Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2023 – 02:35 AM EDT

PDB ID	:	2HHW
Title	:	ddTTP:O6-methyl-guanine pair in the polymerase active site, in the closed conformation
Authors	:	Warren, J.J.; Forsberg, L.J.; Beese, L.S.
Deposited on	:	2006-06-28
Resolution	:	1.88 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:

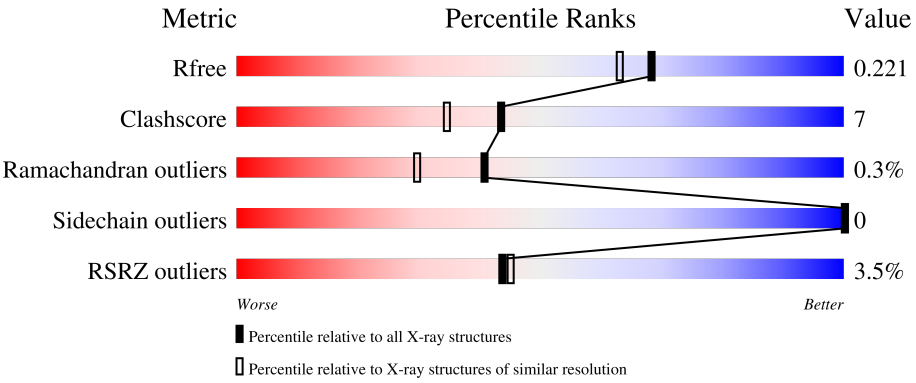
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 1.88 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.



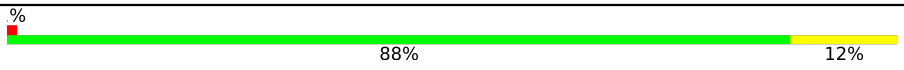
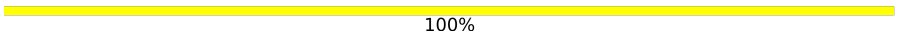
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	B	9	11% 33% 56% 11%
1	E	9	67% 22% 11%
2	C	13	62% 23% 15%
2	F	13	62% 23% 15%
3	A	580	6% 79% 21%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	580	
4	G	2	
4	H	2	

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 11027 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	E	9	Total	C	N	O	P	0	0	0
			177	86	31	52	8			
1	B	9	Total	C	N	O	P	0	0	0
			177	86	31	52	8			

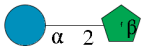
- Molecule 2 is a DNA chain called 5'-D(*CP*AP*TP*(6OG)P*CP*GP*AP*GP*TP*CP*AP*GP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	11	Total	C	N	O	P	0	0	0
			228	109	45	64	10			
2	C	11	Total	C	N	O	P	0	0	0
			228	109	45	64	10			

- 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...

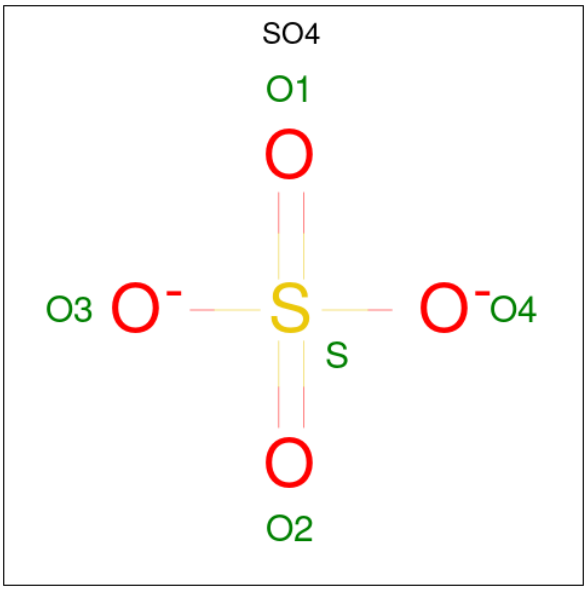
Continued from previous page...

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

- Molecule 5 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

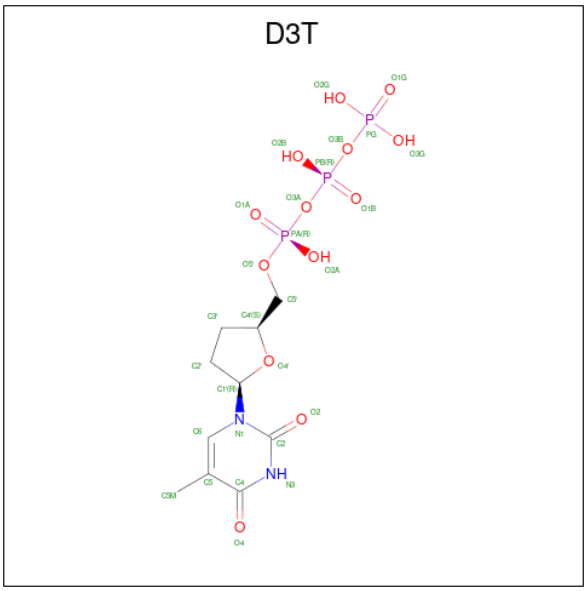
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Mn 1 1	0	0
5	D	1	Total Mn 1 1	0	0

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



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

- Molecule 7 is 2',3'-DIDEOXY-THYMIDINE-5'-TRIPHOSPHATE (three-letter code: D3T) (formula: $C_{10}H_{17}N_2O_{13}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	P	0	0
			28	10	2	13	3		
7	D	1	Total	C	N	O	P	0	0
			28	10	2	13	3		

- Molecule 8 is water.

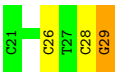
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	E	28	Total	O	0	0
			28	28		
8	F	39	Total	O	0	0
			39	39		
8	B	30	Total	O	0	0
			30	30		
8	C	42	Total	O	0	0
			42	42		
8	A	239	Total	O	0	0
			239	239		
8	D	416	Total	O	0	0
			416	416		

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.

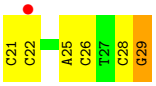
• Molecule 1: 5'-D(*CP*CP*TP*GP*AP*CP*TP*CP*(DDG))-3'

Chain E:



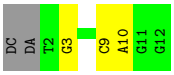
• Molecule 1: 5'-D(*CP*CP*TP*GP*AP*CP*TP*CP*(DDG))-3'

Chain B:



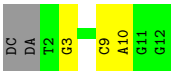
• Molecule 2: 5'-D(*CP*AP*TP*(6OG)P*CP*GP*AP*GP*TP*CP*AP*GP*G)-3'

Chain F:



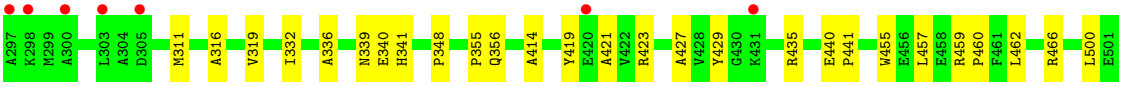
• Molecule 2: 5'-D(*CP*AP*TP*(6OG)P*CP*GP*AP*GP*TP*CP*AP*GP*G)-3'

