Full wwPDB X-ray Structure Validation Report (i)

Sep 3, 2023 – 01:10 AM EDT

PDB ID	:	3RAX
Title	:	Dpo4 extension ternary complex with 3'-terminal primer T base opposite the 1-methylguanine (M1G) lesion
Authors	:	Rechkoblit, O.; Patel, D.J.
Deposited on	:	2011-03-28
Resolution	:	1.89 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:

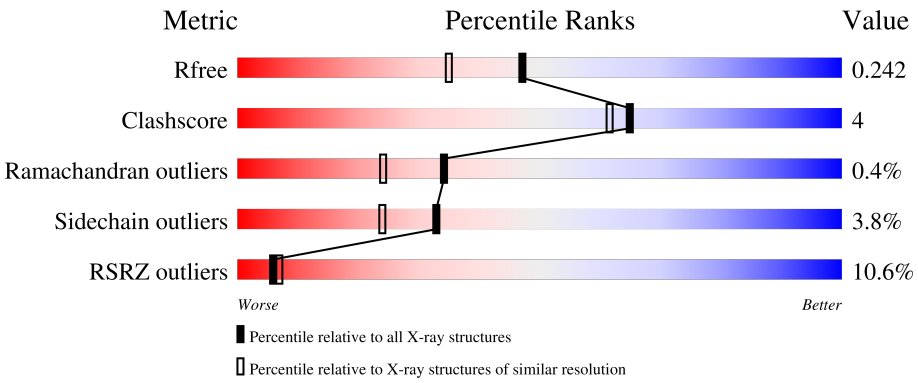
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 1.89 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	341	8% 89% 11% •
1	B	341	13% 91% 8% •
2	D	13	54% 38% 8%
2	H	13	8% 69% 23% 8%
3	E	20	10% 65% 15% 10% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	J	20	15% 60% 20% 5% 15%

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
2	2DT	H	1814	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7302 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 IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	341	2740	1757	472	505	6	0	0	0
1	B	341	2740	1757	472	505	6	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP Q97W02
B	1001	GLY	-	expression tag	UNP Q97W02

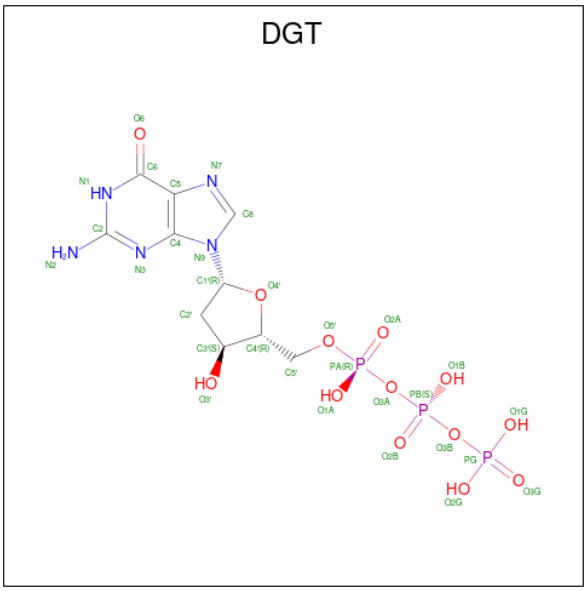
- Molecule 2 is a DNA chain called DNA (5'-D(*GP*TP*TP*GP*GP*AP*TP*GP*GP*TP*AP*GP*(2DT))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	13	270	130	50	78	12	0	0	0
2	H	13	270	130	50	78	12	0	0	0

- Molecule 3 is a DNA chain called DNA (5'-D(*CP*CP*TP*AP*AP*CP*(MG1)P*CP*TP*AP*CP*CP*AP*TP*CP*CP*AP*AP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	E	18	358	173	65	103	17	0	0	0
3	J	17	338	163	63	96	16	0	0	0

- Molecule 4 is 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DGT) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	31	10	5	13	3	0	0
			Total	C	N	O	P		
4	B	1	31	10	5	13	3	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
5	A	3	3	3	0	0
			Total	Ca		
5	B	3	3	3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	259	259	259	0	0
			Total	O		
6	D	24	24	24	0	0
			Total	O		
6	E	22	22	22	0	0
			Total	O		
6	B	170	170	170	0	0
			Total	O		
6	H	22	22	22	0	0

Continued on next page...

Continued from previous page...

