



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 16, 2018 – 07:11 am GMT

PDB ID : 10MH  
Title : TERNARY STRUCTURE OF HHAI METHYLTRANSFERASE WITH ADOHCY AND HEMIMETHYLATED DNA CONTAINING 5,6-DIHYDR O-5-AZACYTOSINE AT THE TARGET  
Authors : Sheikhnejad, G.; Brank, A.; Christman, J.K.; Goddard, A.; Alvarez, E.; Ford Junior, H.; Marquez, V.E.; Marasco, C.J.; Sufirin, J.R.; O’Gara, M.; Cheng, X.  
Deposited on : 1998-08-10  
Resolution : 2.55 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

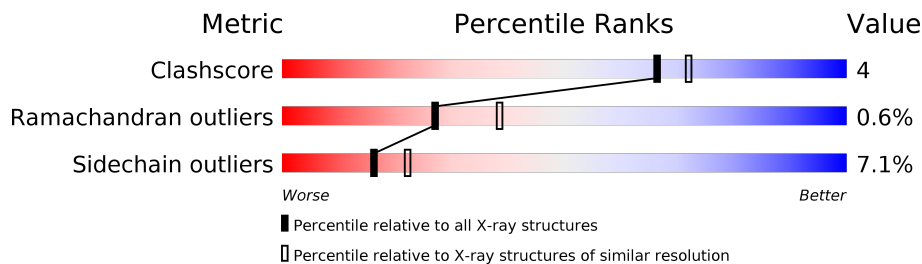
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 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))
Clashscore	122078	1097 (2.56-2.52)
Ramachandran outliers	120005	1087 (2.56-2.52)
Sidechain outliers	119972	1087 (2.56-2.52)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	B	12	
2	C	12	
3	A	327	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 3254 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 DNA chain called DNA (5'-D(P\*CP\*CP\*AP\*TP\*GP\*(5CM)P\*GP\*CP\*TP\*GP\*AP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	B	12	244	116	44	72	12	0	0	0

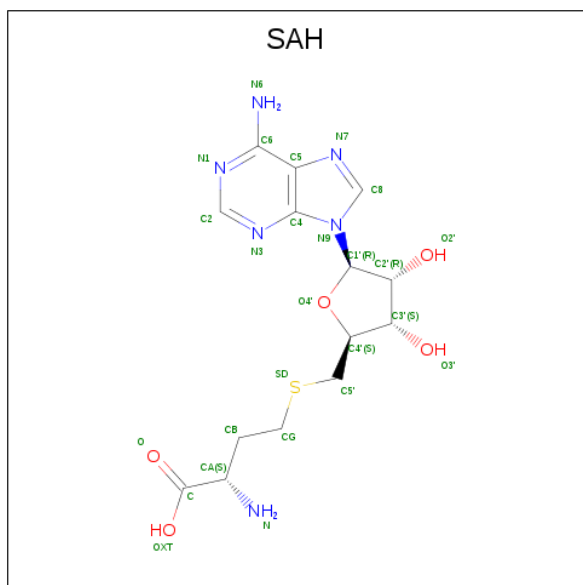
- Molecule 2 is a DNA chain called DNA (5'-D(P\*GP\*TP\*CP\*AP\*GP\*5NCP\*GP\*CP\*AP\*TP\*GP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	12	249	116	49	72	12	0	0	0

- Molecule 3 is a protein called PROTEIN (CYTOSINE-SPECIFIC METHYLTRANSFERASE HHAI).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	327	2606	1662	444	487	13	0	0	0

- Molecule 4 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: C<sub>14</sub>H<sub>20</sub>N<sub>6</sub>O<sub>5</sub>S).



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

- Molecule 5 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	6	Total	O	0	0
			6	6		
5	C	9	Total	O	0	0
			9	9		
5	A	114	Total	O	0	0
			114	114		

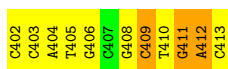
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.


Note EDS was not executed.