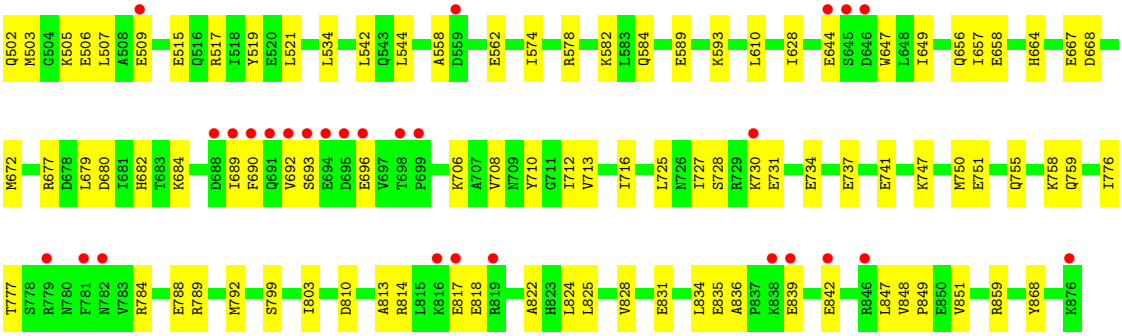
Chain C:



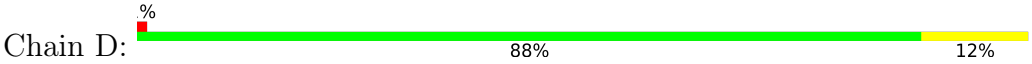
• Molecule 3: DNA Polymerase I

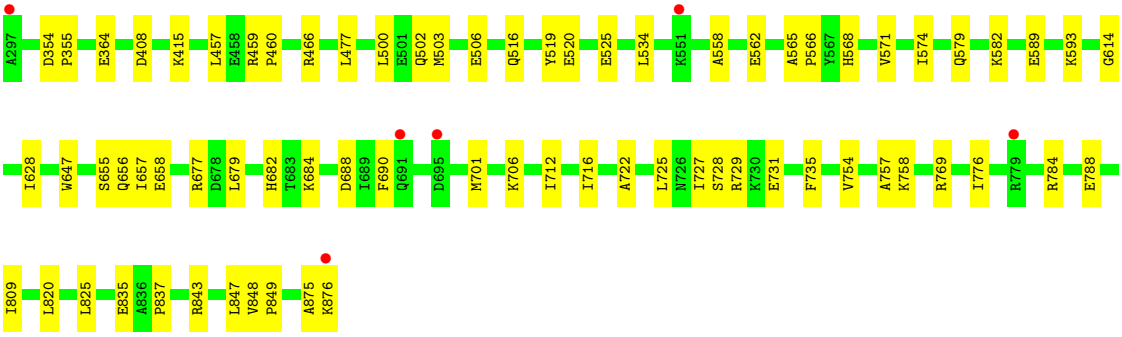
Chain A:





• Molecule 3: DNA Polymerase I





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

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

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	93.83Å 108.90Å 150.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.88 36.84 – 1.88	Depositor EDS
% Data completeness (in resolution range)	90.6 (50.00-1.88) 90.6 (36.84-1.88)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 1.88Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.196 , 0.225 0.192 , 0.221	Depositor DCC
R_{free} test set	5689 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	17.7	Xtriage
Anisotropy	0.204	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 57.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11027	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.35% 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: 6OG, DDG, SO4, FRU, D3T, MN, 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	0.36	0/173	0.81	0/264
1	E	0.45	0/173	0.73	0/264
2	C	0.36	0/229	0.72	0/350
2	F	0.39	0/229	0.71	0/350
3	A	0.28	0/4736	0.53	0/6400
3	D	0.32	0/4736	0.58	0/6400
All	All	0.31	0/10276	0.58	0/14028

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	9	0
1	E	177	0	103	2	0
2	C	228	0	127	1	0
2	F	228	0	127	1	0
3	A	4652	0	4710	77	0
3	D	4652	0	4710	53	0
4	G	23	0	21	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	H	23	0	21	1	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
6	A	10	0	0	0	0
6	D	5	0	0	0	0
7	A	28	0	13	2	0
7	D	28	0	13	3	0
8	A	239	0	0	2	0
8	B	30	0	0	0	0
8	C	42	0	0	0	0
8	D	416	0	0	3	0
8	E	28	0	0	1	0
8	F	39	0	0	0	0
All	All	11027	0	9948	136	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 7.

