wwPDB X-ray Structure Validation Summary Report (i)

Apr 9, 2022 – 08:07 AM EDT

PDB ID	:	4YFU
Title	:	Crystal structure of open Bacillus fragment DNA polymerase bound to DNA and dTTP
Authors	:	Wu, E.Y.
Deposited on	:	2015-02-25
Resolution	:	1.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 following versions of software and data (see references (i)) were used in the production of this report:

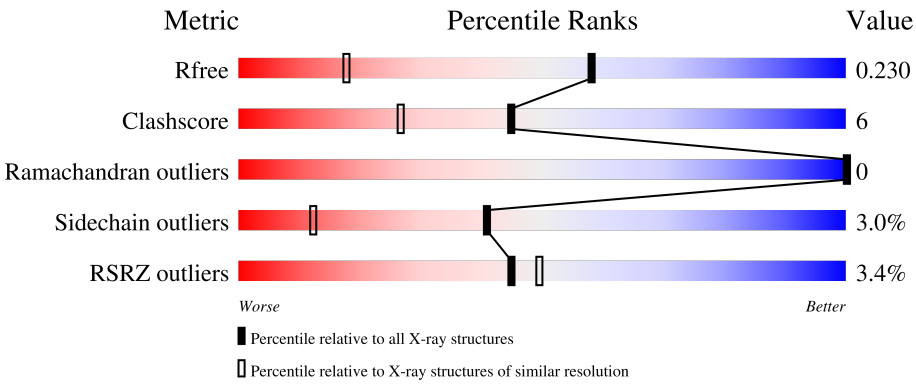
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.27
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 1.50 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	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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	3%, 83%, 15%, •
1	D	580	4%, 83%, 13%, • •
2	B	9	44%, 33%, 22%
2	E	9	44%, 22%, 33%
3	C	14	57%, 21%, 21%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	14	7% 71% 29%
4	G	2	50% 50%
4	H	2	50% 50%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 11744 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	579	4700	2995	812	873	20	0	10	0
1	D	580	4719	3011	811	877	20	0	13	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	297	LYS	-	expression tag	UNP E1C9K5
A	456	GLU	ALA	conflict	UNP E1C9K5
A	505	LYS	GLU	conflict	UNP E1C9K5
A	512	GLY	ARG	conflict	UNP E1C9K5
A	550	THR	SER	conflict	UNP E1C9K5
A	598	ALA	ASP	engineered mutation	UNP E1C9K5
A	710	TYR	PHE	engineered mutation	UNP E1C9K5
A	823	HIS	ARG	conflict	UNP E1C9K5
D	297	LYS	-	expression tag	UNP E1C9K5
D	456	GLU	ALA	conflict	UNP E1C9K5
D	505	LYS	GLU	conflict	UNP E1C9K5
D	512	GLY	ARG	conflict	UNP E1C9K5
D	550	THR	SER	conflict	UNP E1C9K5
D	598	ALA	ASP	engineered mutation	UNP E1C9K5
D	710	TYR	PHE	engineered mutation	UNP E1C9K5
D	823	HIS	ARG	conflict	UNP E1C9K5

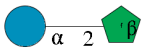
- Molecule 2 is a DNA chain called Primer DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	9	180	88	32	52	8	0	0	0
2	E	9	180	88	32	52	8	0	0	0

- Molecule 3 is a DNA chain called Template DNA.

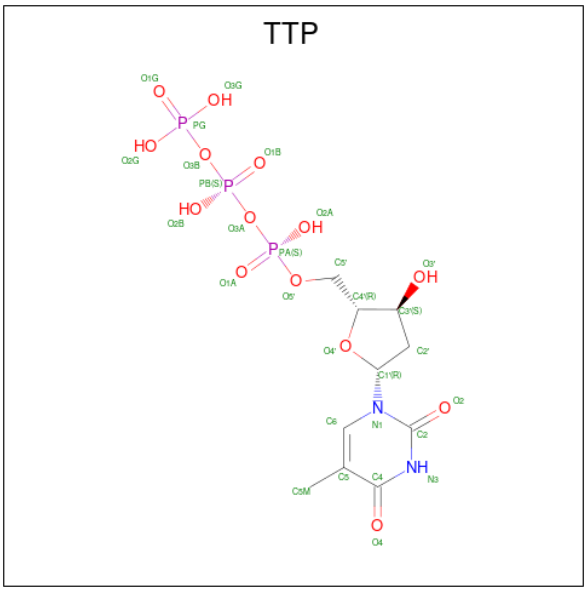
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	11	226	107	43	65	11	0	0	0
3	F	14	288	137	58	80	13	0	0	0

