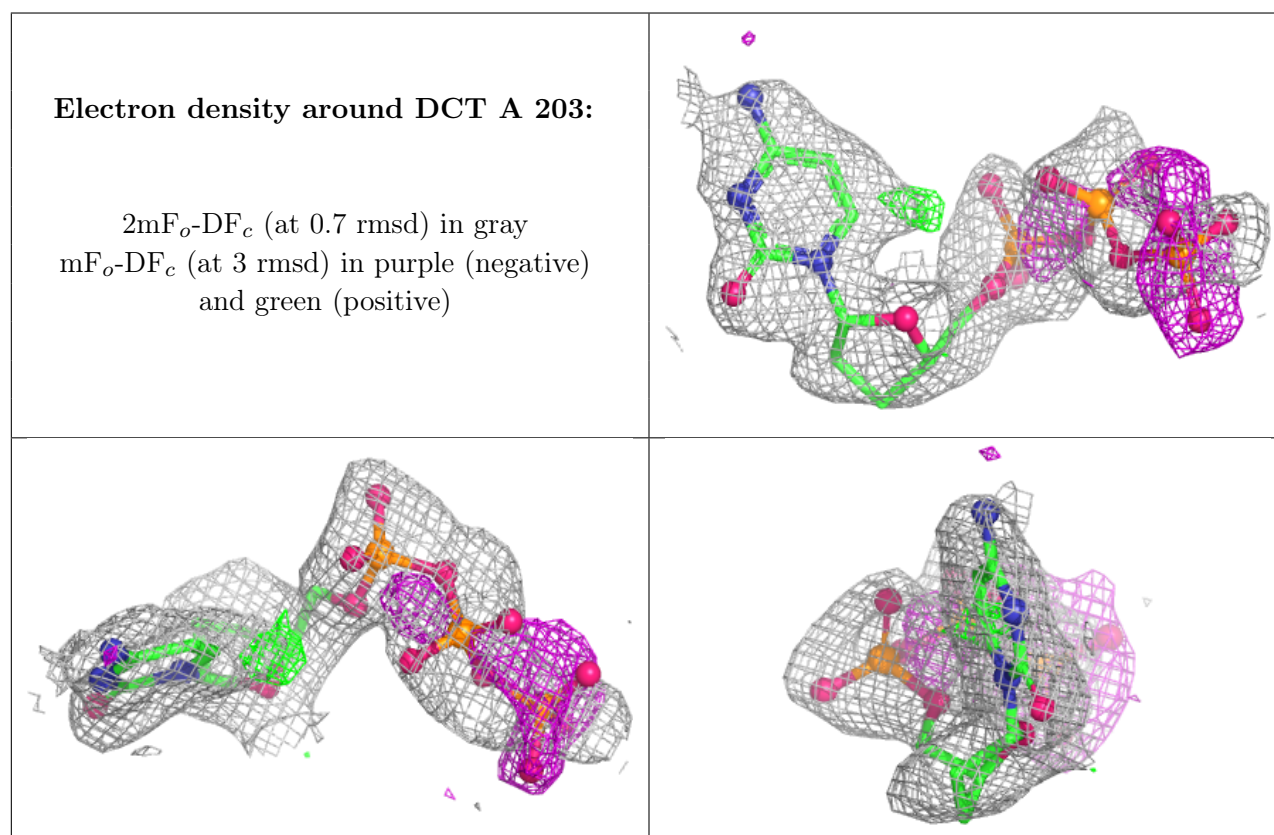


6.4 Ligands ⓘ

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.