All (136) 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:A:558:ALA:O	3:A:562:GLU:HG3	1.90	0.71
3:A:828:VAL:HB	3:A:831:GLU:CG	2.21	0.71
3:D:534:LEU:HD11	3:D:574:ILE:HD13	1.73	0.70
3:D:558:ALA:O	3:D:562:GLU:HG3	1.92	0.69
3:A:842:GLU:HG3	8:A:1066:HOH:O	1.94	0.69
3:D:875:ALA:O	3:D:876:LYS:O	2.13	0.66
3:D:658:GLU:CD	7:D:880:D3T:H2'2	2.17	0.65
3:A:737:GLU:O	3:A:741:GLU:HG3	1.97	0.65
3:D:679:LEU:HD13	3:D:684:LYS:NZ	2.13	0.64
3:A:725:LEU:O	3:A:727:ILE:HG23	1.98	0.64
3:A:789:ARG:HA	3:A:792:MET:HE3	1.79	0.63
3:D:459:ARG:HB3	3:D:460:PRO:HD3	1.80	0.62
3:D:656:GLN:HA	7:D:880:D3T:O2B	1.99	0.62
3:A:658:GLU:CD	7:A:881:D3T:H2'2	2.19	0.62
3:A:730:LYS:O	3:A:734:GLU:HG3	1.99	0.61
3:A:828:VAL:HB	3:A:831:GLU:HG2	1.83	0.61
3:A:459:ARG:HB3	3:A:460:PRO:HD3	1.82	0.61
3:A:462:LEU:O	3:A:466:ARG:HG3	2.01	0.60
3:D:754:VAL:HG12	3:D:758:LYS:HE2	1.84	0.60
3:A:747:LYS:O	3:A:751:GLU:HG3	2.02	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:664:HIS:ND1	3:A:859:ARG:HG3	2.17	0.59
3:A:784:ARG:O	3:A:788:GLU:HG3	2.02	0.59
3:D:843:ARG:HD2	8:D:931:HOH:O	2.02	0.59
3:A:680:ASP:O	3:A:684:LYS:HD3	2.03	0.58
3:A:839:GLU:H	3:A:839:GLU:CD	2.07	0.58
3:D:516:GLN:O	3:D:520:GLU:HG3	2.03	0.58
3:A:848:VAL:HB	3:A:849:PRO:HD3	1.84	0.58
3:A:825:LEU:HD11	3:A:835:GLU:HB3	1.86	0.57
3:D:725:LEU:O	3:D:727:ILE:HG23	2.04	0.57
1:B:26:DC:H1'	3:A:582:LYS:HG3	1.88	0.56
3:A:414:ALA:HB1	3:A:419:TYR:HB3	1.87	0.56
3:D:688:ASP:OD2	3:D:735:PHE:HE1	1.87	0.55
8:E:295:HOH:O	3:D:579:GLN:HG2	2.06	0.55
3:D:848:VAL:HB	3:D:849:PRO:HD3	1.87	0.55
3:A:319:VAL:HB	3:A:336:ALA:HB3	1.88	0.55
1:B:25:DA:H5"	3:A:578:ARG:NH2	2.22	0.54
3:D:825:LEU:HD11	3:D:835:GLU:HB3	1.89	0.54
3:A:500:LEU:HA	3:A:503:MET:HE2	1.90	0.54
3:A:502:GLN:O	3:A:506:GLU:HG3	2.08	0.53
3:D:679:LEU:HD13	3:D:684:LYS:HZ3	1.72	0.53
3:D:847:LEU:C	3:D:847:LEU:HD23	2.28	0.53
3:A:507:LEU:HD22	3:A:584:GLN:HB2	1.90	0.53
3:A:589:GLU:O	3:A:593:LYS:HG3	2.08	0.53
3:A:657:ILE:HG23	3:A:658:GLU:N	2.24	0.52
3:D:565:ALA:HA	3:D:571:VAL:HG21	1.90	0.52
3:A:682:HIS:CE1	3:A:706:LYS:HG3	2.45	0.52
3:A:667:GLU:HG2	3:A:859:ARG:HE	1.75	0.52
3:A:776:ILE:HG23	3:A:777:THR:HG23	1.91	0.51
3:D:593:LYS:HB3	3:D:593:LYS:NZ	2.26	0.51
3:D:754:VAL:CG1	3:D:758:LYS:HE2	2.40	0.50
1:B:25:DA:H2'	1:B:26:DC:C6	2.46	0.50
3:A:644:GLU:HB2	3:A:647:TRP:CG	2.47	0.50
3:D:690:PHE:CD2	3:D:701:MET:HE2	2.46	0.50
3:A:534:LEU:HD11	3:A:574:ILE:HD13	1.93	0.50
3:A:668:ASP:O	3:A:672:MET:HG3	2.11	0.50
3:A:584:GLN:O	3:A:589:GLU:HG3	2.13	0.49
3:A:355:PRO:HG2	3:A:356:GLN:OE1	2.12	0.49
1:B:26:DC:H5'	3:A:578:ARG:HB3	1.94	0.49
3:A:758:LYS:HG2	3:A:776:ILE:HG13	1.94	0.48
3:A:515:GLU:HG2	3:A:519:TYR:CE2	2.48	0.48
3:D:457:LEU:C	3:D:460:PRO:HD2	2.33	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:568:HIS:O	3:D:571:VAL:HG22	2.12	0.48
3:D:589:GLU:O	3:D:593:LYS:HG3	2.12	0.48
3:D:657:ILE:HG23	3:D:658:GLU:N	2.29	0.48
3:D:784:ARG:O	3:D:788:GLU:HG3	2.14	0.48
1:B:21:DC:H2'	1:B:22:DC:C6	2.48	0.48
3:A:755:GLN:O	3:A:759:GLN:HG3	2.14	0.48
3:D:757:ALA:HB3	3:D:776:ILE:HD13	1.95	0.48
3:A:457:LEU:C	3:A:460:PRO:HD2	2.35	0.47
2:C:9:DC:H2'	2:C:10:DA:C8	2.50	0.47
3:D:502:GLN:O	3:D:506:GLU:HG3	2.15	0.47
3:A:693:SER:OG	3:A:696:GLU:HG3	2.15	0.47
3:A:728:SER:HB3	3:A:731:GLU:HG3	1.96	0.47
3:D:677:ARG:HB2	3:D:679:LEU:HG	1.97	0.47
3:D:690:PHE:CG	3:D:701:MET:HE2	2.49	0.47
3:A:610:LEU:HD23	3:A:610:LEU:C	2.35	0.47
1:E:28:DC:H2'	1:E:29:DDG:H8	1.96	0.46
3:A:824:LEU:HD23	3:A:834:LEU:HG	1.98	0.46
3:D:647:TRP:CE2	3:D:837:PRO:HG3	2.51	0.46
3:A:810:ASP:O	3:A:814:ARG:HG2	2.16	0.45
3:A:814:ARG:O	3:A:818:GLU:HG3	2.16	0.45
3:A:340:GLU:HG3	3:A:341:HIS:CD2	2.52	0.45
2:F:9:DC:H2'	2:F:10:DA:C8	2.51	0.45
3:A:429:TYR:O	3:A:435:ARG:HA	2.16	0.45
3:D:457:LEU:O	3:D:460:PRO:HD2	2.17	0.45
3:A:690:PHE:O	3:A:692:VAL:HG13	2.17	0.44
3:A:847:LEU:O	3:A:851:VAL:HG23	2.18	0.44
3:D:500:LEU:HD12	3:D:503:MET:HE3	2.00	0.44
1:B:25:DA:H5"	3:A:578:ARG:CZ	2.47	0.44
3:A:667:GLU:HG2	3:A:859:ARG:NE	2.32	0.44
3:D:354:ASP:HA	3:D:355:PRO:HD3	1.89	0.44
3:D:679:LEU:HD13	3:D:684:LYS:HZ1	1.82	0.44
3:A:710:TYR:HA	3:A:713:VAL:HG22	1.98	0.44
3:A:332:ILE:O	3:A:348:PRO:HG2	2.18	0.44
3:A:813:ALA:O	3:A:817:GLU:HG3	2.17	0.44
3:A:799:SER:O	3:A:803:ILE:HG13	2.18	0.43
3:D:712:ILE:HA	3:D:716:ILE:HG22	2.00	0.43
3:D:466:ARG:HG2	3:D:466:ARG:HH11	1.84	0.43
3:D:614:GLY:HA3	3:D:769:ARG:HD3	2.01	0.43
3:A:507:LEU:C	3:A:507:LEU:HD23	2.39	0.43
3:A:712:ILE:HA	3:A:716:ILE:HG22	1.99	0.43
3:D:477:LEU:HD12	3:D:809:ILE:HD12	1.99	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:505:LYS:O	3:A:509:GLU:HG3	2.19	0.43
1:B:26:DC:C1'	3:A:582:LYS:HG3	2.49	0.43
3:A:311:MET:HA	3:A:339:ASN:HD21	1.84	0.42
3:A:457:LEU:O	3:A:460:PRO:HD2	2.19	0.42
3:A:656:GLN:HA	7:A:881:D3T:O1B	2.19	0.42
3:A:517:ARG:NH2	3:A:521:LEU:HD21	2.34	0.42
3:A:859:ARG:HG2	3:A:859:ARG:HH11	1.83	0.42
1:E:26:DC:H2"	3:D:582:LYS:HG2	2.02	0.42
1:B:28:DC:H2'	1:B:29:DDG:H8	2.01	0.42
3:D:706:LYS:HD2	7:D:880:D3T:H5M2	2.02	0.42
3:D:364:GLU:HG2	8:D:1054:HOH:O	2.19	0.42
3:D:415:LYS:HE3	8:D:1245:HOH:O	2.19	0.42
3:A:423:ARG:NH2	3:A:427:ALA:O	2.53	0.42
3:D:408:ASP:HB2	4:H:2:FRU:H11	2.01	0.42
3:A:507:LEU:HD23	3:A:507:LEU:O	2.19	0.41
3:A:677:ARG:CB	3:A:679:LEU:HD13	2.49	0.41
3:D:728:SER:HB3	3:D:731:GLU:CD	2.40	0.41
3:A:657:ILE:HG23	3:A:658:GLU:H	1.84	0.41
3:A:822:ALA:CB	3:A:836:ALA:HB2	2.51	0.41
3:A:649:ILE:O	3:A:868:TYR:HA	2.20	0.41
3:A:750:MET:HG3	8:A:1051:HOH:O	2.21	0.41
3:D:614:GLY:HA3	3:D:769:ARG:CD	2.50	0.41
3:D:682:HIS:CE1	3:D:706:LYS:HG3	2.55	0.41
3:D:722:ALA:HB2	3:D:729:ARG:HA	2.02	0.41
1:B:26:DC:H2"	3:A:582:LYS:HG3	2.03	0.41
3:D:820:LEU:HD21	3:D:843:ARG:CZ	2.51	0.41
3:A:689:ILE:HD11	3:A:708:VAL:HG11	2.03	0.40
3:D:519:TYR:CD1	3:D:525:GLU:HA	2.56	0.40
3:A:542:LEU:HB2	3:A:544:LEU:HD13	2.03	0.40
3:A:316:ALA:HA	3:A:455:TRP:CZ3	2.56	0.40
3:D:565:ALA:N	3:D:566:PRO:CD	2.85	0.40
3:A:440:GLU:HB3	3:A:441:PRO:HD3	2.04	0.40
3:D:593:LYS:NZ	3:D:593:LYS:CB	2.84	0.40
3:D:655:SER:O	3:D:656:GLN:C	2.60	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	
3	A	578/580 (100%)	559 (97%)	17 (3%)	2 (0%)	41	30
3	D	578/580 (100%)	564 (98%)	13 (2%)	1 (0%)	47	37
All	All	1156/1160 (100%)	1123 (97%)	30 (3%)	3 (0%)	41	30