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



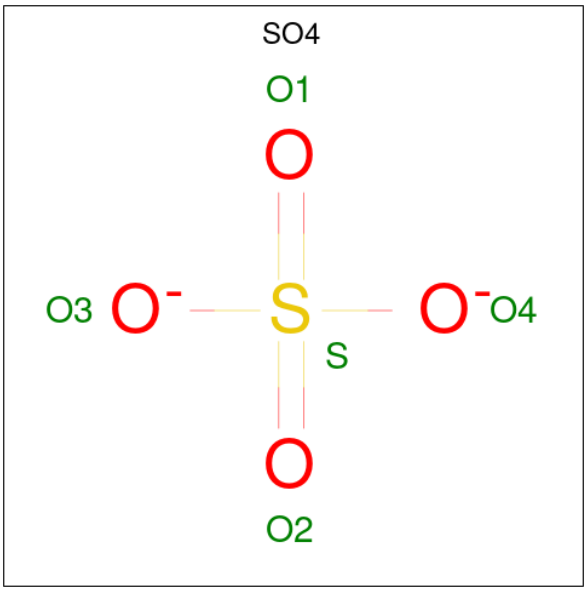
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
4	G	2	23	12	11	0	0	0
4	H	2	23	12	11	0	0	0

- Molecule 5 is THYMIDINE-5'-TRIPHOSPHATE (three-letter code: TTP) (formula: $C_{10}H_{17}N_2O_{14}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
5	A	1	29	10	2	14	3	0	0

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



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

- Molecule 7 is water.

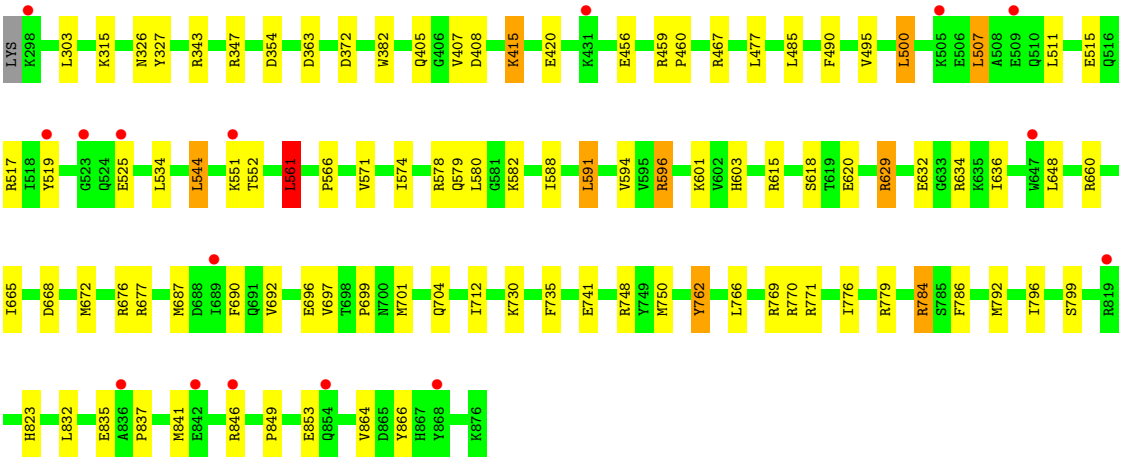
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	556	Total O 556 556	0	0
7	D	622	Total O 622 622	0	0
7	B	37	Total O 37 37	0	0
7	C	50	Total O 50 50	0	0
7	E	35	Total O 35 35	0	0
7	F	71	Total O 71 71	0	0

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.

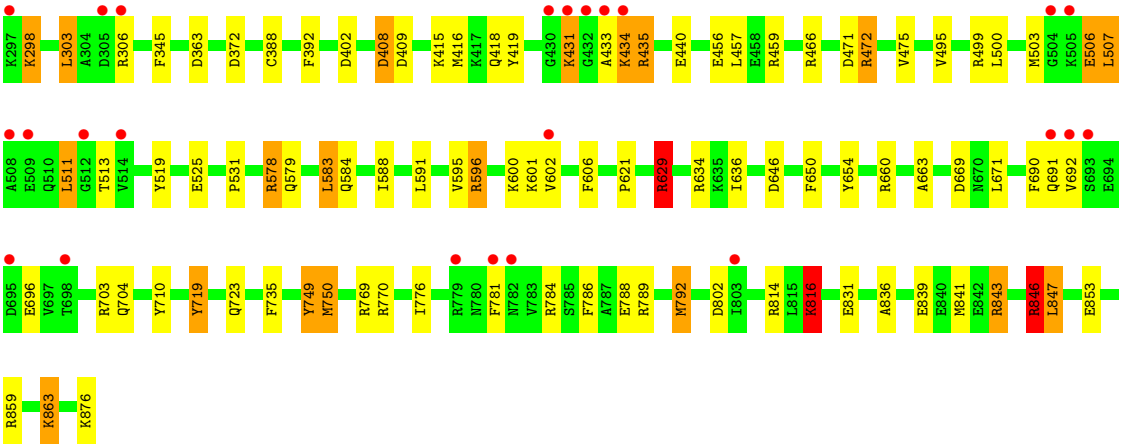
• Molecule 1: DNA polymerase

Chain A:



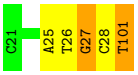
• Molecule 1: DNA polymerase

Chain D:



• Molecule 2: Primer DNA

Chain B:



• Molecule 2: Primer DNA

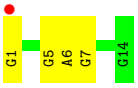
Chain E:

• Molecule 3: Template DNA

Chain C:

• Molecule 3: Template DNA

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	93.23Å 108.10Å 151.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.68 – 1.50 44.68 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.0 (44.68-1.50) 99.0 (44.68-1.50)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.29 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.195 , 0.230 0.195 , 0.230	Depositor DCC
R_{free} test set	11140 reflections (4.62%)	wwPDB-VP
Wilson B-factor (Å^2)	19.6	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 54.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	11744	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% 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, CME, SO4, GLC, 2DT, TTP

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	1.16	4/4803 (0.1%)	1.28	38/6487 (0.6%)
1	D	1.29	11/4828 (0.2%)	1.46	54/6524 (0.8%)
2	B	1.06	1/180 (0.6%)	1.98	4/276 (1.4%)
2	E	1.27	3/180 (1.7%)	1.84	5/276 (1.8%)
3	C	1.06	1/253 (0.4%)	1.32	2/388 (0.5%)
3	F	1.16	1/324 (0.3%)	1.21	3/499 (0.6%)
All	All	1.22	21/10568 (0.2%)	1.39	106/14450 (0.7%)

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	839	GLU	CD-OE1	8.26	1.34	1.25
1	A	382	TRP	CG-CD1	7.13	1.46	1.36
3	C	5	DG	O3'-P	-7.03	1.52	1.61
1	D	853	GLU	CD-OE1	6.72	1.33	1.25
2	B	27	DG	O3'-P	-6.33	1.53	1.61

The worst 5 of 106 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	25	DA	O5'-P-OP2	-18.39	88.63	110.70
1	D	843	ARG	NE-CZ-NH2	-15.91	112.35	120.30
2	B	25	DA	O5'-P-OP1	14.92	128.60	110.70
1	D	578	ARG	NE-CZ-NH2	-14.80	112.90	120.30
1	D	843	ARG	NE-CZ-NH1	14.74	127.67	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	4700	0	4794	54	0
1	D	4719	0	4817	74	0
2	B	180	0	104	3	0
2	E	180	0	104	6	0
3	C	226	0	124	1	0
3	F	288	0	158	1	0
4	G	23	0	21	1	0
4	H	23	0	21	1	0
5	A	29	0	13	2	0
6	D	5	0	0	0	0
7	A	556	0	0	15	1
7	B	37	0	0	0	0
7	C	50	0	0	1	0
7	D	622	0	0	32	1
7	E	35	0	0	1	0
7	F	71	0	0	1	0
All	All	11744	0	10156	133	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 133 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:749:TYR:CE2	1:D:750[A]:MET:CE	2.41	1.04
1:D:418:GLN:HG2	7:D:1402:HOH:O	1.64	0.96
1:D:843:ARG:HD2	7:D:1248:HOH:O	1.66	0.95
1:D:792:MET:HB3	7:D:1460:HOH:O	1.69	0.92
1:A:690:PHE:HZ	1:A:704:GLN:NE2	1.72	0.87

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:1004:HOH:O	7:D:1292:HOH:O[2_745]	1.99	0.21

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	586/580 (101%)	574 (98%)	12 (2%)	0	100	100
1	D	590/580 (102%)	574 (97%)	16 (3%)	0	100	100
All	All	1176/1160 (101%)	1148 (98%)	28 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	504/495 (102%)	489 (97%)	15 (3%)	41	12
1	D	507/495 (102%)	491 (97%)	16 (3%)	39	10
All	All	1011/990 (102%)	980 (97%)	31 (3%)	41	11

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	846	ARG
1	D	776	ILE
1	D	415	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	846	ARG
1	D	513	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	867	HIS
1	D	768	HIS
1	A	823	HIS
1	A	755	GLN
1	D	704	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	$\#\|Z\| > 2$
1	CME	D	388	1	8,9,10	0.74	0	5,9,11	1.21	1 (20%)
2	2DT	B	101	2,3	14,20,21	1.37	2 (14%)	12,28,31	2.06	4 (33%)
1	CME	A	388	1	8,9,10	0.70	0	5,9,11	1.14	0
2	2DT	E	101	2,3	14,20,21	1.59	4 (28%)	12,28,31	2.96	4 (33%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CME	D	388	1	-	1/5/8/10	-
2	2DT	B	101	2,3	-	0/4/18/19	0/2/2/2
1	CME	A	388	1	-	1/5/8/10	-
2	2DT	E	101	2,3	-	0/4/18/19	0/2/2/2

