Full wwPDB X-ray Structure Validation Report (i)

Dec 3, 2023 – 05:30 am GMT

PDB ID	:	2C7R
Title	:	HhaI DNA methyltransferase (T250G mutant) complex with oligonucleotide containing 2-aminopurine as a target base (GPGC:GMGC) and SAH
Authors	:	Daujotyte, D.; Grazulis, S.
Deposited on	:	2005-11-27
Resolution	:	1.90 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (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 (1)) were used in the production of this report:

MolProbity	:	FAILED
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
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 1.90 Å.

There are no overall percentile quality scores available for this entry.

MolProbity failed to run properly; EDS was not executed - the sequence quality summary graphics cannot be shown.

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 3471 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 MODIFICATION METHYLASE HHAI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	327	2612	1666	446	487	13	0	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	250	GLY	THR	engineered mutation	UNP P05102

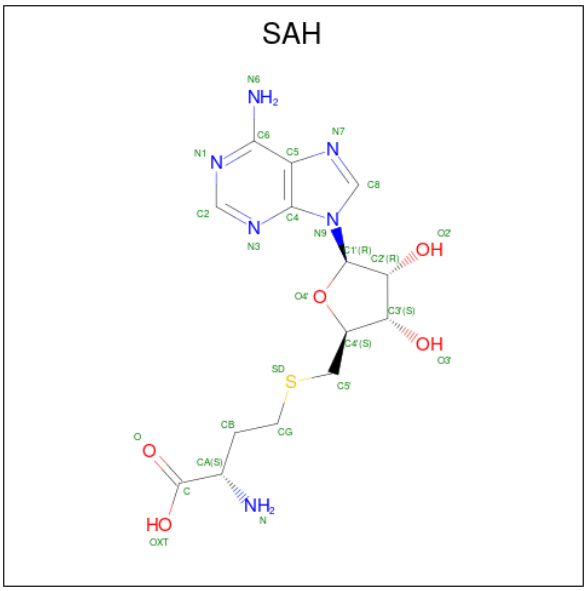
- Molecule 2 is a DNA chain called 5'-D(*G*GP*AP*TP*GP*(5CM)*GP*CP*TP*GP*AP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	12	251	118	48	73	12	0	0	0

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

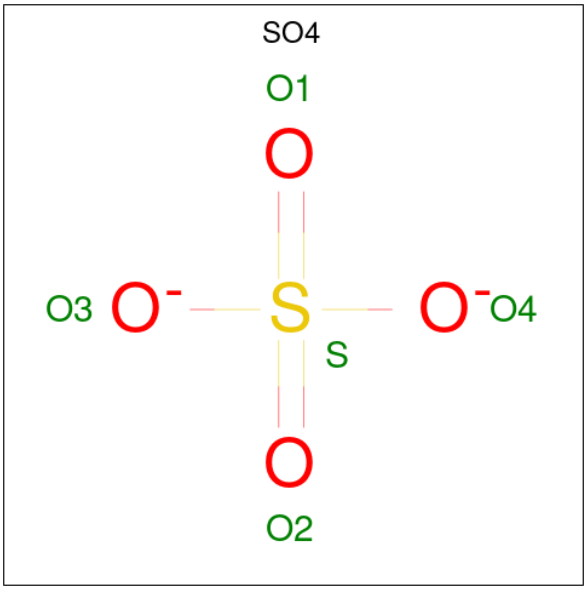
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	12	246	116	46	72	12	0	0	0

- Molecule 4 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: $C_{14}H_{20}N_6O_5S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
4	A	1	26	14	6	5	1	0	0

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



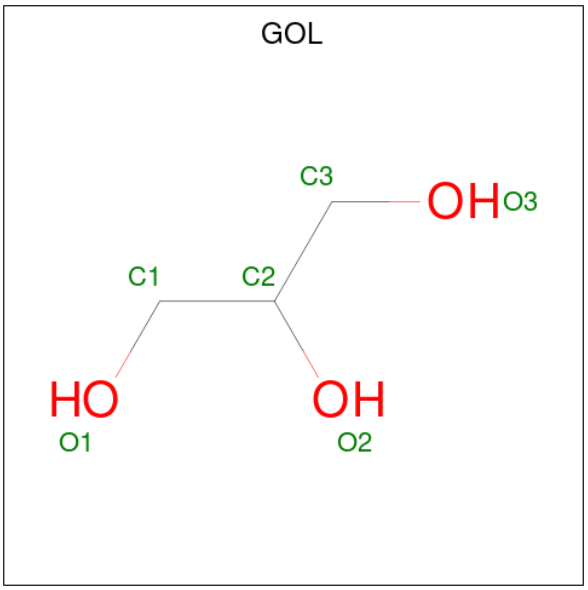
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	O	S			
5	A	1	5	4	1		0	0
5	A	1	Total 5	O 4	S 1		0	0
5	A	1	Total 5	O 4	S 1		0	0

Continued on next page...

