Full wwPDB X-ray Structure Validation Report (i)

Oct 4, 2023 – 05:32 PM EDT

PDB ID	:	6UEU
Title	:	Crystal structure of BF DNA polymerase F710Y mutant bound to tetrahydrofuran and dATP
Authors	:	Wu, E.Y.; Beese, L.S.
Deposited on	:	2019-09-23
Resolution	:	1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

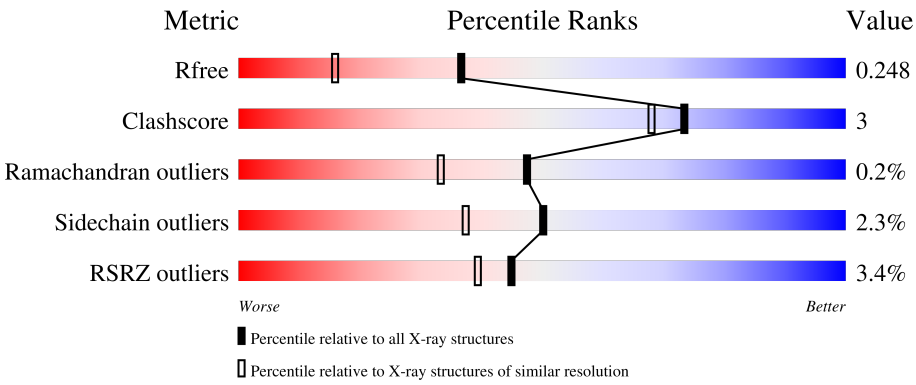
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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

Mol	Chain	Length	Quality of chain
1	A	580	6%, 90%, 9%
1	D	580	%, 90%, 10%
2	B	9	22%, 67%, 11%
2	E	9	11%, 44%, 44%, 11%
3	C	12	50%, 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	12	8% 42% 58%
4	G	2	100%
4	H	2	100%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 11116 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 DNA polymerase I.

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

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	297	ALA	-	expression tag	UNP D9N168
A	710	TYR	PHE	engineered mutation	UNP D9N168
A	713	VAL	PRO	conflict	UNP D9N168
D	297	ALA	-	expression tag	UNP D9N168
D	710	TYR	PHE	engineered mutation	UNP D9N168
D	713	VAL	PRO	conflict	UNP D9N168

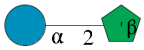
- Molecule 2 is a DNA chain called DNA (5'-D(*CP*GP*AP*TP*CP*AP*CP*GP*(2DT))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	9	179	87	33	51	8	0	0	0
2	E	9	179	87	33	51	8	0	0	0

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

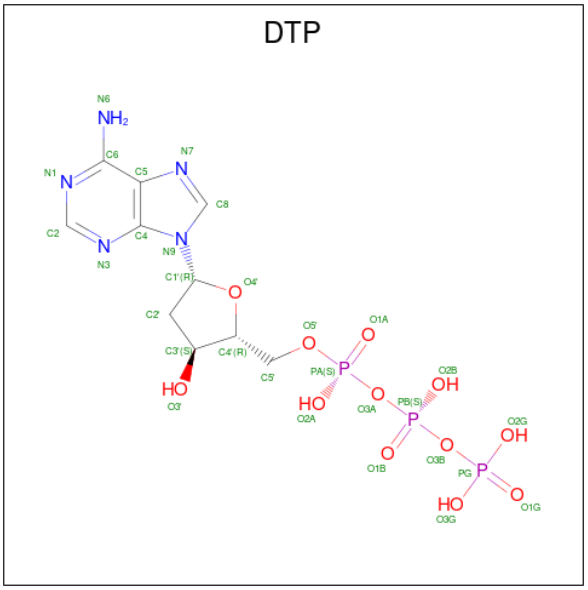
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	12	235	112	43	69	11	0	0	0
3	F	12	235	112	43	69	11	0	0	0

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



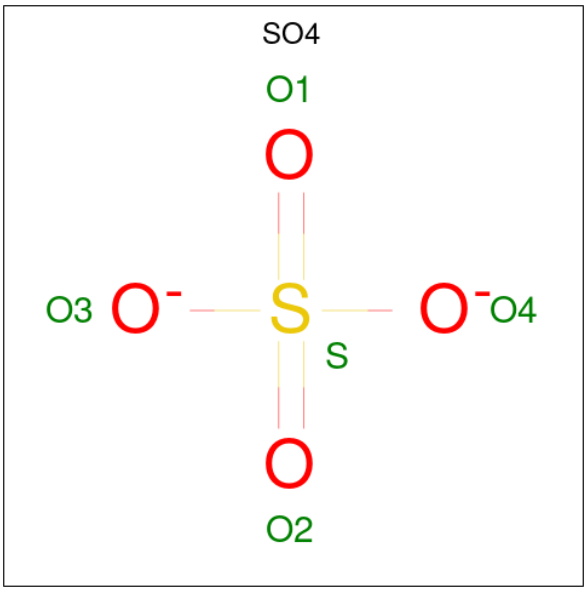
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	O				
4	G	2	23	12	11		0	0	0
			Total	C	O				
4	H	2	23	12	11		0	0	0

- Molecule 5 is 2'-DEOXYADENOSINE 5'-TRIPHOSPHATE (three-letter code: DTP) (formula: $C_{10}H_{16}N_5O_{12}P_3$).



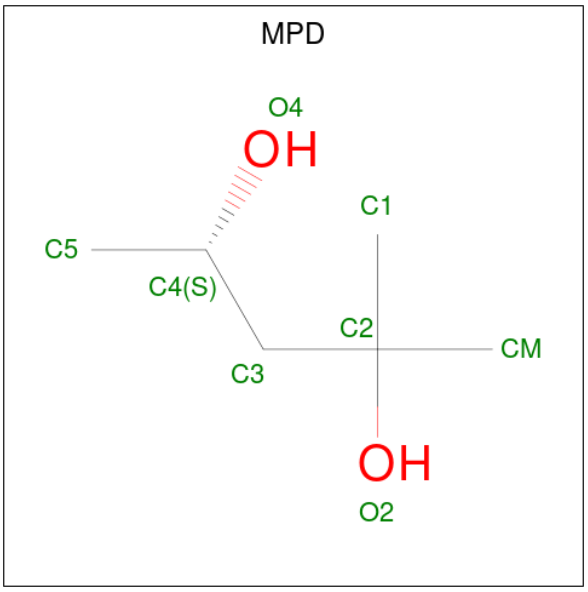
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
5	A	1	30	10	5	12	3	0	0
			Total	C	N	O	P		
5	D	1	30	10	5	12	3	0	0

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



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

- Molecule 7 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
7	D	1	8	6	2	0	0

- Molecule 8 is water.