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	101	2DT	C3'-C2'	-3.05	1.45	1.54
2	E	101	2DT	C4-N3	2.69	1.37	1.33
2	E	101	2DT	C5M-C5	2.42	1.55	1.51
2	B	101	2DT	C4-N3	2.35	1.37	1.33
2	B	101	2DT	C5M-C5	2.22	1.55	1.51

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	101	2DT	C5-C6-N1	-6.88	114.78	122.19
2	E	101	2DT	C4'-O4'-C1'	5.22	114.74	109.81
2	E	101	2DT	C3'-C2'-C1'	3.87	107.25	102.78
2	B	101	2DT	C4'-O4'-C1'	3.82	113.42	109.81
2	B	101	2DT	C2-N3-C4	2.84	117.54	115.14

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	388	CME	SD-CE-CZ-OH
1	D	388	CME	SD-CE-CZ-OH

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	101	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	0.78	0	15,15,17	1.55	3 (20%)
4	FRU	G	2	4	11,12,12	1.89	3 (27%)	10,18,18	1.70	2 (20%)
4	GLC	H	1	4	11,11,12	0.86	0	15,15,17	1.68	3 (20%)
4	FRU	H	2	4	11,12,12	1.36	2 (18%)	10,18,18	1.23	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	-	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 (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	2	FRU	O2-C2	4.92	1.49	1.40
4	H	2	FRU	O2-C2	2.48	1.45	1.40
4	G	2	FRU	O5-C2	2.44	1.47	1.43
4	H	2	FRU	O3-C3	2.39	1.47	1.42
4	G	2	FRU	C1-C2	2.09	1.55	1.52

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	G	2	FRU	O6-C6-C5	-4.37	96.30	111.29
4	H	1	GLC	O5-C1-C2	-4.08	104.47	110.77
4	G	1	GLC	C1-O5-C5	3.46	116.88	112.19
4	H	2	FRU	O6-C6-C5	-2.88	101.41	111.29
4	G	2	FRU	O3-C3-C4	-2.62	104.28	113.32

There are no chirality outliers.

There are no torsion outliers.

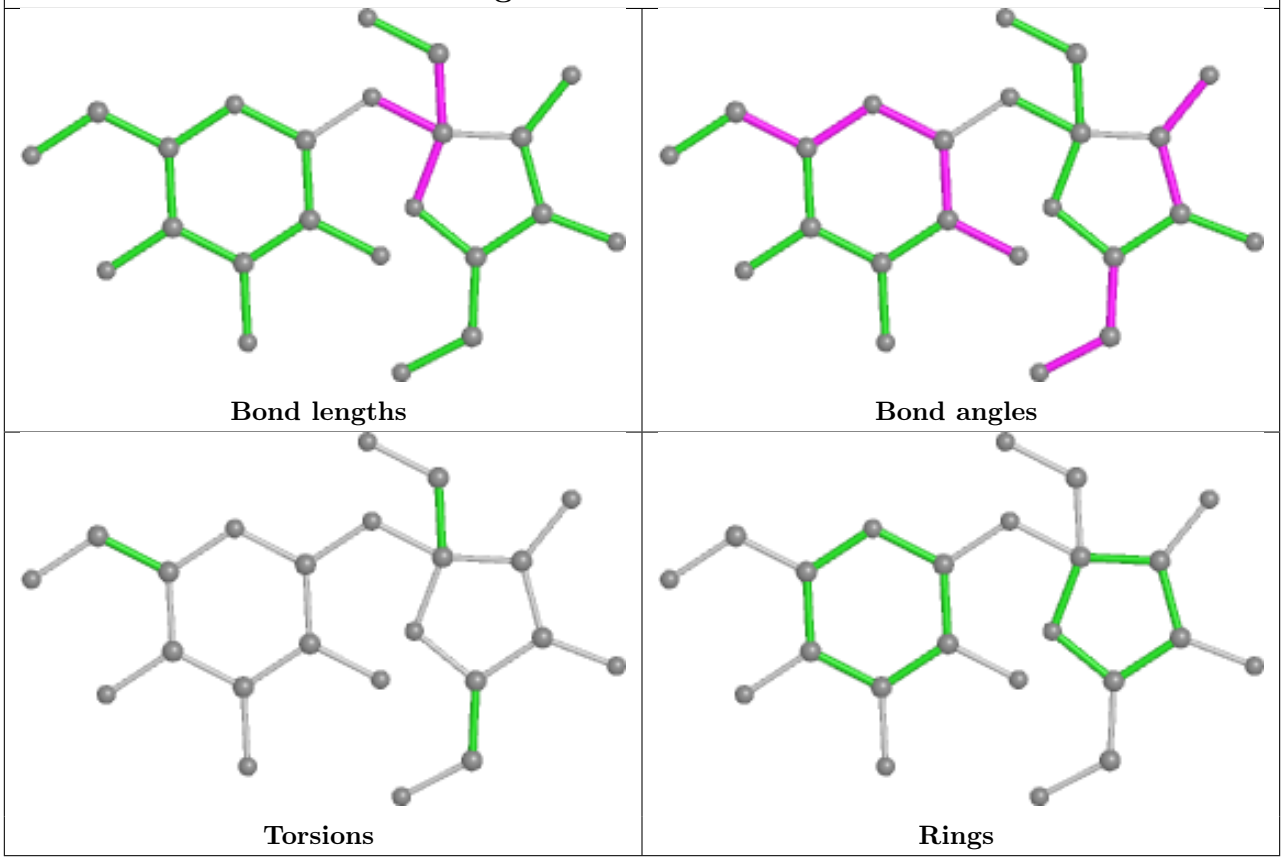
There are no ring outliers.