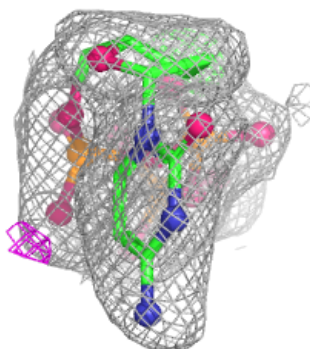
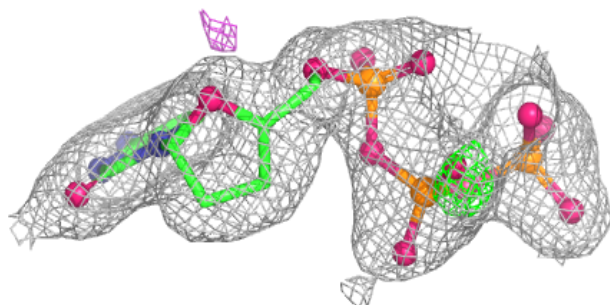
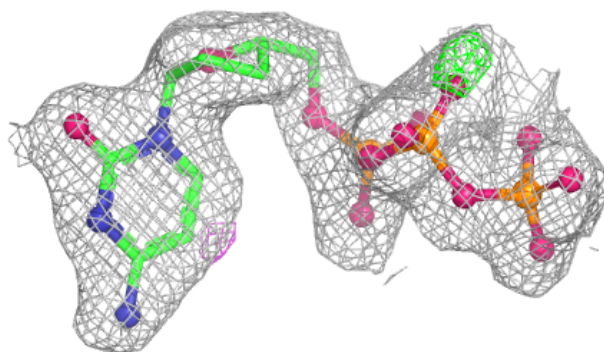
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
5	DCT	A	203	27/27	0.71	0.14	48,54,78,80	0
7	SO4	D	879	5/5	0.76	0.37	92,92,92,92	0
6	MG	D	878	1/1	0.95	0.05	20,20,20,20	0
5	DCT	A	201	27/27	0.96	0.06	22,25,32,32	0
5	DCT	D	202	27/27	0.97	0.07	20,23,33,35	0
6	MG	A	878	1/1	0.97	0.03	24,24,24,24	0

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.

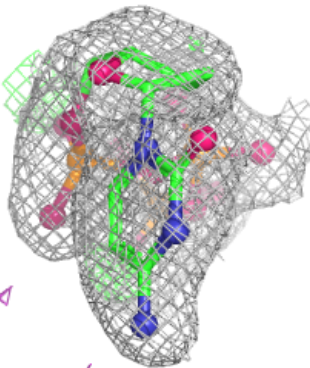
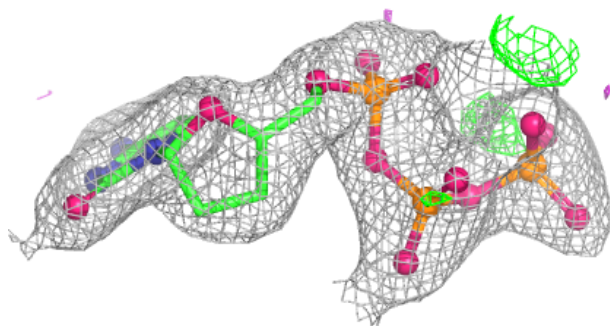
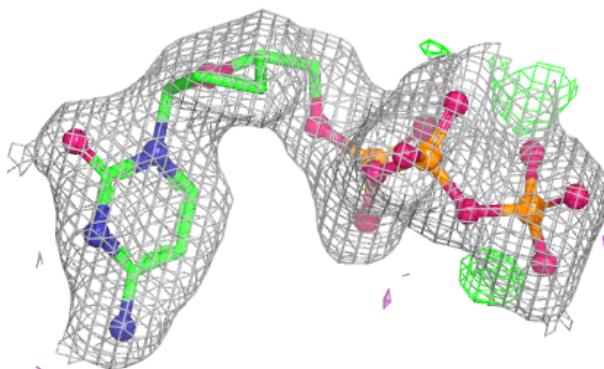


Electron density around DCT A 201:

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

**Electron density around DCT D 202:**

$2mF_o-DF_c$ (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.