Full wwPDB X-ray Structure Validation Report ⓘ

Aug 26, 2023 – 04:53 PM EDT

PDB ID	:	3EZ5
Title	:	Cocrystal structure of Bacillus fragment DNA polymerase I with duplex DNA , dCTP, and zinc (closed form).
Authors	:	Warren, J.J.; Wu, E.Y.; Golosov, A.A.; Karplus, M.; Beese, L.S.
Deposited on	:	2008-10-22
Resolution	:	1.90 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:

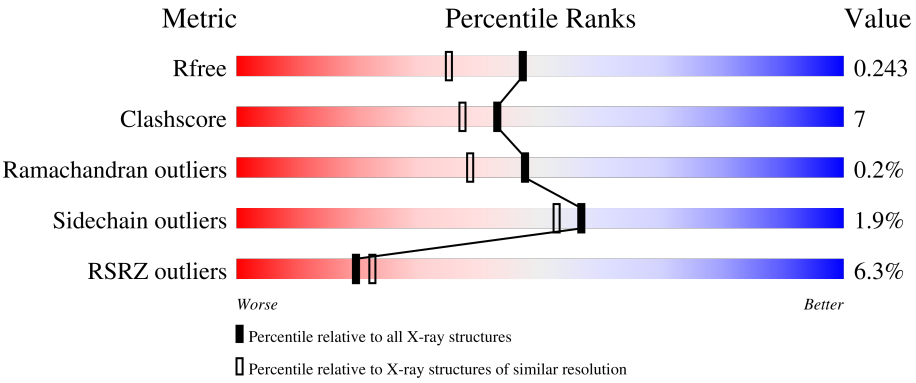
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	11% 83% 16%
1	D	580	2% 88% 11%
2	B	9	44% 33% 22%
2	E	9	44% 33% 22%
3	C	12	8% 50% 42% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	12	50% 50%
4	G	2	50% 50%
4	H	2	50% 50%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	ZN	D	3	-	-	-	X
7	DAD	A	201	X	-	-	-
7	DAD	D	202	X	-	-	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 10822 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
1	A	580	Total	C	N	O	S	0	1	0
			4660	2963	811	869	17			
1	D	579	Total	C	N	O	S	0	1	0
			4652	2958	807	870	17			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	710	TYR	PHE	engineered mutation	PDB 3EZ5
D	710	TYR	PHE	engineered mutation	PDB 3EZ5

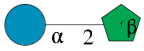
- Molecule 2 is a DNA chain called 5'-D(*DCP*DCP*DTP*DGP*DAP*DCP*DTP*DCP*DG)-3'.

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

- Molecule 3 is a DNA chain called 5'-D(*DAP*DTP*DTP*DCP*DGP*DAP*DGP*DTP*DCP*DAP*DGP*DG)-3'.

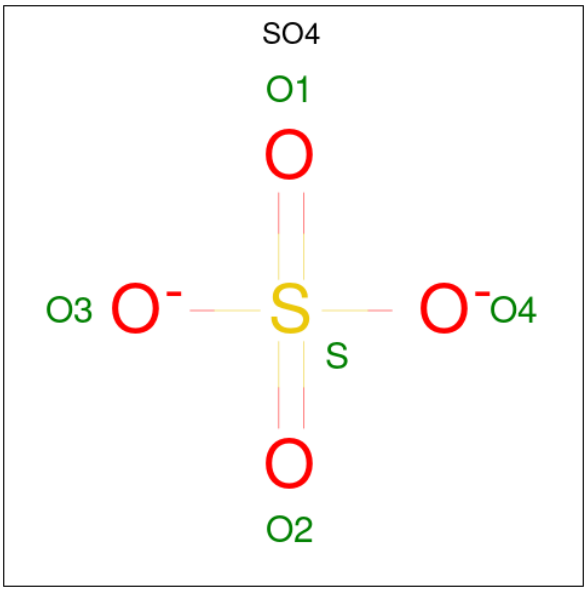
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	12	Total	C	N	O	P	0	0	0
			246	118	47	70	11			
3	F	12	Total	C	N	O	P	0	0	0
			246	118	47	70	11			

- 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 23 12 11	0	0	0
4	H	2	Total C O 23 12 11	0	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

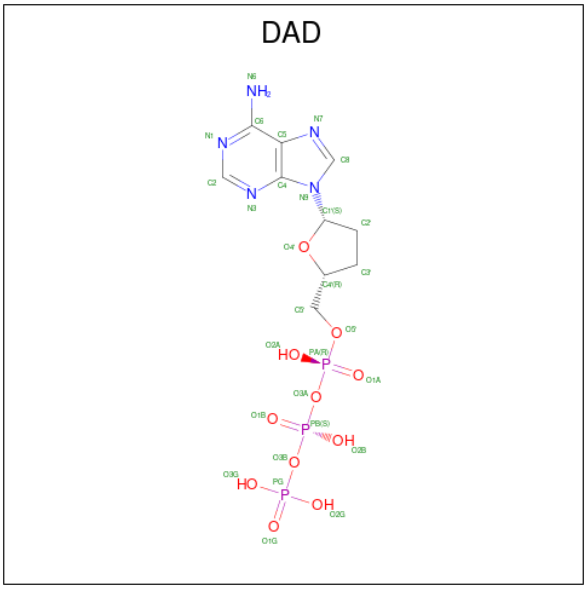


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	D	1	Total O S 5 4 1	0	0

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	2	Total Zn 2 2	0	0
6	D	3	Total Zn 3 3	0	0

- Molecule 7 is 2',3'-DIDEOXYADENOSINE-5'-TRIPHOSPHATE (three-letter code: DAD) (formula: $C_{10}H_{16}N_5O_{11}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	P	0	0
			29	10	5	11	3		
7	D	1	Total	C	N	O	P	0	0
			29	10	5	11	3		

- Molecule 8 is water.

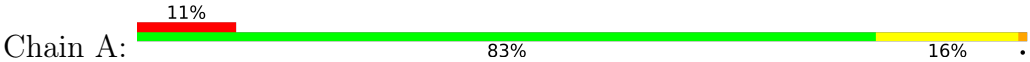
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	154	Total	O	0	0
			154	154		
8	D	290	Total	O	0	0
			290	290		
8	B	20	Total	O	0	0
			20	20		
8	C	20	Total	O	0	0
			20	20		
8	E	19	Total	O	0	0
			19	19		
8	F	37	Total	O	0	0
			37	37		

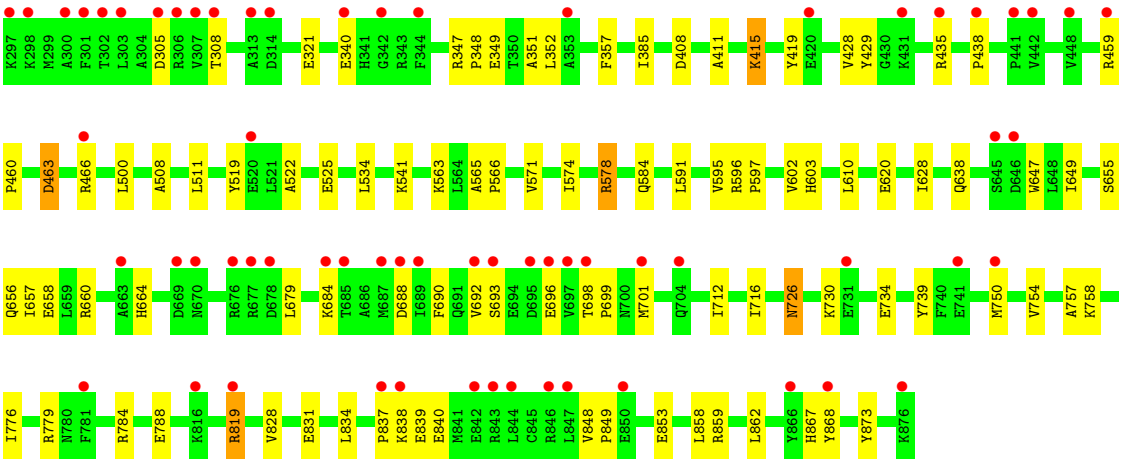
3 Residue-property plots (i)