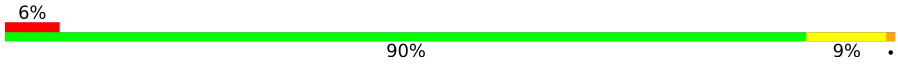
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	271	271	271	0	0
			Total	O		
8	D	484	484	484	0	0
			Total	O		
8	B	13	13	13	0	0
			Total	O		
8	C	36	36	36	0	0
			Total	O		
8	E	20	20	20	0	0
			Total	O		
8	F	41	41	41	0	0

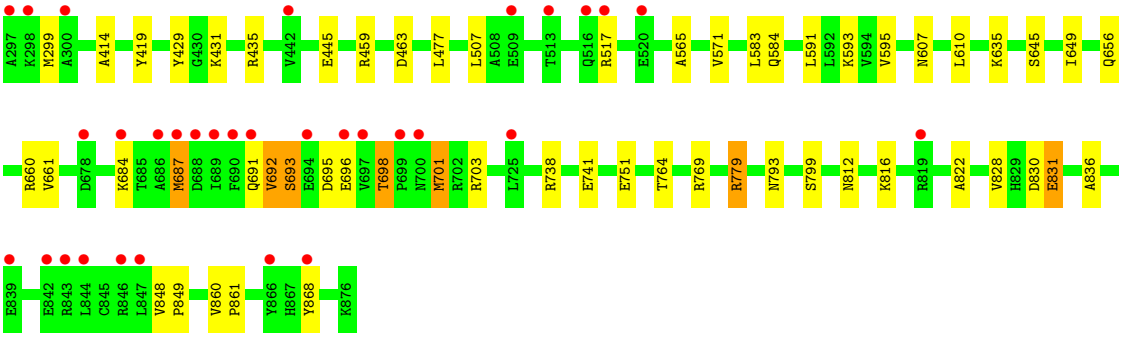
3 Residue-property plots (i)

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: DNA polymerase I

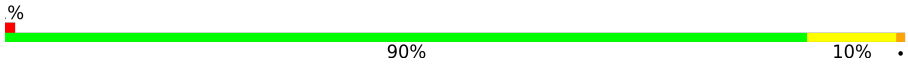
Chain A:

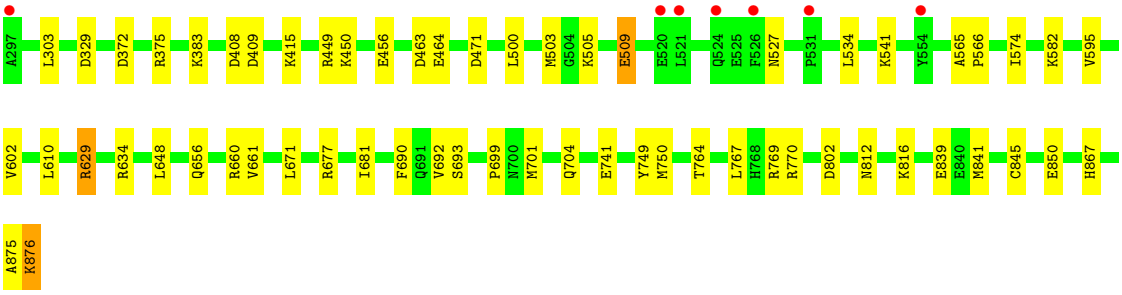




• Molecule 1: DNA polymerase I

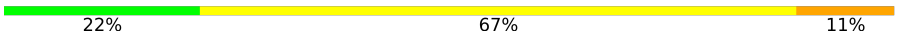
Chain D:

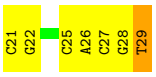




• Molecule 2: DNA (5'-D(*CP*GP*AP*TP*CP*AP*CP*GP*(2DT))-3')

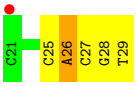
Chain B:





• Molecule 2: DNA (5'-D(*CP*GP*AP*TP*CP*AP*CP*GP*(2DT))-3')

Chain E:

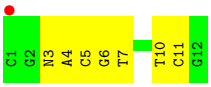


• Molecule 3: DNA (5'-D(*CP*GP*(3DR)P*AP*CP*GP*TP*GP*AP*TP*CP*G)-3')

Chain C:

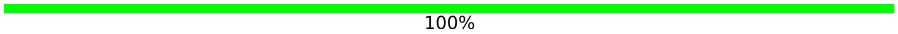
• Molecule 3: DNA (5'-D(*CP*GP*(3DR)P*AP*CP*GP*TP*GP*AP*TP*CP*G)-3')

Chain F:



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

Chain G:



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

Chain H:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	93.88Å 109.07Å 150.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.99 – 1.80 44.99 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.9 (44.99-1.80) 96.9 (44.99-1.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.70 (at 1.79Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.200 , 0.242 0.206 , 0.248	Depositor DCC
R_{free} test set	6922 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	21.4	Xtriage
Anisotropy	0.165	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 56.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11116	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: DTP, FRU, MPD, 3DR, SO4, GLC, 2DT

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	0.63	0/4736	0.62	0/6400
1	D	0.93	7/4736 (0.1%)	0.89	12/6400 (0.2%)
2	B	1.90	5/179 (2.8%)	1.39	2/274 (0.7%)
2	E	2.23	7/179 (3.9%)	1.30	1/274 (0.4%)
3	C	1.91	5/250 (2.0%)	1.54	3/382 (0.8%)
3	F	1.94	6/250 (2.4%)	1.73	6/382 (1.6%)
All	All	0.95	30/10330 (0.3%)	0.87	24/14112 (0.2%)

All (30) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	25	DC	C3'-O3'	-11.60	1.28	1.44
1	D	770	ARG	CB-CG	-8.77	1.28	1.52
2	B	25	DC	C3'-O3'	-8.26	1.33	1.44
3	C	7	DT	C3'-O3'	7.59	1.53	1.44
1	D	845	CYS	CB-SG	-6.96	1.70	1.82
2	E	27	DC	C4-C5	6.83	1.48	1.43
1	D	749	TYR	CE2-CZ	-6.70	1.29	1.38
3	F	7	DT	N1-C2	6.17	1.43	1.38
3	F	5	DC	C2-N3	5.98	1.40	1.35
3	C	5	DC	C5-C6	5.92	1.39	1.34
1	D	770	ARG	CG-CD	5.83	1.66	1.51
3	F	7	DT	C4-C5	5.79	1.50	1.45
2	E	28	DG	N9-C8	5.79	1.42	1.37
2	E	26	DA	C3'-O3'	-5.74	1.36	1.44
3	C	6	DG	N7-C5	5.67	1.42	1.39
1	D	450	LYS	CE-NZ	5.65	1.63	1.49
2	B	28	DG	C8-N7	5.64	1.34	1.30
2	E	28	DG	C3'-C2'	5.58	1.58	1.52
1	D	661	VAL	CB-CG1	-5.57	1.41	1.52
3	F	6	DG	C3'-O3'	-5.48	1.36	1.44

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	383	LYS	CE-NZ	5.46	1.62	1.49
3	F	4	DA	C8-N7	-5.29	1.27	1.31
2	E	27	DC	C3'-C2'	5.22	1.58	1.52
2	E	28	DG	C3'-O3'	5.20	1.50	1.44
3	F	4	DA	C4'-O4'	-5.17	1.39	1.45
3	C	8	DG	C2-N3	5.09	1.36	1.32
3	C	8	DG	C6-N1	5.07	1.43	1.39
2	B	27	DC	C4-N4	5.07	1.38	1.33
2	B	28	DG	C3'-O3'	-5.05	1.37	1.44
2	B	25	DC	C3'-C2'	5.02	1.58	1.52