2 monomers are involved in 2 short contacts:

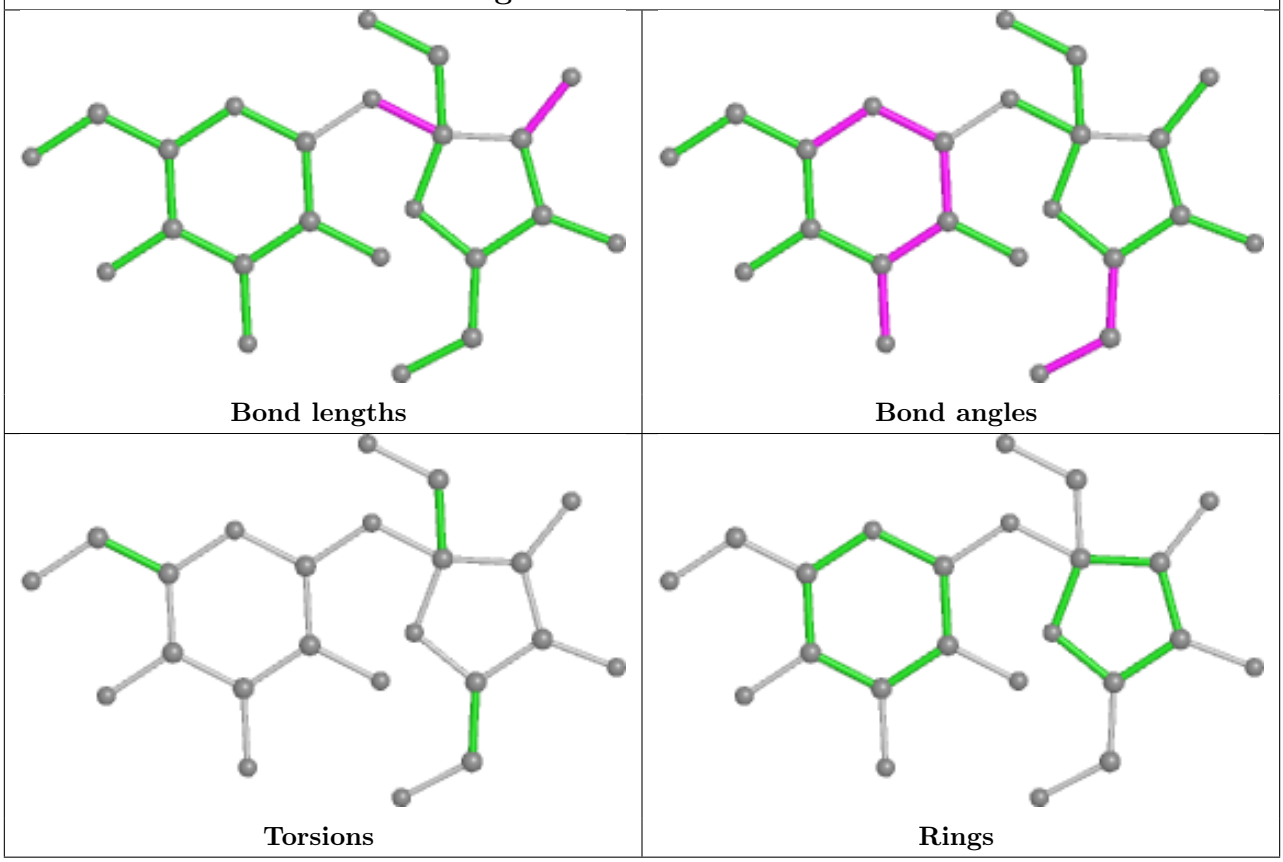
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	2	FRU	1	0
4	H	2	FRU	1	0