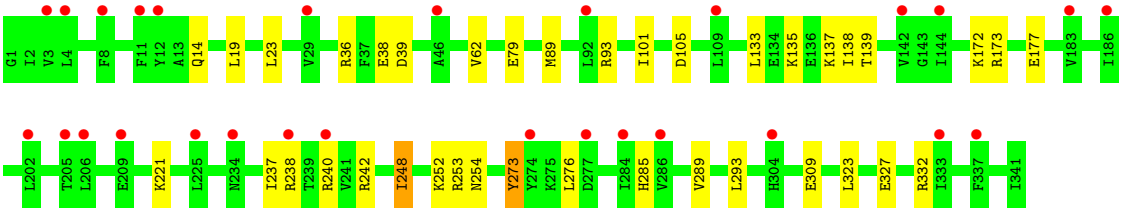
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	J	21	21	21	0	0

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.

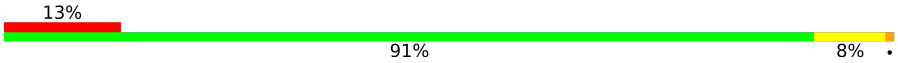
• Molecule 1: DNA polymerase IV

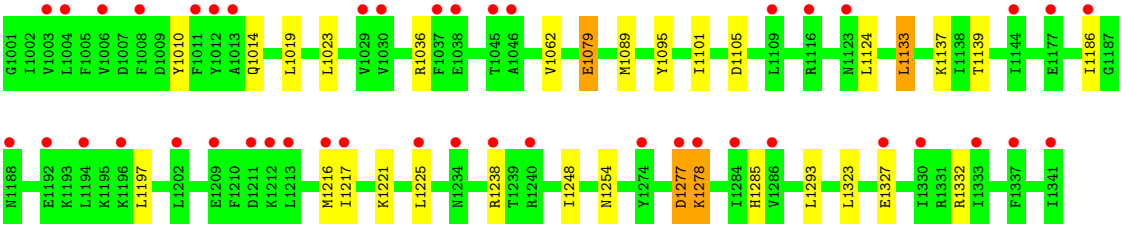
Chain A:



• Molecule 1: DNA polymerase IV

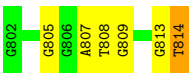
Chain B:





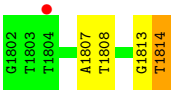
• Molecule 2: DNA (5'-D(*GP*TP*TP*GP*GP*AP*TP*GP*GP*TP*AP*GP*(2DT))-3')

Chain D:



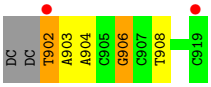
• Molecule 2: DNA (5'-D(*GP*TP*TP*GP*GP*AP*TP*GP*GP*TP*AP*GP*(2DT))-3')

Chain H:



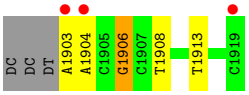
• Molecule 3: DNA (5'-D(*CP*CP*TP*AP*AP*CP*(MG1)P*CP*TP*AP*CP*CP*AP*TP*CP*CP*AP*AP*CP*C)-3')

Chain E:



• Molecule 3: DNA (5'-D(*CP*CP*TP*AP*AP*CP*(MG1)P*CP*TP*AP*CP*CP*AP*TP*CP*CP*AP*AP*CP*C)-3')

Chain J:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	52.88Å 61.65Å 92.68Å 98.50° 103.82° 93.29°	Depositor
Resolution (Å)	20.00 – 1.89 19.98 – 1.89	Depositor EDS
% Data completeness (in resolution range)	95.8 (20.00-1.89) 95.8 (19.98-1.89)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.18 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.197 , 0.230 0.214 , 0.242	Depositor DCC
R_{free} test set	4290 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	34.6	Xtriage
Anisotropy	0.340	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 44.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7302	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.09% 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: CA, 2DT, DGT, MG1

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	0/2779	0.60	0/3731
1	B	0.37	0/2779	0.55	0/3731
2	D	0.76	0/282	1.42	6/436 (1.4%)
2	H	0.65	0/282	1.29	3/436 (0.7%)
3	E	0.75	0/373	1.37	5/568 (0.9%)
3	J	0.66	0/351	1.24	2/534 (0.4%)
All	All	0.48	0/6846	0.80	16/9436 (0.2%)

