wwPDB X-ray Structure Validation Summary Report (i)

Dec 16, 2023 – 09:36 am GMT

PDB ID	:	4BZB
Title	:	Crystal structure of the tetrameric dGTP-bound SAMHD1 mutant catalytic core
Authors	:	Ji, X.; Yang, H.; Wu, Y.; Yan, J.; Mehrens, J.; DeLucia, M.; Hao, C.; Gronenborn, A.M.; Skowronski, J.; Ahn, J.; Xiong, Y.
Deposited on	:	2013-07-25
Resolution	:	1.83 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 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:

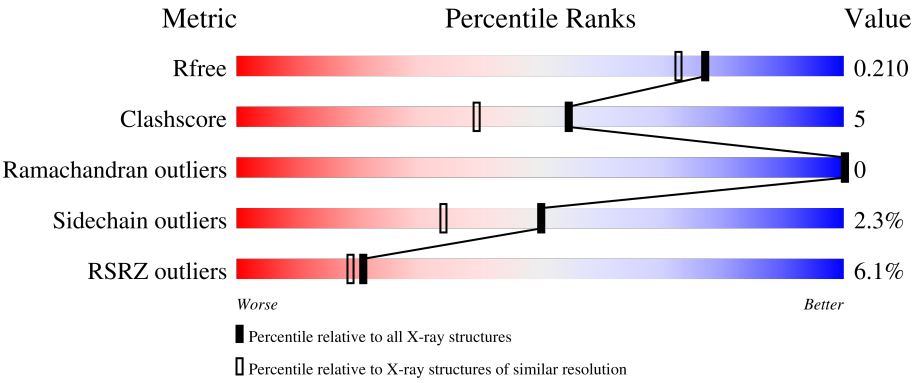
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.83 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	4003 (1.86-1.82)
Clashscore	141614	4233 (1.86-1.82)
Ramachandran outliers	138981	4185 (1.86-1.82)
Sidechain outliers	138945	4186 (1.86-1.82)
RSRZ outliers	127900	3957 (1.86-1.82)

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	550	6% 78% 9% • 13%
1	B	550	4% 79% 8% 13%
1	C	550	7% 77% 9% • 13%
1	D	550	4% 79% 7% • 13%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 17207 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 DEOXYNUCLEOSIDE TRIPHOSPHATE TRIPHOSPHOHYDROLASE SAMHD1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	481	Total	C	N	O	S	0	7	0
			3997	2554	698	724	21			
1	B	481	Total	C	N	O	S	0	4	0
			3975	2543	694	718	20			
1	C	481	Total	C	N	O	S	0	5	0
			3981	2546	695	719	21			
1	D	481	Total	C	N	O	S	0	7	0
			3997	2554	698	724	21			

There are 152 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	77	MET	-	expression tag	UNP Q9Y3Z3
A	78	GLY	-	expression tag	UNP Q9Y3Z3
A	79	SER	-	expression tag	UNP Q9Y3Z3
A	80	SER	-	expression tag	UNP Q9Y3Z3
A	81	HIS	-	expression tag	UNP Q9Y3Z3
A	82	HIS	-	expression tag	UNP Q9Y3Z3
A	83	HIS	-	expression tag	UNP Q9Y3Z3
A	84	HIS	-	expression tag	UNP Q9Y3Z3
A	85	HIS	-	expression tag	UNP Q9Y3Z3
A	86	HIS	-	expression tag	UNP Q9Y3Z3
A	87	SER	-	expression tag	UNP Q9Y3Z3
A	88	SER	-	expression tag	UNP Q9Y3Z3
A	89	GLY	-	expression tag	UNP Q9Y3Z3
A	90	LEU	-	expression tag	UNP Q9Y3Z3
A	91	VAL	-	expression tag	UNP Q9Y3Z3
A	92	PRO	-	expression tag	UNP Q9Y3Z3
A	93	ARG	-	expression tag	UNP Q9Y3Z3
A	94	GLY	-	expression tag	UNP Q9Y3Z3
A	95	SER	-	expression tag	UNP Q9Y3Z3
A	96	HIS	-	expression tag	UNP Q9Y3Z3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	97	MET	-	expression tag	UNP Q9Y3Z3
A	98	ALA	-	expression tag	UNP Q9Y3Z3
A	99	SER	-	expression tag	UNP Q9Y3Z3
A	100	MET	-	expression tag	UNP Q9Y3Z3
A	101	THR	-	expression tag	UNP Q9Y3Z3
A	102	GLY	-	expression tag	UNP Q9Y3Z3
A	103	GLY	-	expression tag	UNP Q9Y3Z3
A	104	GLN	-	expression tag	UNP Q9Y3Z3
A	105	GLN	-	expression tag	UNP Q9Y3Z3
A	106	MET	-	expression tag	UNP Q9Y3Z3
A	107	GLY	-	expression tag	UNP Q9Y3Z3
A	108	ARG	-	expression tag	UNP Q9Y3Z3
A	109	ASP	-	expression tag	UNP Q9Y3Z3
A	110	PRO	-	expression tag	UNP Q9Y3Z3
A	111	ASN	-	expression tag	UNP Q9Y3Z3
A	112	SER	-	expression tag	UNP Q9Y3Z3
A	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
A	207	ASN	ASP	engineered mutation	UNP Q9Y3Z3
B	77	MET	-	expression tag	UNP Q9Y3Z3
B	78	GLY	-	expression tag	UNP Q9Y3Z3
B	79	SER	-	expression tag	UNP Q9Y3Z3
B	80	SER	-	expression tag	UNP Q9Y3Z3
B	81	HIS	-	expression tag	UNP Q9Y3Z3
B	82	HIS	-	expression tag	UNP Q9Y3Z3
B	83	HIS	-	expression tag	UNP Q9Y3Z3
B	84	HIS	-	expression tag	UNP Q9Y3Z3
B	85	HIS	-	expression tag	UNP Q9Y3Z3
B	86	HIS	-	expression tag	UNP Q9Y3Z3
B	87	SER	-	expression tag	UNP Q9Y3Z3
B	88	SER	-	expression tag	UNP Q9Y3Z3
B	89	GLY	-	expression tag	UNP Q9Y3Z3
B	90	LEU	-	expression tag	UNP Q9Y3Z3
B	91	VAL	-	expression tag	UNP Q9Y3Z3
B	92	PRO	-	expression tag	UNP Q9Y3Z3
B	93	ARG	-	expression tag	UNP Q9Y3Z3
B	94	GLY	-	expression tag	UNP Q9Y3Z3
B	95	SER	-	expression tag	UNP Q9Y3Z3
B	96	HIS	-	expression tag	UNP Q9Y3Z3
B	97	MET	-	expression tag	UNP Q9Y3Z3
B	98	ALA	-	expression tag	UNP Q9Y3Z3
B	99	SER	-	expression tag	UNP Q9Y3Z3
B	100	MET	-	expression tag	UNP Q9Y3Z3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	101	THR	-	expression tag	UNP Q9Y3Z3
B	102	GLY	-	expression tag	UNP Q9Y3Z3
B	103	GLY	-	expression tag	UNP Q9Y3Z3
B	104	GLN	-	expression tag	UNP Q9Y3Z3
B	105	GLN	-	expression tag	UNP Q9Y3Z3
B	106	MET	-	expression tag	UNP Q9Y3Z3
B	107	GLY	-	expression tag	UNP Q9Y3Z3
B	108	ARG	-	expression tag	UNP Q9Y3Z3
B	109	ASP	-	expression tag	UNP Q9Y3Z3
B	110	PRO	-	expression tag	UNP Q9Y3Z3
B	111	ASN	-	expression tag	UNP Q9Y3Z3
B	112	SER	-	expression tag	UNP Q9Y3Z3
B	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
B	207	ASN	ASP	engineered mutation	UNP Q9Y3Z3
C	77	MET	-	expression tag	UNP Q9Y3Z3
C	78	GLY	-	expression tag	UNP Q9Y3Z3
C	79	SER	-	expression tag	UNP Q9Y3Z3
C	80	SER	-	expression tag	UNP Q9Y3Z3
C	81	HIS	-	expression tag	UNP Q9Y3Z3
C	82	HIS	-	expression tag	UNP Q9Y3Z3
C	83	HIS	-	expression tag	UNP Q9Y3Z3
C	84	HIS	-	expression tag	UNP Q9Y3Z3
C	85	HIS	-	expression tag	UNP Q9Y3Z3
C	86	HIS	-	expression tag	UNP Q9Y3Z3
C	87	SER	-	expression tag	UNP Q9Y3Z3
C	88	SER	-	expression tag	UNP Q9Y3Z3
C	89	GLY	-	expression tag	UNP Q9Y3Z3
C	90	LEU	-	expression tag	UNP Q9Y3Z3
C	91	VAL	-	expression tag	UNP Q9Y3Z3
C	92	PRO	-	expression tag	UNP Q9Y3Z3
C	93	ARG	-	expression tag	UNP Q9Y3Z3
C	94	GLY	-	expression tag	UNP Q9Y3Z3
C	95	SER	-	expression tag	UNP Q9Y3Z3
C	96	HIS	-	expression tag	UNP Q9Y3Z3
C	97	MET	-	expression tag	UNP Q9Y3Z3
C	98	ALA	-	expression tag	UNP Q9Y3Z3
C	99	SER	-	expression tag	UNP Q9Y3Z3
C	100	MET	-	expression tag	UNP Q9Y3Z3
C	101	THR	-	expression tag	UNP Q9Y3Z3
C	102	GLY	-	expression tag	UNP Q9Y3Z3
C	103	GLY	-	expression tag	UNP Q9Y3Z3
C	104	GLN	-	expression tag	UNP Q9Y3Z3

Continued on next page...

Continued from previous page...

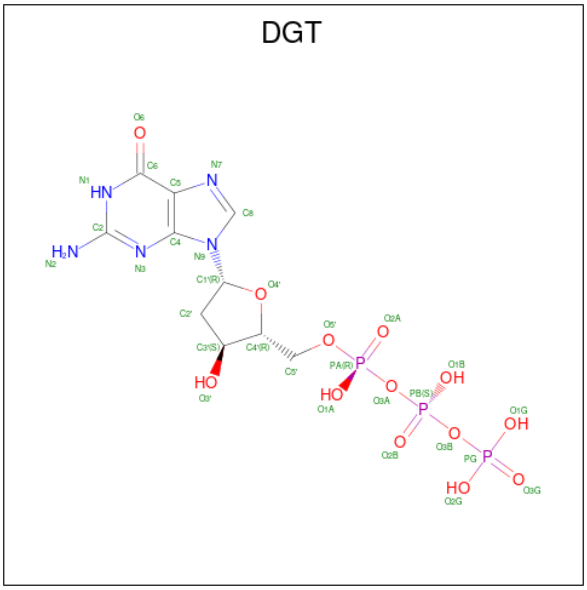
Chain	Residue	Modelled	Actual	Comment	Reference
C	105	GLN	-	expression tag	UNP Q9Y3Z3
C	106	MET	-	expression tag	UNP Q9Y3Z3
C	107	GLY	-	expression tag	UNP Q9Y3Z3
C	108	ARG	-	expression tag	UNP Q9Y3Z3
C	109	ASP	-	expression tag	UNP Q9Y3Z3
C	110	PRO	-	expression tag	UNP Q9Y3Z3
C	111	ASN	-	expression tag	UNP Q9Y3Z3
C	112	SER	-	expression tag	UNP Q9Y3Z3
C	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
C	207	ASN	ASP	engineered mutation	UNP Q9Y3Z3
D	77	MET	-	expression tag	UNP Q9Y3Z3
D	78	GLY	-	expression tag	UNP Q9Y3Z3
D	79	SER	-	expression tag	UNP Q9Y3Z3
D	80	SER	-	expression tag	UNP Q9Y3Z3
D	81	HIS	-	expression tag	UNP Q9Y3Z3
D	82	HIS	-	expression tag	UNP Q9Y3Z3
D	83	HIS	-	expression tag	UNP Q9Y3Z3
D	84	HIS	-	expression tag	UNP Q9Y3Z3
D	85	HIS	-	expression tag	UNP Q9Y3Z3
D	86	HIS	-	expression tag	UNP Q9Y3Z3
D	87	SER	-	expression tag	UNP Q9Y3Z3
D	88	SER	-	expression tag	UNP Q9Y3Z3
D	89	GLY	-	expression tag	UNP Q9Y3Z3
D	90	LEU	-	expression tag	UNP Q9Y3Z3
D	91	VAL	-	expression tag	UNP Q9Y3Z3
D	92	PRO	-	expression tag	UNP Q9Y3Z3
D	93	ARG	-	expression tag	UNP Q9Y3Z3
D	94	GLY	-	expression tag	UNP Q9Y3Z3
D	95	SER	-	expression tag	UNP Q9Y3Z3
D	96	HIS	-	expression tag	UNP Q9Y3Z3
D	97	MET	-	expression tag	UNP Q9Y3Z3
D	98	ALA	-	expression tag	UNP Q9Y3Z3
D	99	SER	-	expression tag	UNP Q9Y3Z3
D	100	MET	-	expression tag	UNP Q9Y3Z3
D	101	THR	-	expression tag	UNP Q9Y3Z3
D	102	GLY	-	expression tag	UNP Q9Y3Z3
D	103	GLY	-	expression tag	UNP Q9Y3Z3
D	104	GLN	-	expression tag	UNP Q9Y3Z3
D	105	GLN	-	expression tag	UNP Q9Y3Z3
D	106	MET	-	expression tag	UNP Q9Y3Z3
D	107	GLY	-	expression tag	UNP Q9Y3Z3
D	108	ARG	-	expression tag	UNP Q9Y3Z3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	109	ASP	-	expression tag	UNP Q9Y3Z3
D	110	PRO	-	expression tag	UNP Q9Y3Z3
D	111	ASN	-	expression tag	UNP Q9Y3Z3
D	112	SER	-	expression tag	UNP Q9Y3Z3
D	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
D	207	ASN	ASP	engineered mutation	UNP Q9Y3Z3

- Molecule 2 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
2	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	C	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	C	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	C	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	D	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	D	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	D	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Mg	0	0
			2	2		
3	B	2	Total	Mg	0	0
			2	2		
3	C	2	Total	Mg	0	0
			2	2		
3	D	2	Total	Mg	0	0
			2	2		

- Molecule 4 is water.