- Molecule 1: DNA (5'-D(P\*CP\*CP\*AP\*TP\*GP\*(5CM)P\*GP\*CP\*TP\*GP\*AP\*C)-3')

Chain B: 




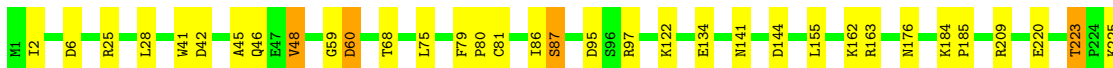
- Molecule 2: DNA (5'-D(P\*GP\*TP\*CP\*AP\*GP\*5NCP\*GP\*CP\*AP\*TP\*GP\*G)-3')

Chain C: 



- Molecule 3: PROTEIN (CYTOSINE-SPECIFIC METHYLTRANSFERASE HHAI)

Chain A: 



## 4 Data and refinement statistics

Xtrriage (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, $\alpha$ , $\beta$ , $\gamma$	99.86Å 99.86Å 325.20Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.30 – 2.55	Depositor
% Data completeness (in resolution range)	76.6 (29.30-2.55)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	0.205 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	3254	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

## 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: SAH, 5CM, 5NC

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	B	3.45	26/249 (10.4%)	4.66	85/379 (22.4%)
2	C	3.70	32/257 (12.5%)	3.94	62/393 (15.8%)
3	A	0.60	1/2661 (0.0%)	0.79	1/3586 (0.0%)
All	All	1.53	59/3167 (1.9%)	1.95	148/4358 (3.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1
2	C	0	3
3	A	0	1
All	All	0	5

All (59) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	422	DG	C2'-C1'	-22.18	1.30	1.52
2	C	422	DG	O4'-C1'	21.37	1.67	1.42
1	B	412	DA	O3'-P	16.77	1.81	1.61
2	C	422	DG	P-O5'	13.05	1.72	1.59
2	C	429	DC	C5'-C4'	12.53	1.65	1.51
1	B	406	DG	N9-C4	12.12	1.47	1.38
1	B	412	DA	N9-C4	10.21	1.44	1.37
1	B	408	DG	C5'-C4'	9.57	1.61	1.51
2	C	431	DT	C1'-N1	9.33	1.61	1.49
1	B	412	DA	N7-C5	9.27	1.44	1.39
2	C	428	DG	P-O5'	-9.23	1.50	1.59
1	B	412	DA	N9-C8	8.72	1.44	1.37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	403	DC	C2-N3	8.70	1.42	1.35
1	B	405	DT	C5-C6	8.57	1.40	1.34
1	B	402	DC	N3-C4	8.35	1.39	1.33
2	C	429	DC	P-O5'	8.34	1.68	1.59
1	B	404	DA	N7-C5	-8.17	1.34	1.39
2	C	422	DG	C5'-C4'	8.06	1.60	1.51
1	B	406	DG	C5'-C4'	7.78	1.59	1.51
2	C	428	DG	C5'-C4'	7.69	1.59	1.51
1	B	411	DG	N3-C4	-7.32	1.30	1.35
1	B	413	DC	P-O5'	7.25	1.67	1.59
1	B	402	DC	P-O5'	7.22	1.67	1.59
2	C	423	DT	N1-C6	-7.09	1.33	1.38
1	B	410	DT	O3'-P	7.06	1.69	1.61
2	C	429	DC	O3'-P	7.04	1.69	1.61
2	C	431	DT	C5'-C4'	7.03	1.59	1.51
2	C	433	DG	C2-N3	6.92	1.38	1.32
1	B	406	DG	C3'-O3'	6.80	1.52	1.44
2	C	425	DA	O3'-P	-6.72	1.53	1.61
2	C	425	DA	P-O5'	6.58	1.66	1.59
1	B	412	DA	N3-C4	6.57	1.38	1.34
1	B	408	DG	N7-C5	6.46	1.43	1.39
2	C	431	DT	N1-C2	6.42	1.43	1.38
1	B	409	DC	C5'-C4'	6.38	1.58	1.51
1	B	413	DC	N1-C6	6.25	1.40	1.37
2	C	430	DA	C5'-C4'	6.24	1.58	1.51
2	C	424	DC	O3'-P	-6.19	1.53	1.61
2	C	430	DA	O3'-P	6.01	1.68	1.61
1	B	408	DG	N9-C8	6.00	1.42	1.37
1	B	405	DT	C5'-C4'	5.99	1.57	1.51
2	C	431	DT	C2'-C1'	5.99	1.58	1.52
2	C	426	DG	C5'-C4'	5.96	1.57	1.51
2	C	432	DG	N7-C5	5.96	1.42	1.39
2	C	432	DG	O3'-P	5.91	1.68	1.61
1	B	410	DT	C5-C7	5.74	1.53	1.50
2	C	431	DT	P-O5'	5.69	1.65	1.59
2	C	432	DG	C6-N1	5.62	1.43	1.39
1	B	412	DA	C5'-C4'	-5.31	1.45	1.51
1	B	408	DG	C6-N1	5.31	1.43	1.39
2	C	425	DA	C4'-O4'	-5.29	1.39	1.45
2	C	428	DG	C4'-C3'	-5.20	1.47	1.52
3	A	163	ARG	CD-NE	5.18	1.55	1.46
2	C	424	DC	C4'-O4'	-5.16	1.39	1.45