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	7	DT	O5'-P-OP2	-11.32	95.51	105.70
1	D	409	ASP	CB-CG-OD1	9.49	126.84	118.30
1	D	677	ARG	NE-CZ-NH2	-8.72	115.94	120.30
3	F	4	DA	O5'-P-OP2	-8.64	97.92	105.70
1	D	449	ARG	NE-CZ-NH2	8.22	124.41	120.30
1	D	660	ARG	NE-CZ-NH1	-7.94	116.33	120.30
1	D	449	ARG	NE-CZ-NH1	-7.90	116.35	120.30
1	D	610	LEU	CB-CG-CD2	6.89	122.72	111.00
3	C	5	DC	O4'-C1'-C2'	6.84	111.37	105.90
3	F	5	DC	O4'-C1'-C2'	6.58	111.17	105.90
3	F	6	DG	O4'-C1'-C2'	6.36	110.99	105.90
1	D	463	ASP	CB-CG-OD2	-6.30	112.63	118.30
2	B	27	DC	O4'-C1'-N1	6.17	112.32	108.00
1	D	463	ASP	CB-CG-OD1	6.14	123.83	118.30
1	D	372	ASP	CB-CG-OD2	5.96	123.66	118.30
1	D	375	ARG	NE-CZ-NH2	-5.95	117.33	120.30
2	B	25	DC	O4'-C4'-C3'	-5.94	102.12	104.50
3	F	7	DT	OP1-P-OP2	5.73	128.19	119.60
3	C	7	DT	O4'-C1'-N1	-5.61	104.08	108.00
2	E	25	DC	O4'-C4'-C3'	-5.38	102.35	104.50
3	F	4	DA	O4'-C4'-C3'	5.38	109.23	106.00
1	D	629	ARG	NE-CZ-NH2	-5.11	117.74	120.30
3	C	11	DC	C4'-C3'-C2'	5.07	107.67	103.10
1	D	409	ASP	CB-CG-OD2	-5.01	113.79	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4652	0	4710	30	0
1	D	4652	0	4710	33	0
2	B	179	0	103	3	0
2	E	179	0	103	1	0
3	C	235	0	133	0	0
3	F	235	0	133	1	0
4	G	23	0	21	0	0
4	H	23	0	21	0	0
5	A	30	0	11	1	0
5	D	30	0	11	0	0
6	D	5	0	0	0	0
7	D	8	0	14	0	0
8	A	271	0	0	11	0
8	B	13	0	0	1	0
8	C	36	0	0	0	0
8	D	484	0	0	13	0
8	E	20	0	0	0	0
8	F	41	0	0	0	0
All	All	11116	0	9970	67	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 3.

All (67) 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:D:534:LEU:HD11	1:D:574:ILE:HD13	1.63	0.80
1:D:741:GLU:OE2	8:D:1001:HOH:O	2.10	0.68
5:A:901:DTP:H4'	2:B:29:2DT:H2'	1.78	0.66
1:D:329:ASP:OD1	8:D:1002:HOH:O	2.14	0.65
1:A:693:SER:OG	8:A:1001:HOH:O	2.14	0.64
1:D:750:MET:HG2	8:D:1401:HOH:O	1.98	0.62
1:D:699:PRO:O	8:D:1004:HOH:O	2.16	0.61
2:B:26:DA:OP2	8:B:101:HOH:O	2.16	0.60
1:A:635:LYS:NZ	8:A:1012:HOH:O	2.34	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:692:VAL:HG21	1:D:701:MET:HE1	1.82	0.60
1:D:509:GLU:OE2	8:D:1005:HOH:O	2.17	0.59
2:B:21:DC:H2"	2:B:22:DG:C8	2.43	0.54
1:A:751:GLU:OE1	8:A:1002:HOH:O	2.19	0.53
1:A:607:ASN:ND2	1:A:610:LEU:HD12	2.25	0.52
1:D:634:ARG:NH2	1:D:876:LYS:HA	2.25	0.52
1:A:779:ARG:NH1	8:A:1020:HOH:O	2.44	0.50
1:D:595:VAL:HG12	1:D:602:VAL:CG1	2.42	0.49
1:A:584:GLN:HG3	8:A:1057:HOH:O	2.12	0.48
1:A:507:LEU:HD21	1:A:583:LEU:HB2	1.95	0.48
1:D:850:GLU:OE2	8:D:1006:HOH:O	2.20	0.47
1:A:692:VAL:HG21	1:A:696:GLU:HB3	1.97	0.47
1:A:738:ARG:O	1:A:741:GLU:HG2	2.15	0.47
1:D:750:MET:CG	8:D:1401:HOH:O	2.58	0.47
1:D:527:ASN:ND2	8:D:1021:HOH:O	2.39	0.47
1:A:463:ASP:OD2	8:A:1003:HOH:O	2.21	0.47
1:D:764:THR:HA	1:D:769:ARG:O	2.15	0.47
1:D:415:LYS:NZ	8:D:1024:HOH:O	2.40	0.46
1:D:812:ASN:O	1:D:816:LYS:HG2	2.16	0.46
1:A:698:THR:OG1	1:A:701:MET:SD	2.73	0.46
1:A:830:ASP:OD1	1:A:831:GLU:HG2	2.16	0.46
1:D:534:LEU:HD11	1:D:574:ILE:CD1	2.41	0.46
1:D:839:GLU:OE1	8:D:1007:HOH:O	2.20	0.45
1:A:591:LEU:O	1:A:595:VAL:HG23	2.15	0.45
1:A:684:LYS:HA	1:A:687:MET:HG3	1.99	0.45
1:D:456:GLU:HG2	8:D:1272:HOH:O	2.16	0.45
1:D:595:VAL:HG12	1:D:602:VAL:HG13	1.98	0.45
1:D:656:GLN:OE1	1:D:681:ILE:HG22	2.17	0.45
1:A:565:ALA:HA	1:A:571:VAL:HG21	1.97	0.45
1:A:703:ARG:O	8:A:1004:HOH:O	2.21	0.45
1:D:767:LEU:HG	1:D:802:ASP:HB3	1.98	0.45
1:A:848:VAL:HB	1:A:849:PRO:HD3	2.00	0.43
1:A:764:THR:HA	1:A:769:ARG:O	2.18	0.43
1:D:471:ASP:OD1	1:D:471:ASP:N	2.51	0.43
1:D:634:ARG:HH22	1:D:876:LYS:HA	1.84	0.43
1:A:649:ILE:O	1:A:868:TYR:HA	2.19	0.43
1:A:860:VAL:HB	1:A:861:PRO:HD2	2.00	0.42
1:A:584:GLN:HG3	8:A:1242:HOH:O	2.18	0.42
1:D:500:LEU:HD12	1:D:503:MET:HE3	2.01	0.42
1:D:464:GLU:OE1	8:D:1008:HOH:O	2.22	0.42
1:A:828:VAL:HB	1:A:831:GLU:HG3	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:648:LEU:HD12	1:D:841:MET:HG3	2.01	0.42
1:D:867:HIS:HB2	1:D:875:ALA:O	2.20	0.42
1:D:408:ASP:CB	8:D:1020:HOH:O	2.67	0.42
1:A:593:LYS:NZ	8:A:1036:HOH:O	2.53	0.41
1:D:671:LEU:HD12	1:D:671:LEU:HA	1.85	0.41
1:A:429:TYR:O	1:A:435:ARG:HA	2.20	0.41
3:F:10:DT:H2"	3:F:11:DC:C6	2.56	0.41
1:A:414:ALA:HB1	1:A:419:TYR:HB3	2.03	0.41
1:D:582:LYS:HG2	2:E:26:DA:H2"	2.02	0.41
1:D:690:PHE:CD2	1:D:701:MET:HE2	2.56	0.41
1:A:793:ASN:HB2	8:A:1122:HOH:O	2.21	0.41
1:D:565:ALA:N	1:D:566:PRO:CD	2.84	0.41
1:D:500:LEU:HA	1:D:503:MET:HE2	2.03	0.41
1:A:812:ASN:ND2	8:A:1040:HOH:O	2.54	0.40
1:A:822:ALA:HB2	1:A:836:ALA:HB2	2.03	0.40
1:A:656:GLN:HB2	1:A:660:ARG:NH1	2.36	0.40
1:A:661:VAL:HG11	1:A:799:SER:HB2	2.04	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	578/580 (100%)	559 (97%)	17 (3%)	2 (0%)	41 27
1	D	578/580 (100%)	566 (98%)	12 (2%)	0	100 100
All	All	1156/1160 (100%)	1125 (97%)	29 (2%)	2 (0%)	47 33