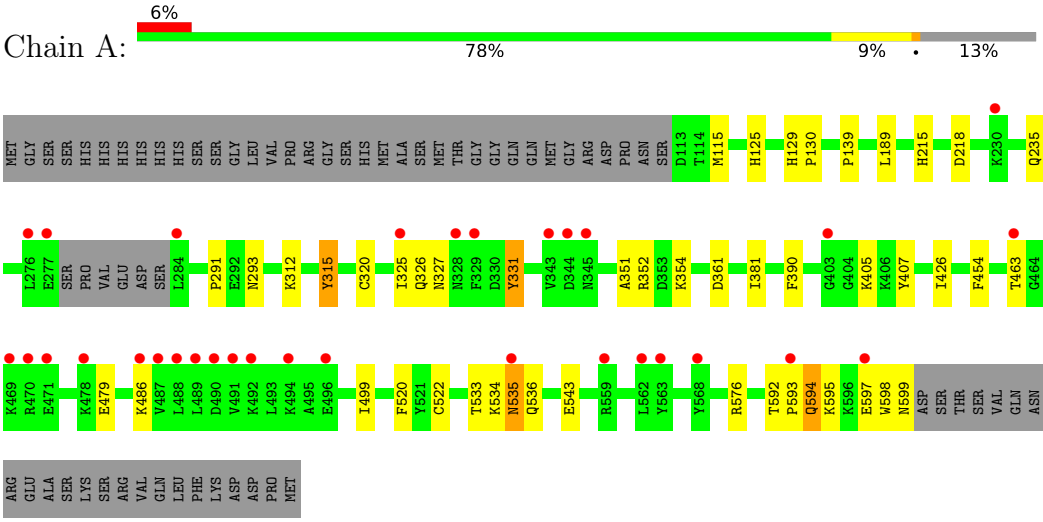
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	228	Total	O	0	0
			228	228		
4	B	271	Total	O	0	0
			271	271		
4	C	156	Total	O	0	0
			156	156		
4	D	222	Total	O	0	0
			222	222		

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.

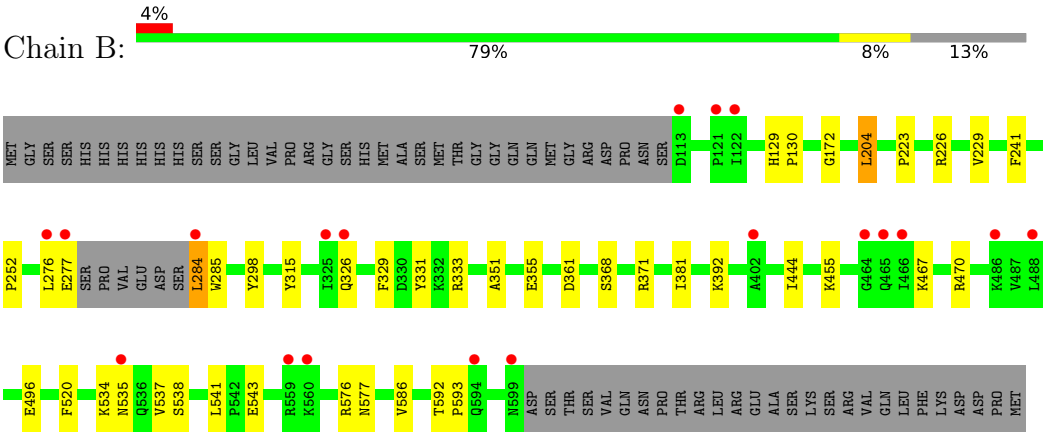
• Molecule 1: DEOXYNUCLEOSIDE TRIPHOSPHATE TRIPHOSPHOHYDROLASE SAMHD1

Chain A:



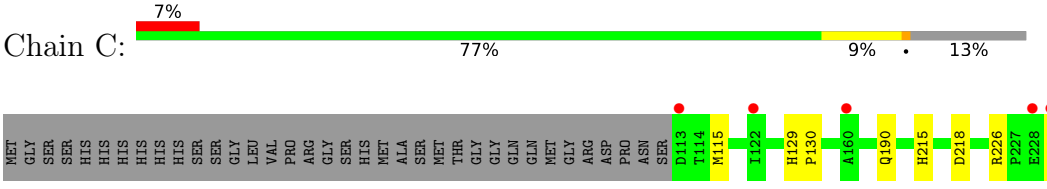
• Molecule 1: DEOXYNUCLEOSIDE TRIPHOSPHATE TRIPHOSPHOHYDROLASE SAMHD1

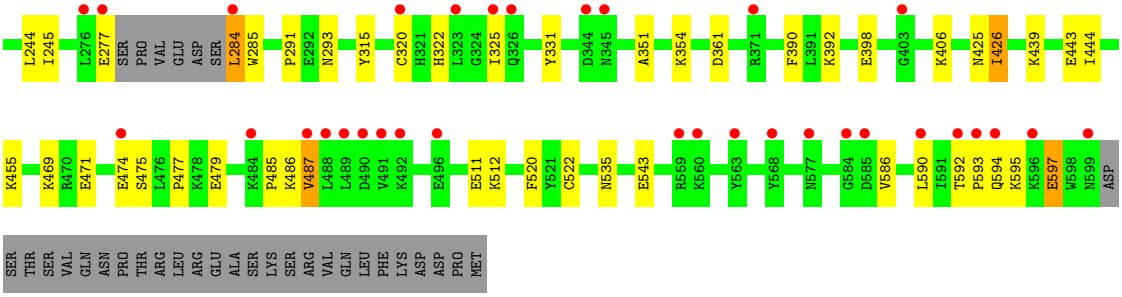
Chain B:



• Molecule 1: DEOXYNUCLEOSIDE TRIPHOSPHATE TRIPHOSPHOHYDROLASE SAMHD1

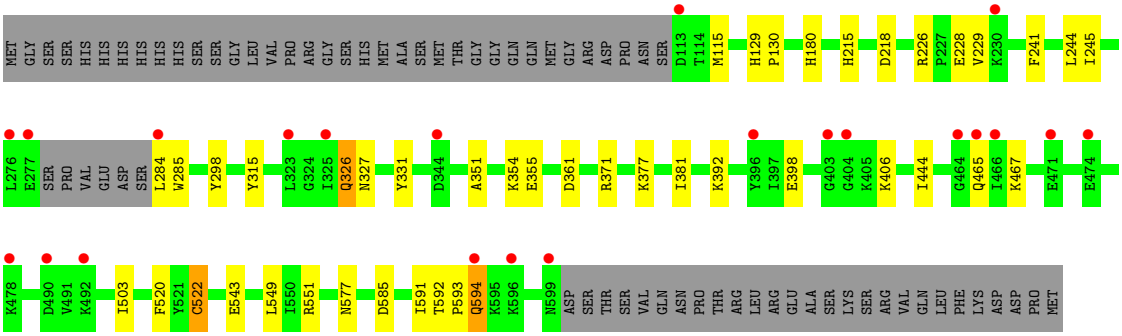
Chain C:





• Molecule 1: DEOXYNUCLEOSIDE TRIPHOSPHATE TRIPHOSPHOHYDROLASE SAMHD1

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	87.20Å 146.75Å 98.72Å 90.00° 114.42° 90.00°	Depositor
Resolution (Å)	48.73 – 1.83 48.68 – 1.83	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.73-1.83) 99.7 (48.68-1.83)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.41 (at 1.83Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.189 , 0.206 0.193 , 0.210	Depositor DCC
R_{free} test set	9923 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	31.9	Xtriage
Anisotropy	0.268	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 40.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.015 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	17207	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.45% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

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

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, DGT

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	0.32	0/4091	0.54	0/5522
1	B	0.30	0/4069	0.53	0/5492
1	C	0.31	0/4075	0.54	0/5500
1	D	0.35	1/4091 (0.0%)	0.56	1/5522 (0.0%)
All	All	0.32	1/16326 (0.0%)	0.54	1/22036 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	354	LYS	N-CA	-5.75	1.34	1.46

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	354	LYS	N-CA-CB	-6.75	98.45	110.60

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3997	0	3964	42	0
1	B	3975	0	3952	42	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3981	0	3956	59	0
1	D	3997	0	3964	32	0
2	A	93	0	36	0	0
2	B	93	0	36	0	0
2	C	93	0	36	0	0
2	D	93	0	36	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	2	0	0	0	0
4	A	228	0	0	3	0
4	B	271	0	0	9	0
4	C	156	0	0	2	0
4	D	222	0	0	5	0
All	All	17207	0	15980	168	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 5.

The worst 5 of 168 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:C:439:LYS:HE2	1:C:443:GLU:HG2	1.28	1.14
1:C:439:LYS:NZ	1:C:443:GLU:HG3	1.62	1.12
1:D:585:ASP:OD1	1:D:592:THR:HG21	1.49	1.11
1:C:439:LYS:HE2	1:C:443:GLU:CG	1.89	1.03
1:C:474:GLU:O	1:C:477:PRO:HD2	1.69	0.93

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	484/550 (88%)	479 (99%)	5 (1%)	0	100	100
1	B	481/550 (88%)	475 (99%)	6 (1%)	0	100	100
1	C	482/550 (88%)	475 (98%)	7 (2%)	0	100	100
1	D	484/550 (88%)	478 (99%)	6 (1%)	0	100	100
All	All	1931/2200 (88%)	1907 (99%)	24 (1%)	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	434/488 (89%)	420 (97%)	14 (3%)	39	21
1	B	431/488 (88%)	426 (99%)	5 (1%)	71	61
1	C	432/488 (88%)	421 (98%)	11 (2%)	47	31
1	D	434/488 (89%)	421 (97%)	13 (3%)	41	23
All	All	1731/1952 (89%)	1688 (98%)	43 (2%)	50	31

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	594	GLN
1	D	361	ASP
1	C	597	GLU
1	D	180[B]	HIS
1	D	465	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	233	HIS
1	D	233	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)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 20 ligands modelled in this entry, 8 are monoatomic - leaving 12 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	$\#\|Z\| > 2$
2	DGT	C	800	3	26,33,33	0.82	1 (3%)	32,52,52	1.16	4 (12%)
2	DGT	B	800	3	26,33,33	0.89	1 (3%)	32,52,52	1.19	3 (9%)
2	DGT	C	900	3	26,33,33	0.84	1 (3%)	32,52,52	1.28	4 (12%)
2	DGT	A	800	3	26,33,33	0.86	1 (3%)	32,52,52	1.20	3 (9%)
2	DGT	A	900	3	26,33,33	0.84	1 (3%)	32,52,52	1.25	4 (12%)
2	DGT	D	700	3	26,33,33	0.88	1 (3%)	32,52,52	1.28	6 (18%)
2	DGT	B	700	3	26,33,33	0.88	1 (3%)	32,52,52	1.27	6 (18%)
2	DGT	C	700	3	26,33,33	0.85	1 (3%)	32,52,52	1.34	5 (15%)
2	DGT	B	900	3	26,33,33	0.83	1 (3%)	32,52,52	1.23	3 (9%)
2	DGT	A	700	3	26,33,33	0.87	1 (3%)	32,52,52	1.29	5 (15%)
2	DGT	D	900	3	26,33,33	0.87	1 (3%)	32,52,52	1.17	3 (9%)
2	DGT	D	800	3	26,33,33	0.81	1 (3%)	32,52,52	1.14	3 (9%)

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

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	DGT	C	800	3	-	3/18/34/34	0/3/3/3
2	DGT	B	800	3	-	2/18/34/34	0/3/3/3
2	DGT	C	900	3	-	5/18/34/34	0/3/3/3
2	DGT	A	800	3	-	2/18/34/34	0/3/3/3
2	DGT	A	900	3	-	1/18/34/34	0/3/3/3
2	DGT	D	700	3	-	3/18/34/34	0/3/3/3
2	DGT	B	700	3	-	3/18/34/34	0/3/3/3
2	DGT	C	700	3	-	3/18/34/34	0/3/3/3
2	DGT	B	900	3	-	4/18/34/34	0/3/3/3
2	DGT	A	700	3	-	3/18/34/34	0/3/3/3
2	DGT	D	900	3	-	3/18/34/34	0/3/3/3
2	DGT	D	800	3	-	2/18/34/34	0/3/3/3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	800	DGT	C6-N1	-3.04	1.33	1.37
2	A	800	DGT	C6-N1	-2.61	1.34	1.37
2	B	700	DGT	C6-N1	-2.45	1.34	1.37
2	D	900	DGT	C6-N1	-2.43	1.34	1.37
2	A	900	DGT	C6-N1	-2.36	1.34	1.37

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	900	DGT	PA-O3A-PB	-3.10	122.18	132.83
2	A	700	DGT	PB-O3B-PG	-3.00	122.54	132.83
2	C	900	DGT	PA-O3A-PB	-2.93	122.77	132.83
2	C	700	DGT	PA-O3A-PB	-2.88	122.94	132.83
2	C	700	DGT	PB-O3B-PG	-2.86	123.00	132.83

There are no chirality outliers.