Continued from previous page...

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

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	229	229	229	0	0
			Total	O		
7	C	21	21	21	0	0
			Total	O		
7	D	21	21	21	0	0

MolProbity failed to run properly; EDS was not executed - this section is therefore empty.

3 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	95.62Å 95.62Å 315.76Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.50 – 1.90	Depositor
% Data completeness (in resolution range)	99.8 (29.50-1.90)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.199 , 0.217	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3471	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

4 Model quality (i)

4.1 Standard geometry (i)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts (i)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles (i)

4.3.1 Protein backbone (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA (i)

MolProbity failed to run properly - this section is therefore empty.

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

2 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	5CM	C	407	3,2	17,21,22	0.74	1 (5%)	24,30,33	1.63	4 (16%)
3	2PR	D	427	3	17,23,24	1.11	2 (11%)	16,33,36	1.74	3 (18%)

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	5CM	C	407	3,2	-	0/7/21/22	0/2/2/2
3	2PR	D	427	3	-	1/3/21/22	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	427	2PR	C6-N1	3.39	1.39	1.32
2	C	407	5CM	C4-N3	2.48	1.38	1.34
3	D	427	2PR	C8-N7	-2.20	1.30	1.34

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	407	5CM	C5-C4-N3	-5.22	116.04	121.67
3	D	427	2PR	C6-N1-C2	4.81	125.33	116.43
2	C	407	5CM	C4-N3-C2	4.50	126.77	120.69
3	D	427	2PR	C2-N3-C4	-3.02	111.91	115.36
2	C	407	5CM	N1-C2-N3	-2.34	114.54	118.81
3	D	427	2PR	C2'-C1'-N9	-2.31	108.95	114.27
2	C	407	5CM	O2-C2-N1	2.00	123.03	118.89

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	427	2PR	C4'-C5'-O5'-P

There are no ring outliers.

No monomer is involved in short contacts.

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

13 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	GOL	A	1340	-	5,5,5	0.92	0	5,5,5	0.40	0
5	SO4	A	1332	-	4,4,4	1.87	2 (50%)	6,6,6	0.90	0
4	SAH	A	1328	-	24,28,28	1.27	4 (16%)	25,40,40	2.41	4 (16%)
6	GOL	A	1338	-	5,5,5	0.93	0	5,5,5	0.41	0
6	GOL	A	1339	-	5,5,5	0.92	0	5,5,5	0.40	0
5	SO4	A	1331	-	4,4,4	1.90	2 (50%)	6,6,6	0.87	0
6	GOL	A	1336	-	5,5,5	0.92	0	5,5,5	0.40	0
5	SO4	A	1334	-	4,4,4	1.86	2 (50%)	6,6,6	0.90	0
5	SO4	A	1333	-	4,4,4	1.88	2 (50%)	6,6,6	0.89	0
6	GOL	A	1337	-	5,5,5	0.90	0	5,5,5	0.40	0
5	SO4	A	1335	-	4,4,4	1.93	2 (50%)	6,6,6	0.97	0
5	SO4	A	1330	-	4,4,4	1.88	2 (50%)	6,6,6	0.92	0
5	SO4	A	1329	-	4,4,4	1.82	2 (50%)	6,6,6	0.89	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	GOL	A	1340	-	-	4/4/4/4	-
6	GOL	A	1338	-	-	2/4/4/4	-
4	SAH	A	1328	-	-	3/11/31/31	0/3/3/3
6	GOL	A	1339	-	-	2/4/4/4	-
6	GOL	A	1336	-	-	4/4/4/4	-
6	GOL	A	1337	-	-	2/4/4/4	-

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1328	SAH	C4-N3	3.26	1.40	1.35
5	A	1331	SO4	O1-S	3.16	1.63	1.46
5	A	1335	SO4	O1-S	3.16	1.63	1.46
5	A	1330	SO4	O1-S	3.12	1.62	1.46
5	A	1333	SO4	O1-S	3.10	1.62	1.46
5	A	1332	SO4	O1-S	3.09	1.62	1.46
5	A	1334	SO4	O1-S	3.08	1.62	1.46
5	A	1329	SO4	O1-S	2.96	1.62	1.46
4	A	1328	SAH	C8-N7	-2.52	1.30	1.34
4	A	1328	SAH	C2'-C1'	-2.50	1.50	1.53
5	A	1335	SO4	O3-S	-2.19	1.29	1.47
5	A	1333	SO4	O3-S	-2.09	1.30	1.47
5	A	1329	SO4	O3-S	-2.08	1.30	1.47
5	A	1331	SO4	O3-S	-2.07	1.30	1.47
5	A	1332	SO4	O3-S	-2.07	1.30	1.47
4	A	1328	SAH	O-C	2.06	1.28	1.22
5	A	1330	SO4	O3-S	-2.06	1.31	1.47
5	A	1334	SO4	O3-S	-2.05	1.31	1.47