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

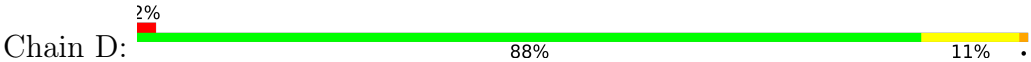
Chain A:

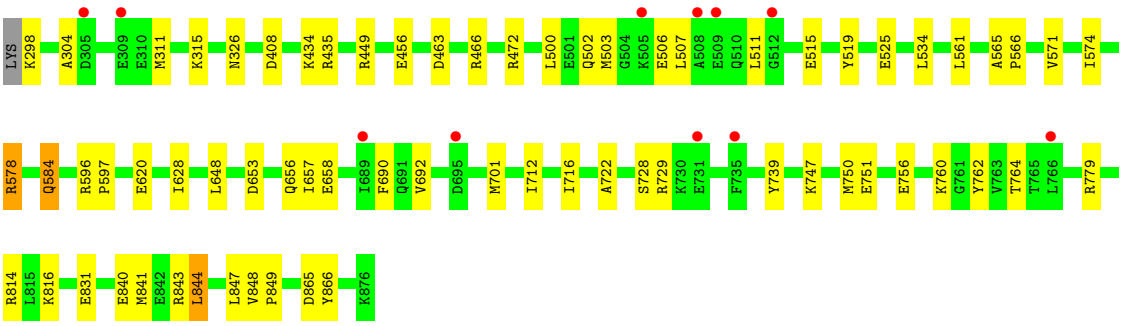




• Molecule 1: DNA polymerase I

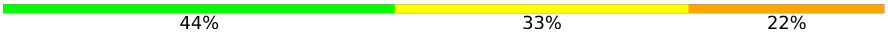
Chain D:

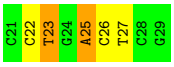




• Molecule 2: 5'-D(*DCP*DCP*DTP*DGP*DAP*DCP*DTP*DCP*DG)-3'

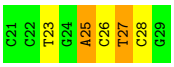
Chain B:





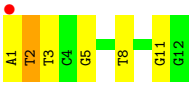
• Molecule 2: 5'-D(*DCP*DCP*DTP*DGP*DAP*DCP*DTP*DCP*DG)-3'

Chain E:



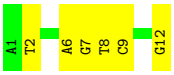
• Molecule 3: 5'-D(*DAP*DTP*DTP*DCP*DGP*DAP*DGP*DTP*DCP*DAP*DGP*DG)-3
'

Chain C:



• Molecule 3: 5'-D(*DAP*DTP*DTP*DCP*DGP*DAP*DGP*DTP*DCP*DAP*DGP*DG)-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	92.55Å 108.57Å 149.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.83 – 1.90 46.83 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.83-1.90) 100.0 (46.83-1.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.86 (at 1.90Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.211 , 0.244 0.210 , 0.243	Depositor DCC
R_{free} test set	5010 reflections (4.21%)	wwPDB-VP
Wilson B-factor (Å^2)	32.8	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 45.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10822	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.12% 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: FRU, GLC, ZN, SO4, DAD

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	0.44	0/4747	0.58	1/6415 (0.0%)
1	D	0.57	0/4739	0.65	1/6404 (0.0%)
2	B	1.02	0/197	1.95	7/301 (2.3%)
2	E	1.06	0/197	1.97	8/301 (2.7%)
3	C	1.07	0/276	1.71	7/425 (1.6%)
3	F	1.08	0/276	1.66	6/425 (1.4%)
All	All	0.58	0/10432	0.82	30/14271 (0.2%)

There are no bond length outliers.

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	25	DA	O4'-C4'-C3'	-9.93	100.04	106.00
3	C	2	DT	O4'-C1'-N1	9.39	114.57	108.00
3	C	8	DT	O4'-C1'-N1	-9.20	101.56	108.00
2	E	27	DT	O4'-C4'-C3'	-8.45	100.93	106.00
2	E	23	DT	O4'-C1'-N1	-8.07	102.35	108.00
3	F	2	DT	O4'-C1'-N1	8.07	113.65	108.00
2	E	23	DT	N3-C4-O4	7.25	124.25	119.90
3	F	8	DT	O4'-C1'-N1	-7.22	102.94	108.00
2	B	25	DA	O5'-P-OP2	-7.01	99.39	105.70
3	C	2	DT	C1'-O4'-C4'	-6.80	103.30	110.10
2	B	23	DT	O4'-C1'-N1	-6.67	103.33	108.00
1	A	578	ARG	NE-CZ-NH2	-6.63	116.98	120.30
3	F	12	DG	O4'-C1'-N9	6.55	112.58	108.00
3	C	3	DT	N3-C4-O4	6.50	123.80	119.90
2	E	23	DT	C5-C4-O4	-6.36	120.45	124.90
2	B	25	DA	C5'-C4'-C3'	6.28	125.40	114.10
2	B	27	DT	O4'-C4'-C3'	-6.18	102.03	104.50
3	F	7	DG	O4'-C1'-N9	-6.13	103.71	108.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	11	DG	O4'-C1'-N9	-6.00	103.80	108.00
2	B	27	DT	O4'-C1'-N1	5.87	112.11	108.00
3	F	9	DC	N1-C2-O2	-5.79	115.43	118.90
2	E	25	DA	C5'-C4'-C3'	5.77	124.48	114.10
3	C	5	DG	O4'-C4'-C3'	-5.58	102.27	104.50
2	B	23	DT	N3-C2-O2	-5.42	119.05	122.30
3	C	3	DT	C5-C4-O4	-5.37	121.14	124.90
2	E	28	DC	O4'-C4'-C3'	-5.32	102.37	104.50
3	F	6	DA	O4'-C1'-C2'	5.30	110.14	105.90
2	E	25	DA	O5'-P-OP2	-5.26	100.97	105.70
2	E	25	DA	N1-C6-N6	5.21	121.73	118.60
1	D	578	ARG	NE-CZ-NH2	-5.08	117.76	120.30