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	422	DG	O3'-P	5.09	1.67	1.61
2	C	424	DC	N1-C2	-5.07	1.35	1.40
2	C	431	DT	O4'-C1'	5.04	1.48	1.42
2	C	430	DA	N9-C8	5.02	1.41	1.37
1	B	413	DC	O4'-C1'	5.00	1.48	1.42

All (148) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	406	DG	O4'-C1'-C2'	-20.20	89.74	105.90
1	B	403	DC	O4'-C4'-C3'	-19.17	94.50	106.00
1	B	409	DC	O4'-C1'-C2'	-19.17	90.56	105.90
1	B	402	DC	O4'-C4'-C3'	-18.21	95.07	106.00
2	C	433	DG	O4'-C1'-C2'	-17.95	91.54	105.90
2	C	432	DG	O4'-C4'-C3'	-17.38	95.57	106.00
2	C	431	DT	O4'-C1'-N1	15.79	119.05	108.00
1	B	404	DA	O4'-C1'-C2'	-15.73	93.31	105.90
2	C	422	DG	N9-C1'-C2'	15.51	142.06	112.60
1	B	411	DG	P-O3'-C3'	14.96	137.66	119.70
1	B	403	DC	O4'-C1'-C2'	-14.68	94.16	105.90
1	B	413	DC	O4'-C4'-C3'	-13.80	97.72	106.00
1	B	404	DA	O4'-C1'-N9	-13.34	98.66	108.00
2	C	422	DG	O4'-C1'-N9	-13.24	98.73	108.00
2	C	431	DT	O4'-C1'-C2'	-12.04	96.27	105.90
2	C	424	DC	P-O3'-C3'	11.80	133.87	119.70
1	B	409	DC	N3-C4-C5	-11.40	117.34	121.90
2	C	428	DG	O3'-P-O5'	-11.28	82.58	104.00
1	B	405	DT	C4-C5-C7	-11.13	112.32	119.00
1	B	412	DA	O4'-C4'-C3'	10.99	112.59	106.00
1	B	404	DA	P-O3'-C3'	10.91	132.79	119.70
1	B	404	DA	C2-N3-C4	-10.87	105.17	110.60
1	B	413	DC	O4'-C1'-N1	10.81	115.57	108.00
1	B	406	DG	C4'-C3'-C2'	-10.59	93.57	103.10
1	B	410	DT	N3-C2-O2	-10.53	115.98	122.30
2	C	430	DA	O4'-C1'-N9	10.51	115.35	108.00
1	B	404	DA	O5'-P-OP2	-10.32	96.41	105.70
1	B	405	DT	O4'-C1'-C2'	-10.31	97.65	105.90
2	C	423	DT	N3-C2-O2	-9.93	116.34	122.30
1	B	404	DA	O4'-C4'-C3'	-9.91	100.05	106.00
2	C	423	DT	O4'-C1'-N1	9.69	114.78	108.00
1	B	410	DT	O4'-C1'-C2'	-9.62	98.20	105.90
2	C	423	DT	C4-C5-C6	9.38	123.63	118.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	429	DC	O4'-C1'-N1	9.18	114.43	108.00
1	B	403	DC	N3-C4-N4	9.05	124.33	118.00
1	B	403	DC	P-O3'-C3'	8.97	130.47	119.70
1	B	411	DG	O4'-C1'-C2'	-8.97	98.73	105.90
2	C	424	DC	O4'-C1'-C2'	-8.93	98.76	105.90
1	B	408	DG	O4'-C1'-N9	8.85	114.20	108.00
2	C	423	DT	C6-C5-C7	-8.79	117.63	122.90
1	B	405	DT	P-O3'-C3'	8.60	130.02	119.70
1	B	410	DT	C4-C5-C6	8.60	123.16	118.00
1	B	404	DA	C4-C5-C6	8.52	121.26	117.00
2	C	429	DC	O3'-P-O5'	-8.48	87.88	104.00
1	B	403	DC	C5-C4-N4	-8.15	114.49	120.20
1	B	405	DT	C4-C5-C6	8.07	122.84	118.00
1	B	406	DG	O4'-C1'-N9	7.97	113.58	108.00
2	C	425	DA	O4'-C1'-C2'	-7.96	99.53	105.90
1	B	412	DA	N1-C2-N3	-7.89	125.36	129.30
2	C	423	DT	N1-C2-N3	7.86	119.31	114.60
2	C	422	DG	O4'-C4'-C3'	-7.83	101.30	106.00
1	B	402	DC	C5-C4-N4	-7.80	114.74	120.20
2	C	428	DG	C8-N9-C4	-7.80	103.28	106.40
1	B	410	DT	C6-N1-C2	-7.58	117.51	121.30