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	299	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	691	GLN

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	495/495 (100%)	480 (97%)	15 (3%)	41 27
1	D	495/495 (100%)	487 (98%)	8 (2%)	62 54
All	All	990/990 (100%)	967 (98%)	23 (2%)	50 37

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

Mol	Chain	Res	Type
1	A	431	LYS
1	A	445	GLU
1	A	459	ARG
1	A	477	LEU
1	A	517	ARG
1	A	645	SER
1	A	687	MET
1	A	692	VAL
1	A	693	SER
1	A	695	ASP
1	A	698	THR
1	A	701	MET
1	A	779	ARG
1	A	816	LYS
1	A	831	GLU
1	D	303	LEU
1	D	505	LYS
1	D	509	GLU
1	D	541	LYS
1	D	629	ARG
1	D	693	SER
1	D	704	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	876	LYS

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	$\#\|Z\| > 2$
3	3DR	F	3	3	8,11,12	5.64	5 (62%)	9,14,17	1.59	1 (11%)
2	2DT	B	29	2,3	17,20,21	4.07	12 (70%)	22,28,31	2.11	7 (31%)
3	3DR	C	3	3	8,11,12	6.08	4 (50%)	9,14,17	1.17	2 (22%)
2	2DT	E	29	2,3	17,20,21	3.53	10 (58%)	22,28,31	1.56	2 (9%)

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
3	3DR	F	3	3	-	0/3/15/16	0/1/1/1
2	2DT	B	29	2,3	-	0/7/18/19	0/2/2/2
3	3DR	C	3	3	-	0/3/15/16	0/1/1/1
2	2DT	E	29	2,3	-	0/7/18/19	0/2/2/2

All (31) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	3	3DR	C2'-C3'	-12.67	1.30	1.52
3	F	3	3DR	C2'-C3'	-11.46	1.32	1.52
3	C	3	3DR	O4'-C4'	-9.90	1.28	1.44
3	F	3	3DR	O4'-C4'	-9.48	1.28	1.44
2	B	29	2DT	O4'-C4'	8.22	1.61	1.44
2	B	29	2DT	C6-C5	8.04	1.47	1.34
2	E	29	2DT	C6-C5	7.10	1.46	1.34
2	E	29	2DT	O4'-C4'	6.28	1.57	1.44
2	E	29	2DT	C2'-C1'	-5.27	1.39	1.51
2	B	29	2DT	C2-N1	5.21	1.46	1.38
2	E	29	2DT	O4'-C1'	5.02	1.53	1.42
2	B	29	2DT	C2'-C1'	-5.01	1.40	1.51
3	C	3	3DR	C3'-C4'	4.34	1.64	1.53
2	B	29	2DT	O4'-C1'	4.18	1.51	1.42
2	B	29	2DT	C4-N3	4.03	1.46	1.38
3	F	3	3DR	O4'-C1'	3.84	1.54	1.42
3	C	3	3DR	O4'-C1'	3.72	1.54	1.42
2	E	29	2DT	C6-N1	3.72	1.44	1.38
2	B	29	2DT	C3'-C2'	3.71	1.64	1.54
2	B	29	2DT	C6-N1	3.64	1.44	1.38
3	F	3	3DR	C3'-C4'	3.56	1.62	1.53
2	E	29	2DT	C2-N1	3.42	1.43	1.38
2	B	29	2DT	C3'-C4'	-3.25	1.35	1.52
2	E	29	2DT	C4-N3	3.21	1.44	1.38
2	E	29	2DT	C2-N3	3.14	1.43	1.38
2	E	29	2DT	O4-C4	-2.97	1.17	1.23
2	E	29	2DT	C3'-C4'	-2.97	1.36	1.52
2	B	29	2DT	C2-N3	2.88	1.43	1.38
2	B	29	2DT	C4-C5	2.15	1.48	1.44
3	F	3	3DR	C2'-C1'	2.15	1.57	1.51
2	B	29	2DT	C5M-C5	2.06	1.55	1.50

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	29	2DT	C4'-O4'-C1'	-5.46	104.65	109.81
2	B	29	2DT	C3'-C2'-C1'	4.33	107.79	102.78
2	E	29	2DT	C3'-C2'-C1'	4.31	107.77	102.78
2	E	29	2DT	C4'-O4'-C1'	-3.53	106.48	109.81
2	B	29	2DT	N3-C2-N1	3.29	119.25	114.89
3	F	3	3DR	C1'-C2'-C3'	3.19	106.81	103.20
2	B	29	2DT	O4'-C4'-C5'	-2.91	104.73	109.52
2	B	29	2DT	C6-N1-C2	-2.49	118.78	121.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	29	2DT	C6-C5-C4	2.33	119.98	118.03
3	C	3	3DR	C1'-C2'-C3'	2.31	105.81	103.20
2	B	29	2DT	C4-N3-C2	-2.25	124.44	127.35
3	C	3	3DR	C1'-O4'-C4'	-2.18	104.96	108.48

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	B	29	2DT	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	0.62	0	15,15,17	0.78	0
4	FRU	G	2	4	11,12,12	0.62	0	10,18,18	0.72	0
4	GLC	H	1	4	11,11,12	0.78	0	15,15,17	0.91	0
4	FRU	H	2	4	11,12,12	0.70	0	10,18,18	0.89	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

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	FRU	H	2	4	-	0/5/24/24	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

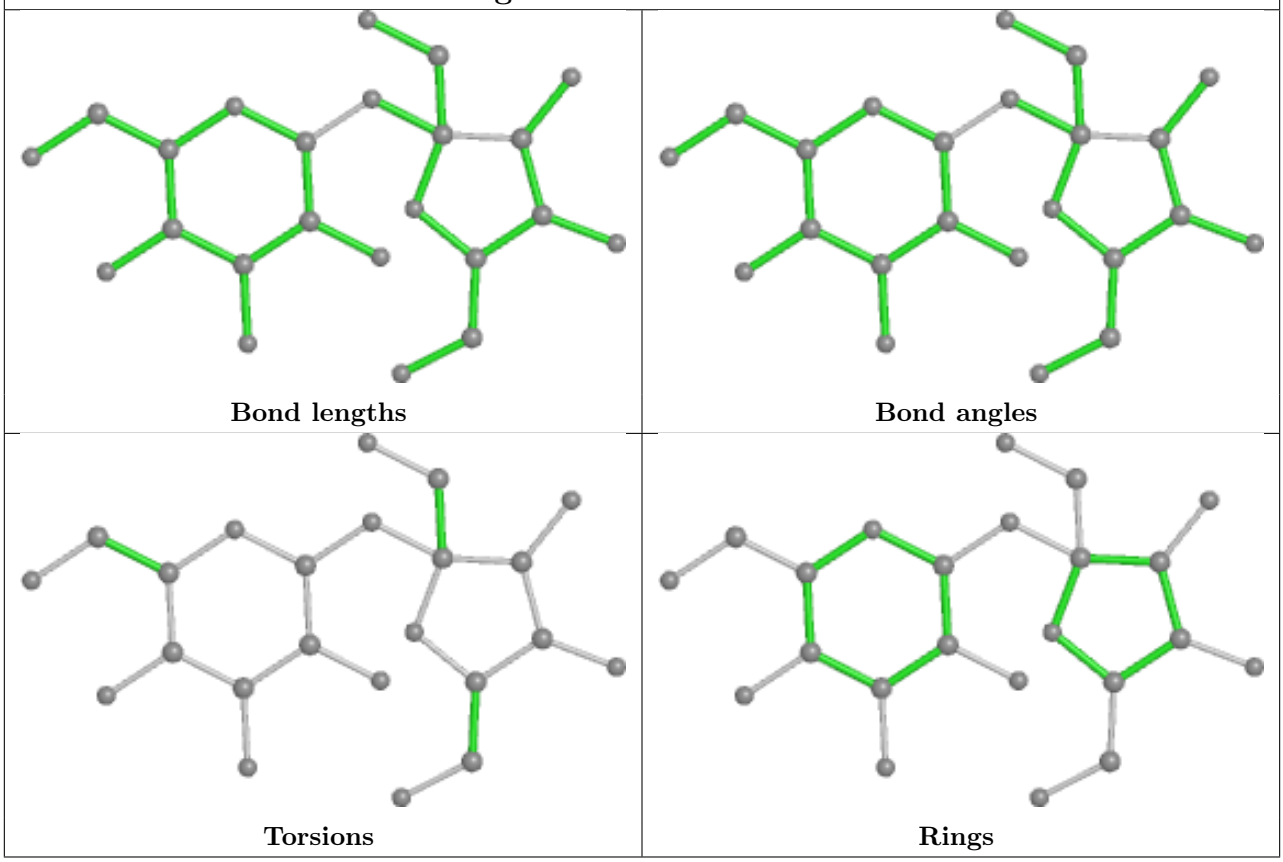
There are no torsion outliers.