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	628	ILE
3	A	628	ILE
3	A	421	ALA

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	A	495/495 (100%)	495 (100%)	0	100	100
3	D	495/495 (100%)	495 (100%)	0	100	100
All	All	990/990 (100%)	990 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

4 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	$\#\|Z\| > 2$
2	6OG	F	3	2	18,25,26	1.38	2 (11%)	20,36,39	2.80	5 (25%)
1	DDG	E	29	2,1	17,23,24	1.08	2 (11%)	15,33,36	0.76	1 (6%)
1	DDG	B	29	2,1	17,23,24	1.05	2 (11%)	15,33,36	0.80	1 (6%)
2	6OG	C	3	2	18,25,26	1.38	2 (11%)	20,36,39	2.83	5 (25%)

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	6OG	F	3	2	-	0/5/23/24	0/3/3/3
1	DDG	E	29	2,1	-	0/3/18/19	0/3/3/3
1	DDG	B	29	2,1	-	0/3/18/19	0/3/3/3
2	6OG	C	3	2	-	0/5/23/24	0/3/3/3

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	3	6OG	C6-N1	4.95	1.40	1.31
2	C	3	6OG	C6-N1	4.93	1.40	1.31
1	B	29	DDG	C8-N7	-2.36	1.31	1.35
1	E	29	DDG	C8-N7	-2.32	1.31	1.35
1	E	29	DDG	C5-C6	-2.30	1.42	1.47
1	B	29	DDG	C5-C6	-2.29	1.42	1.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	3	6OG	C8-N7	-2.22	1.30	1.34
2	C	3	6OG	C8-N7	-2.17	1.30	1.34

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	3	6OG	O6-C6-C5	8.14	127.64	116.01
2	F	3	6OG	O6-C6-C5	8.01	127.46	116.01
2	C	3	6OG	C5-C6-N1	-6.31	111.25	123.26
2	F	3	6OG	C5-C6-N1	-6.27	111.32	123.26
2	C	3	6OG	C2-N1-C6	5.63	125.13	116.08
2	F	3	6OG	C2-N1-C6	5.56	125.01	116.08
2	F	3	6OG	C2-N3-C4	-3.05	111.87	115.36
2	C	3	6OG	C2-N3-C4	-2.99	111.94	115.36
2	C	3	6OG	N3-C2-N1	-2.52	123.87	127.22
2	F	3	6OG	N3-C2-N1	-2.51	123.88	127.22
1	B	29	DDG	O6-C6-C5	2.30	128.87	124.37
1	E	29	DDG	O6-C6-C5	2.16	128.60	124.37

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	3.28	4 (36%)	15,15,17	1.58	3 (20%)
4	FRU	G	2	4	11,12,12	1.48	2 (18%)	10,18,18	0.73	0
4	GLC	H	1	4	11,11,12	3.18	4 (36%)	15,15,17	1.51	3 (20%)
4	FRU	H	2	4	11,12,12	1.55	2 (18%)	10,18,18	0.84	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	-	0/5/24/24	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 (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	1	GLC	C2-C3	9.50	1.66	1.52
4	H	1	GLC	C2-C3	9.26	1.66	1.52
4	H	2	FRU	O2-C2	3.78	1.47	1.40
4	G	2	FRU	O2-C2	3.66	1.47	1.40
4	H	1	GLC	O5-C5	2.92	1.49	1.43
4	G	1	GLC	O5-C1	2.87	1.48	1.43
4	H	2	FRU	C1-C2	2.80	1.56	1.52
4	G	1	GLC	O5-C5	2.57	1.48	1.43
4	G	1	GLC	C4-C5	2.54	1.58	1.53
4	G	2	FRU	C1-C2	2.53	1.56	1.52
4	H	1	GLC	C4-C5	2.41	1.58	1.53
4	H	1	GLC	O5-C1	2.31	1.47	1.43

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	G	1	GLC	C1-O5-C5	3.95	117.55	112.19
4	H	1	GLC	C1-O5-C5	3.78	117.31	112.19
4	G	1	GLC	C1-C2-C3	-3.02	105.95	109.67
4	H	1	GLC	C1-C2-C3	-2.71	106.34	109.67
4	H	1	GLC	O3-C3-C2	-2.33	105.54	109.99
4	G	1	GLC	O3-C3-C2	-2.21	105.75	109.99

There are no chirality outliers.

There are no torsion outliers.

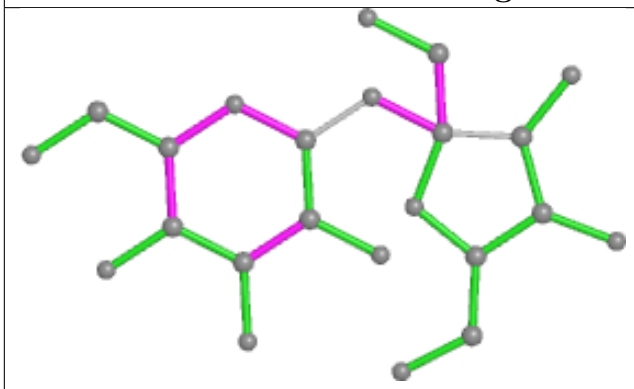
There are no ring outliers.