2	C	433	DG	N3-C4-C5	-7.56	124.82	128.60
1	B	404	DA	O3'-P-O5'	-7.53	89.69	104.00
2	C	428	DG	O4'-C1'-C2'	-7.45	99.94	105.90
1	B	405	DT	N1-C2-N3	7.44	119.06	114.60
2	C	432	DG	C4'-C3'-C2'	-7.37	96.47	103.10
2	C	429	DC	N1-C2-O2	7.36	123.32	118.90
2	C	426	DG	P-O5'-C5'	7.21	132.44	120.90
1	B	404	DA	C5-C6-N1	-7.21	114.10	117.70
2	C	423	DT	N3-C4-C5	-7.14	110.92	115.20
2	C	429	DC	N3-C2-O2	-7.12	116.91	121.90
1	B	409	DC	O4'-C1'-N1	7.11	112.98	108.00
2	C	432	DG	C2-N3-C4	7.10	115.45	111.90
1	B	413	DC	N1-C2-O2	7.04	123.12	118.90
1	B	402	DC	C6-N1-C2	6.95	123.08	120.30
1	B	406	DG	N3-C4-N9	6.91	130.15	126.00
1	B	405	DT	O3'-P-O5'	-6.86	90.97	104.00
2	C	433	DG	O4'-C4'-C3'	-6.84	101.76	104.50
1	B	412	DA	C2-N3-C4	6.82	114.01	110.60
2	C	432	DG	N1-C2-N2	6.80	122.32	116.20
1	B	409	DC	P-O3'-C3'	6.78	127.83	119.70
2	C	432	DG	N3-C2-N2	-6.75	115.18	119.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	422	DG	C8-N9-C4	-6.74	103.70	106.40
1	B	406	DG	C6-C5-N7	-6.70	126.38	130.40
2	C	426	DG	C8-N9-C4	-6.63	103.75	106.40
2	C	433	DG	C2-N3-C4	6.59	115.19	111.90
2	C	429	DC	P-O3'-C3'	6.45	127.44	119.70
1	B	402	DC	N1-C2-N3	-6.39	114.73	119.20
1	B	405	DT	O5'-P-OP2	6.38	118.35	110.70
2	C	422	DG	N9-C4-C5	6.38	107.95	105.40
1	B	410	DT	C4'-C3'-C2'	-6.36	97.38	103.10
2	C	426	DG	N9-C4-C5	6.36	107.94	105.40
1	B	409	DC	C2-N3-C4	6.31	123.06	119.90
1	B	406	DG	P-O5'-C5'	6.23	130.87	120.90
1	B	412	DA	C4-C5-C6	-6.22	113.89	117.00
2	C	429	DC	N3-C4-C5	-6.20	119.42	121.90
1	B	404	DA	N1-C2-N3	6.19	132.40	129.30
1	B	402	DC	N3-C4-N4	6.17	122.32	118.00
1	B	406	DG	N3-C4-C5	-6.15	125.52	128.60
1	B	404	DA	N1-C6-N6	6.06	122.24	118.60
1	B	413	DC	C1'-O4'-C4'	6.05	116.15	110.10
1	B	405	DT	N3-C4-C5	-6.05	111.57	115.20
2	C	430	DA	N9-C4-C5	5.97	108.19	105.80
1	B	412	DA	C1'-O4'-C4'	-5.93	104.17	110.10
1	B	412	DA	O4'-C1'-N9	5.92	112.15	108.00
2	C	426	DG	C5-N7-C8	-5.88	101.36	104.30
1	B	402	DC	O4'-C1'-N1	-5.88	103.89	108.00
2	C	423	DT	O4'-C4'-C3'	5.88	109.53	106.00
2	C	422	DG	C5-N7-C8	-5.86	101.37	104.30
1	B	412	DA	C4'-C3'-C2'	-5.86	97.83	103.10
2	C	430	DA	P-O5'-C5'	5.86	130.27	120.90
1	B	404	DA	N9-C1'-C2'	5.86	123.73	112.60
2	C	428	DG	P-O5'-C5'	-5.80	111.61	120.90
2	C	422	DG	N1-C2-N3	-5.80	120.42	123.90
3	A	75	LEU	CA-CB-CG	5.78	128.60	115.30
1	B	409	DC	C4'-C3'-C2'	-5.77	97.90	103.10
2	C	425	DA	O4'-C4'-C3'	-5.74	102.20	104.50
2	C	426	DG	N1-C2-N3	-5.74	120.46	123.90
1	B	405	DT	N3-C4-O4	5.73	123.34	119.90
1	B	403	DC	O3'-P-O5'	5.72	114.87	104.00
1	B	404	DA	C8-N9-C4	-5.69	103.52	105.80
1	B	405	DT	C4'-C3'-C2'	-5.69	97.98	103.10
2	C	422	DG	N1-C2-N2	5.69	121.32	116.20
2	C	426	DG	N1-C2-N2	5.66	121.30	116.20