There are no ring outliers.

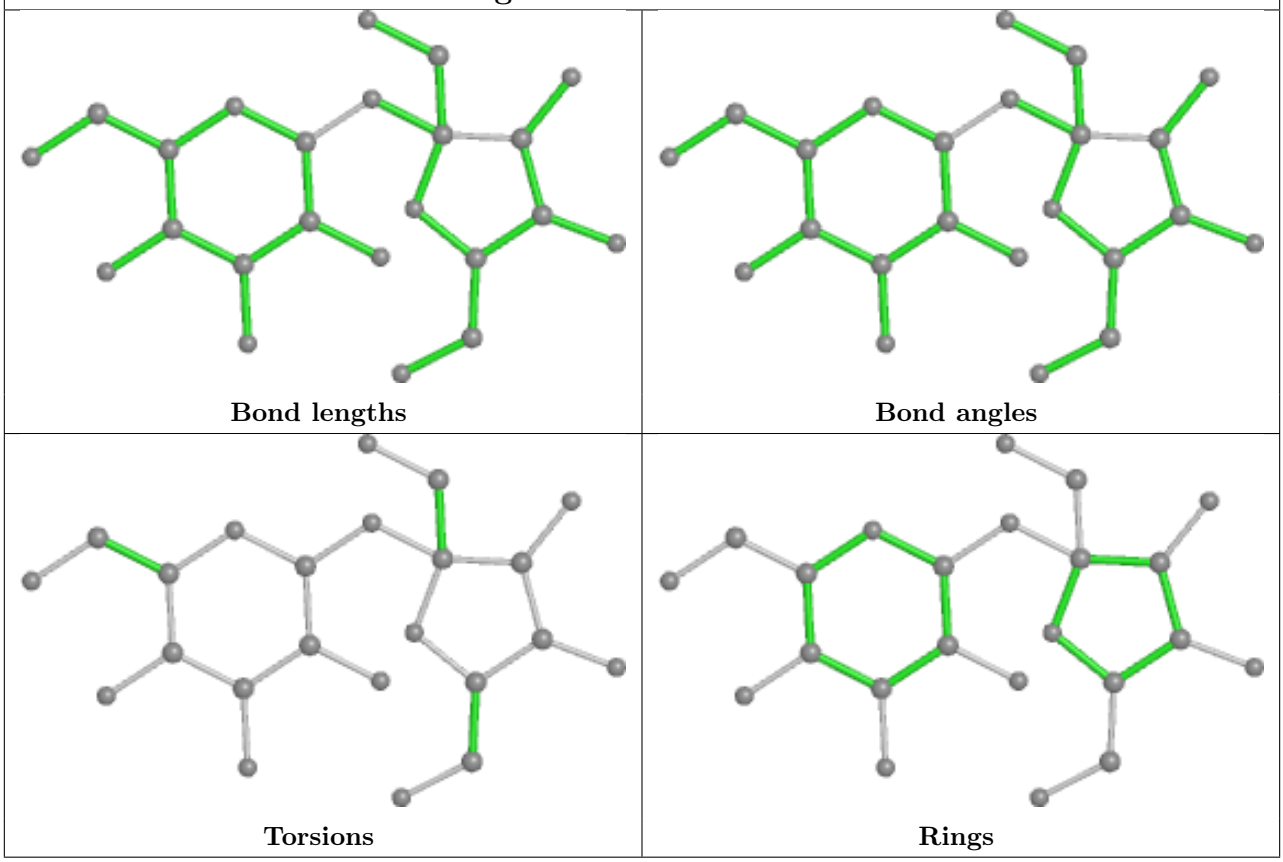
No monomer is involved in short contacts.

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

Oligosaccharide Chain G



Oligosaccharide Chain H



5.6 Ligand geometry (i)

4 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	$\#\|Z\| > 2$
5	DTP	A	901	-	26,32,32	3.27	10 (38%)	30,50,50	1.42	4 (13%)
7	MPD	D	904	-	7,7,7	0.62	0	9,10,10	0.84	0
5	DTP	D	901	-	26,32,32	3.01	7 (26%)	30,50,50	2.01	6 (20%)
6	SO4	D	903	-	4,4,4	0.31	0	6,6,6	0.40	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the

Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	DTP	A	901	-	-	4/18/34/34	0/3/3/3
5	DTP	D	901	-	-	2/18/34/34	0/3/3/3
7	MPD	D	904	-	-	1/5/5/5	-

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	901	DTP	C2'-C3'	-10.28	1.25	1.52
5	D	901	DTP	C2'-C3'	-10.16	1.26	1.52
5	D	901	DTP	O4'-C4'	-7.42	1.28	1.45
5	A	901	DTP	O4'-C4'	-6.96	1.29	1.45
5	A	901	DTP	C1'-N9	-4.94	1.34	1.49
5	A	901	DTP	C3'-C4'	4.53	1.65	1.53
5	D	901	DTP	C1'-N9	-4.34	1.36	1.49
5	A	901	DTP	C2-N3	4.08	1.38	1.32
5	A	901	DTP	O4'-C1'	4.04	1.51	1.42
5	D	901	DTP	O4'-C1'	4.01	1.51	1.42
5	A	901	DTP	O3'-C3'	3.20	1.50	1.43
5	A	901	DTP	C6-N6	2.90	1.44	1.34
5	A	901	DTP	C2-N1	2.83	1.39	1.33
5	D	901	DTP	C3'-C4'	2.83	1.60	1.53
5	D	901	DTP	C6-N6	2.67	1.43	1.34
5	D	901	DTP	C5-C4	-2.44	1.34	1.40
5	A	901	DTP	C5-C4	-2.44	1.34	1.40

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	D	901	DTP	O4'-C1'-C2'	-5.37	96.10	106.25
5	D	901	DTP	C5-C6-N6	5.20	128.25	120.35
5	A	901	DTP	N3-C2-N1	-4.82	121.14	128.68
5	D	901	DTP	N6-C6-N1	-3.75	110.79	118.57
5	D	901	DTP	N3-C2-N1	-3.36	123.42	128.68
5	A	901	DTP	N6-C6-N1	-2.28	113.84	118.57
5	A	901	DTP	C5-C6-N6	2.23	123.75	120.35
5	D	901	DTP	O3'-C3'-C4'	-2.21	101.66	110.10
5	A	901	DTP	PB-O3B-PG	-2.20	125.28	132.83
5	D	901	DTP	C3'-C2'-C1'	2.05	107.67	102.54