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

Oligosaccharide Chain G



Oligosaccharide Chain H



5.6 Ligand geometry (i)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
6	SO4	D	901	-	4,4,4	0.39	0	6,6,6	1.66	3 (50%)
5	TTP	A	901	-	23,30,30	0.91	2 (8%)	29,47,47	2.22	7 (24%)

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	TTP	A	901	-	-	4/19/34/34	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	901	TTP	C5-C4	2.32	1.46	1.41
5	A	901	TTP	C2-N3	-2.02	1.34	1.38

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	901	TTP	C2-N3-C4	8.60	122.40	115.14
5	A	901	TTP	C5-C6-N1	-4.03	117.85	122.19
5	A	901	TTP	C5M-C5-C6	2.78	124.56	118.68
5	A	901	TTP	C5M-C5-C4	-2.62	116.68	121.37
5	A	901	TTP	O3G-PG-O3B	-2.50	96.24	104.64

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	901	TTP	C5'-O5'-PA-O3A
5	A	901	TTP	C5'-O5'-PA-O1A
5	A	901	TTP	PB-O3A-PA-O2A
5	A	901	TTP	PB-O3A-PA-O1A

There are no ring outliers.

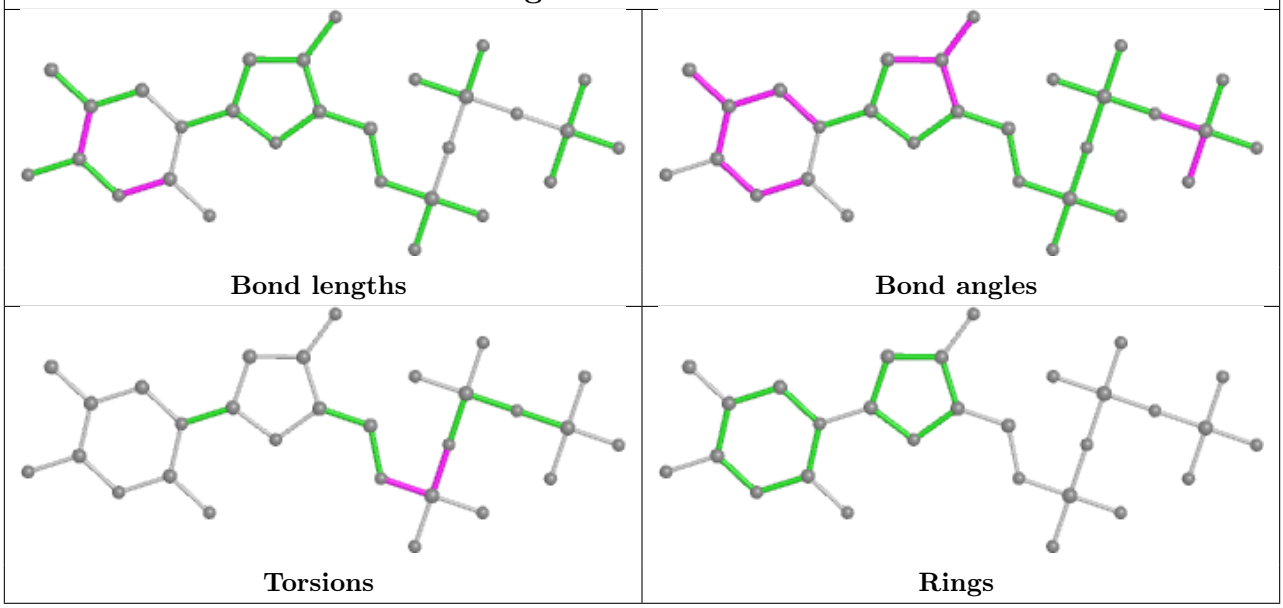
1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	901	TTP	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 TTP A 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	578/580 (99%)	-0.04	16 (2%)	53	57	14, 27, 45, 67	3 (0%)
1	D	579/580 (99%)	0.15	24 (4%)	37	41	11, 24, 49, 89	1 (0%)
2	B	8/9 (88%)	-0.41	0	100	100	22, 26, 30, 37	0
2	E	8/9 (88%)	-0.40	0	100	100	17, 24, 31, 51	0
3	C	11/14 (78%)	-0.44	0	100	100	17, 22, 37, 45	0
3	F	14/14 (100%)	0.09	1 (7%)	16	17	14, 23, 41, 46	0
All	All	1198/1206 (99%)	0.04	41 (3%)	45	49	11, 25, 47, 89	4 (0%)