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	408	DG	O4'-C1'-C2'	-5.66	101.38	105.90
1	B	402	DC	N1-C2-O2	5.61	122.27	118.90
2	C	428	DG	C5-C6-O6	-5.60	125.24	128.60
1	B	404	DA	P-O5'-C5'	5.58	129.82	120.90
2	C	429	DC	O5'-P-OP2	5.52	117.32	110.70
1	B	402	DC	C4-C5-C6	-5.52	114.64	117.40
1	B	402	DC	C4'-C3'-C2'	5.49	108.04	103.10
2	C	430	DA	O3'-P-O5'	-5.47	93.61	104.00
1	B	413	DC	O4'-C1'-C2'	-5.47	101.52	105.90
1	B	410	DT	N1-C2-N3	5.45	117.87	114.60
1	B	408	DG	C2-N3-C4	5.45	114.62	111.90
1	B	411	DG	OP2-P-O3'	5.40	117.09	105.20
2	C	428	DG	N7-C8-N9	5.39	115.80	113.10
2	C	433	DG	C5-C6-N1	5.36	114.18	111.50
2	C	423	DT	C1'-O4'-C4'	-5.32	104.78	110.10
1	B	409	DC	N1-C2-O2	5.27	122.06	118.90
2	C	431	DT	P-O3'-C3'	5.25	126.00	119.70
1	B	411	DG	C8-N9-C4	-5.24	104.30	106.40
1	B	409	DC	C1'-O4'-C4'	-5.23	104.87	110.10
2	C	432	DG	N9-C4-C5	5.21	107.48	105.40
1	B	410	DT	C3'-C2'-C1'	5.18	108.72	102.50
2	C	433	DG	N1-C6-O6	-5.18	116.79	119.90
1	B	404	DA	N9-C4-C5	5.11	107.84	105.80
2	C	424	DC	N3-C4-N4	5.11	121.58	118.00
1	B	409	DC	C5-C6-N1	5.08	123.54	121.00
2	C	431	DT	C4-C5-C6	5.07	121.04	118.00
1	B	412	DA	P-O3'-C3'	5.05	125.77	119.70
1	B	406	DG	C4-N9-C1'	5.05	133.07	126.50
1	B	410	DT	C6-C5-C7	-5.02	119.89	122.90
2	C	422	DG	C1'-O4'-C4'	-5.01	105.09	110.10
2	C	433	DG	C8-N9-C4	-5.00	104.40	106.40