5 of 34 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	700	DGT	O4'-C4'-C5'-O5'
2	A	700	DGT	C3'-C4'-C5'-O5'

Continued on next page...

Continued from previous page...

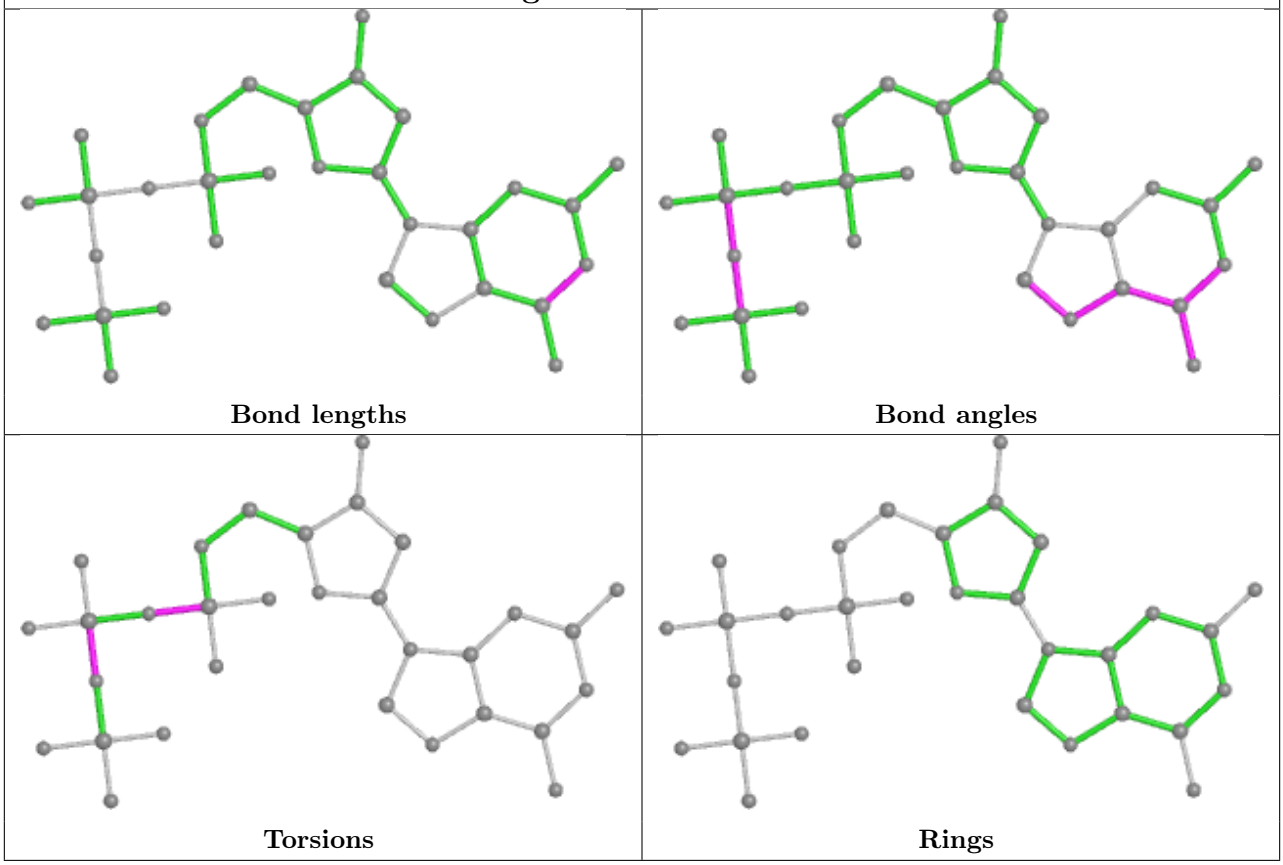
Mol	Chain	Res	Type	Atoms
2	B	700	DGT	C3'-C4'-C5'-O5'
2	C	700	DGT	C3'-C4'-C5'-O5'
2	D	700	DGT	C3'-C4'-C5'-O5'

There are no ring outliers.

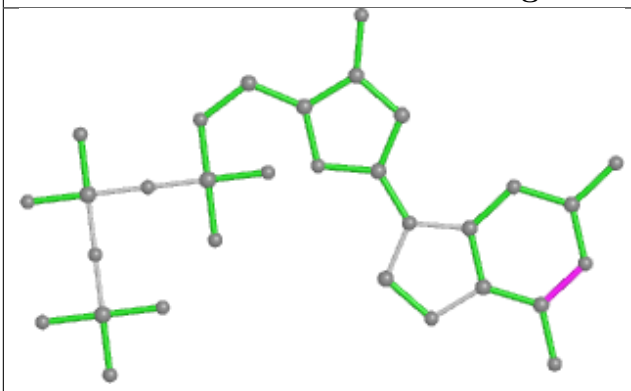
No monomer is involved in short contacts.

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.