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	4660	0	4720	81	0
1	D	4652	0	4709	50	0
2	B	177	0	101	3	0
2	E	177	0	101	3	0
3	C	246	0	137	3	0
3	F	246	0	137	0	0
4	G	23	0	21	2	0
4	H	23	0	21	1	0
5	A	10	0	0	1	0
5	D	5	0	0	0	0
6	A	2	0	0	0	0
6	D	3	0	0	0	0
7	A	29	0	12	2	0
7	D	29	0	12	1	0
8	A	154	0	0	5	0
8	B	20	0	0	0	0
8	C	20	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	D	290	0	0	8	0
8	E	19	0	0	0	0
8	F	37	0	0	0	0
All	All	10822	0	9971	138	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 (138) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:408:ASP:HB2	4:G:2:FRU:H11	1.19	1.16
1:A:819:ARG:HH11	1:A:819:ARG:HG3	1.01	1.15
1:A:819:ARG:HG3	1:A:819:ARG:NH1	1.82	0.87
1:A:522:ALA:O	1:A:541:LYS:HE2	1.74	0.87
1:D:653[A]:ASP:OD1	1:D:831:GLU:HB2	1.75	0.86
1:A:739:TYR:HD2	8:A:917:HOH:O	1.58	0.86
1:D:750:MET:HG3	8:D:938:HOH:O	1.83	0.77
1:D:561:LEU:HB3	1:D:571:VAL:HG13	1.66	0.76
1:A:408:ASP:HB2	4:G:2:FRU:C1	2.08	0.75
1:A:656:GLN:HB2	1:A:660:ARG:NH1	2.01	0.75
1:D:561:LEU:O	1:D:571:VAL:HG11	1.87	0.74
1:D:456:GLU:HG2	8:D:929:HOH:O	1.88	0.72
1:A:519:TYR:CD2	1:A:525:GLU:HG2	2.25	0.71
1:D:578:ARG:NH2	2:E:25:DA:H5"	2.05	0.70
1:A:679:LEU:HD13	1:A:684:LYS:NZ	2.06	0.70
3:C:1:DA:H2"	3:C:2:DT:O5'	1.92	0.69
1:A:739:TYR:CD2	8:A:917:HOH:O	2.39	0.69
1:D:690:PHE:CG	1:D:701:MET:HE3	2.29	0.68
1:A:819:ARG:HH11	1:A:819:ARG:CG	1.92	0.65
1:A:656:GLN:HB2	1:A:660:ARG:HH12	1.62	0.64
1:D:653[B]:ASP:OD1	1:D:865:ASP:O	2.17	0.62
1:A:656:GLN:HA	7:A:201:DAD:O1B	2.00	0.61
1:D:690:PHE:CB	1:D:701:MET:HE3	2.31	0.61
1:D:298:LYS:HB3	1:D:449:ARG:NH2	2.16	0.60
1:D:534:LEU:HD11	1:D:574:ILE:HD13	1.82	0.60
1:A:784:ARG:O	1:A:788:GLU:HG3	2.01	0.60
1:A:596[B]:ARG:NH1	8:A:920:HOH:O	2.35	0.59
1:A:828:VAL:HB	1:A:831:GLU:HG3	1.85	0.59
1:A:411:ALA:O	1:A:415:LYS:HG2	2.03	0.58
1:D:690:PHE:O	1:D:692:VAL:HG13	2.02	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:848:VAL:HB	1:A:849:PRO:HD3	1.84	0.58
1:D:690:PHE:CD2	1:D:701:MET:HE3	2.38	0.58
1:A:305:ASP:HB3	1:A:347:ARG:HH12	1.68	0.57
1:A:692:VAL:HG21	1:A:701:MET:HE1	1.85	0.57
1:A:459:ARG:HB2	1:A:460:PRO:HD3	1.86	0.57
1:A:657:ILE:HG23	1:A:658:GLU:N	2.20	0.56
1:D:656:GLN:HA	7:D:202:DAD:O1B	2.05	0.56
1:A:563:LYS:HG2	1:A:726:ASN:HD21	1.69	0.56
8:A:884:HOH:O	3:C:2:DT:H71	2.05	0.56
3:C:2:DT:H6	8:C:328:HOH:O	1.88	0.56
1:D:840:GLU:O	1:D:844:LEU:HD22	2.07	0.55
1:D:692:VAL:HG21	1:D:701:MET:HE1	1.88	0.55
1:D:712:ILE:HA	1:D:716:ILE:HG22	1.89	0.55
1:D:408:ASP:HB2	4:H:2:FRU:H11	1.88	0.55
1:D:739:TYR:CD2	1:D:739:TYR:C	2.81	0.53
1:A:664:HIS:CE1	1:A:859:ARG:HH11	2.25	0.53
1:A:838:LYS:HG3	1:A:839:GLU:OE2	2.08	0.53
1:D:519:TYR:CD2	1:D:525:GLU:HG2	2.43	0.53
1:A:349:GLU:H	1:A:349:GLU:CD	2.12	0.53
1:A:754:VAL:O	1:A:758:LYS:HG3	2.10	0.51
1:A:858:LEU:HD12	1:A:862:LEU:HD21	1.92	0.51
1:D:722:ALA:HB2	1:D:729:ARG:HA	1.93	0.50
1:D:304:ALA:HB2	1:D:311:MET:HE2	1.93	0.50
1:A:591:LEU:O	1:A:595:VAL:HG23	2.12	0.50
1:A:351:ALA:HB1	1:A:357:PHE:CD2	2.47	0.49
1:A:578:ARG:NH2	2:B:25:DA:H5"	2.27	0.49
1:A:839:GLU:H	1:A:839:GLU:CD	2.16	0.49
1:A:679:LEU:HD13	1:A:684:LYS:HZ1	1.76	0.49
1:A:690:PHE:CG	1:A:701:MET:HE3	2.49	0.48
1:D:298:LYS:NZ	8:D:961:HOH:O	2.46	0.48
1:A:658:GLU:CD	7:A:201:DAD:H2'1	2.34	0.48
1:D:648:LEU:HD12	1:D:841:MET:HG3	1.96	0.48
1:A:610:LEU:HD23	1:A:610:LEU:C	2.34	0.48
2:B:22:DC:H1'	2:B:23:DT:H5'	1.96	0.48
1:A:308:THR:HA	5:A:1:SO4:O4	2.14	0.48
1:A:596[B]:ARG:NH1	1:A:603:HIS:HD2	2.11	0.48
1:A:837:PRO:HG2	1:A:840:GLU:HG3	1.96	0.48
1:D:762:TYR:HE1	1:D:764:THR:CG2	2.27	0.48
1:A:508:ALA:HB2	1:A:584:GLN:HE22	1.80	0.47
1:D:298:LYS:HB3	1:D:449:ARG:HH22	1.78	0.47
1:A:348:PRO:HD2	1:A:349:GLU:OE2	2.14	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:500:LEU:HD21	1:A:591:LEU:HD23	1.96	0.47
1:D:456:GLU:CG	8:D:929:HOH:O	2.52	0.47
1:D:584:GLN:HA	1:D:584:GLN:HE21	1.80	0.47
1:A:690:PHE:CE2	1:A:701:MET:HB3	2.50	0.47
1:D:816:LYS:HD3	1:D:816:LYS:HA	1.66	0.47
1:D:843:ARG:HD2	8:D:965:HOH:O	2.14	0.46
1:D:690:PHE:HB2	1:D:701:MET:CE	2.45	0.46
1:A:459:ARG:N	1:A:460:PRO:CD	2.79	0.46
1:A:660:ARG:HH11	1:A:660:ARG:HG3	1.79	0.46
1:D:657:ILE:HG23	1:D:658:GLU:N	2.31	0.46
1:D:747:LYS:HE2	1:D:751:GLU:OE1	2.16	0.46
1:D:750:MET:CG	8:D:938:HOH:O	2.54	0.46
1:A:849:PRO:O	1:A:853:GLU:HG3	2.16	0.45
1:A:693:SER:OG	1:A:696:GLU:HG3	2.16	0.45
1:A:595:VAL:HG22	1:A:602:VAL:HG13	1.99	0.45
1:D:463:ASP:OD2	1:D:466:ARG:NH2	2.50	0.45
1:A:415:LYS:HE2	1:A:419:TYR:O	2.16	0.45
1:D:326:ASN:HD22	1:D:620:GLU:CD	2.20	0.45
1:A:819:ARG:NH1	1:A:819:ARG:CG	2.63	0.45
1:D:866:TYR:N	1:D:866:TYR:CD2	2.85	0.45
1:A:657:ILE:HG23	1:A:658:GLU:H	1.80	0.45
1:A:647:TRP:CE2	1:A:837:PRO:HD3	2.52	0.44
1:A:730:LYS:O	1:A:734:GLU:HG3	2.16	0.44
1:A:664:HIS:CE1	1:A:859:ARG:NH1	2.85	0.44
1:A:757:ALA:HB3	1:A:776:ILE:HD13	1.99	0.44
1:A:757:ALA:CB	1:A:776:ILE:HD13	2.48	0.44
1:D:503:MET:O	1:D:507:LEU:HB2	2.18	0.44
1:A:595:VAL:HG22	1:A:602:VAL:CG1	2.47	0.44
1:A:828:VAL:HB	1:A:831:GLU:CG	2.47	0.44
1:A:638:GLN:HG3	1:A:873:TYR:CG	2.52	0.44
1:D:565:ALA:N	1:D:566:PRO:CD	2.80	0.44
1:A:428:VAL:HG12	1:A:438:PRO:HG3	2.00	0.43
1:A:664:HIS:O	1:A:859:ARG:HD2	2.18	0.43
1:A:690:PHE:CB	1:A:701:MET:HE3	2.48	0.43
1:A:834:LEU:HD12	1:A:834:LEU:N	2.33	0.43
1:A:867:HIS:CE1	8:A:7:HOH:O	2.71	0.43
2:E:25:DA:H2'	2:E:26:DC:C6	2.54	0.43
1:A:698:THR:HB	1:A:699:PRO:HD2	2.00	0.43
1:D:472:ARG:NH1	8:D:989:HOH:O	2.51	0.43
1:A:750:MET:O	1:A:754:VAL:HG23	2.19	0.43
1:A:415:LYS:HD3	1:A:415:LYS:HA	1.75	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:596[B]:ARG:HH12	1:A:603:HIS:HB2	1.84	0.42
1:A:565:ALA:HA	1:A:571:VAL:CG2	2.50	0.42
1:A:463:ASP:HA	1:A:466:ARG:NH1	2.34	0.42
1:A:649:ILE:O	1:A:868:TYR:HA	2.19	0.42
1:A:534:LEU:HD11	1:A:574:ILE:HD13	2.01	0.42
1:D:315:LYS:HD2	8:D:988:HOH:O	2.18	0.42
1:A:596[B]:ARG:NH1	1:A:603:HIS:CD2	2.88	0.42
1:D:848:VAL:HB	1:D:849:PRO:HD3	2.01	0.42
2:E:26:DC:H2'	2:E:27:DT:C6	2.55	0.42
1:A:565:ALA:HA	1:A:571:VAL:HG21	2.01	0.42
1:A:596[A]:ARG:HA	1:A:597:PRO:HD3	1.82	0.42
1:D:756:GLU:OE2	1:D:760:LYS:HE2	2.20	0.41
1:D:840:GLU:O	1:D:844:LEU:CD2	2.68	0.41
1:A:828:VAL:O	1:A:828:VAL:HG12	2.21	0.41
1:D:596:ARG:HA	1:D:597:PRO:HD3	1.90	0.41
1:A:321:GLU:OE2	1:A:429:TYR:OH	2.35	0.41
1:A:712:ILE:HA	1:A:716:ILE:HG22	2.03	0.41
1:D:502:GLN:O	1:D:506:GLU:HG3	2.21	0.41
2:B:25:DA:H2'	2:B:26:DC:C6	2.56	0.41
1:D:690:PHE:CD2	1:D:701:MET:CE	3.04	0.41
1:A:500:LEU:HD12	1:A:500:LEU:HA	1.94	0.40
1:A:565:ALA:N	1:A:566:PRO:CD	2.83	0.40
1:D:814:ARG:NH2	1:D:847:LEU:HD13	2.36	0.40
1:A:352:LEU:HD22	1:A:385:ILE:HD13	2.04	0.40
1:D:515:GLU:HG2	1:D:519:TYR:CE2	2.57	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	579/580 (100%)	560 (97%)	18 (3%)	1 (0%)	47 38