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	D	901	DTP	O4'-C4'-C5'-O5'
5	A	901	DTP	O4'-C4'-C5'-O5'
5	A	901	DTP	C3'-C4'-C5'-O5'
5	D	901	DTP	C3'-C4'-C5'-O5'
7	D	904	MPD	O2-C2-C3-C4
5	A	901	DTP	PA-O3A-PB-O3B
5	A	901	DTP	PA-O3A-PB-O1B

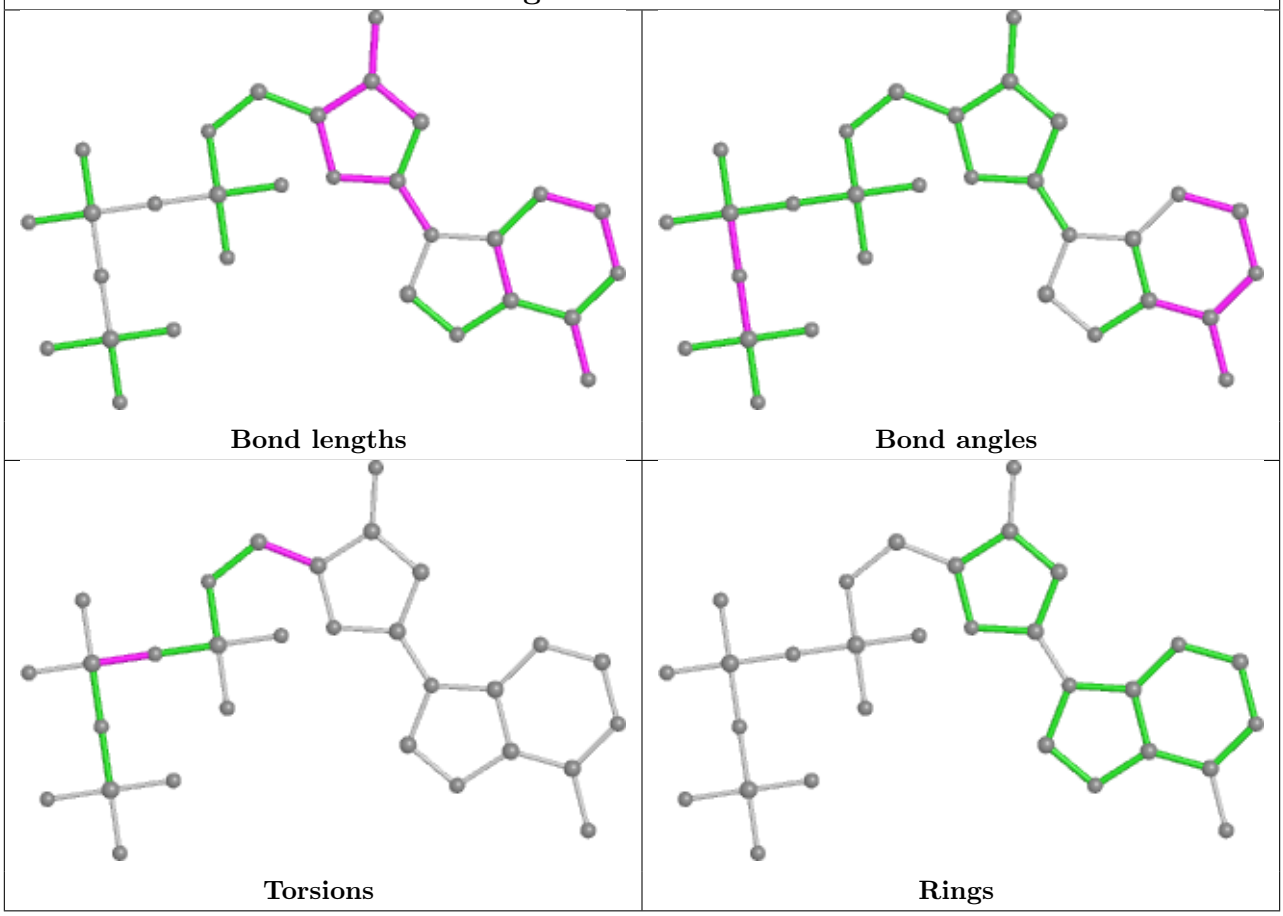
There are no ring outliers.

1 monomer is involved in 1 short contact:

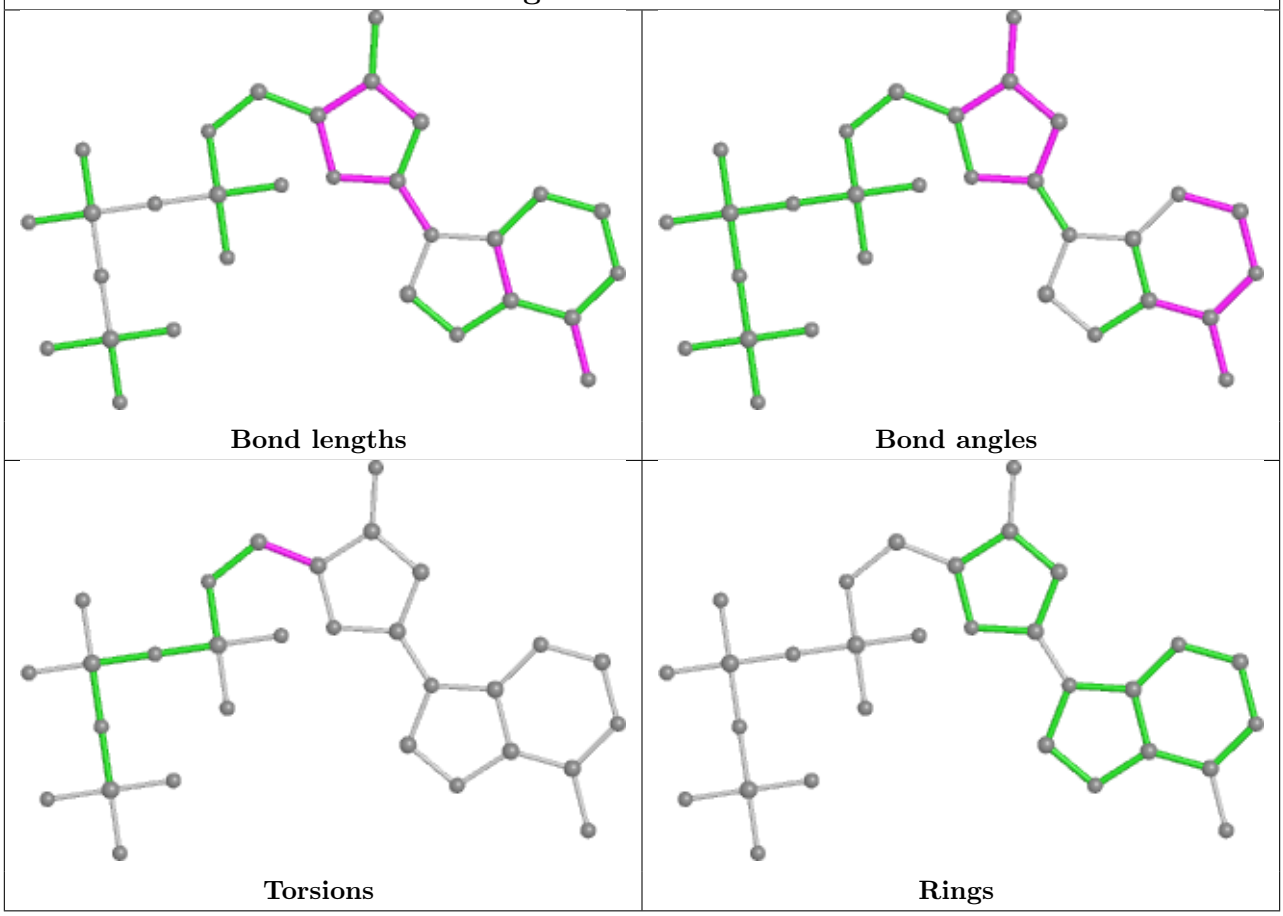
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	901	DTP	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand DTP A 901