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	902	DT	O4'-C4'-C3'	7.39	110.43	106.00
3	E	902	DT	P-O3'-C3'	7.02	128.13	119.70
2	H	1813	DG	O4'-C1'-N9	6.87	112.81	108.00
3	E	902	DT	C5'-C4'-O4'	6.61	121.85	109.30
3	J	1913	DT	P-O3'-C3'	6.30	127.26	119.70
2	H	1807	DA	O4'-C1'-N9	-6.25	103.62	108.00
3	E	902	DT	C5'-C4'-C3'	6.18	125.23	114.10
2	D	809	DG	P-O3'-C3'	6.09	127.01	119.70
2	D	808	DT	C1'-O4'-C4'	-5.73	104.37	110.10
2	H	1808	DT	C1'-O4'-C4'	-5.70	104.40	110.10
2	D	805	DG	O4'-C1'-N9	5.57	111.90	108.00
3	J	1908	DT	C1'-O4'-C4'	-5.49	104.61	110.10
2	D	813	DG	O4'-C1'-N9	5.27	111.69	108.00
2	D	808	DT	O4'-C1'-N1	5.14	111.60	108.00
2	D	807	DA	O4'-C1'-N9	-5.10	104.43	108.00
3	E	908	DT	C1'-O4'-C4'	-5.08	105.03	110.10

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	2740	0	2883	28	0
1	B	2740	0	2880	20	0
2	D	270	0	150	4	0
2	H	270	0	150	3	0
3	E	358	0	205	10	0
3	J	338	0	193	4	0
4	A	31	0	12	0	0
4	B	31	0	12	0	0
5	A	3	0	0	0	0
5	B	3	0	0	0	0
6	A	259	0	0	8	0
6	B	170	0	0	3	0
6	D	24	0	0	0	0
6	E	22	0	0	1	0
6	H	22	0	0	0	0
6	J	21	0	0	2	0
All	All	7302	0	6485	54	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 4.

All (54) 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)
6:A:567:HOH:O	2:D:814:2DT:H3'2	1.02	1.18
6:B:107:HOH:O	2:H:1814:2DT:H3'2	1.56	1.05
1:B:1014:GLN:HE22	1:B:1139:THR:H	1.11	0.95
1:A:101:ILE:O	1:A:240:ARG:HG3	1.71	0.89
1:A:14:GLN:HE22	1:A:139:THR:H	1.20	0.87
2:D:814:2DT:HN3	3:E:906:MG1:HM11	1.40	0.86
3:E:902:DT:H4'	3:E:903:DA:OP1	1.77	0.82
2:D:814:2DT:N3	3:E:906:MG1:HM11	1.98	0.78

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:814:2DT:HN3	3:E:906:MG1:CM1	2.01	0.73
1:A:133:LEU:HD13	6:A:520:HOH:O	1.88	0.72
1:A:89:MET:SD	1:A:101:ILE:HD12	2.36	0.66
2:H:1814:2DT:O2	3:J:1906:MG1:HM11	1.95	0.66
1:A:79:GLU:H	1:A:79:GLU:CD	2.04	0.60
1:B:1248:ILE:HG22	6:J:77:HOH:O	2.05	0.57
1:A:137:LYS:HD2	1:B:1137:LYS:HD2	1.87	0.56
1:B:1285:HIS:HD2	6:B:2:HOH:O	1.88	0.56
1:A:133:LEU:HD12	1:A:138:ILE:O	2.05	0.56
1:A:248:ILE:HG23	1:A:332:ARG:NH1	2.21	0.56
1:A:14:GLN:NE2	1:A:139:THR:H	1.99	0.55
1:B:1089:MET:SD	1:B:1101:ILE:HD12	2.47	0.55
1:A:133:LEU:HD23	1:B:1133:LEU:HB3	1.88	0.54
1:A:137:LYS:HE2	6:A:633:HOH:O	2.07	0.54
1:B:1036:ARG:NH2	1:B:1254:ASN:OD1	2.40	0.54
1:B:1293:LEU:HD22	6:J:609:HOH:O	2.08	0.54
1:A:62:VAL:CG2	3:E:903:DA:H4'	2.37	0.53
1:B:1248:ILE:HG23	1:B:1332:ARG:NH1	2.24	0.53
1:B:1079:GLU:CD	1:B:1079:GLU:H	2.13	0.52
1:A:36:ARG:NH2	1:A:254:ASN:OD1	2.43	0.52
6:B:107:HOH:O	2:H:1814:2DT:C3'	2.31	0.52
1:A:137:LYS:CD	1:B:1137:LYS:HD2	2.41	0.51
1:B:1293:LEU:HD13	3:J:1904:DA:C4	2.46	0.50
1:B:1217:ILE:HD12	1:B:1221:LYS:HB3	1.92	0.50
1:A:248:ILE:HG22	6:E:46:HOH:O	2.11	0.49
1:A:93:ARG:NH2	6:A:586:HOH:O	2.47	0.47
1:B:1186:ILE:HD11	1:B:1225:LEU:HD21	1.97	0.47
1:A:173:ARG:NH1	1:A:177:GLU:OE1	2.45	0.47
1:A:293:LEU:HD11	3:E:904:DA:C4	2.51	0.46
1:B:1062:VAL:CG2	3:J:1903:DA:H4'	2.46	0.46
1:A:62:VAL:HG21	3:E:903:DA:H4'	1.98	0.45
1:A:62:VAL:HG21	3:E:903:DA:O5'	2.16	0.45
1:B:1133:LEU:O	1:B:1137:LYS:HD3	2.16	0.44
1:A:285:HIS:HD2	6:A:577:HOH:O	2.00	0.43
1:B:1293:LEU:HD13	3:J:1904:DA:C2	2.54	0.43
1:B:1095:TYR:CD1	1:B:1124:LEU:HD11	2.54	0.43
1:A:248:ILE:HG21	6:A:540:HOH:O	2.18	0.42
1:A:289:VAL:HB	1:A:332:ARG:HB2	2.00	0.42
1:B:1277:ASP:O	1:B:1278:LYS:HB2	2.20	0.41
1:B:1197:LEU:HD21	1:B:1216:MET:CE	2.50	0.41
1:A:172:LYS:HD3	6:A:726:HOH:O	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:273:TYR:HA	1:A:276:LEU:HD12	2.02	0.41
1:A:248:ILE:CG2	6:A:540:HOH:O	2.68	0.41
1:A:293:LEU:HD11	3:E:904:DA:C5	2.54	0.41
1:A:38:GLU:O	1:A:39:ASP:HB2	2.21	0.41
3:E:902:DT:H2"	3:E:903:DA:H5'	2.01	0.41