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	D	578/580 (100%)	566 (98%)	11 (2%)	1 (0%)	47 38
All	All	1157/1160 (100%)	1126 (97%)	29 (2%)	2 (0%)	47 38

All (2) Ramachandran outliers are listed below:

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

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	496/496 (100%)	485 (98%)	11 (2%)	52 47
1	D	496/496 (100%)	488 (98%)	8 (2%)	62 60
All	All	992/992 (100%)	973 (98%)	19 (2%)	57 53

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	340	GLU
1	A	415	LYS
1	A	435	ARG
1	A	463	ASP
1	A	511	LEU
1	A	620	GLU
1	A	655	SER
1	A	688	ASP
1	A	726	ASN
1	A	779	ARG
1	A	819	ARG
1	D	434	LYS
1	D	435	ARG
1	D	500	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	511	LEU
1	D	584	GLN
1	D	728	SER
1	D	779	ARG
1	D	844	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	584	GLN
1	A	724	ASN
1	A	726	ASN
1	D	502	GLN
1	D	584	GLN
1	D	823	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

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	0.47	0	15,15,17	0.97	0
4	FRU	G	2	4	11,12,12	0.44	0	10,18,18	0.93	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	GLC	H	1	4	11,11,12	0.59	0	15,15,17	1.15	1 (6%)
4	FRU	H	2	4	11,12,12	0.60	0	10,18,18	1.04	1 (10%)

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	-	4/5/24/24	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

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	H	1	GLC	C1-O5-C5	2.34	115.36	112.19
4	H	2	FRU	O6-C6-C5	-2.07	104.20	111.29

There are no chirality outliers.

All (4) 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-O2
4	G	2	FRU	O1-C1-C2-O5
4	G	2	FRU	C4-C5-C6-O6

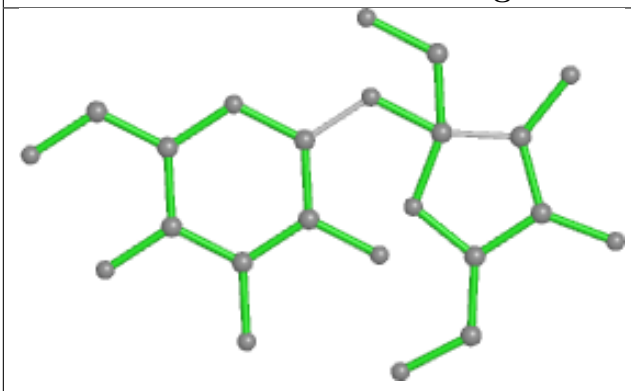
There are no ring outliers.

2 monomers are involved in 3 short contacts:

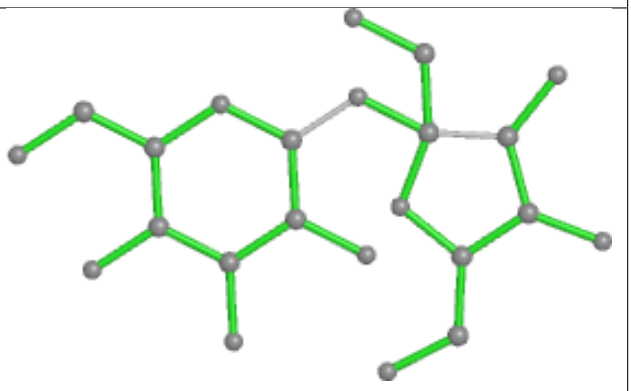
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	2	FRU	2	0
4	H	2	FRU	1	0