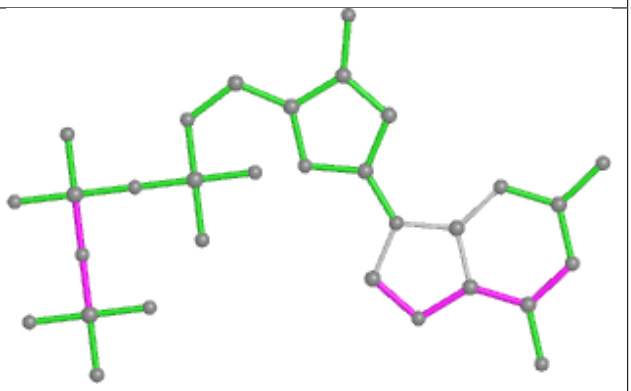
Ligand DGT C 800



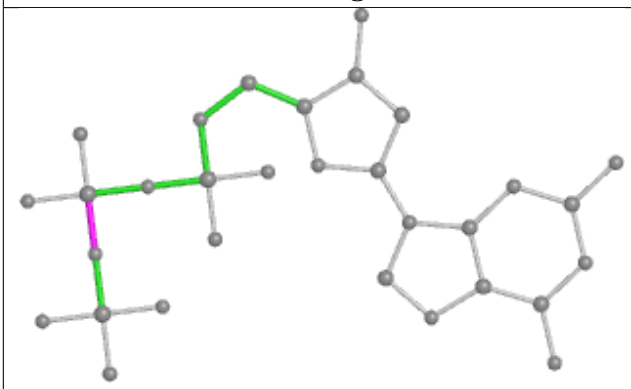
Ligand DGT B 800



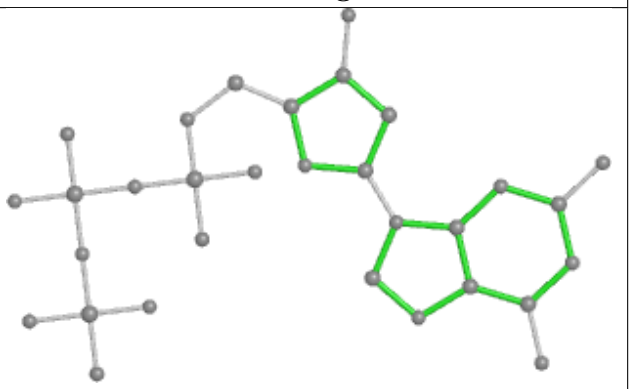
Bond lengths



Bond angles

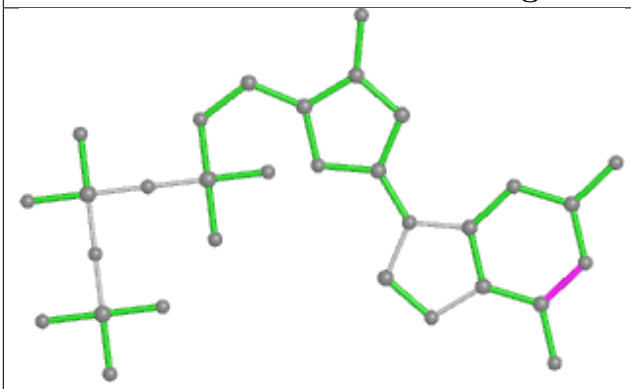


Torsions

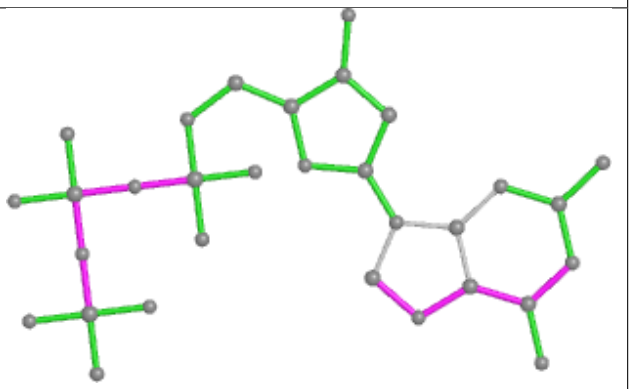


Rings

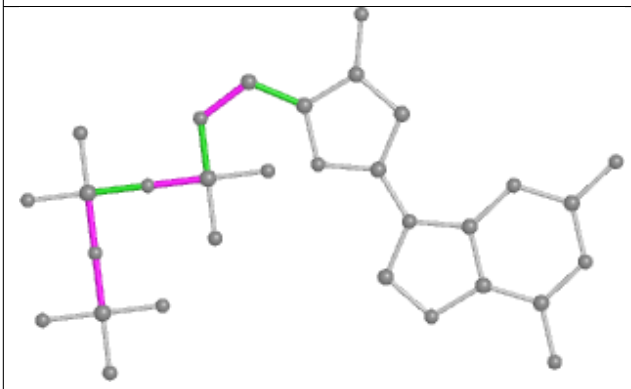
Ligand DGT C 900



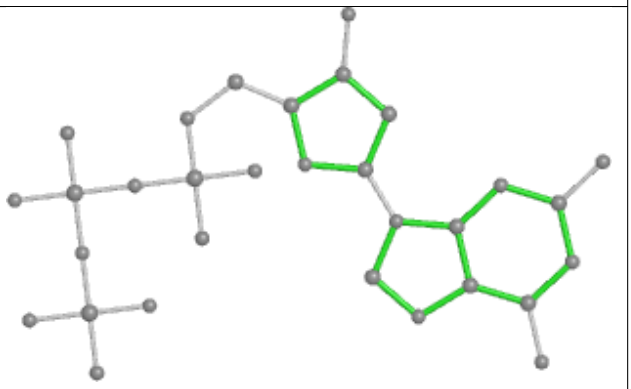
Bond lengths



Bond angles

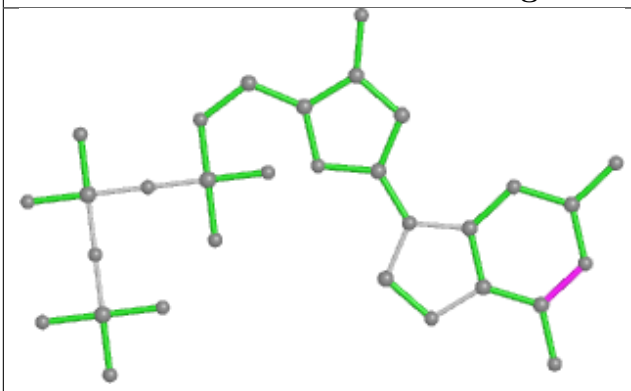


Torsions

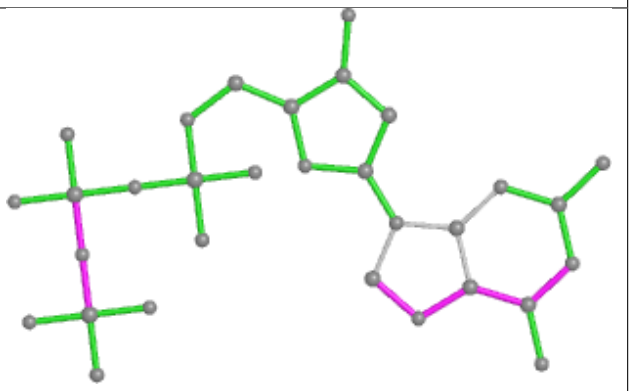


Rings

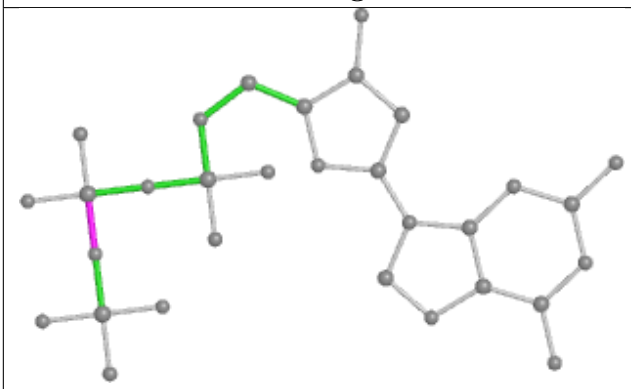
Ligand DGT A 800



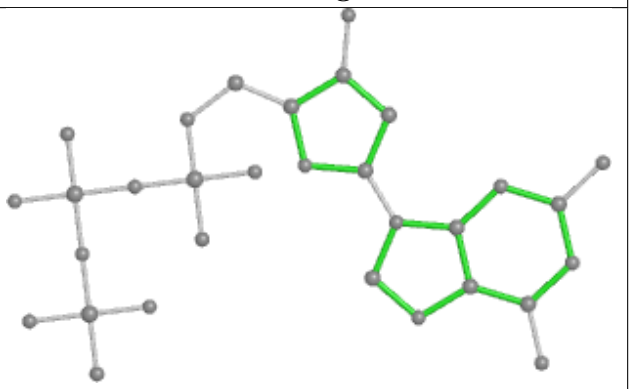
Bond lengths



Bond angles

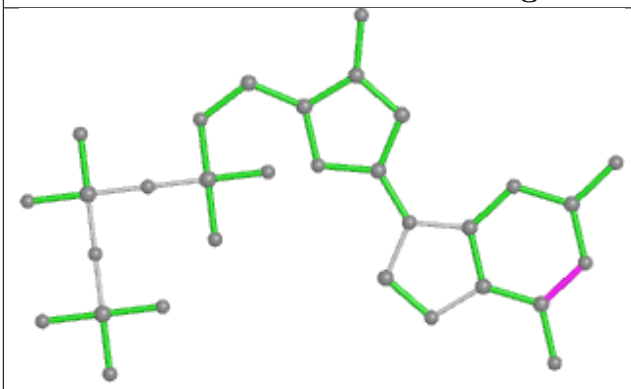


Torsions

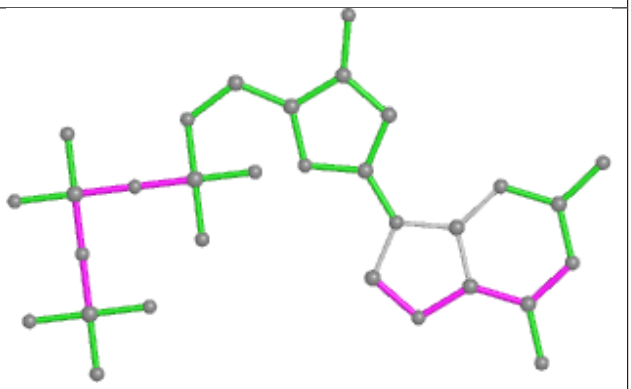


Rings

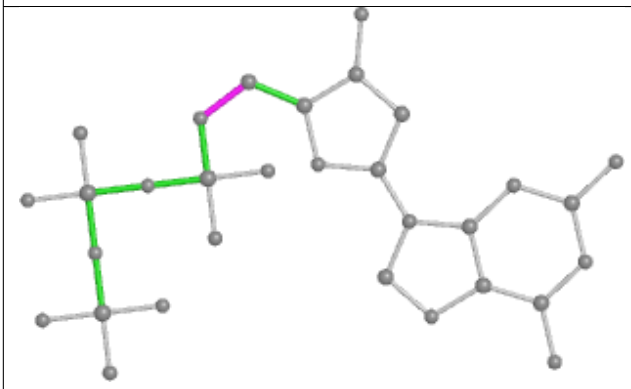
Ligand DGT A 900



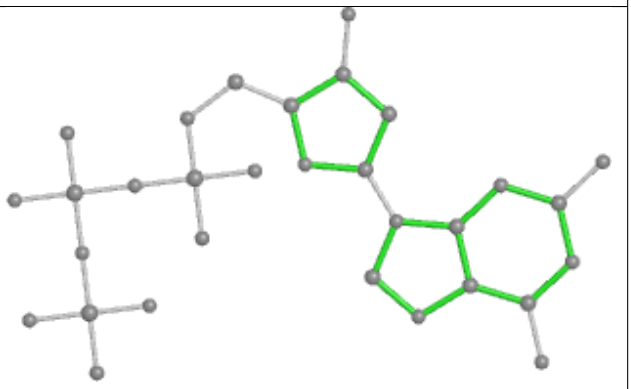
Bond lengths



Bond angles

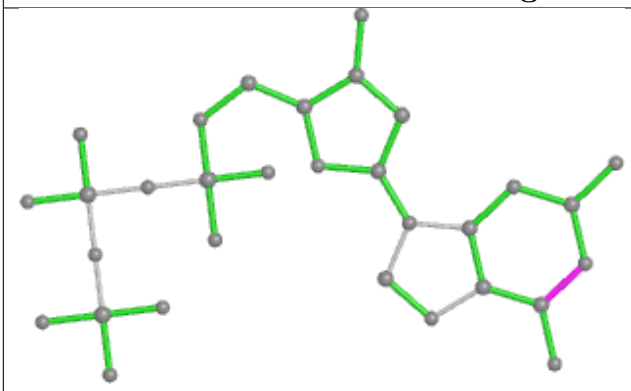


Torsions

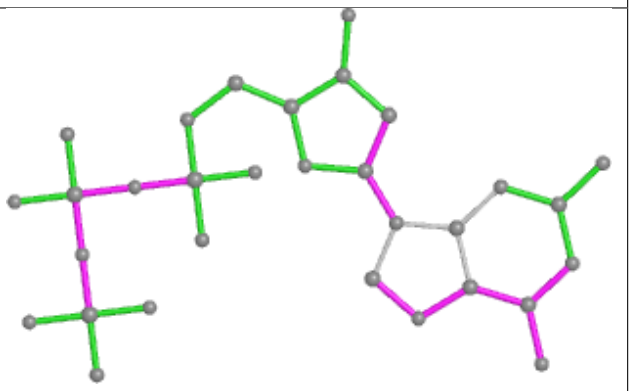


Rings

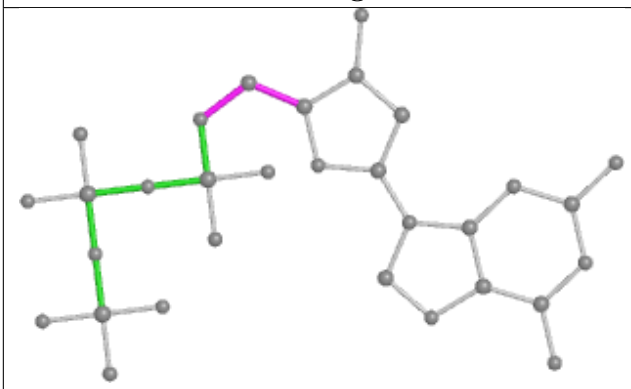
Ligand DGT D 700



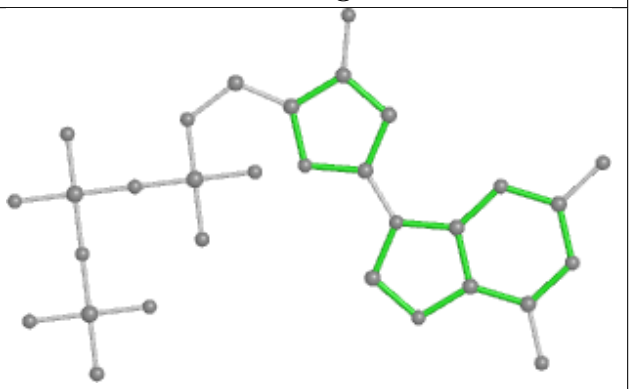
Bond lengths



Bond angles

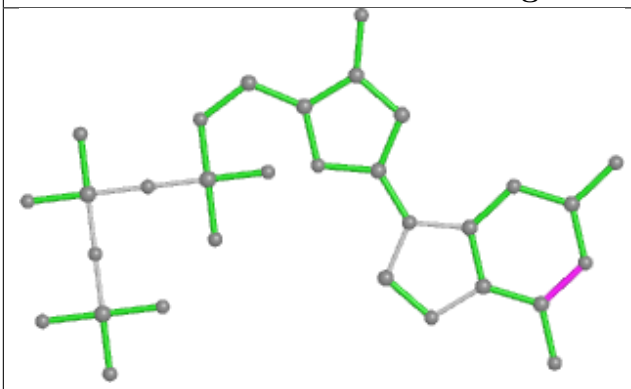


Torsions

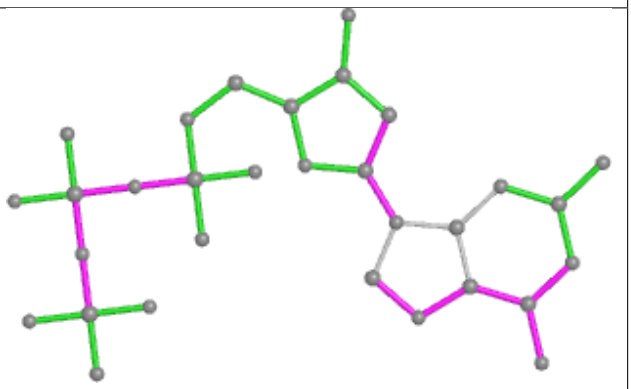


Rings

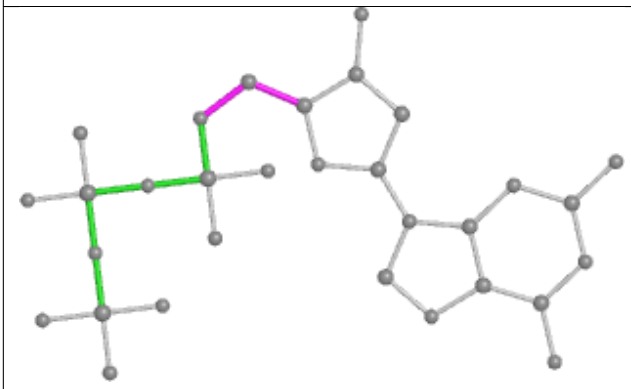
Ligand DGT B 700



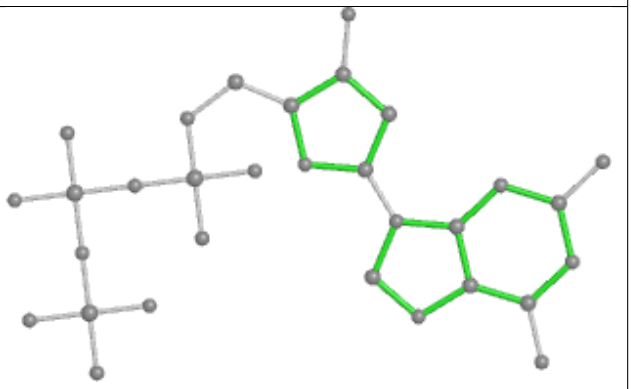
Bond lengths



Bond angles

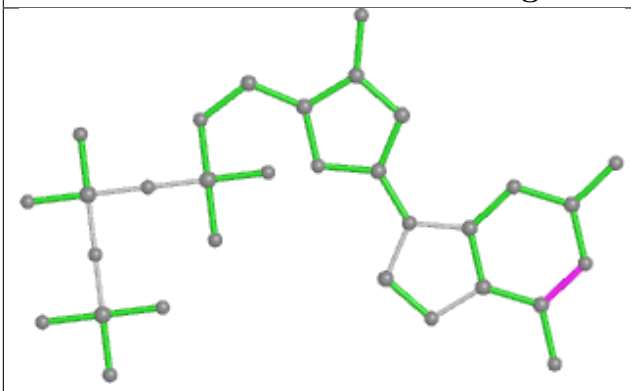


Torsions