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	432	GLY	6.3
1	D	602[A]	VAL	5.3
1	D	434	LYS	4.1
1	D	693	SER	4.1
1	A	523	GLY	3.8

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

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	CME	A	388	10/11	0.95	0.10	19,22,55,68	0
1	CME	D	388	10/11	0.96	0.09	18,26,45,49	0
2	2DT	B	101	19/20	0.98	0.08	20,22,28,32	0
2	2DT	E	101	19/20	0.98	0.07	16,18,22,22	0

6.3 Carbohydrates (i)

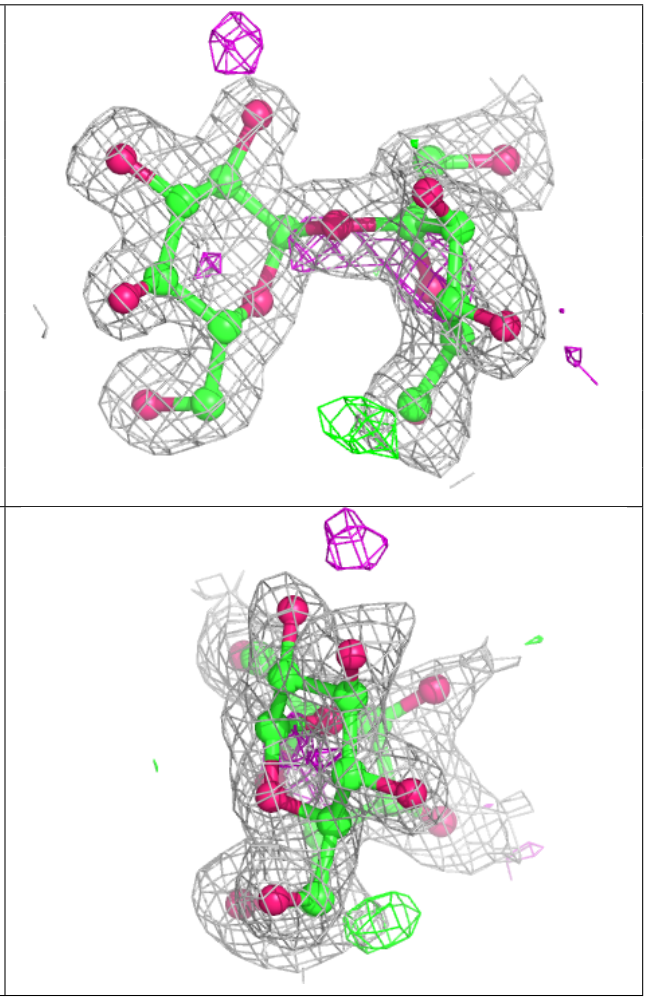
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	FRU	G	2	12/12	0.79	0.16	29,34,37,43	0
4	GLC	H	1	11/12	0.81	0.15	18,20,22,24	11
4	FRU	H	2	12/12	0.81	0.18	15,19,20,20	12
4	GLC	G	1	11/12	0.83	0.18	35,37,41,45	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.

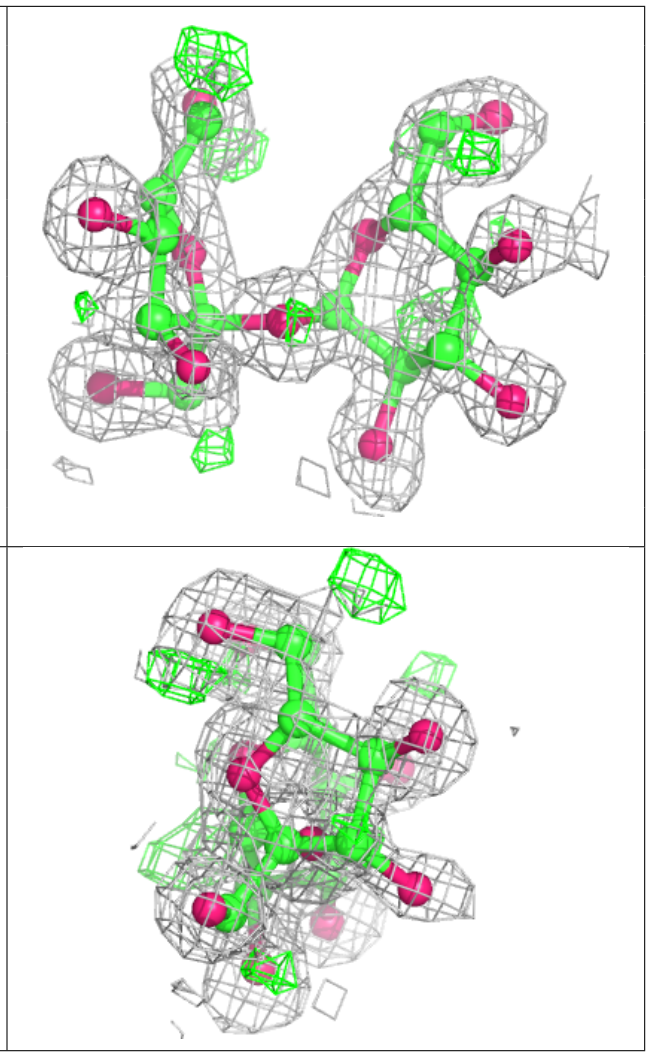
Electron density around Chain G:

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)

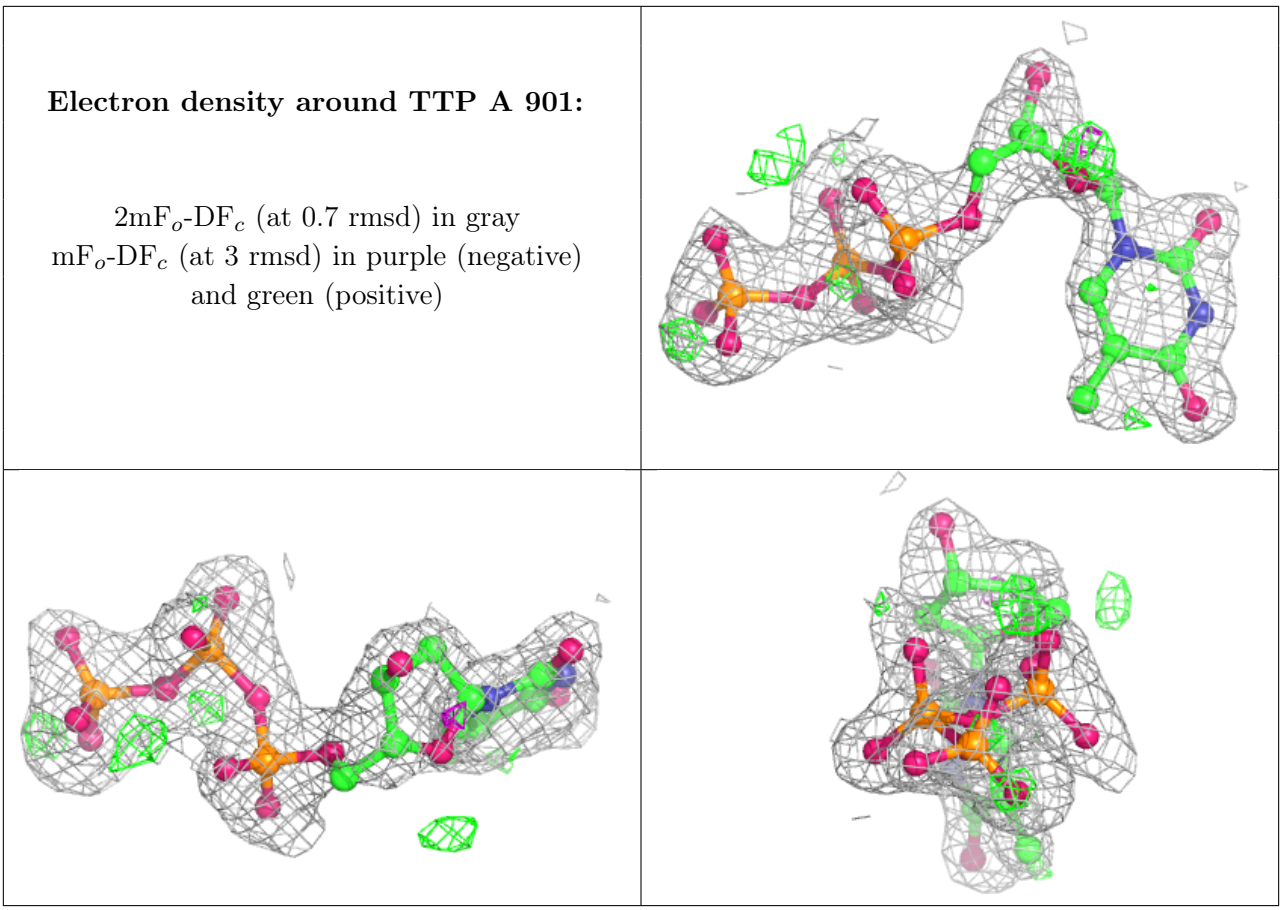
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	TTP	A	901	29/29	0.93	0.11	28,35,43,43	29
6	SO4	D	901	5/5	0.97	0.10	36,38,50,50	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 TTP 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)



6.5 Other polymers (i)

There are no such residues in this entry.