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

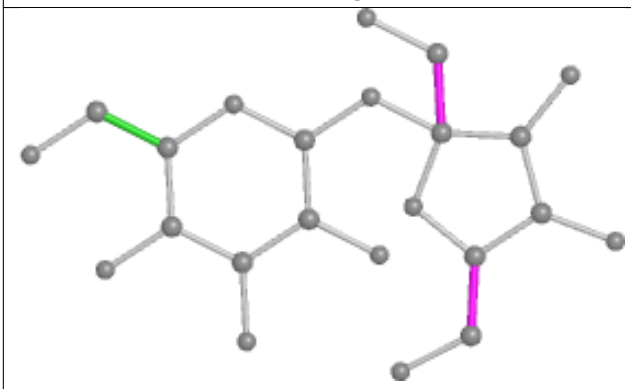
Oligosaccharide Chain G



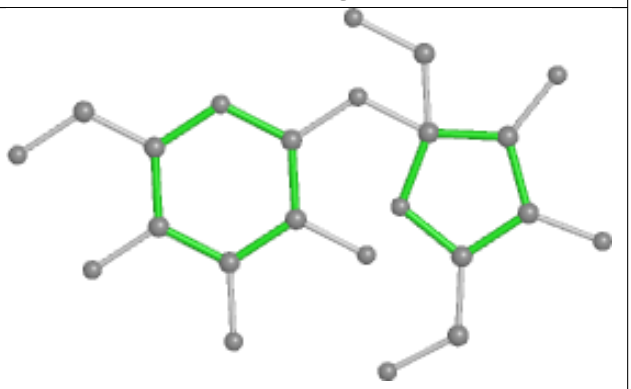
Bond lengths



Bond angles

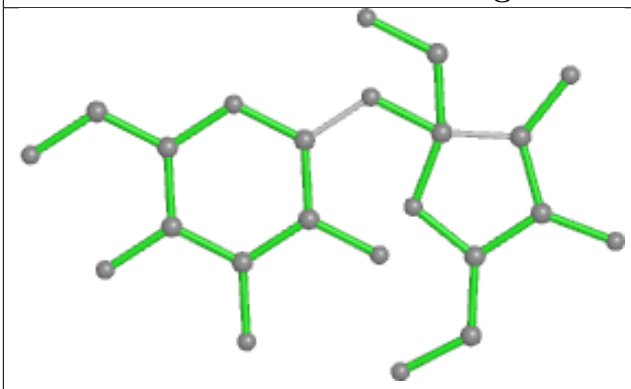


Torsions

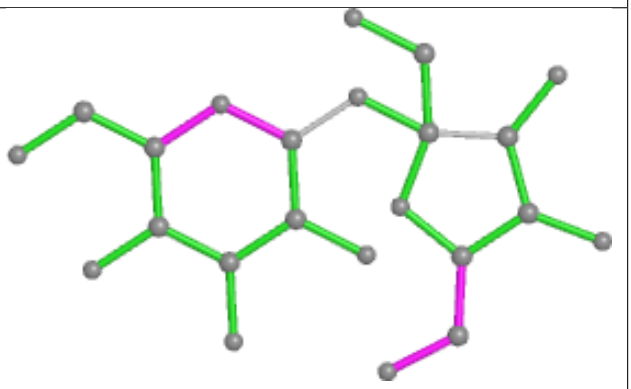


Rings

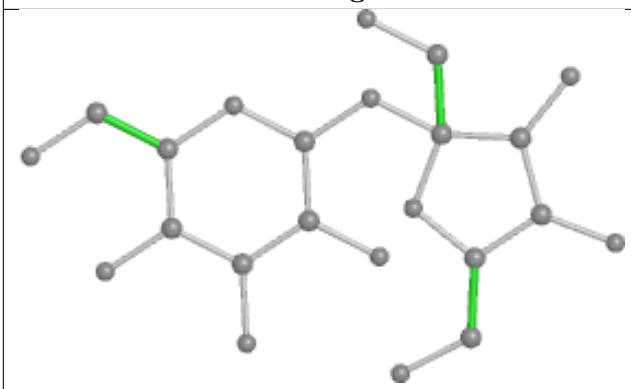
Oligosaccharide Chain H



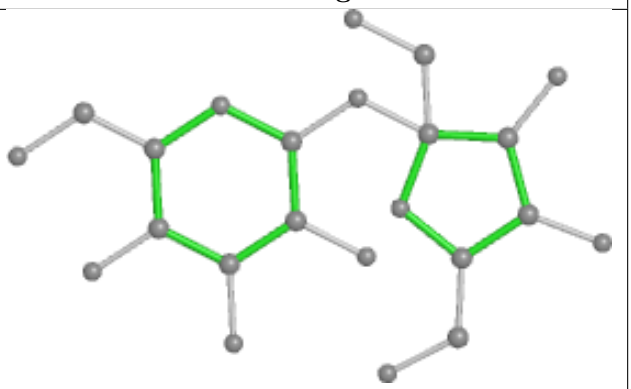
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 5 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	DAD	A	201	6	25,31,31	2.47	8 (32%)	26,48,48	1.90	5 (19%)
5	SO4	A	2	-	4,4,4	0.15	0	6,6,6	0.27	0
7	DAD	D	202	6	25,31,31	2.25	7 (28%)	26,48,48	1.92	9 (34%)
5	SO4	D	878	-	4,4,4	0.23	0	6,6,6	0.51	0
5	SO4	A	1	-	4,4,4	0.18	0	6,6,6	0.26	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	DAD	D	202	6	2/2/5/5	6/18/31/31	0/3/3/3
7	DAD	A	201	6	2/2/5/5	3/18/31/31	0/3/3/3

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	201	DAD	PG-O2G	6.35	1.79	1.54
7	A	201	DAD	PG-O3G	5.71	1.76	1.54
7	D	202	DAD	PG-O3G	5.58	1.76	1.54
7	D	202	DAD	PG-O2G	5.40	1.75	1.54
7	A	201	DAD	PB-O2B	4.67	1.77	1.55
7	D	202	DAD	PB-O2B	4.63	1.77	1.55
7	A	201	DAD	PA-O2A	4.55	1.76	1.55
7	D	202	DAD	PA-O2A	4.37	1.75	1.55
7	A	201	DAD	PG-O1G	3.00	1.60	1.50
7	A	201	DAD	C5-C4	2.56	1.47	1.40
7	A	201	DAD	PB-O1B	2.55	1.60	1.50
7	D	202	DAD	PG-O1G	2.48	1.58	1.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	201	DAD	PA-O1A	2.26	1.58	1.50
7	D	202	DAD	C5-C4	2.05	1.46	1.40
7	D	202	DAD	PB-O1B	2.03	1.58	1.50

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	201	DAD	C2'-C1'-N9	-5.72	101.71	112.48
7	D	202	DAD	C2'-C1'-N9	-5.00	103.07	112.48
7	A	201	DAD	N3-C2-N1	-4.36	121.86	128.68
7	D	202	DAD	N3-C2-N1	-3.13	123.79	128.68
7	D	202	DAD	C3'-C2'-C1'	2.78	106.00	102.78
7	D	202	DAD	C4-C5-N7	-2.66	106.63	109.40
7	D	202	DAD	O4'-C1'-C2'	-2.64	103.81	106.67
7	D	202	DAD	C4'-O4'-C1'	2.63	112.29	109.81
7	A	201	DAD	PA-O3A-PB	-2.48	124.33	132.83
7	D	202	DAD	PA-O3A-PB	-2.42	124.52	132.83
7	A	201	DAD	O2G-PG-O3B	2.38	112.63	104.64
7	A	201	DAD	C2-N1-C6	2.27	122.63	118.75
7	D	202	DAD	C5-C6-N6	2.03	123.44	120.35
7	D	202	DAD	PB-O3B-PG	-2.02	125.90	132.83

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
7	A	201	DAD	C1'
7	A	201	DAD	C4'
7	D	202	DAD	C1'
7	D	202	DAD	C4'

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	D	202	DAD	O4'-C4'-C5'-O5'
7	D	202	DAD	PB-O3B-PG-O1G
7	A	201	DAD	PG-O3B-PB-O2B
7	D	202	DAD	PB-O3A-PA-O2A
7	A	201	DAD	PB-O3A-PA-O2A
7	D	202	DAD	PG-O3B-PB-O2B
7	D	202	DAD	PG-O3B-PB-O1B
7	A	201	DAD	PG-O3B-PB-O1B
7	D	202	DAD	PB-O3A-PA-O1A