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1328	SAH	OXT-C-O	-9.40	102.76	124.09
4	A	1328	SAH	O4'-C1'-C2'	-5.11	99.45	106.93
4	A	1328	SAH	C5-C6-N6	2.96	124.85	120.35
4	A	1328	SAH	CB-CA-N	2.85	117.63	110.17

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	1336	GOL	C1-C2-C3-O3
6	A	1337	GOL	O1-C1-C2-C3
6	A	1338	GOL	O1-C1-C2-C3
6	A	1340	GOL	C1-C2-C3-O3
6	A	1340	GOL	O2-C2-C3-O3
4	A	1328	SAH	OXT-C-CA-CB
6	A	1338	GOL	O1-C1-C2-O2
6	A	1339	GOL	O1-C1-C2-C3
6	A	1340	GOL	O1-C1-C2-C3
6	A	1336	GOL	O2-C2-C3-O3
6	A	1337	GOL	O1-C1-C2-O2

Continued on next page...

Continued from previous page...

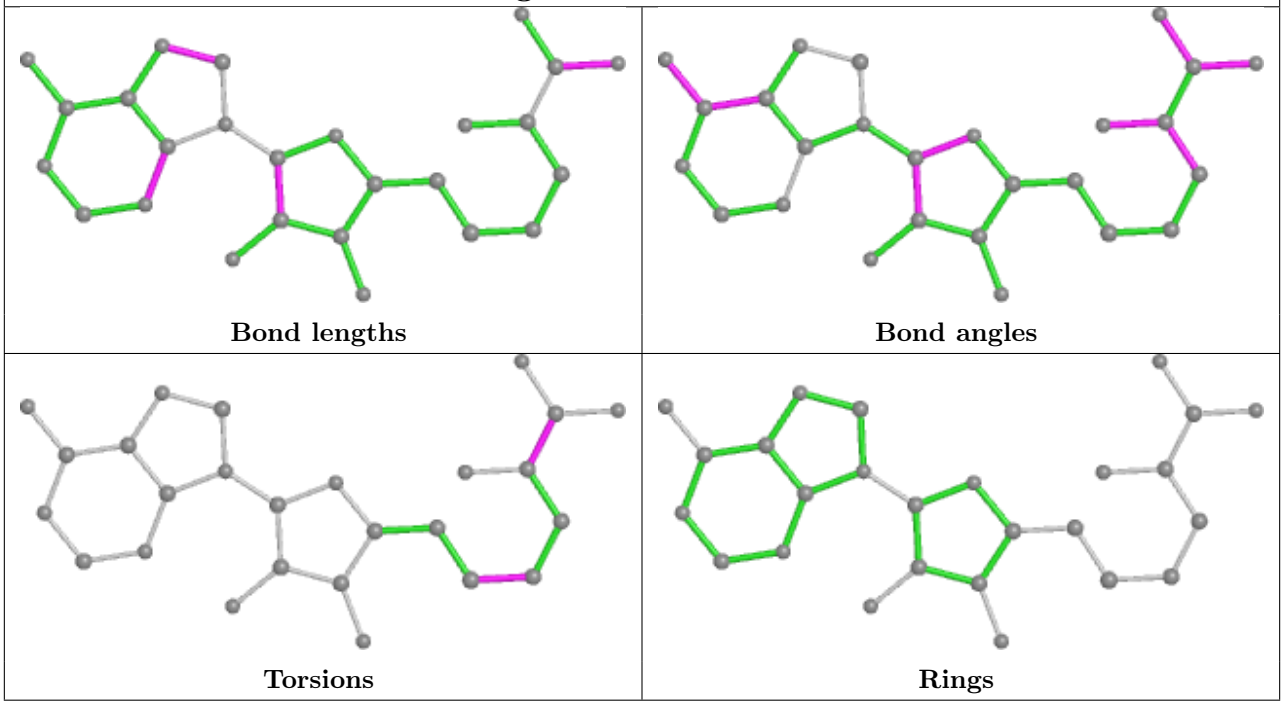
Mol	Chain	Res	Type	Atoms
6	A	1339	GOL	O1-C1-C2-O2
6	A	1340	GOL	O1-C1-C2-O2
4	A	1328	SAH	OXT-C-CA-N
6	A	1336	GOL	O1-C1-C2-O2
6	A	1336	GOL	O1-C1-C2-C3
4	A	1328	SAH	CB-CG-SD-C5'

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 SAH A 1328



4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	221:GLN	C	222:THR	N	1.16
1	A	220:GLU	C	221:GLN	N	1.09

5 Fit of model and data (i)

5.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

5.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

5.4 Ligands (i)

EDS was not executed - this section is therefore empty.

5.5 Other polymers (i)

EDS was not executed - this section is therefore empty.