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	265	TYR	Sidechain
1	B	409	DC	Sidechain
2	C	428	DG	Sidechain
2	C	432	DG	Sidechain
2	C	433	DG	Sidechain

## 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	B	244	0	137	1	0
2	C	249	0	134	8	0
3	A	2606	0	2587	22	0
4	A	26	0	19	0	0
5	A	114	0	0	3	0
5	B	6	0	0	0	0
5	C	9	0	0	2	0
All	All	3254	0	2877	26	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 (26) 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)
2:C:422:DG:C1'	2:C:422:DG:O4'	1.67	1.13
3:A:275:HIS:NE2	5:A:470:HOH:O	2.19	0.63
3:A:258:ILE:HG23	3:A:259:PHE:CD2	2.39	0.57
3:A:223:THR:HB	3:A:225:LYS:HG2	1.87	0.56
5:C:486:HOH:O	3:A:162:LYS:HD3	2.05	0.55
3:A:42:ASP:O	3:A:46:GLN:HG3	2.06	0.55
2:C:428:DG:N7	3:A:255:GLY:HA2	2.22	0.55
3:A:95:ASP:OD1	3:A:97:ARG:HD3	2.07	0.54
3:A:155:LEU:HD12	3:A:155:LEU:C	2.28	0.53
2:C:428:DG:O4'	3:A:87:SER:HB3	2.08	0.53
3:A:248:ALA:HB1	5:A:469:HOH:O	2.09	0.53
3:A:162:LYS:HG2	3:A:162:LYS:O	2.09	0.52
3:A:80:PRO:HA	5:A:353:HOH:O	2.09	0.52
2:C:427:5NC:H6	3:A:81:CYS:SG	2.54	0.47
5:C:452:HOH:O	3:A:122:LYS:HB3	2.15	0.46
2:C:427:5NC:N5	3:A:81:CYS:HB3	2.31	0.46
1:B:411:DG:H2''	1:B:412:DA:C8	2.51	0.46
2:C:422:DG:N9	2:C:422:DG:O4'	2.44	0.43
2:C:430:DA:O5'	2:C:430:DA:H2'	2.19	0.43
2:C:425:DA:H4'	3:A:162:LYS:HE3	2.00	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:231:ILE:HD13	3:A:235:GLY:HA3	1.99	0.42
3:A:79:PHE:HB2	3:A:80:PRO:HD2	2.01	0.42
3:A:141:ASN:HA	3:A:141:ASN:HD22	1.66	0.41
3:A:45:ALA:O	3:A:48:VAL:HG23	2.21	0.41
3:A:184:LYS:HA	3:A:185:PRO:HD3	1.80	0.40
3:A:59:GLY:O	3:A:60:ASP:C	2.60	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	A	325/327 (99%)	307 (94%)	16 (5%)	2 (1%)	27 38