Ligand DTP D 901



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	580/580 (100%)	0.30	32 (5%) 25 20	17, 38, 65, 87	0
1	D	580/580 (100%)	-0.04	7 (1%) 79 76	10, 22, 48, 63	0
2	B	8/9 (88%)	-0.17	0 100 100	24, 34, 56, 57	0
2	E	8/9 (88%)	0.20	1 (12%) 3 2	16, 33, 60, 67	0
3	C	11/12 (91%)	-0.32	0 100 100	19, 28, 58, 69	0
3	F	11/12 (91%)	-0.08	1 (9%) 9 7	14, 23, 65, 75	0
All	All	1198/1202 (99%)	0.12	41 (3%) 45 39	10, 31, 58, 87	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	699	PRO	5.8
1	D	297	ALA	4.6
1	A	690	PHE	4.2
1	A	846	ARG	4.1
1	A	687	MET	4.1
1	A	297	ALA	3.6
1	A	520	GLU	3.6
1	A	696	GLU	3.5
1	A	517	ARG	3.4
1	A	842	GLU	3.3
1	A	513	THR	3.2
1	A	697	VAL	3.2
1	A	300	ALA	3.1
1	A	866	TYR	3.0
1	D	520	GLU	2.8
1	A	819	ARG	2.8
1	A	725	LEU	2.7
1	D	521	LEU	2.7
1	A	684	LYS	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	847	LEU	2.6
1	A	700	ASN	2.6
1	A	442	VAL	2.6
1	A	298	LYS	2.5
1	A	689	ILE	2.5
1	D	526	PHE	2.4
1	A	691	GLN	2.4
1	A	686	ALA	2.4
2	E	21	DC	2.4
1	A	678	ASP	2.4
1	A	843	ARG	2.4
1	A	694	GLU	2.3
1	A	844	LEU	2.3
1	A	516	GLN	2.3
1	D	531	PRO	2.2
1	D	524	GLN	2.2
1	D	554	TYR	2.2
1	A	868	TYR	2.1
3	F	1	DC	2.1
1	A	509	GLU	2.1
1	A	688	ASP	2.0
1	A	839	GLU	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
2	2DT	B	29	19/20	0.95	0.12	18,25,31,32	0
3	3DR	C	3	11/12	0.96	0.09	27,33,40,42	0
3	3DR	F	3	11/12	0.96	0.09	18,26,37,38	0
2	2DT	E	29	19/20	0.99	0.10	12,16,18,20	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

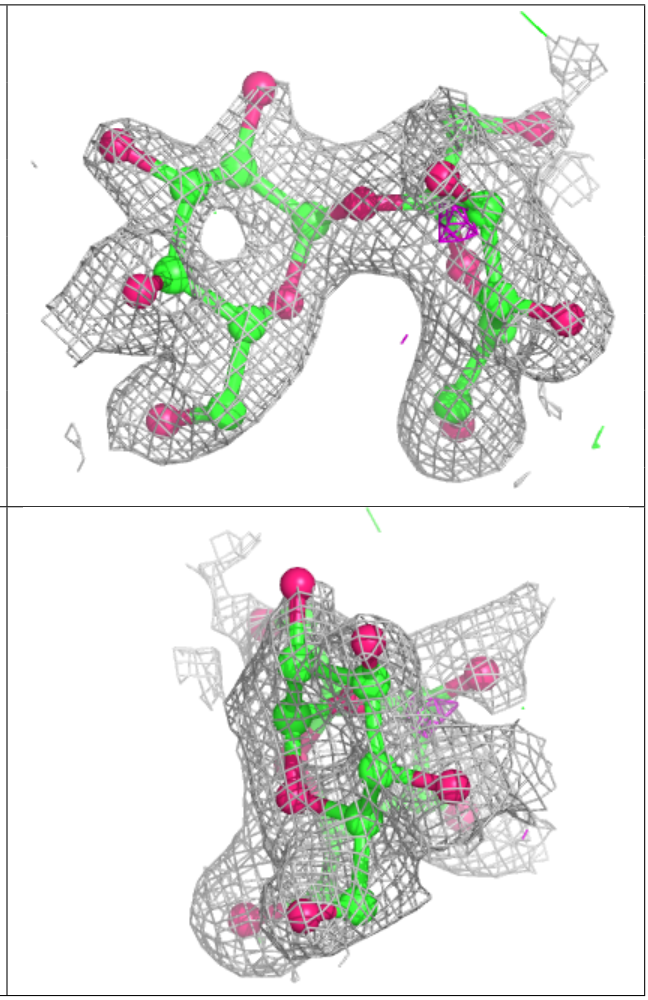
labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	GLC	G	1	11/12	0.81	0.22	43,56,59,68	0
4	GLC	H	1	11/12	0.83	0.19	26,40,49,49	0
4	FRU	G	2	12/12	0.86	0.14	34,51,57,59	0
4	FRU	H	2	12/12	0.90	0.15	21,34,41,42	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.

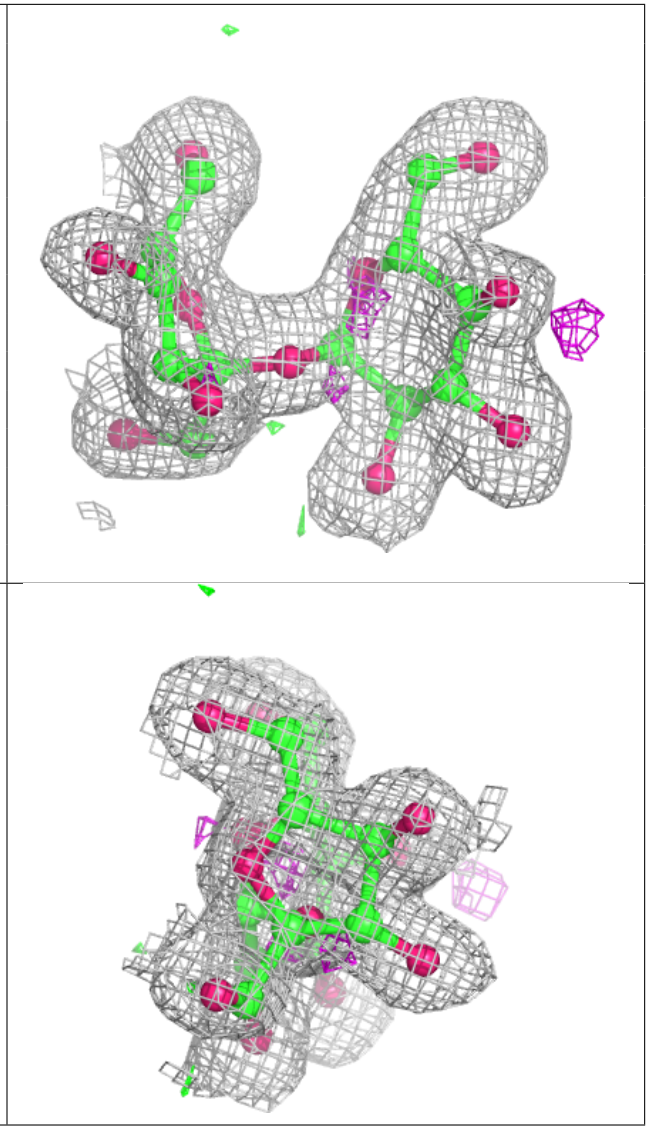
Electron density around Chain G:

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 (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.