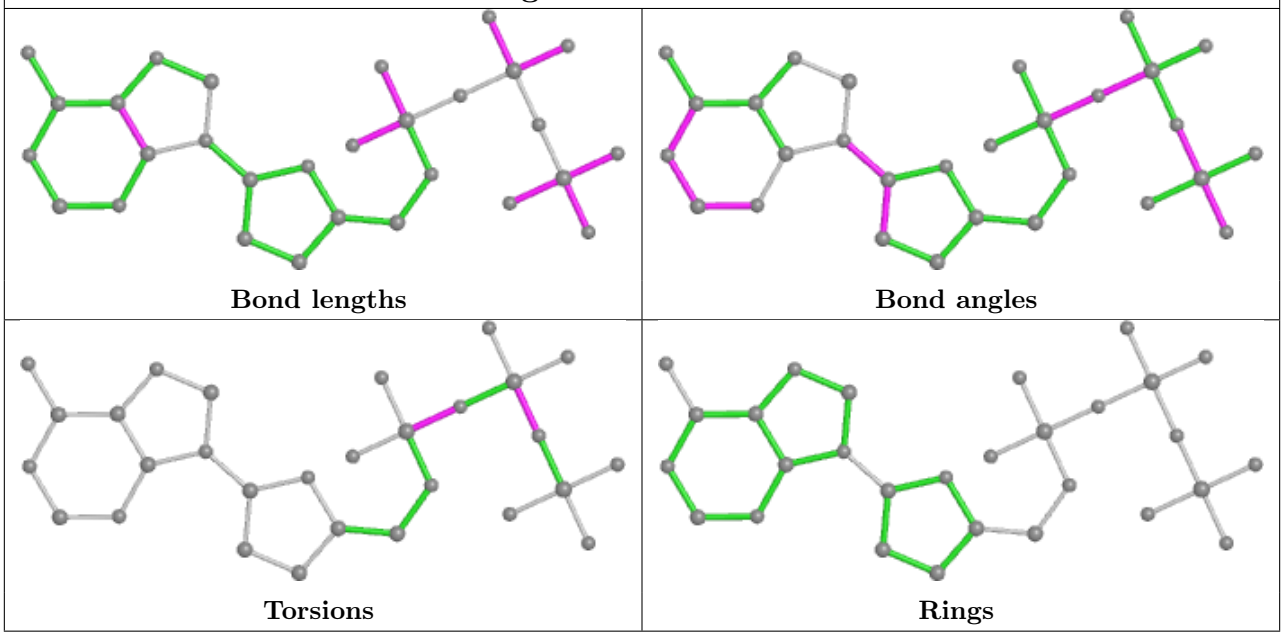
There are no ring outliers.

3 monomers are involved in 4 short contacts:

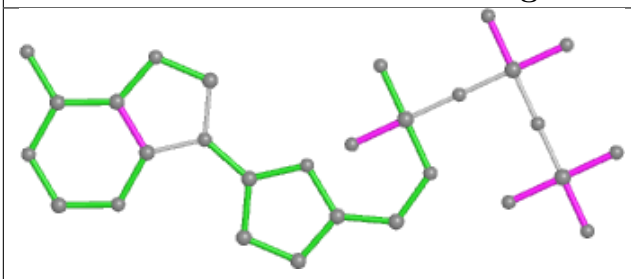
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	201	DAD	2	0
7	D	202	DAD	1	0
5	A	1	SO4	1	0

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.

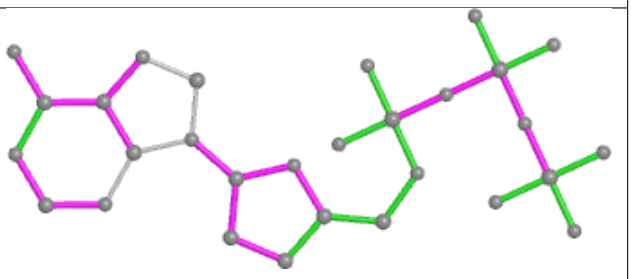
Ligand DAD A 201



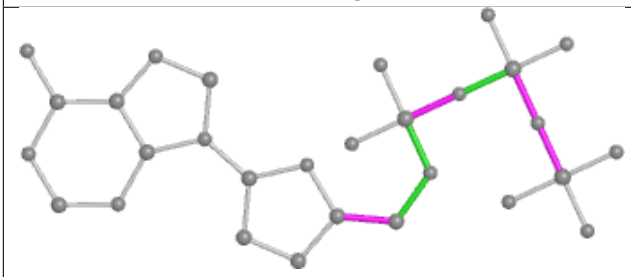
Ligand DAD D 202



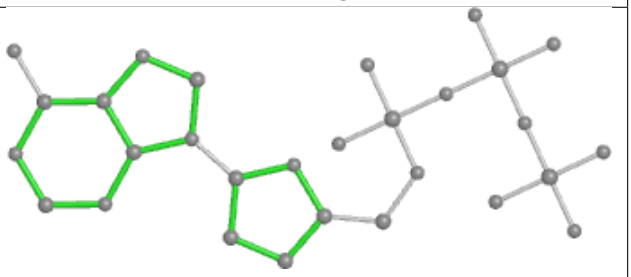
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	A	580/580 (100%)	0.68	64 (11%) 5 6	23, 41, 61, 70	0
1	D	579/580 (99%)	0.32	11 (1%) 66 69	17, 30, 48, 60	0
2	B	9/9 (100%)	-0.23	0 100 100	26, 28, 43, 54	0
2	E	9/9 (100%)	-0.28	0 100 100	20, 28, 40, 50	0
3	C	12/12 (100%)	0.18	1 (8%) 11 13	24, 30, 67, 87	0
3	F	12/12 (100%)	-0.05	0 100 100	19, 26, 56, 73	0
All	All	1201/1202 (99%)	0.48	76 (6%) 20 22	17, 35, 58, 87	0

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	298	LYS	5.2
1	A	677	ARG	4.8
1	A	819	ARG	4.7
1	A	693	SER	4.4
1	A	696	GLU	4.0
1	A	678	ASP	3.5
1	A	297	LYS	3.5
1	A	685	THR	3.4
1	A	353	ALA	3.4
1	A	306	ARG	3.3
3	C	1	DA	3.2
1	A	876	LYS	3.2
1	A	846	ARG	3.1
1	A	684	LYS	3.1
1	A	695	ASP	3.1
1	A	442	VAL	3.0
1	D	689	ILE	3.0
1	D	505	LYS	3.0
1	A	300	ALA	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	698	THR	2.9
1	A	781	PHE	2.9
1	A	842	GLU	2.9
1	A	431	LYS	2.9
1	A	692	VAL	2.9
1	A	466	ARG	2.8
1	A	843	ARG	2.8
1	A	868	TYR	2.8
1	A	303	LEU	2.8
1	A	646	ASP	2.8
1	D	512	GLY	2.7
1	D	509	GLU	2.7
1	A	301	PHE	2.7
1	A	459	ARG	2.6
1	A	307	VAL	2.6
1	A	816	LYS	2.6
1	A	435	ARG	2.5
1	A	847	LEU	2.5
1	A	687	MET	2.5
1	A	645	SER	2.5
1	A	689	ILE	2.4
1	A	697	VAL	2.4
1	A	850	GLU	2.4
1	D	731	GLU	2.4
1	A	314	ASP	2.4
1	A	837	PRO	2.4
1	D	766	LEU	2.4
1	A	342	GLY	2.4
1	A	441	PRO	2.4
1	A	344	PHE	2.4
1	A	688	ASP	2.3
1	A	838	LYS	2.3
1	A	676	ARG	2.3
1	A	750	MET	2.3
1	D	735	PHE	2.3
1	A	741	GLU	2.3
1	A	520	GLU	2.3
1	A	844	LEU	2.2
1	A	305	ASP	2.2
1	A	302	THR	2.2
1	A	731	GLU	2.2
1	D	695	ASP	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	508	ALA	2.2
1	D	309	GLU	2.2
1	A	308	THR	2.1
1	A	420	GLU	2.1
1	A	701	MET	2.1
1	A	313	ALA	2.1
1	A	663	ALA	2.1
1	A	866	TYR	2.0
1	A	438	PRO	2.0
1	A	448	VAL	2.0
1	A	669	ASP	2.0
1	A	704	GLN	2.0
1	D	305	ASP	2.0
1	A	670	ASN	2.0
1	A	340	GLU	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