1 monomer is involved in 1 short contact:

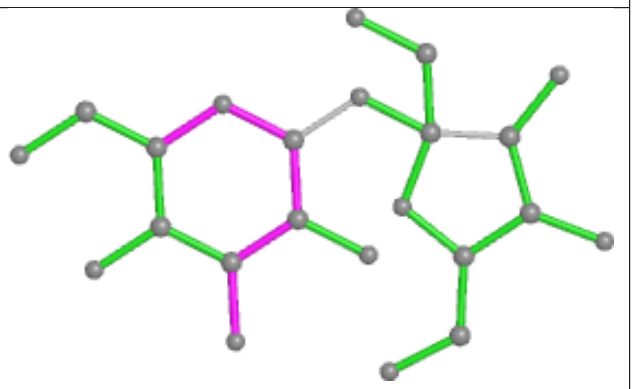
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	2	FRU	1	0

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

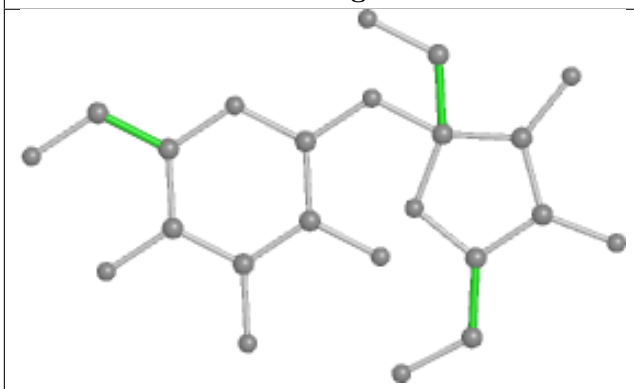
Oligosaccharide Chain G



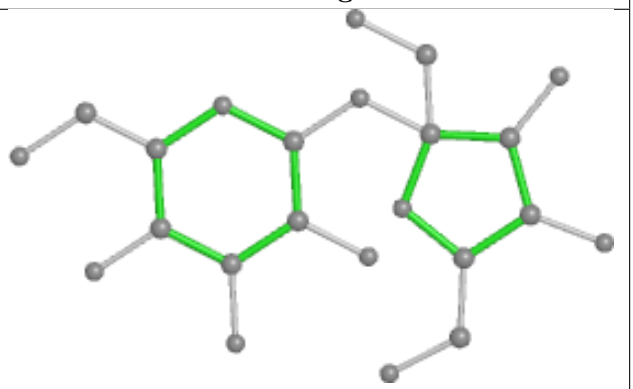
Bond lengths



Bond angles

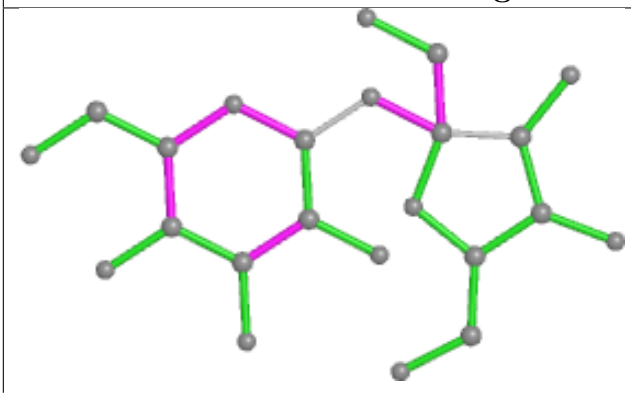


Torsions

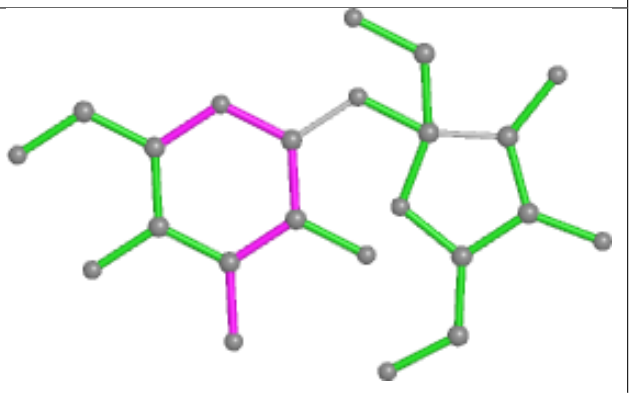


Rings

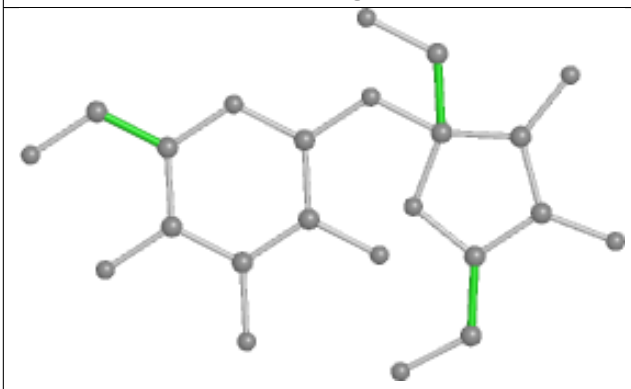
Oligosaccharide Chain H



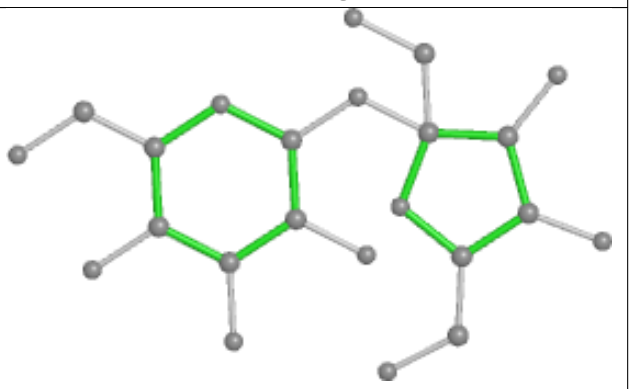
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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	$\#\|Z\| > 2$
7	D3T	D	880	5	25,29,29	1.51	4 (16%)	35,45,45	2.17	12 (34%)
6	SO4	D	879	-	4,4,4	0.21	0	6,6,6	0.08	0
6	SO4	A	880	-	4,4,4	0.26	0	6,6,6	0.05	0
7	D3T	A	881	5	25,29,29	1.55	4 (16%)	35,45,45	2.37	16 (45%)
6	SO4	A	879	-	4,4,4	0.26	0	6,6,6	0.04	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
7	D3T	A	881	5	-	2/22/31/31	0/2/2/2
7	D3T	D	880	5	-	4/22/31/31	0/2/2/2

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	881	D3T	PG-O1G	3.38	1.61	1.50
7	A	881	D3T	C4-C5	-3.23	1.39	1.44
7	D	880	D3T	C4-C5	-3.13	1.39	1.44
7	A	881	D3T	C6-C5	3.01	1.39	1.34
7	A	881	D3T	PB-O1B	3.00	1.61	1.50
7	D	880	D3T	C6-C5	2.97	1.39	1.34
7	D	880	D3T	PB-O1B	2.93	1.61	1.50
7	D	880	D3T	PA-O1A	2.93	1.61	1.50