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	339/341 (99%)	332 (98%)	7 (2%)	0	100 100
1	B	339/341 (99%)	332 (98%)	4 (1%)	3 (1%)	17 7
All	All	678/682 (99%)	664 (98%)	11 (2%)	3 (0%)	34 24

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1277	ASP
1	B	1010	TYR
1	B	1278	LYS

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	299/299 (100%)	284 (95%)	15 (5%)	24	15
1	B	299/299 (100%)	291 (97%)	8 (3%)	44	38
All	All	598/598 (100%)	575 (96%)	23 (4%)	33	24

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

Mol	Chain	Res	Type
1	A	19	LEU
1	A	23	LEU
1	A	105	ASP
1	A	135	LYS
1	A	221	LYS
1	A	237	ILE
1	A	238	ARG
1	A	242	ARG
1	A	248	ILE
1	A	252	LYS
1	A	253	ARG
1	A	273	TYR
1	A	309	GLU
1	A	323	LEU
1	A	327	GLU
1	B	1019	LEU
1	B	1023	LEU
1	B	1079	GLU
1	B	1105	ASP
1	B	1133	LEU
1	B	1238	ARG
1	B	1323	LEU
1	B	1327	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	14	GLN
1	A	285	HIS
1	B	1014	GLN
1	B	1285	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)

4 non-standard protein/DNA/RNA residues are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	2DT	D	814	2,3	17,20,21	0.48	0	22,28,31	1.10	2 (9%)
2	2DT	H	1814	2,3	17,20,21	0.51	0	22,28,31	1.10	2 (9%)
3	MG1	E	906	2,3	18,25,26	1.27	2 (11%)	19,37,40	1.25	2 (10%)
3	MG1	J	1906	2,3	18,25,26	1.17	2 (11%)	19,37,40	1.28	3 (15%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	2DT	D	814	2,3	-	4/7/18/19	0/2/2/2
2	2DT	H	1814	2,3	-	2/7/18/19	0/2/2/2
3	MG1	E	906	2,3	-	0/3/21/22	0/3/3/3
3	MG1	J	1906	2,3	-	0/3/21/22	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	906	MG1	C6-N1	-3.52	1.33	1.39
3	J	1906	MG1	C5-C6	-2.77	1.39	1.47
3	E	906	MG1	C5-C6	-2.62	1.39	1.47
3	J	1906	MG1	C6-N1	-2.60	1.34	1.39

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	J	1906	MG1	C5-C6-N1	3.77	119.57	113.90
3	E	906	MG1	C5-C6-N1	3.75	119.54	113.90
2	D	814	2DT	O4'-C1'-C2'	2.74	109.64	106.67
3	J	1906	MG1	C8-N7-C5	2.39	107.55	102.99
3	J	1906	MG1	O6-C6-C5	-2.37	120.00	124.19
3	E	906	MG1	C8-N7-C5	2.18	107.14	102.99
2	D	814	2DT	C2'-C3'-C4'	2.06	106.58	102.72
2	H	1814	2DT	C5M-C5-C4	2.01	120.98	118.77
2	H	1814	2DT	C3'-C2'-C1'	2.01	105.10	102.78