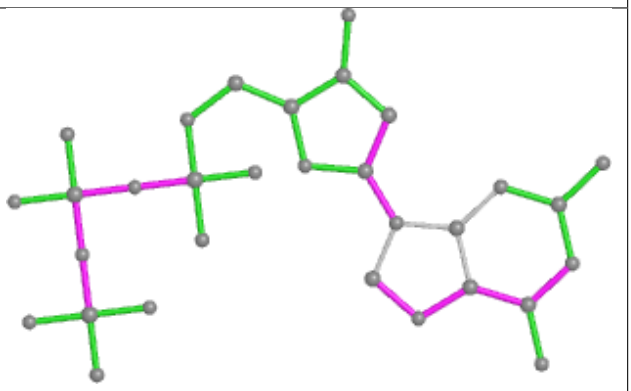


Rings

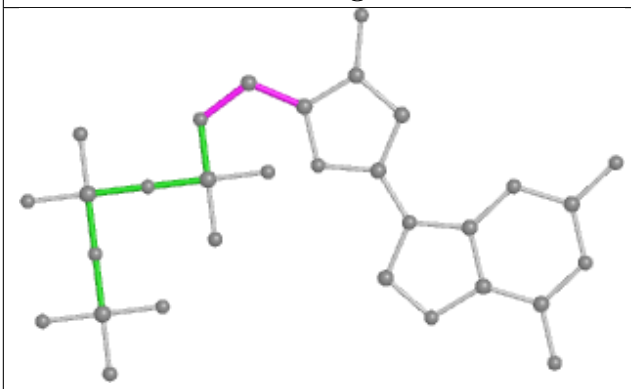
Ligand DGT C 700



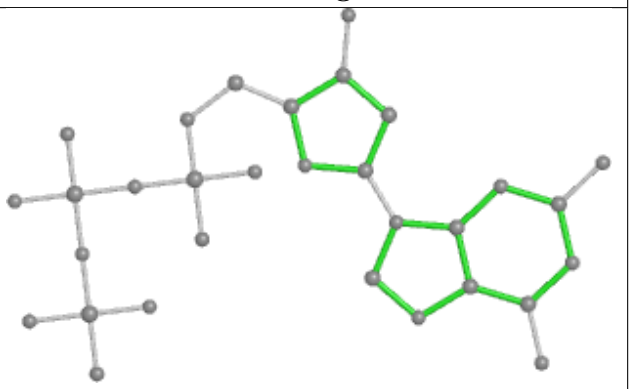
Bond lengths



Bond angles

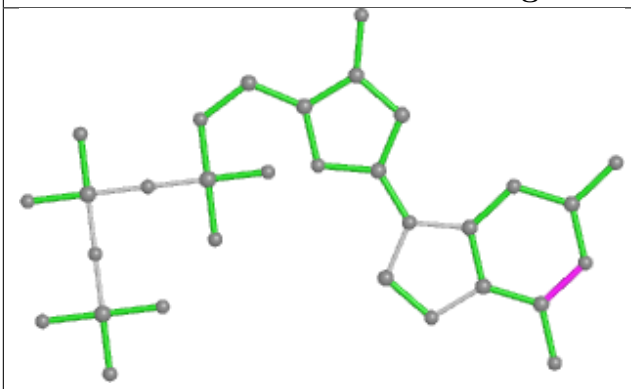


Torsions

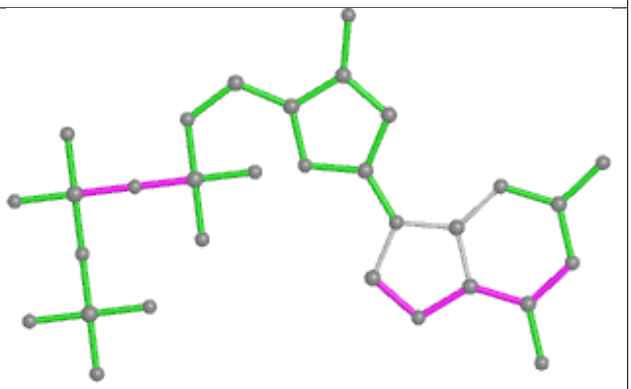


Rings

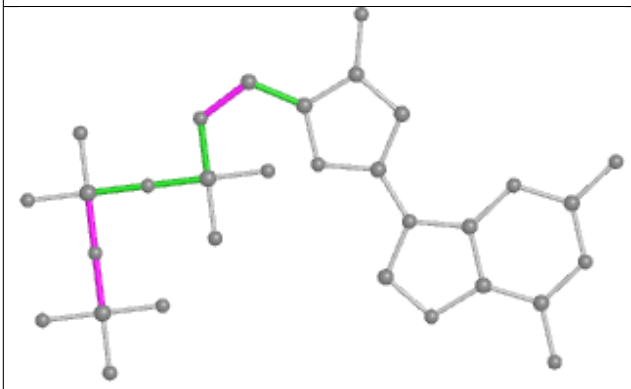
Ligand DGT B 900



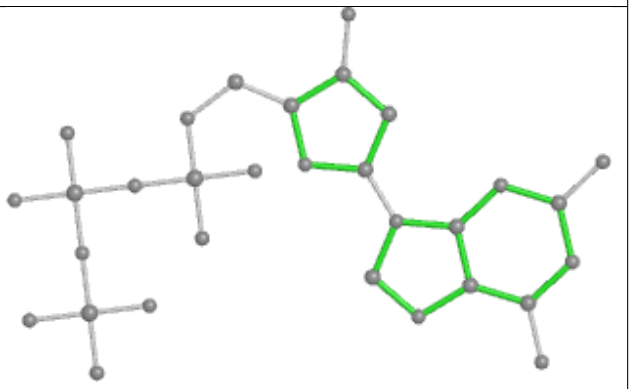
Bond lengths



Bond angles

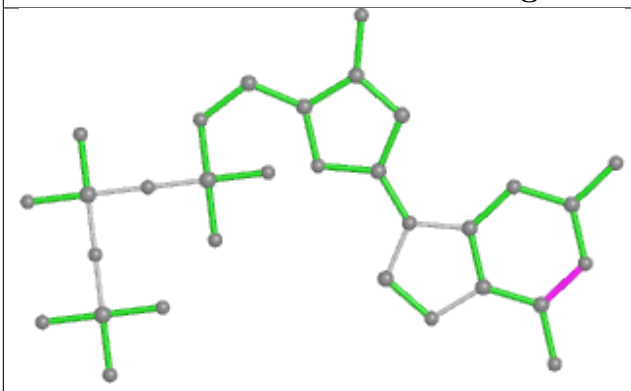


Torsions

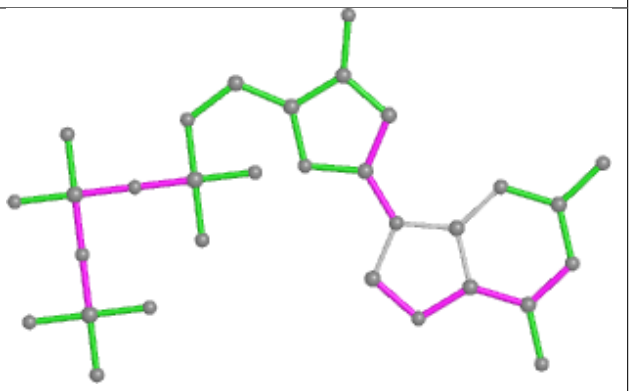


Rings

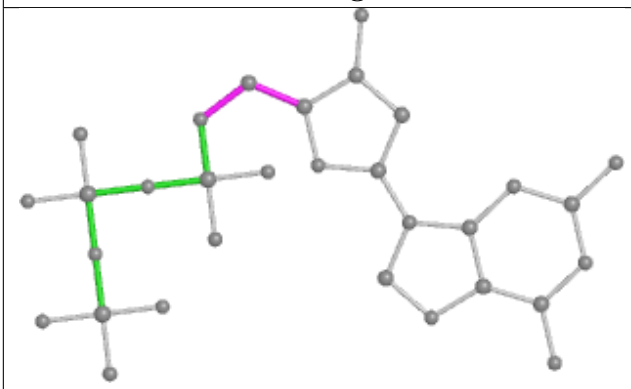
Ligand DGT A 700



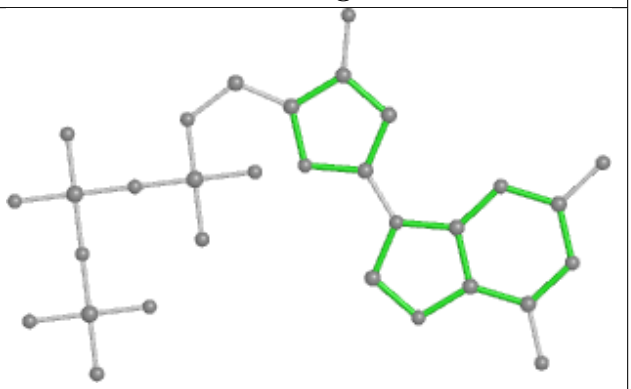
Bond lengths



Bond angles

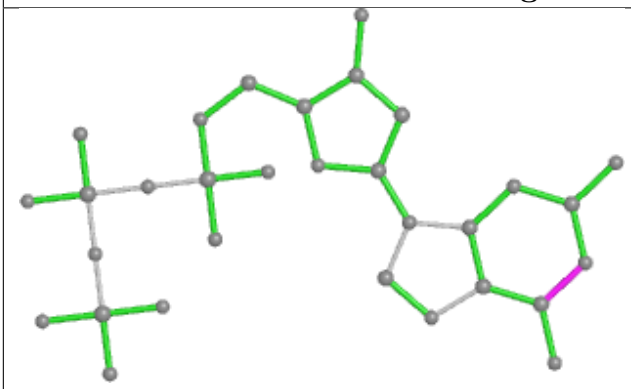


Torsions

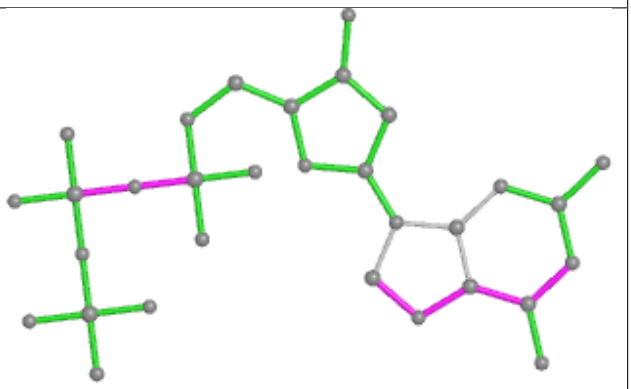


Rings

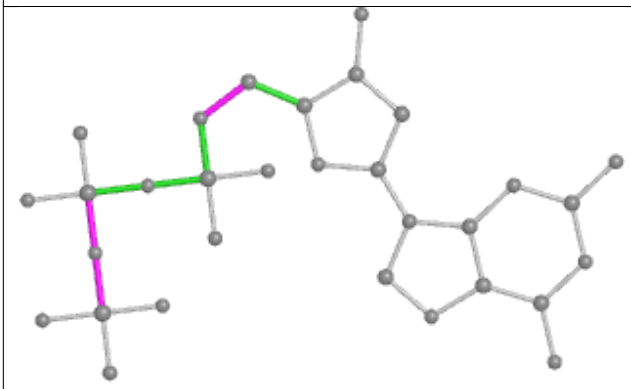
Ligand DGT D 900



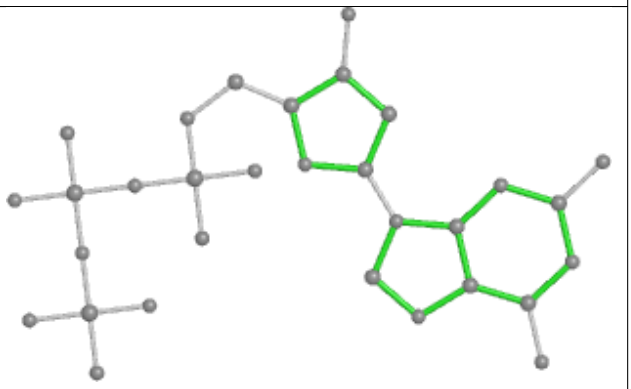
Bond lengths



Bond angles

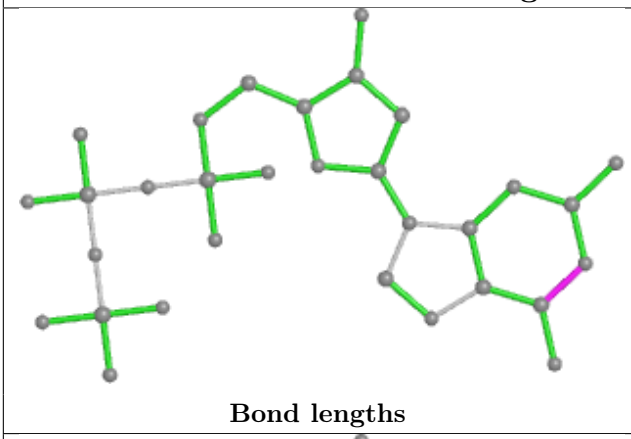


Torsions

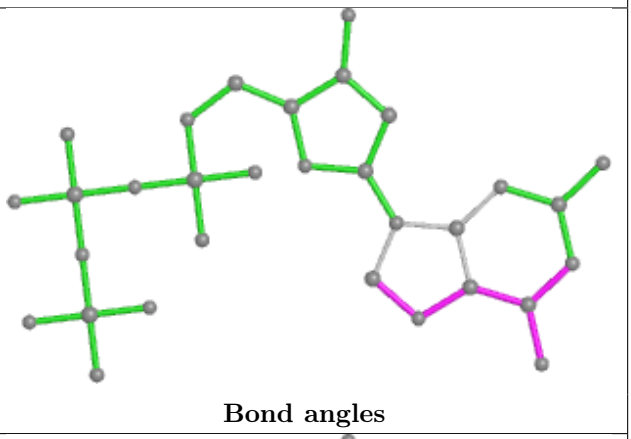


Rings

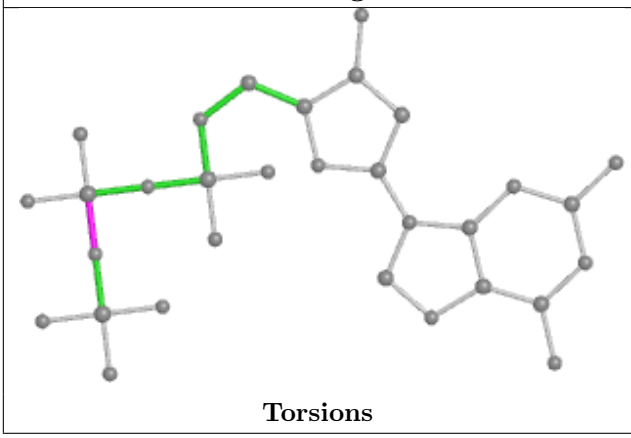
Ligand DGT D 800



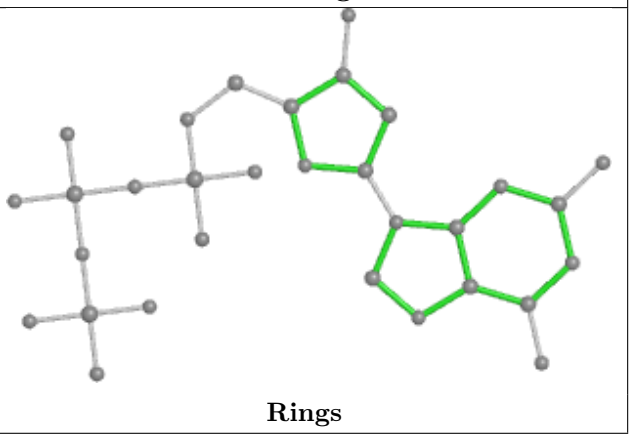
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	481/550 (87%)	0.28	34 (7%)	16	14	24, 41, 68, 89	0
1	B	481/550 (87%)	0.07	23 (4%)	30	28	20, 34, 61, 78	0
1	C	481/550 (87%)	0.35	39 (8%)	12	11	24, 42, 69, 94	0
1	D	481/550 (87%)	0.05	22 (4%)	32	29	19, 34, 58, 75	0
All	All	1924/2200 (87%)	0.19	118 (6%)	21	19	19, 38, 64, 94	0