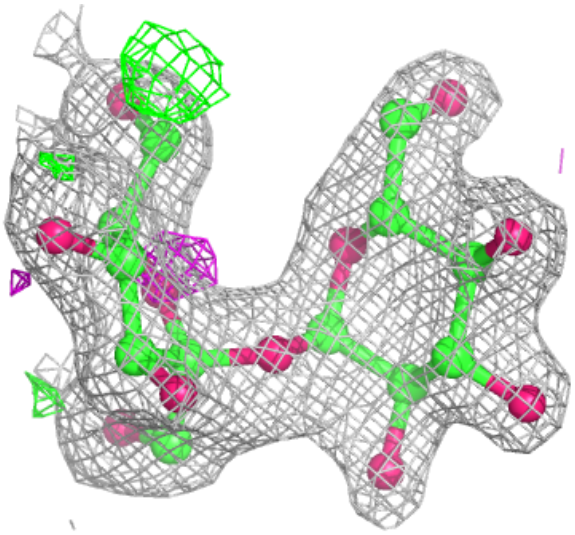
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

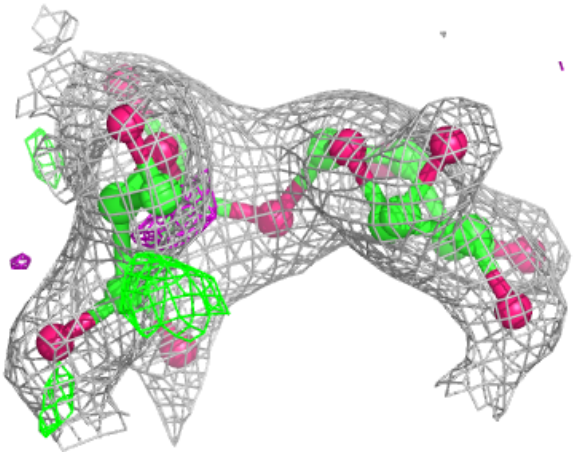
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	FRU	G	2	12/12	0.85	0.15	49,51,53,55	0
4	GLC	G	1	11/12	0.91	0.17	47,51,51,51	0
4	GLC	H	1	11/12	0.92	0.12	34,35,37,37	0
4	FRU	H	2	12/12	0.93	0.13	30,34,35,36	0

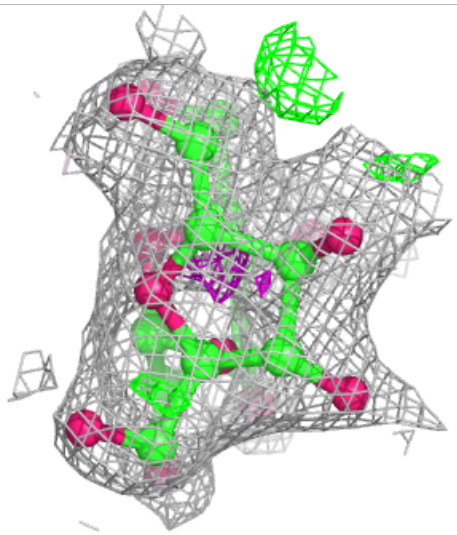
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

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)

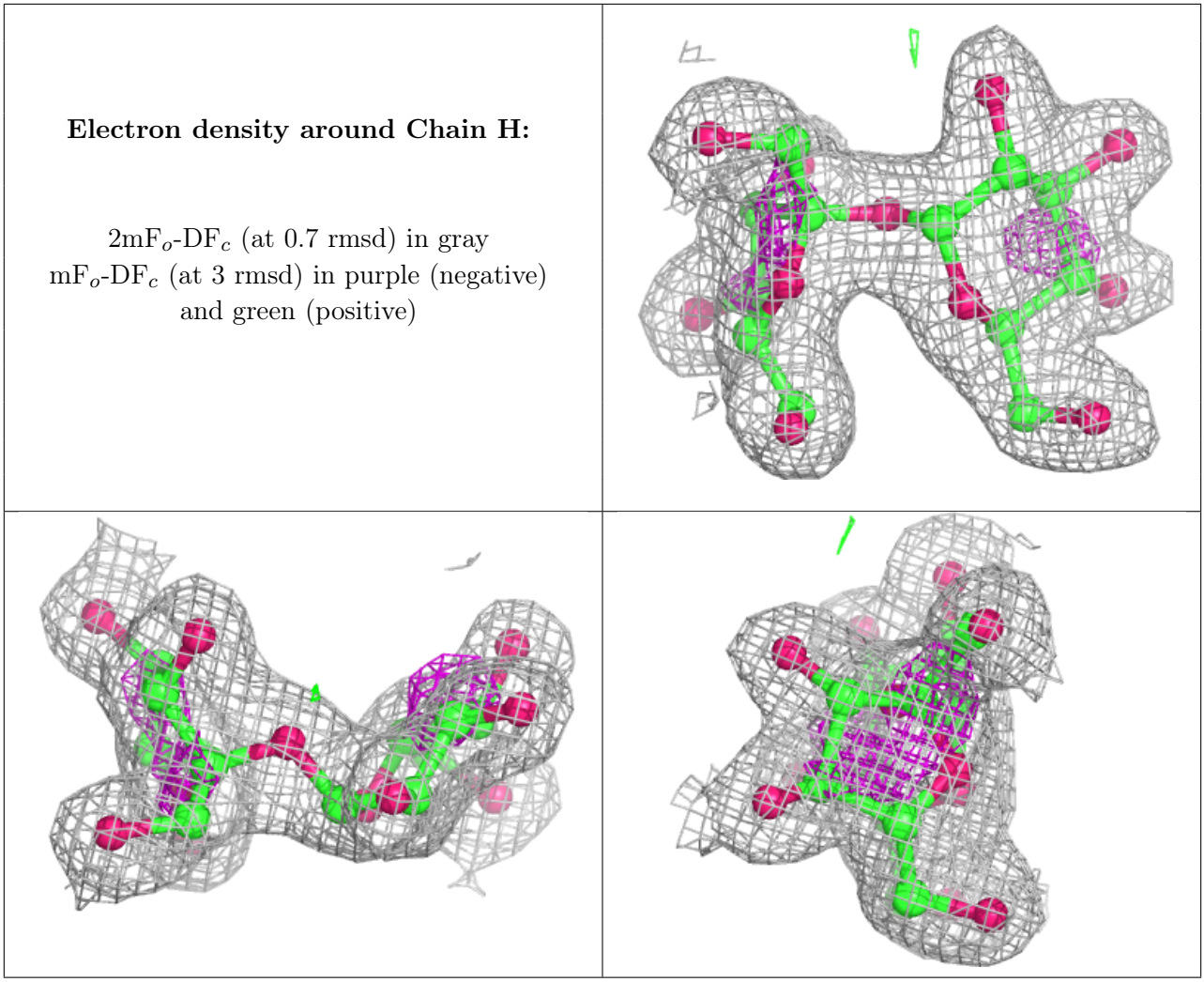






Electron density around Chain H:

$2mF_o$-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	ZN	D	3	1/1	0.40	0.40	46,46,46,46	1
6	ZN	A	5	1/1	0.90	0.13	32,32,32,32	1
5	SO4	A	1	5/5	0.92	0.24	80,81,81,81	0
7	DAD	A	201	29/29	0.94	0.11	32,35,42,44	0
5	SO4	D	878	5/5	0.96	0.15	46,46,49,50	0
6	ZN	D	2	1/1	0.96	0.23	29,29,29,29	1
5	SO4	A	2	5/5	0.97	0.14	45,45,47,47	0
7	DAD	D	202	29/29	0.98	0.11	16,23,30,32	0
6	ZN	A	4	1/1	0.99	0.08	29,29,29,29	1

Continued on next page...