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	814	2DT	O4'-C4'-C5'-O5'
2	D	814	2DT	C3'-C4'-C5'-O5'
2	D	814	2DT	O4'-C1'-N1-C6
2	D	814	2DT	O4'-C1'-N1-C2
2	H	1814	2DT	O4'-C4'-C5'-O5'
2	H	1814	2DT	O4'-C1'-N1-C6

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	814	2DT	4	0
2	H	1814	2DT	3	0
3	E	906	MG1	3	0
3	J	1906	MG1	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 6 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
4	DGT	A	414	5	26,33,33	1.04	3 (11%)	32,52,52	2.32	8 (25%)
4	DGT	B	1414	5	26,33,33	0.99	3 (11%)	32,52,52	2.53	9 (28%)

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	DGT	A	414	5	-	3/18/34/34	0/3/3/3
4	DGT	B	1414	5	-	3/18/34/34	0/3/3/3

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1414	DGT	C5-C6	-2.79	1.41	1.47
4	A	414	DGT	C5-C6	-2.59	1.42	1.47
4	A	414	DGT	C6-N1	2.24	1.41	1.37
4	A	414	DGT	C5-C4	-2.23	1.37	1.43
4	B	1414	DGT	C5-C4	-2.06	1.37	1.43
4	B	1414	DGT	C8-N7	-2.02	1.31	1.35

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	1414	DGT	O2G-PG-O3G	-6.40	85.62	110.68
4	B	1414	DGT	PB-O3B-PG	-5.91	112.55	132.83
4	A	414	DGT	PB-O3B-PG	-5.60	113.61	132.83
4	B	1414	DGT	O1G-PG-O3G	-5.18	90.41	110.68
4	A	414	DGT	O2G-PG-O3G	-5.00	91.10	110.68
4	A	414	DGT	O1G-PG-O3G	-4.86	91.65	110.68
4	B	1414	DGT	O2G-PG-O3B	4.65	120.22	104.64
4	A	414	DGT	O2G-PG-O1G	4.36	124.32	107.64
4	B	1414	DGT	O1G-PG-O3B	4.30	119.06	104.64
4	A	414	DGT	O1G-PG-O3B	3.93	117.80	104.64
4	B	1414	DGT	O3B-PG-O3G	-3.91	89.48	111.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	414	DGT	O2G-PG-O3B	3.89	117.67	104.64
4	B	1414	DGT	C2'-C1'-N9	-3.58	106.01	114.27
4	A	414	DGT	C2'-C1'-N9	-3.56	106.06	114.27
4	A	414	DGT	O3B-PG-O3G	-3.50	91.76	111.19
4	B	1414	DGT	O2G-PG-O1G	3.37	120.50	107.64
4	B	1414	DGT	PA-O3A-PB	-2.96	122.68	132.83

There are no chirality outliers.

All (6) torsion outliers are listed below:

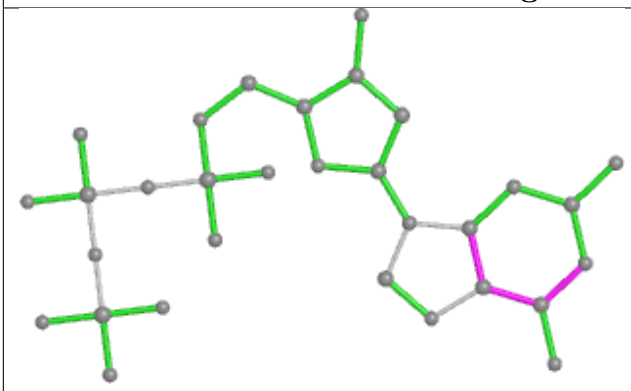
Mol	Chain	Res	Type	Atoms
4	A	414	DGT	PB-O3B-PG-O2G
4	A	414	DGT	PA-O3A-PB-O1B
4	B	1414	DGT	PA-O3A-PB-O1B
4	A	414	DGT	PB-O3B-PG-O3G
4	B	1414	DGT	PB-O3B-PG-O3G
4	B	1414	DGT	PB-O3B-PG-O2G

There are no ring outliers.