The worst 5 of 118 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	488	LEU	8.3
1	A	466	ILE	7.0
1	D	465	GLN	6.6
1	C	488	LEU	6.3
1	B	488	LEU	6.1

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

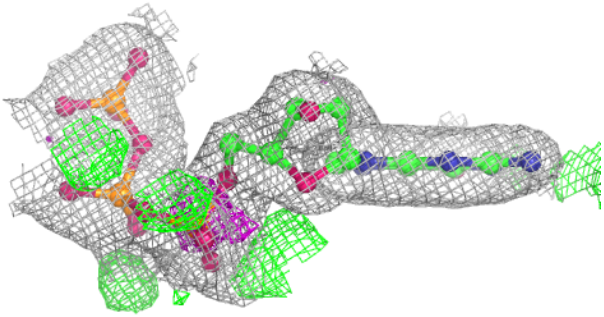
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

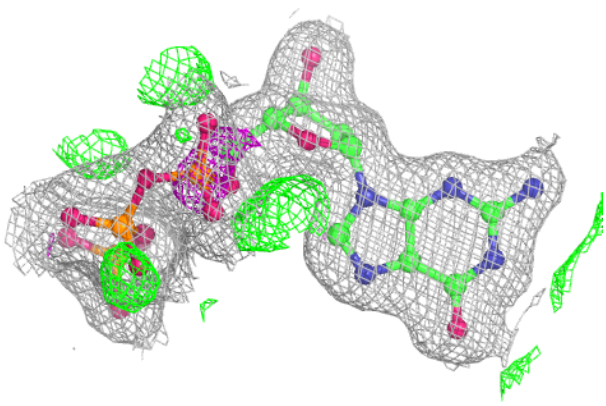
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	DGT	A	700	31/31	0.94	0.07	29,35,43,44	0
2	DGT	C	700	31/31	0.95	0.09	30,36,43,45	0
2	DGT	D	700	31/31	0.95	0.09	25,28,35,36	0
2	DGT	B	700	31/31	0.96	0.07	23,28,33,34	0
3	MG	A	750	1/1	0.96	0.12	42,42,42,42	0
3	MG	C	850	1/1	0.96	0.03	27,27,27,27	0
3	MG	D	850	1/1	0.96	0.03	24,24,24,24	0
2	DGT	B	900	31/31	0.97	0.08	24,25,29,31	0
2	DGT	D	800	31/31	0.97	0.11	22,22,24,24	0
2	DGT	A	800	31/31	0.97	0.10	24,27,29,30	0
3	MG	A	850	1/1	0.97	0.02	28,28,28,28	0
2	DGT	C	800	31/31	0.97	0.11	25,28,31,31	0
3	MG	D	750	1/1	0.97	0.08	32,32,32,32	0
2	DGT	C	900	31/31	0.97	0.10	28,29,36,38	0
3	MG	B	850	1/1	0.98	0.03	24,24,24,24	0
3	MG	C	750	1/1	0.98	0.03	44,44,44,44	0
2	DGT	A	900	31/31	0.98	0.09	26,27,33,35	0
2	DGT	B	800	31/31	0.98	0.11	21,22,26,27	0
3	MG	B	750	1/1	0.98	0.03	30,30,30,30	0
2	DGT	D	900	31/31	0.99	0.10	23,24,29,30	0

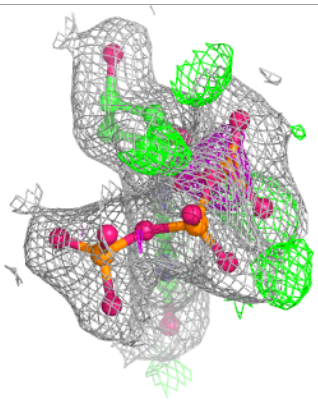
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around DGT A 700:

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

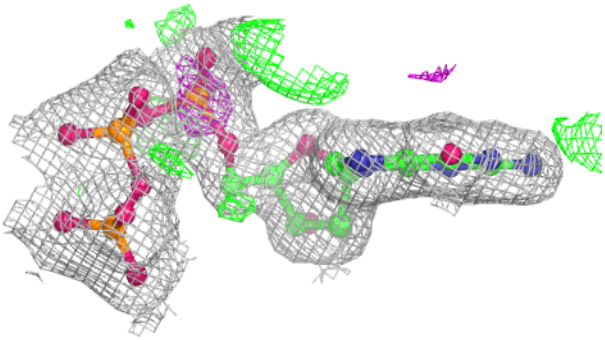


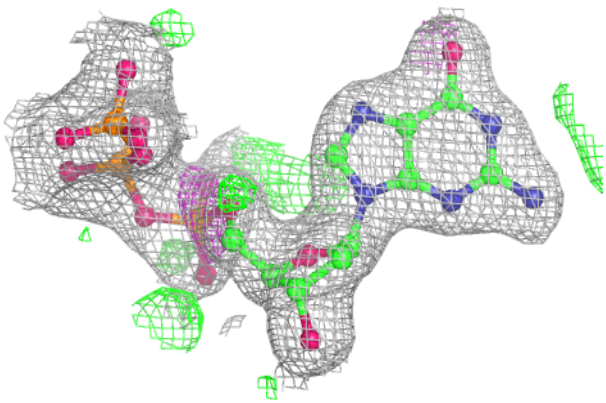


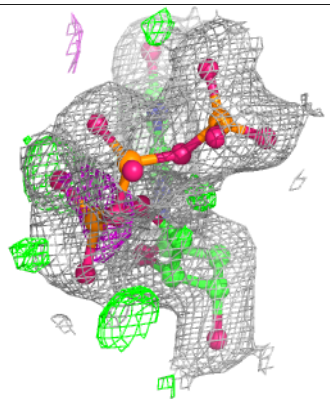


Electron density around DGT C 700:

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

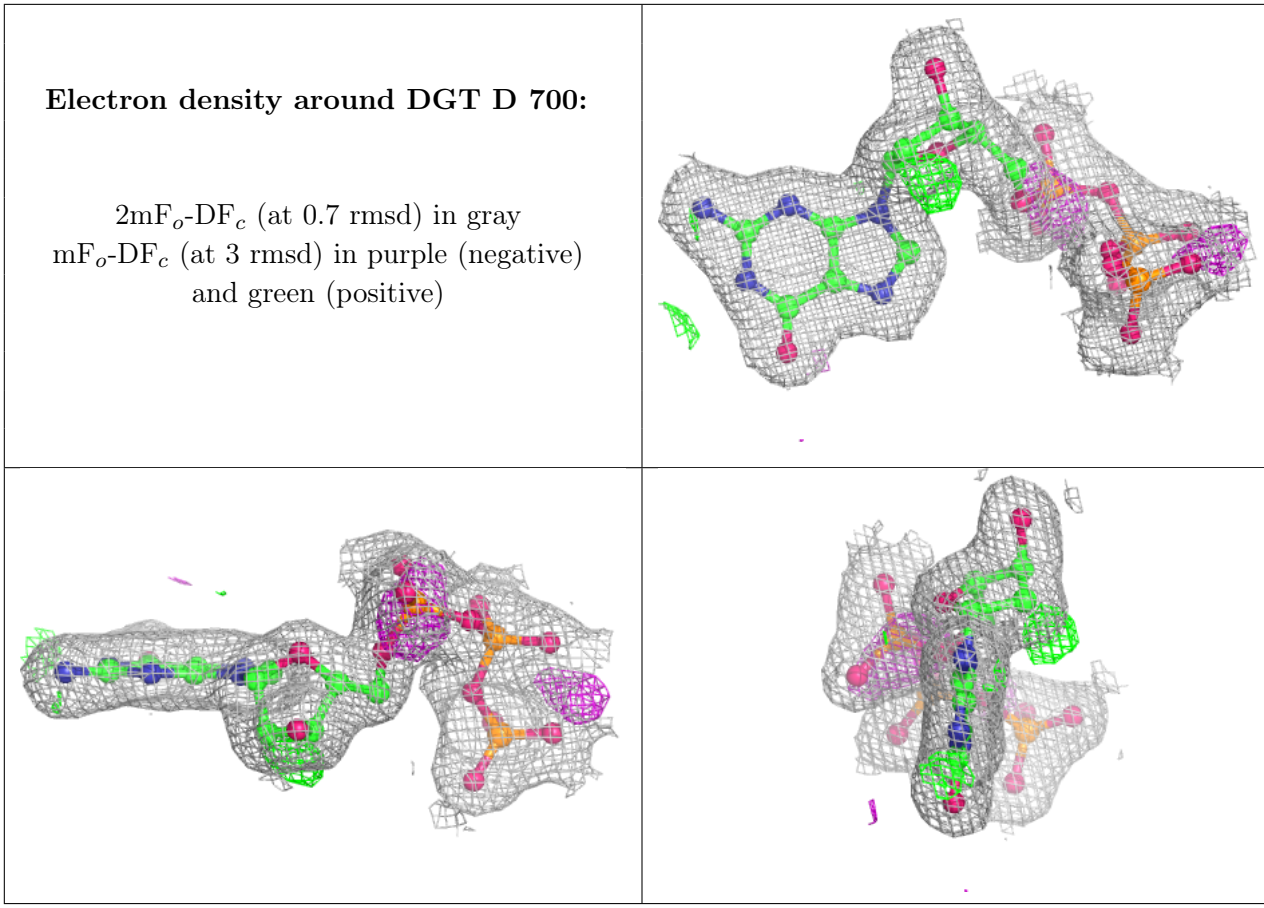






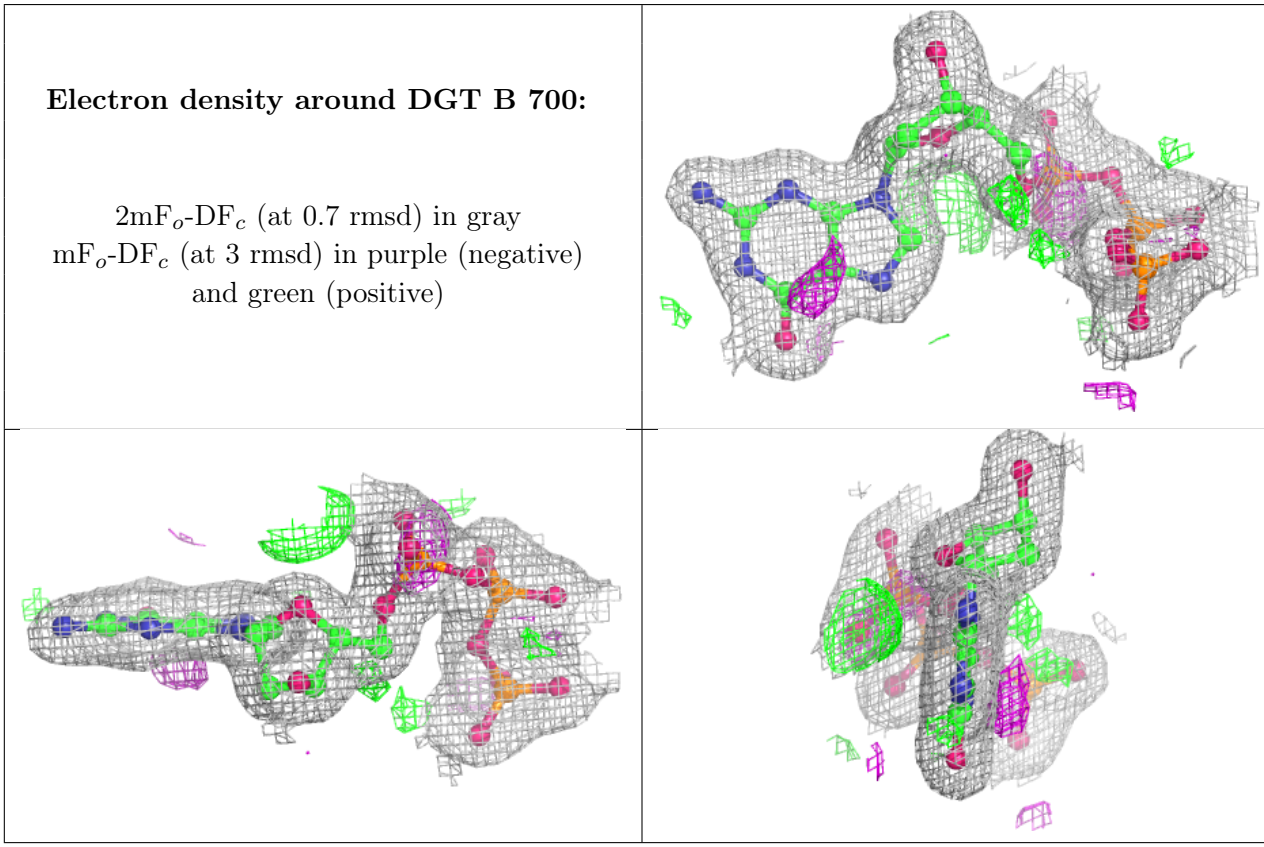
Electron density around DGT D 700:

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



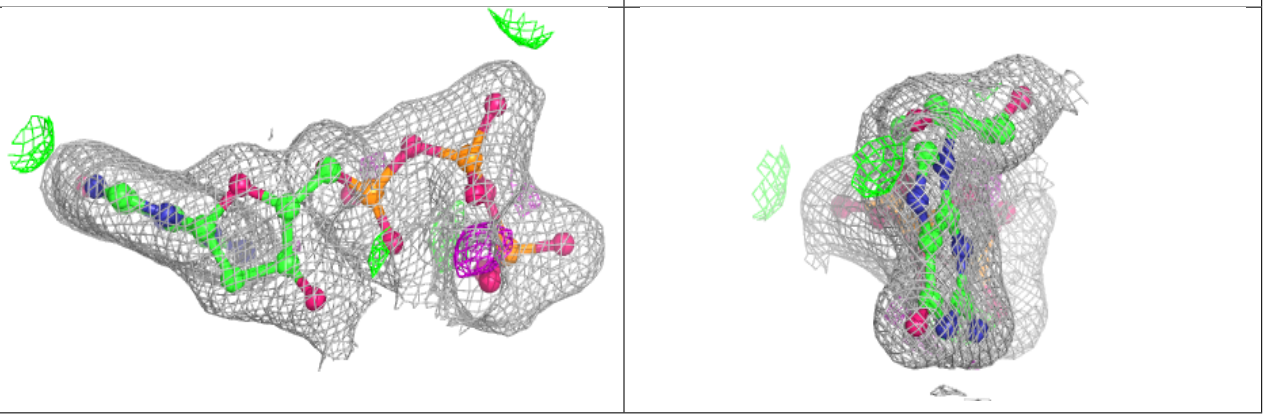
Electron density around DGT B 700:

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



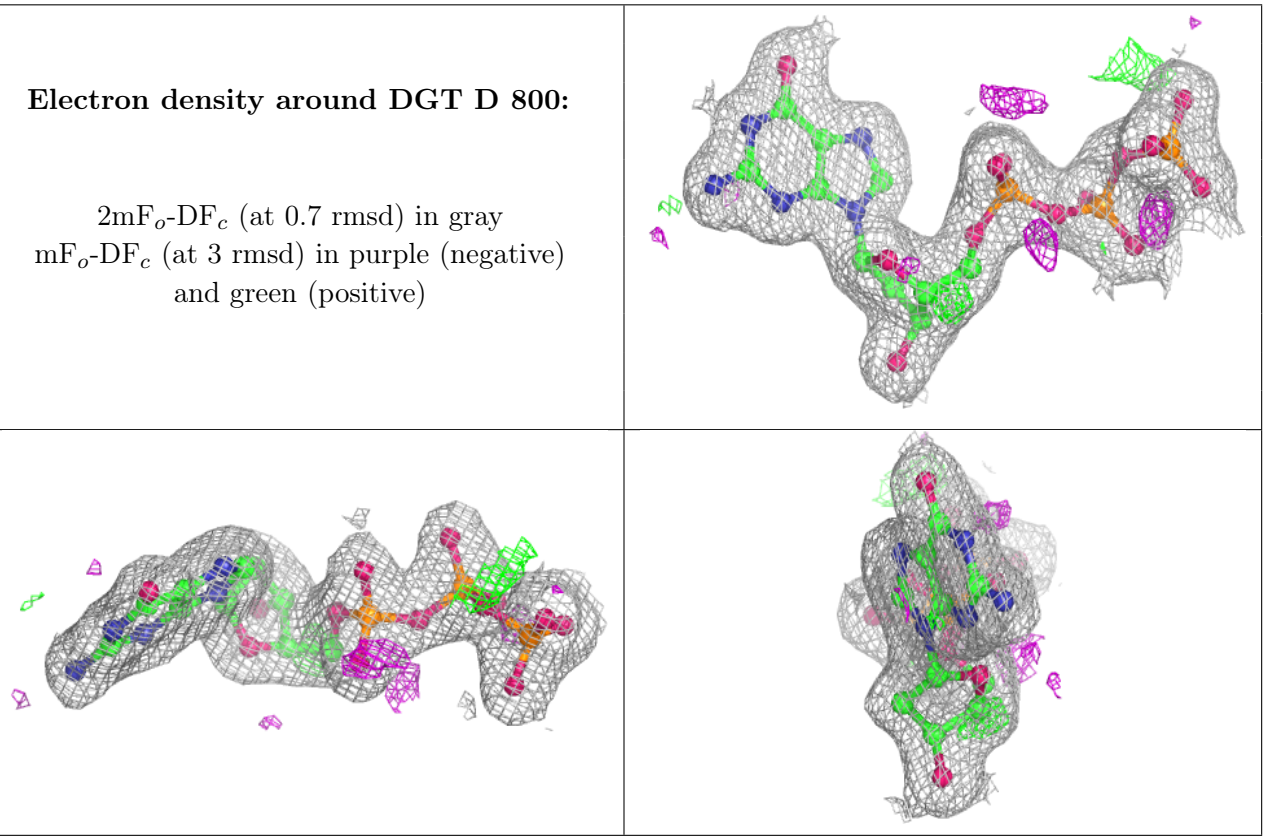
Electron density around DGT B 900:

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



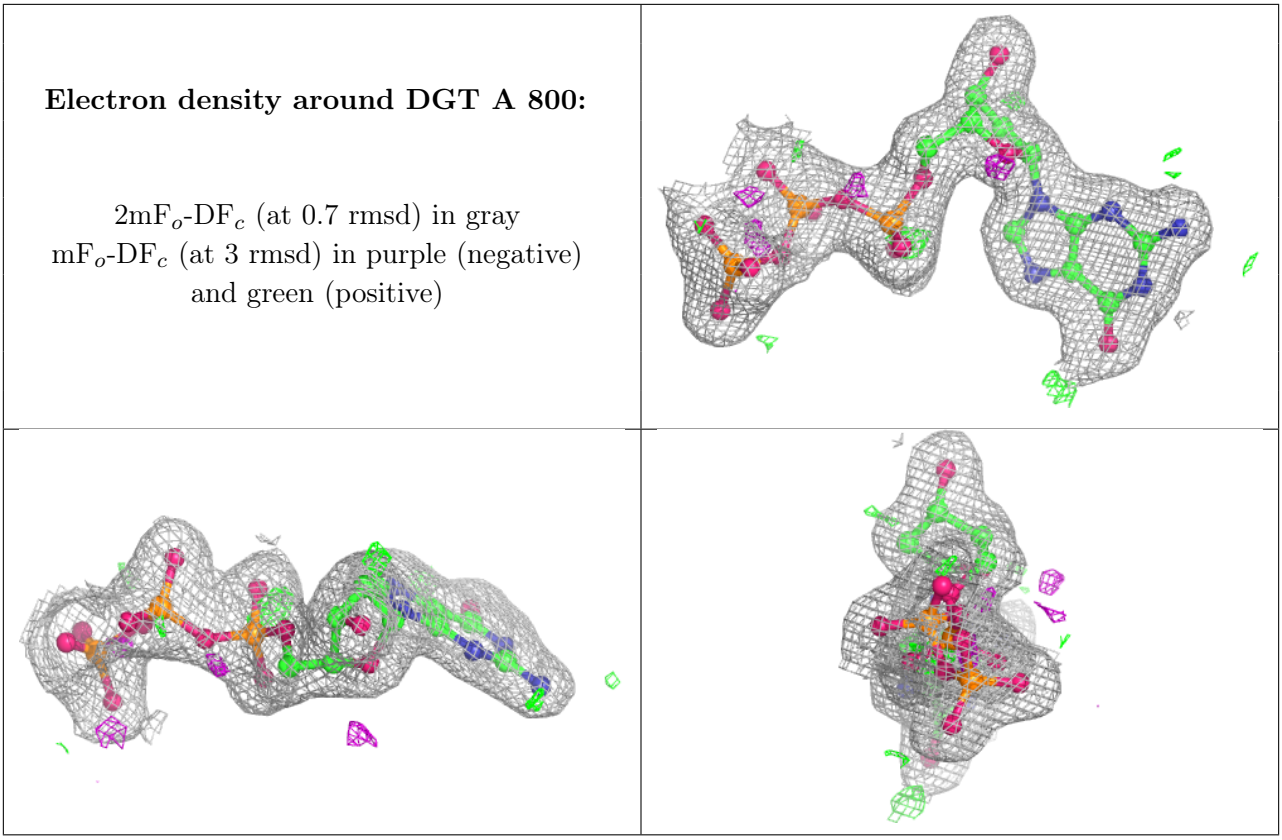
Electron density around DGT D 800:

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



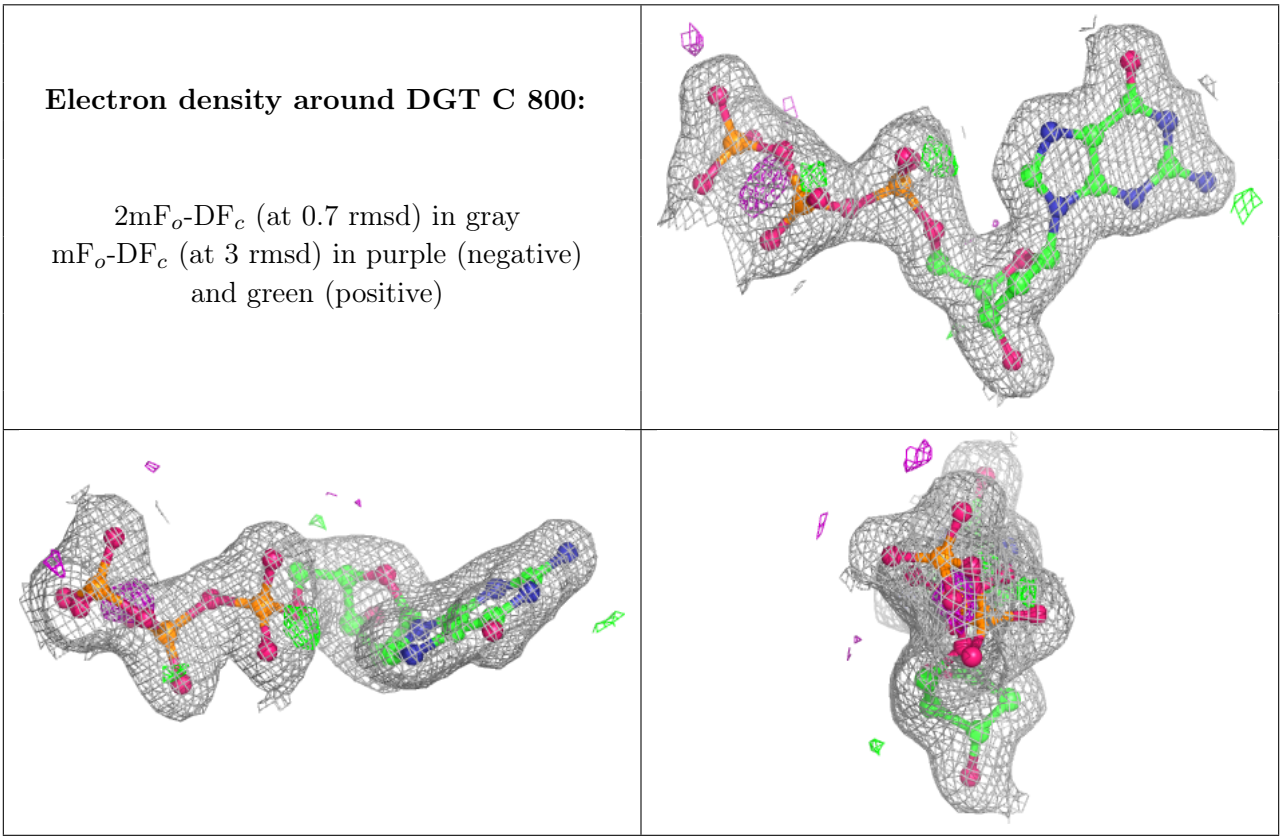
Electron density around DGT A 800:

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



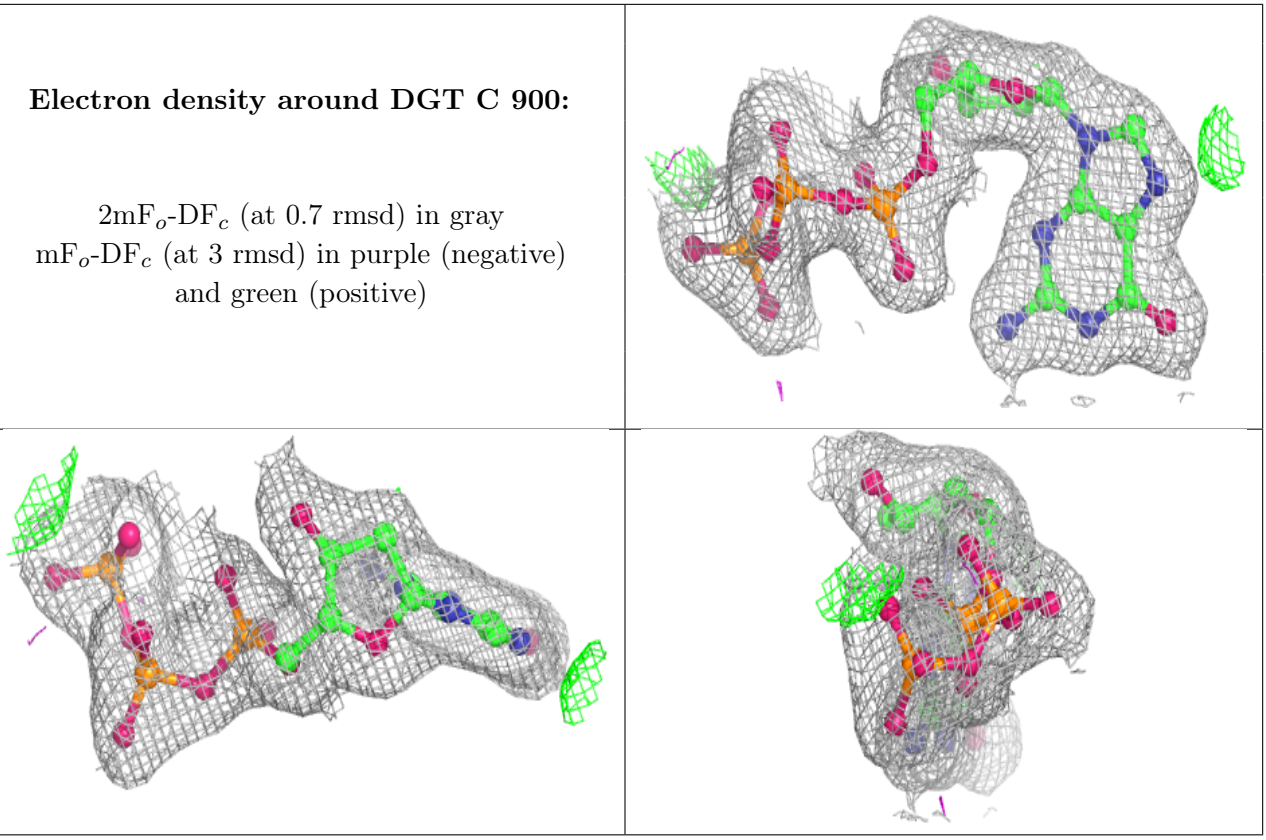
Electron density around DGT C 800:

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



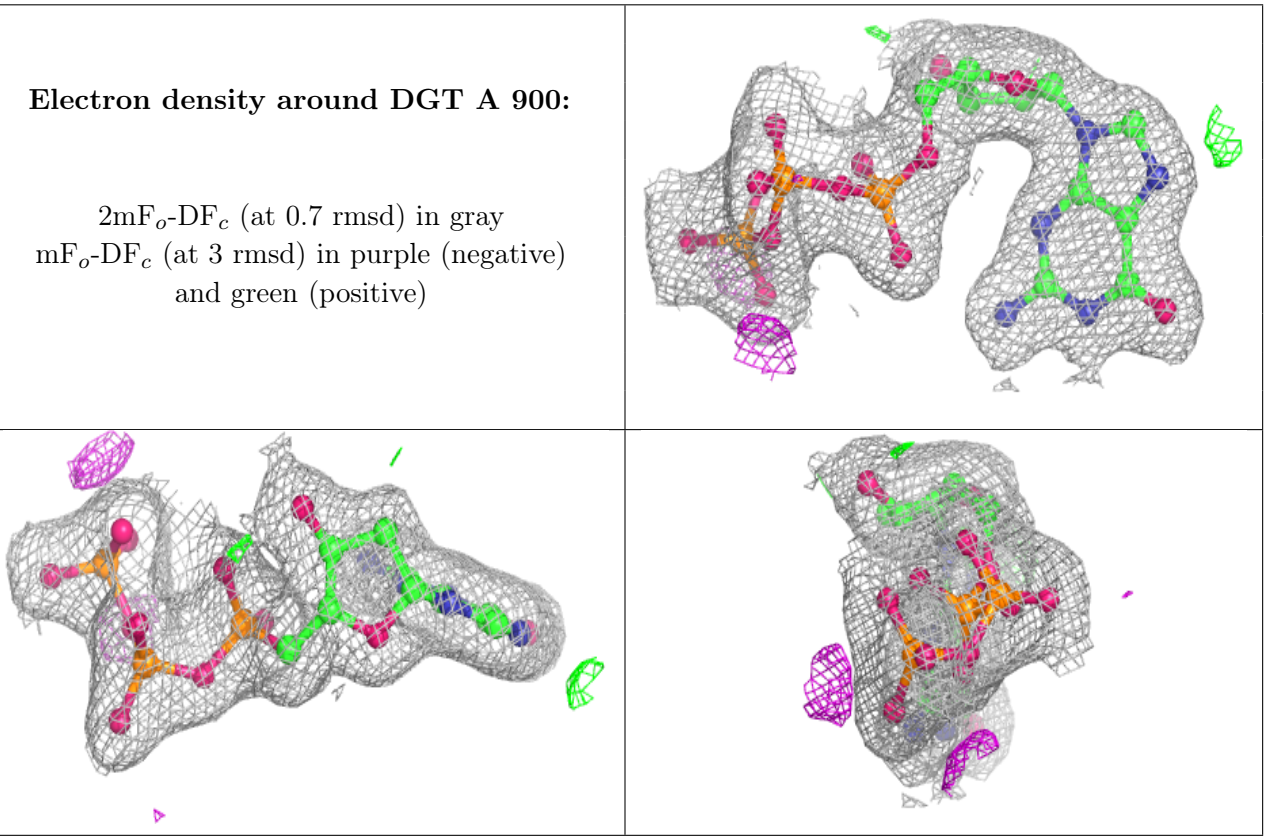
Electron density around DGT C 900:

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



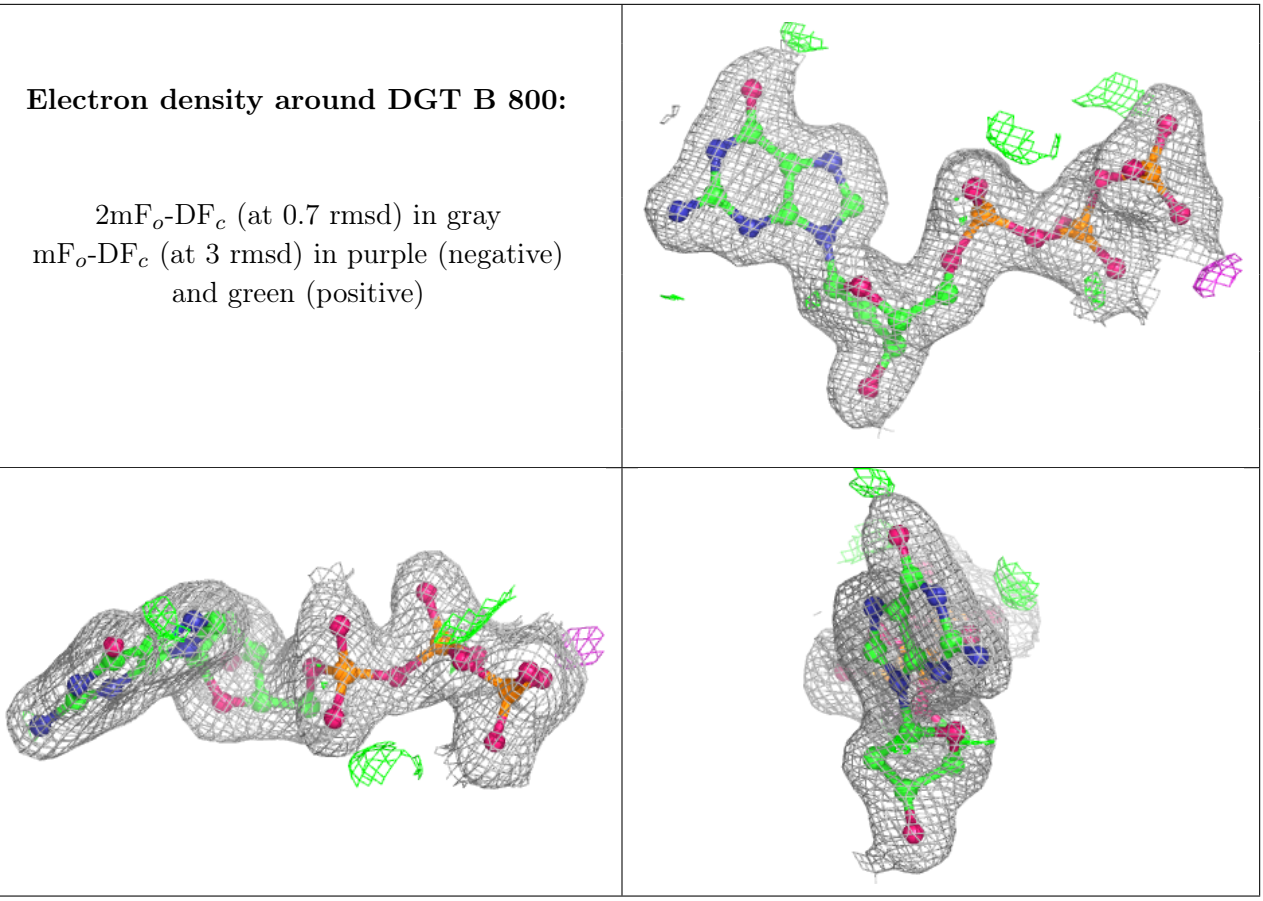
Electron density around DGT A 900:

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



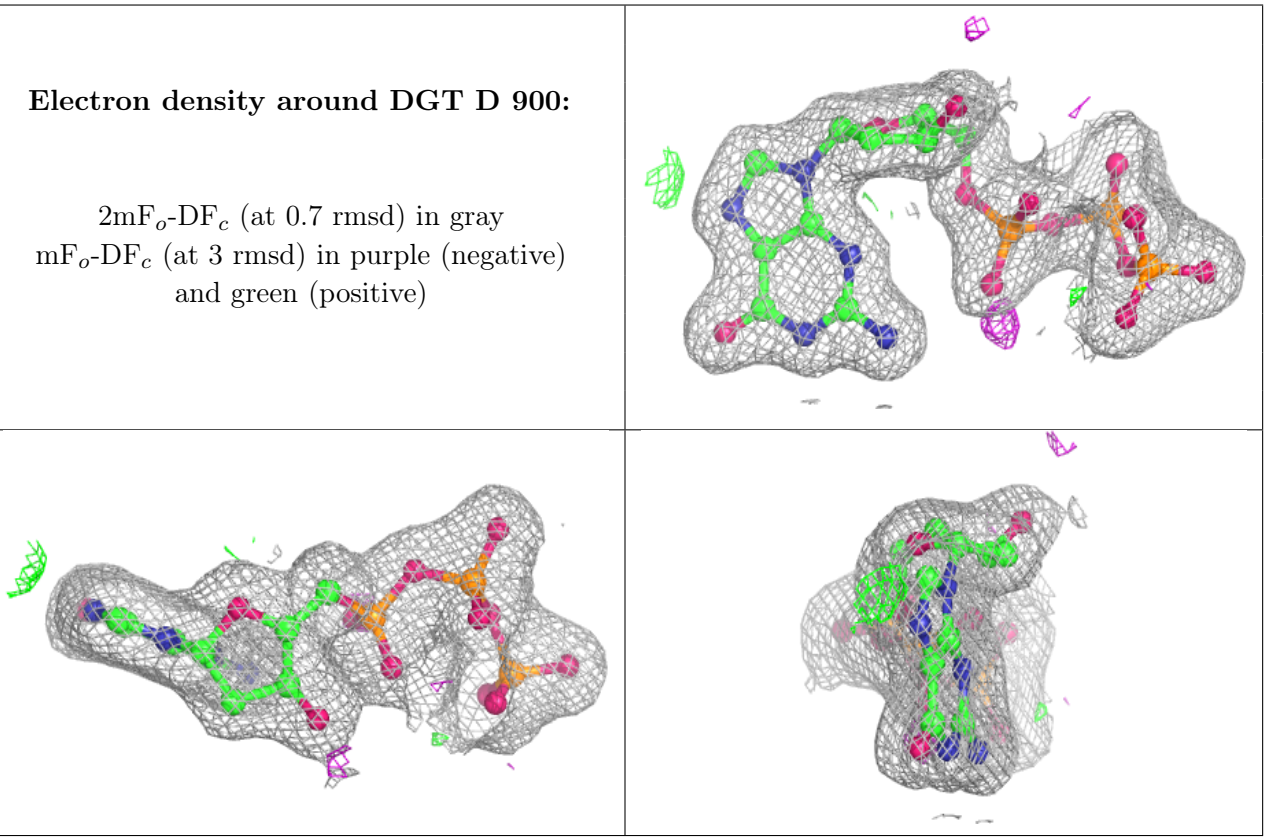
Electron density around DGT B 800:

$2mF_o$-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 D 900:

$2mF_o$-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.