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	881	D3T	C4-N3-C2	-5.92	119.69	127.35
7	A	881	D3T	C5-C4-N3	5.79	120.25	115.31
7	D	880	D3T	C4-N3-C2	-5.75	119.91	127.35
7	D	880	D3T	C5-C4-N3	5.62	120.11	115.31
7	D	880	D3T	O4-C4-C5	-4.10	120.15	124.90
7	A	881	D3T	O4-C4-C5	-4.05	120.21	124.90
7	A	881	D3T	N3-C2-N1	3.77	119.90	114.89
7	D	880	D3T	N3-C2-N1	3.66	119.74	114.89
7	A	881	D3T	PB-O3B-PG	-3.49	120.86	132.83
7	A	881	D3T	C5-C6-N1	-3.26	119.98	123.34
7	D	880	D3T	C5-C6-N1	-3.17	120.07	123.34
7	A	881	D3T	C2'-C3'-C4'	-3.11	96.91	102.72
7	D	880	D3T	C2'-C3'-C4'	-3.07	96.97	102.72
7	A	881	D3T	PB-O3A-PA	-2.87	122.98	132.83
7	A	881	D3T	C6-C5-C4	2.63	120.23	118.03
7	D	880	D3T	C6-C5-C4	2.56	120.17	118.03
7	A	881	D3T	C3'-C2'-C1'	-2.53	99.86	102.78
7	D	880	D3T	PB-O3A-PA	-2.48	124.31	132.83
7	A	881	D3T	C4'-O4'-C1'	-2.33	107.60	109.81
7	D	880	D3T	C4'-O4'-C1'	-2.29	107.64	109.81
7	A	881	D3T	C5M-C5-C6	-2.29	119.80	122.85

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	D	880	D3T	O2-C2-N1	-2.24	119.81	122.79
7	A	881	D3T	O3G-PG-O3B	2.22	112.09	104.64
7	A	881	D3T	O2-C2-N1	-2.18	119.89	122.79
7	D	880	D3T	C5M-C5-C6	-2.14	119.99	122.85
7	A	881	D3T	C2'-C1'-N1	-2.10	108.42	112.40
7	D	880	D3T	C3'-C2'-C1'	-2.07	100.39	102.78
7	A	881	D3T	O2B-PB-O1B	2.05	122.39	112.24

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	D	880	D3T	PB-O3B-PG-O1G
7	D	880	D3T	PB-O3B-PG-O2G
7	D	880	D3T	PB-O3A-PA-O2A
7	D	880	D3T	PG-O3B-PB-O1B
7	A	881	D3T	PB-O3A-PA-O1A
7	A	881	D3T	PG-O3B-PB-O1B

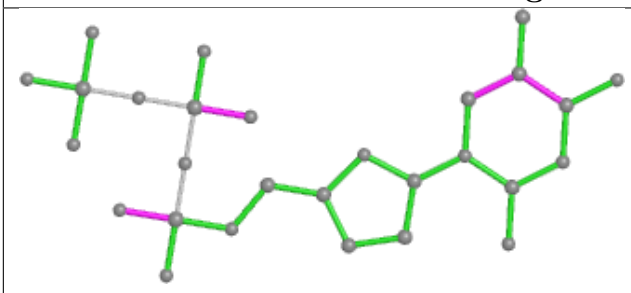
There are no ring outliers.

2 monomers are involved in 5 short contacts:

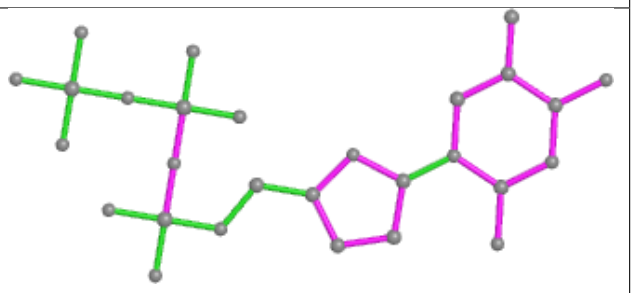
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	D	880	D3T	3	0
7	A	881	D3T	2	0

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 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.