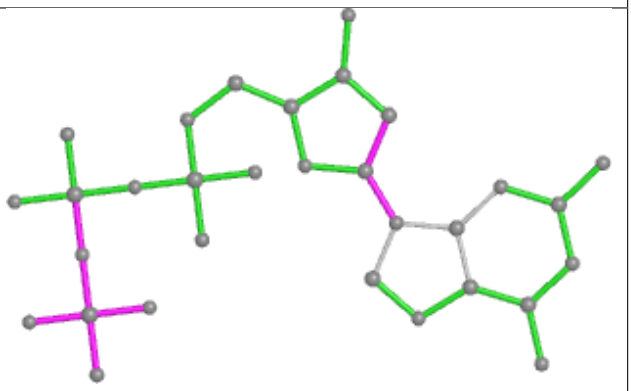
No monomer is involved in short contacts.

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

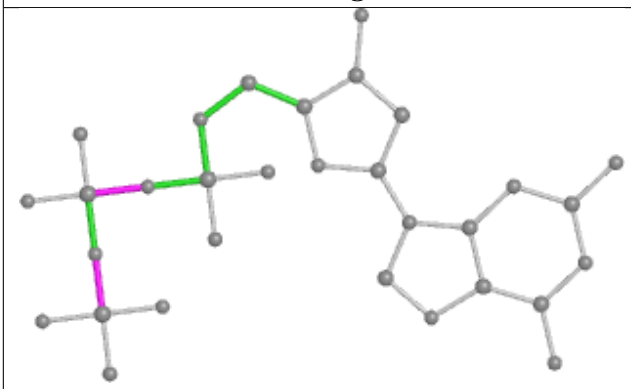
Ligand DGT A 414



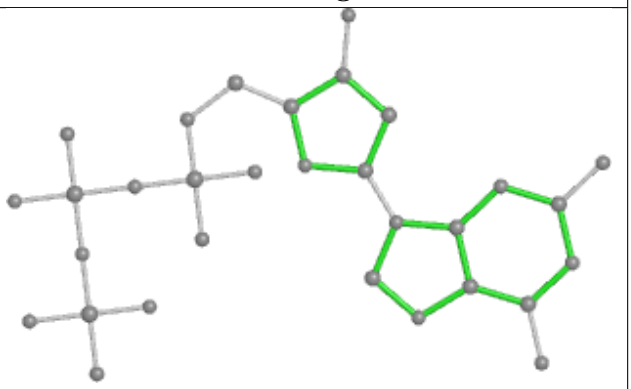
Bond lengths



Bond angles

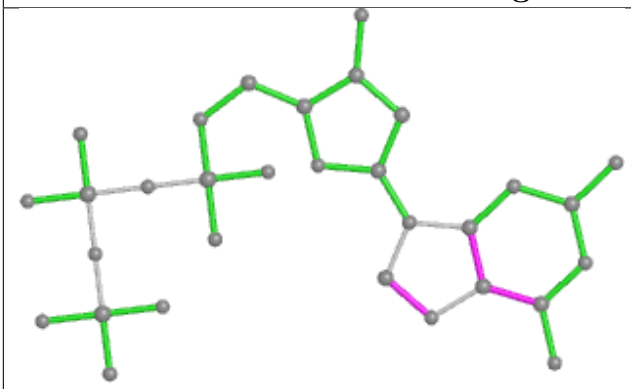


Torsions

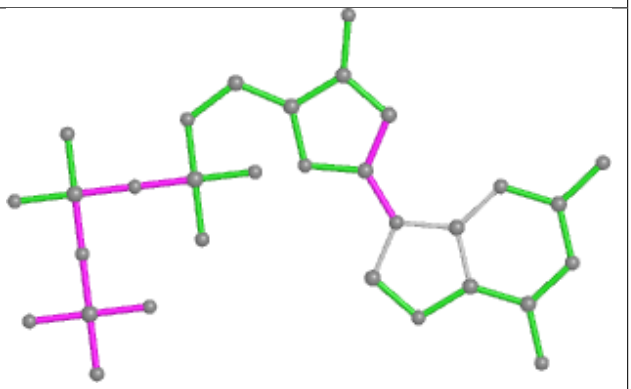


Rings

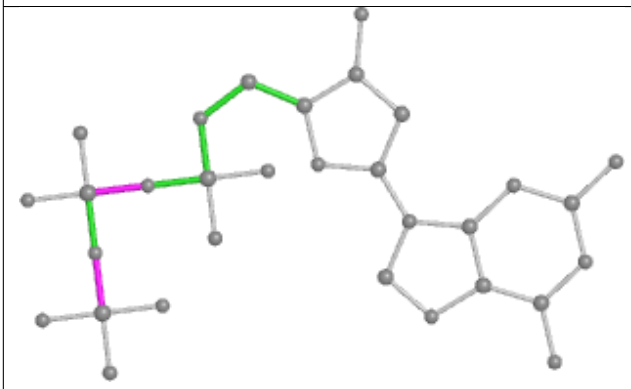
Ligand DGT B 1414



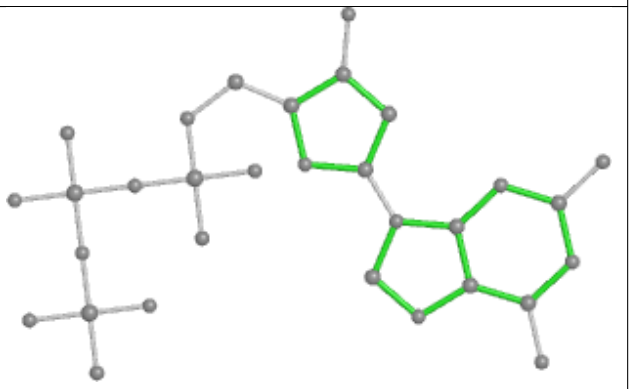
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (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	341/341 (100%)	0.53	28 (8%) 11 13	20, 35, 52, 61	0
1	B	341/341 (100%)	0.77	44 (12%) 3 3	29, 44, 73, 92	0
2	D	12/13 (92%)	-0.00	0 100 100	32, 40, 84, 89	0
2	H	12/13 (92%)	0.15	1 (8%) 11 13	41, 46, 94, 96	0
3	E	17/20 (85%)	0.24	2 (11%) 4 5	34, 47, 89, 92	0
3	J	16/20 (80%)	0.59	3 (18%) 1 1	37, 63, 90, 98	0
All	All	739/748 (98%)	0.62	78 (10%) 6 7	20, 41, 73, 98	0

All (78) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1225	LEU	5.0
1	B	1209	GLU	4.7
1	B	1217	ILE	4.5
1	B	1037	PHE	4.2
1	A	225	LEU	4.2
1	A	234	ASN	4.1
1	B	1278	LYS	3.7
1	B	1202	LEU	3.6
1	B	1004	LEU	3.6
1	A	202	LEU	3.5
3	E	902	DT	3.5
1	B	1286	VAL	3.4
1	B	1234	ASN	3.3
1	B	1030	VAL	3.3
1	B	1013	ALA	3.2
1	B	1341	ILE	3.1
1	A	286	VAL	3.1
3	E	919	DC	3.0