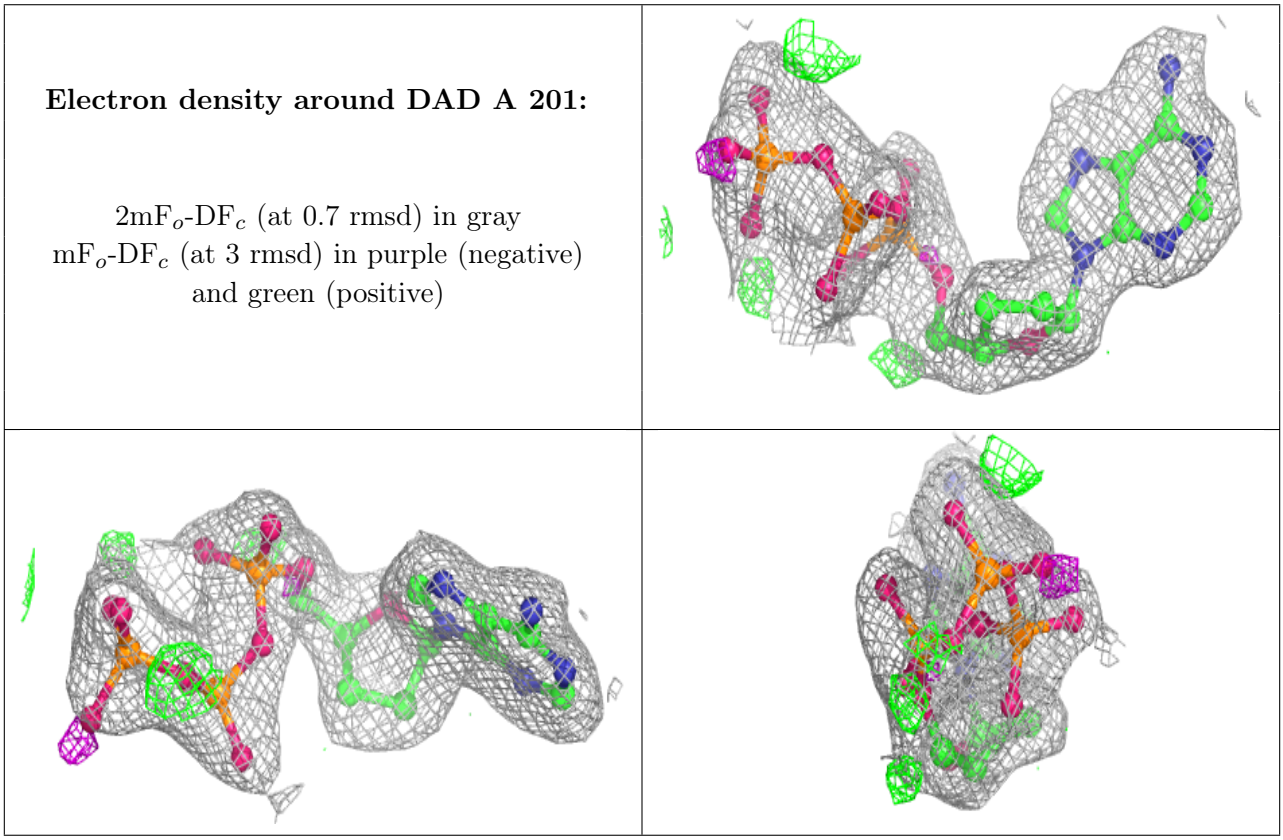
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	ZN	D	1	1/1	0.99	0.04	29,29,29,29	1

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.

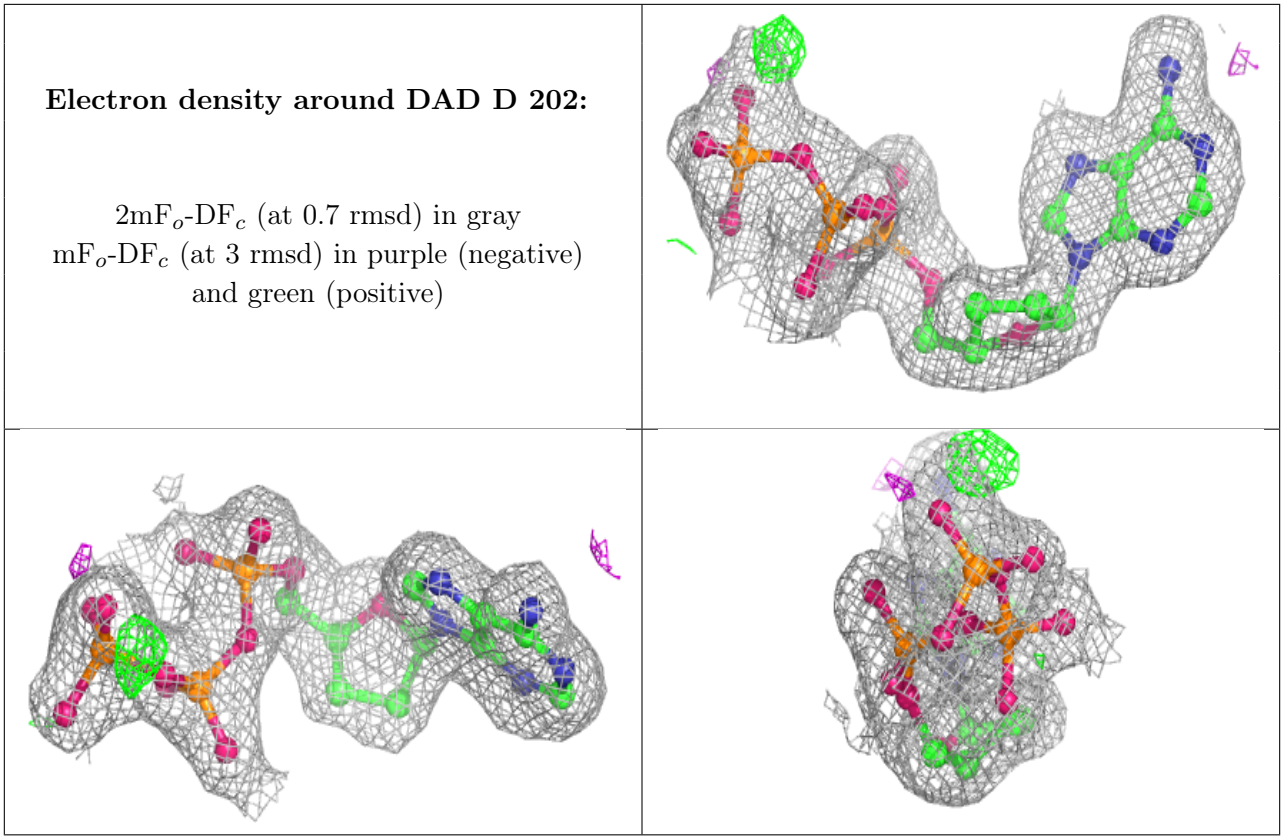
Electron density around DAD A 201:

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



Electron density around DAD 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.