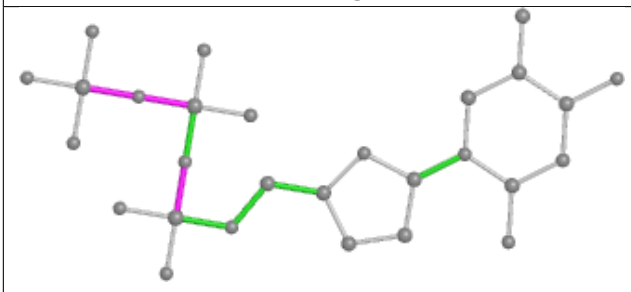
Ligand D3T D 880



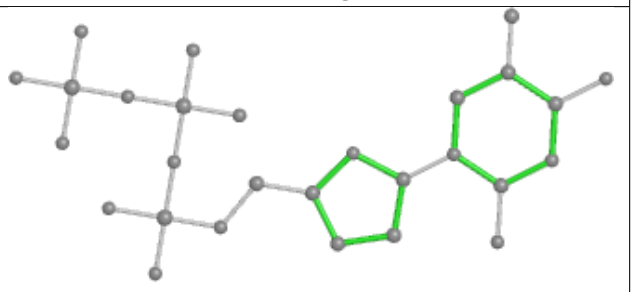
Bond lengths



Bond angles

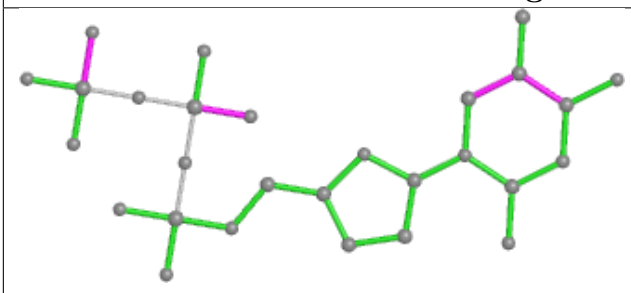


Torsions

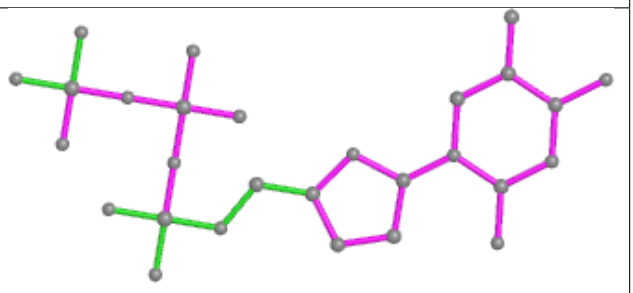


Rings

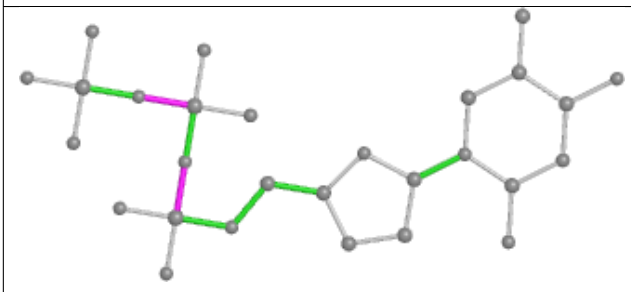
Ligand D3T A 881



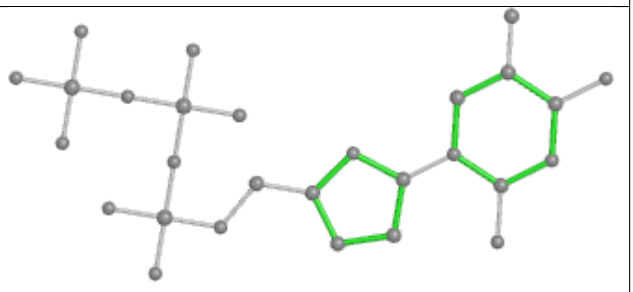
Bond lengths



Bond angles



Torsions



Rings

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, 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	B	8/9 (88%)	0.34	1 (12%) 3 4	16, 21, 52, 64	0
1	E	8/9 (88%)	0.02	0 100 100	11, 19, 39, 63	0
2	C	10/13 (76%)	-0.41	0 100 100	14, 19, 33, 50	0
2	F	10/13 (76%)	-0.17	0 100 100	10, 14, 30, 51	0
3	A	580/580 (100%)	0.36	35 (6%) 21 23	13, 26, 48, 67	0
3	D	580/580 (100%)	-0.10	6 (1%) 82 83	6, 17, 33, 49	0
All	All	1196/1204 (99%)	0.12	42 (3%) 44 45	6, 22, 44, 67	0

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	297	ALA	14.2
3	D	297	ALA	7.7
3	A	645	SER	6.8
3	A	298	LYS	5.6
3	A	695	ASP	5.2
3	D	691	GLN	4.8
3	D	876	LYS	4.7
3	A	842	GLU	4.0
3	A	819	ARG	3.9
3	A	696	GLU	3.9
3	A	846	ARG	3.6
3	A	816	LYS	3.6
3	A	692	VAL	3.6
3	A	876	LYS	3.1
3	A	691	GLN	3.1
3	D	779	ARG	3.1
3	A	646	ASP	3.0
3	A	300	ALA	3.0
3	A	781	PHE	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	A	817	GLU	2.9
3	A	420	GLU	2.8
3	A	779	ARG	2.8
3	A	690	PHE	2.8
3	D	695	ASP	2.6
3	A	838	LYS	2.6
3	A	689	ILE	2.6
3	A	305	ASP	2.5
3	A	693	SER	2.5
3	A	644	GLU	2.5
3	A	509	GLU	2.5
3	A	694	GLU	2.4
3	A	839	GLU	2.4
3	A	431	LYS	2.3
3	A	303	LEU	2.3
3	A	688	ASP	2.3
3	D	551	LYS	2.2
3	A	698	THR	2.2
3	A	699	PRO	2.2
3	A	559	ASP	2.1
3	A	782	ASN	2.1
3	A	730	LYS	2.1
1	B	22	DC	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, 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	6OG	C	3	23/24	0.96	0.09	18,20,28,30	0
2	6OG	F	3	23/24	0.97	0.10	11,14,20,22	0
1	DDG	B	29	21/22	0.98	0.14	15,16,18,18	0
1	DDG	E	29	21/22	0.98	0.12	4,8,12,12	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,

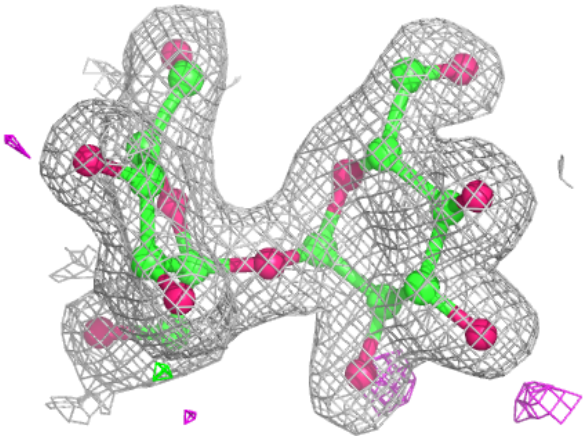
median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

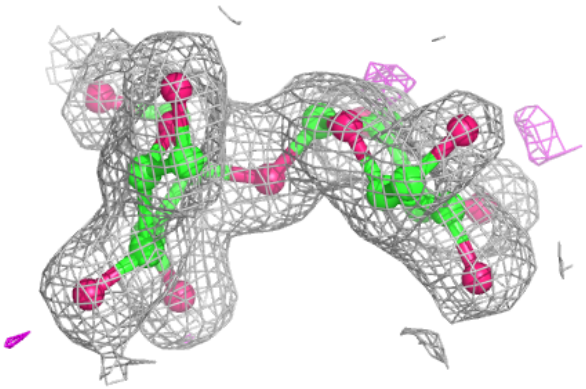
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	FRU	G	2	12/12	0.92	0.10	25,28,28,30	0
4	GLC	G	1	11/12	0.93	0.12	28,28,28,28	0
4	FRU	H	2	12/12	0.96	0.08	17,18,19,20	0
4	GLC	H	1	11/12	0.97	0.09	14,18,19,20	0

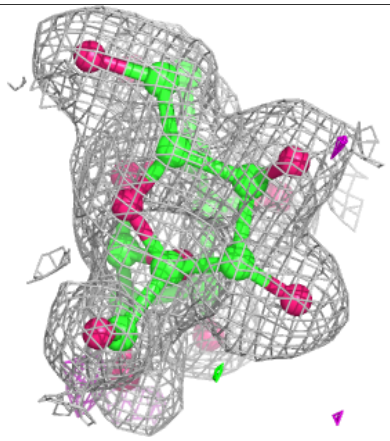
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

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)