1	B	1144	ILE	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	J	1919	DC	3.0
1	A	29	VAL	3.0
1	B	1116	ARG	2.9
1	B	1046	ALA	2.9
1	B	1011	PHE	2.9
1	B	1006	VAL	2.9
1	A	11	PHE	2.8
1	B	1277	ASP	2.8
1	B	1238	ARG	2.8
1	A	333	ILE	2.8
1	A	209	GLU	2.8
1	B	1194	LEU	2.8
1	A	277	ASP	2.7
1	A	205	THR	2.7
1	A	284	ILE	2.7
1	B	1274	TYR	2.7
1	B	1038	GLU	2.7
1	B	1240	ARG	2.7
1	B	1109	LEU	2.7
1	B	1211	ASP	2.7
1	B	1123	ASN	2.6
1	A	3	VAL	2.6
1	A	142	VAL	2.6
1	A	12	TYR	2.5
1	A	144	ILE	2.5
1	A	186	ILE	2.5
1	A	337	PHE	2.5
1	B	1213	LEU	2.5
1	A	4	LEU	2.5
1	A	8	PHE	2.5
1	A	240	ARG	2.4
1	A	183	VAL	2.4
1	B	1012	TYR	2.4
1	B	1337	PHE	2.4
3	J	1903	DA	2.4
1	B	1186	ILE	2.3
1	B	1177	GLU	2.3
1	B	1008	PHE	2.3
1	B	1216	MET	2.3
1	B	1284	ILE	2.3
1	A	206	LEU	2.3
3	J	1904	DA	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	46	ALA	2.2
2	H	1804	DT	2.2
1	B	1029	VAL	2.2
1	B	1333	ILE	2.1
1	A	109	LEU	2.1
1	B	1327	GLU	2.1
1	B	1212	LYS	2.1
1	B	1003	VAL	2.1
1	A	92	LEU	2.1
1	A	274	TYR	2.0
1	A	304	HIS	2.0
1	B	1045	THR	2.0
1	B	1196	LYS	2.0
1	B	1188	ASN	2.0
1	B	1330	ILE	2.0
1	A	238	ARG	2.0
1	B	1192	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	H	1814	19/20	0.76	0.47	62,73,80,80	0
2	2DT	D	814	19/20	0.81	0.44	50,66,73,73	0
3	MG1	E	906	23/24	0.95	0.10	33,36,38,41	0
3	MG1	J	1906	23/24	0.97	0.08	35,39,43,44	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