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	60	ASP
3	A	144	ASP

### 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
3	A	283/283 (100%)	263 (93%)	20 (7%)	16 22

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	2	ILE
3	A	6	ASP
3	A	25	ARG
3	A	28	LEU
3	A	41	TRP
3	A	48	VAL
3	A	68	THR
3	A	86	ILE
3	A	87	SER
3	A	134	GLU
3	A	176	ASN
3	A	209	ARG
3	A	220	GLU
3	A	223	THR
3	A	231	ILE
3	A	245	ARG
3	A	251	LEU
3	A	266	LEU
3	A	270	LYS
3	A	276	PRO

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

Mol	Chain	Res	Type
3	A	63	GLN
3	A	65	ASN
3	A	123	ASN
3	A	141	ASN
3	A	237	GLN
3	A	297	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](#)

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
1	5CM	B	407	1,2	14,21,22	1.81	4 (28%)	18,30,33	3.13	10 (55%)
2	5NC	C	427	2	14,20,21	2.98	5 (35%)	16,28,31	1.92	6 (37%)

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	5CM	B	407	1,2	-	0/3/21/22	0/2/2/2
2	5NC	C	427	2	-	0/3/21/22	0/2/2/2

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	407	5CM	C6-C5	-3.31	1.31	1.40
1	B	407	5CM	O4'-C4'	-2.39	1.39	1.45
2	C	427	5NC	O3'-C3'	2.02	1.47	1.43
2	C	427	5NC	C4-N3	2.34	1.39	1.35
2	C	427	5NC	C2'-C1'	2.43	1.59	1.52
1	B	407	5CM	O5'-C5'	2.76	1.48	1.44
1	B	407	5CM	C5'-C4'	3.14	1.61	1.51
2	C	427	5NC	C6-N1	5.06	1.46	1.35
2	C	427	5NC	C6-N5	8.29	1.52	1.32

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	407	5CM	C4'-O4'-C1'	-7.26	91.80	109.43
1	B	407	5CM	O4'-C4'-C3'	-5.05	93.67	105.69
1	B	407	5CM	O4'-C1'-C2'	-4.46	97.74	106.25
2	C	427	5NC	N5-C4-N3	-3.59	122.01	125.77
1	B	407	5CM	C2'-C3'-C4'	-3.30	95.77	102.76
2	C	427	5NC	O4'-C1'-C2'	-2.99	100.55	106.25
1	B	407	5CM	C5-C4-N3	-2.45	117.33	121.23

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	407	5CM	C5-C6-N1	-2.11	119.87	122.15
1	B	407	5CM	C6-C5-C4	2.12	121.53	116.30
2	C	427	5NC	O5'-C5'-C4'	2.16	116.41	109.00
2	C	427	5NC	N4-C4-N3	2.43	121.06	117.25
2	C	427	5NC	C2'-C1'-N1	2.62	120.38	114.27
1	B	407	5CM	N4-C4-N3	2.92	121.20	117.01
2	C	427	5NC	C2-N3-C4	3.09	122.10	115.27
1	B	407	5CM	C2'-C1'-N1	3.90	123.36	114.27
1	B	407	5CM	O5'-C5'-C4'	4.27	123.66	109.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	427	5NC	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand is 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	SAH	A	328	-	20,28,28	1.05	1 (5%)	19,40,40	1.85	5 (26%)

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	SAH	A	328	-	-	0/7/31/31	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	328	SAH	C5'-C4'	2.01	1.60	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	328	SAH	C4'-O4'-C1'	-4.51	105.12	109.83
4	A	328	SAH	O4'-C4'-C3'	-3.14	98.94	105.15
4	A	328	SAH	C4'-C5'-SD	2.13	121.34	113.71
4	A	328	SAH	C1'-N9-C4	2.45	130.87	126.64
4	A	328	SAH	C4-C5-N7	2.71	112.03	109.41

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	412:DA	O3'	413:DC	P	1.81

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.