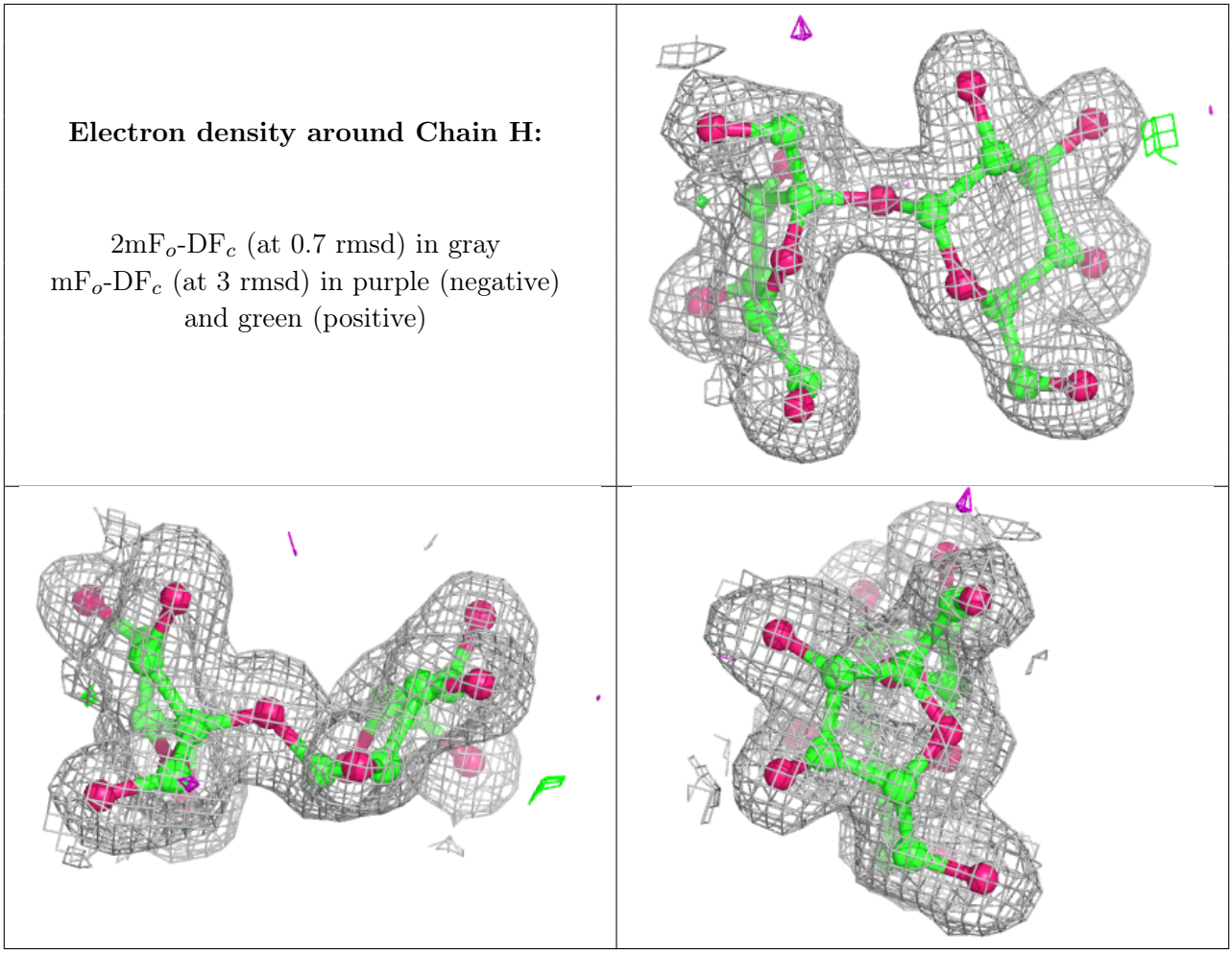






Electron density around Chain H:

$2mF_o$-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, 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.

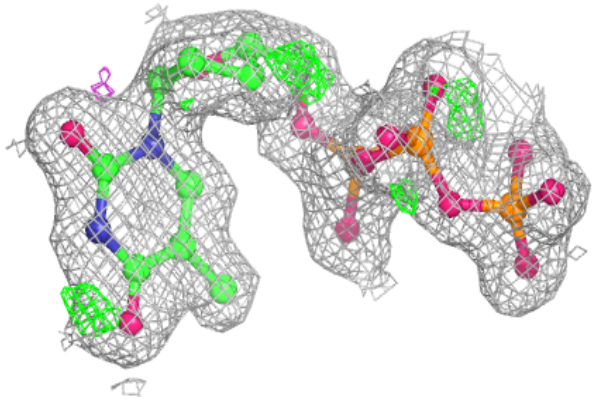
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	SO4	A	879	5/5	0.78	0.37	94,95,95,95	0
6	SO4	A	880	5/5	0.84	0.34	97,97,98,98	0
6	SO4	D	879	5/5	0.91	0.27	76,77,78,78	0
7	D3T	A	881	28/28	0.96	0.12	17,19,24,25	0
7	D3T	D	880	28/28	0.98	0.13	9,11,14,15	0
5	MN	D	878	1/1	1.00	0.07	10,10,10,10	0
5	MN	A	878	1/1	1.00	0.06	21,21,21,21	0

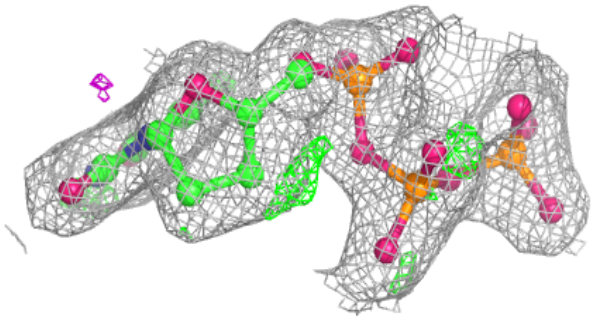
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different

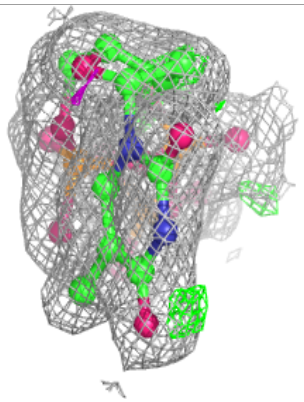
orientation to approximate a three-dimensional view.

Electron density around D3T A 881:

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

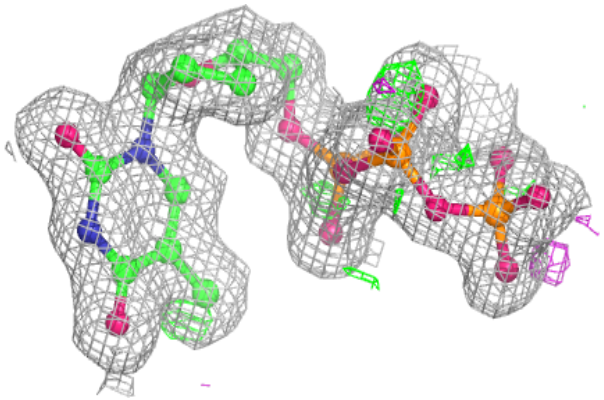


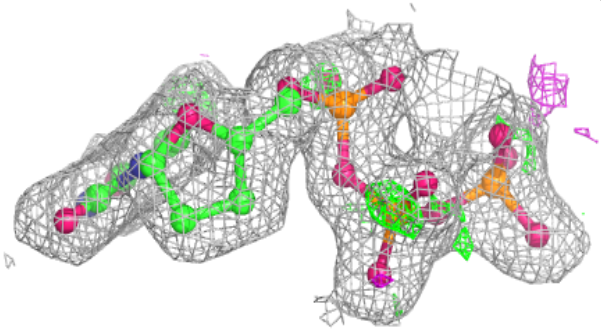


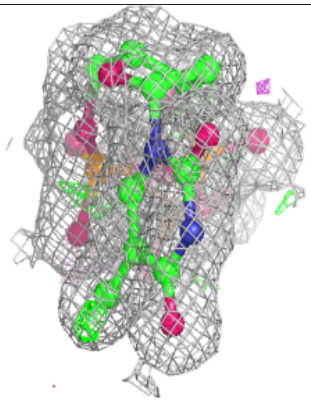


Electron density around D3T D 880:

$2mF_o$-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.