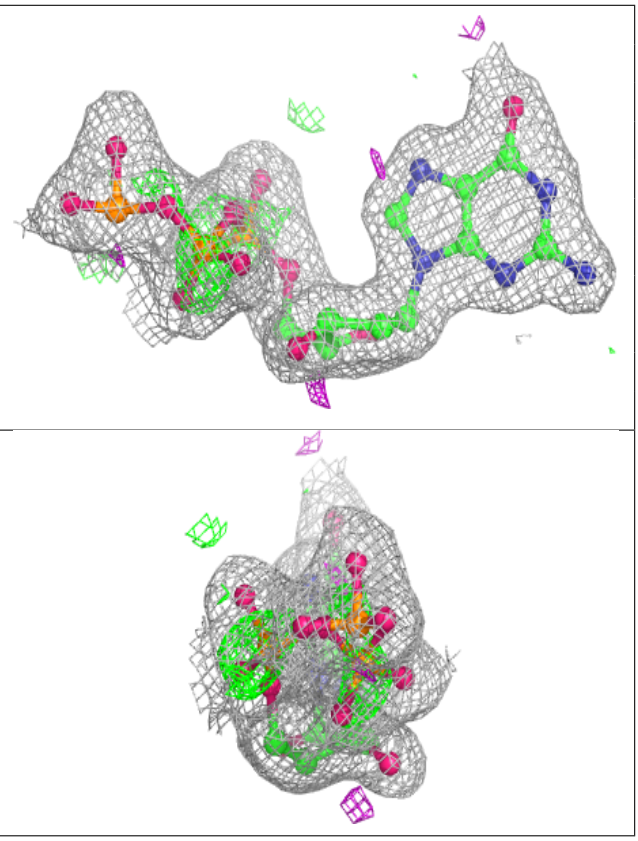
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	DGT	B	1414	31/31	0.97	0.12	28,30,36,42	0
5	CA	B	1417	1/1	0.97	0.07	42,42,42,42	0
5	CA	B	1415	1/1	0.98	0.03	29,29,29,29	0
5	CA	B	1416	1/1	0.98	0.07	26,26,26,26	0
4	DGT	A	414	31/31	0.98	0.12	21,24,27,29	0
5	CA	A	416	1/1	0.99	0.10	18,18,18,18	0
5	CA	A	417	1/1	1.00	0.03	36,36,36,36	0
5	CA	A	415	1/1	1.00	0.05	20,20,20,20	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.

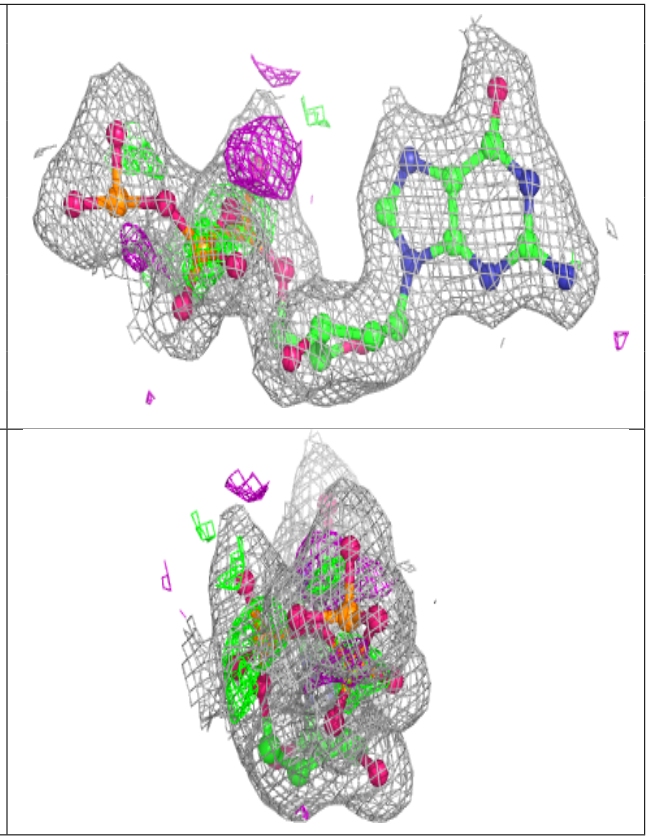
Electron density around DGT B 1414:

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 DGT A 414:

$2mF_o$-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.