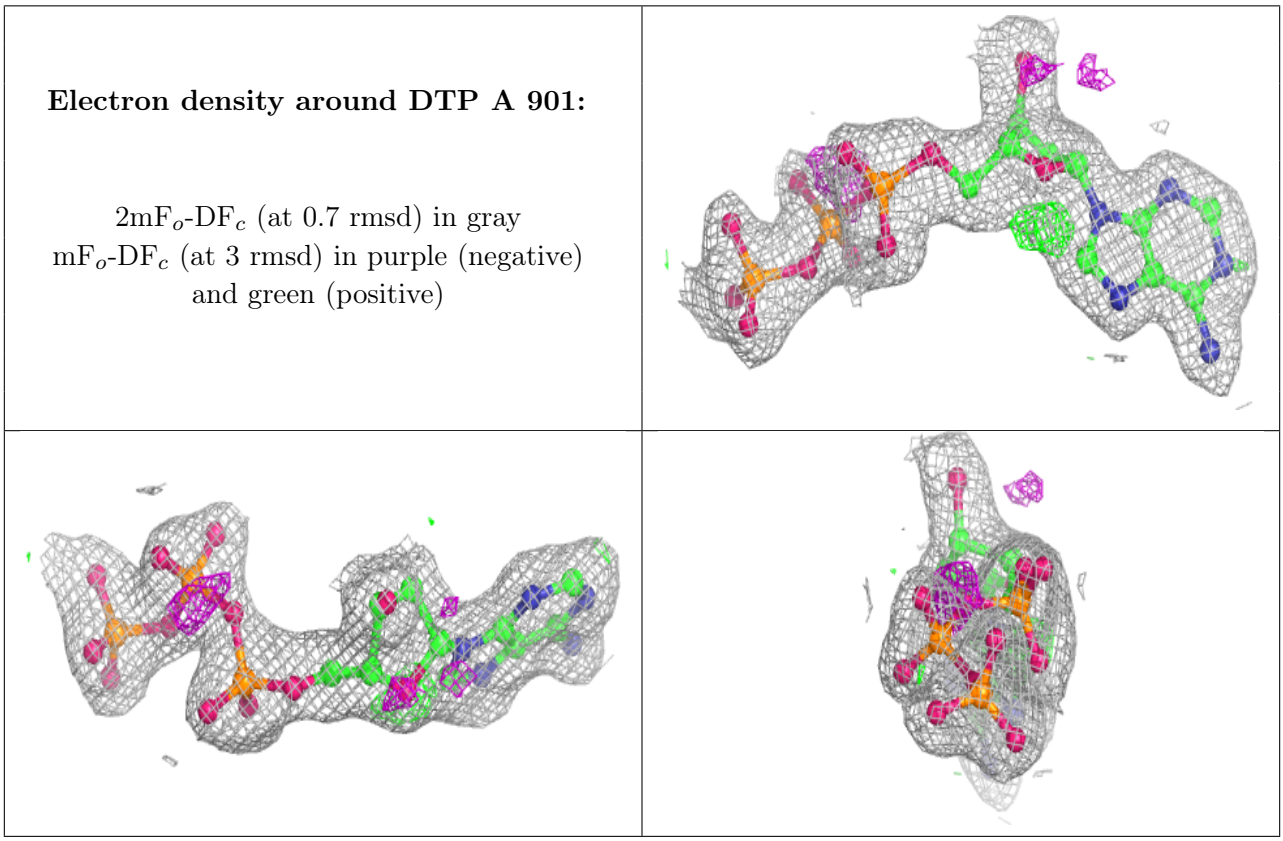
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	MPD	D	904	8/8	0.86	0.18	31,34,39,40	0
5	DTP	A	901	30/30	0.94	0.10	23,34,54,59	0
5	DTP	D	901	30/30	0.98	0.10	15,20,26,32	0
6	SO4	D	903	5/5	0.99	0.07	25,30,36,36	0

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.

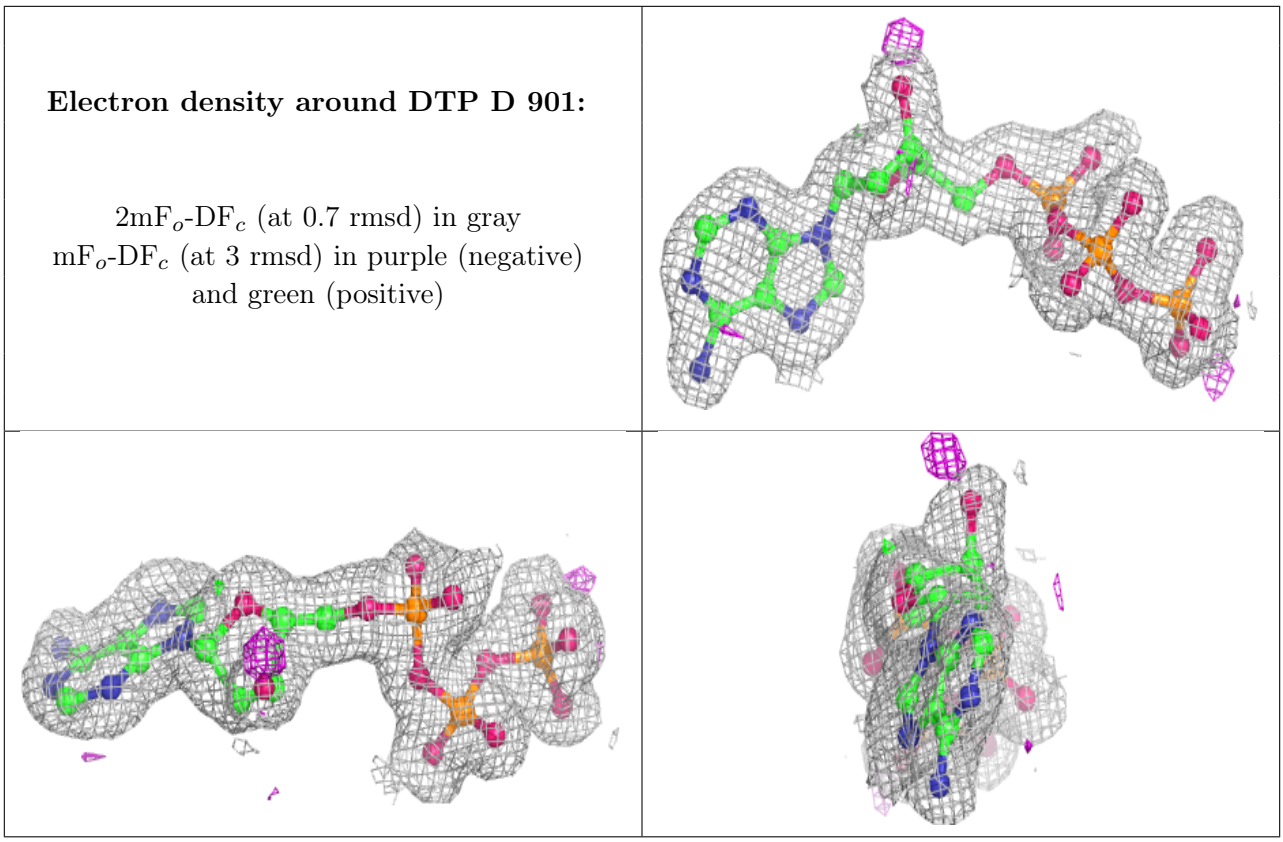
Electron density around DTP A 901:

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



Electron density around DTP D 901:

$2mF_o$-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.