Full wwPDB X-ray Structure Validation Report (i)

Feb 27, 2014 – 08:04 PM GMT

PDB ID : 4DMG
Title : Thermus thermophilus m5C1942 methyltransferase RlmO
Authors : Larsen, H.G.; Rasmussen, A.; Giessing, A.M.B.; Jogl, G.; Kirpekar, F.
Deposited on : 2012-02-07
Resolution : 1.70 Å(reported)

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 1.70 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	Percentile Ranks	Value
Rfree		0.202
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		0.9%
RSRZ outliers		5.7%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	2456 (1.70-1.70)
Clashscore	79885	2929 (1.70-1.70)
Ramachandran outliers	78287	2878 (1.70-1.70)
Sidechain outliers	78261	2878 (1.70-1.70)
RSRZ outliers	66119	2456 (1.70-1.70)

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 >=3, 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	393	
1	B	393	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6937 atoms, of which 0 are hydrogen and 0 are deuterium.

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.
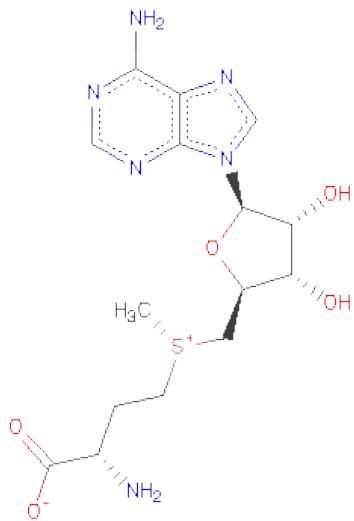
- Molecule 1 is a protein called Putative uncharacterized protein TTHA1493.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	385	Total 3103	C 1984	N 572	O 542	S 5	0	4	0
1	B	385	Total 3092	C 1975	N 571	O 541	S 5	0	2	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	386	LEU	-	EXPRESSION TAG	UNP Q5SI81
A	387	GLU	-	EXPRESSION TAG	UNP Q5SI81
A	388	HIS	-	EXPRESSION TAG	UNP Q5SI81
A	389	HIS	-	EXPRESSION TAG	UNP Q5SI81
A	390	HIS	-	EXPRESSION TAG	UNP Q5SI81
A	391	HIS	-	EXPRESSION TAG	UNP Q5SI81
A	392	HIS	-	EXPRESSION TAG	UNP Q5SI81
A	393	HIS	-	EXPRESSION TAG	UNP Q5SI81
B	386	LEU	-	EXPRESSION TAG	UNP Q5SI81
B	387	GLU	-	EXPRESSION TAG	UNP Q5SI81
B	388	HIS	-	EXPRESSION TAG	UNP Q5SI81
B	389	HIS	-	EXPRESSION TAG	UNP Q5SI81
B	390	HIS	-	EXPRESSION TAG	UNP Q5SI81
B	391	HIS	-	EXPRESSION TAG	UNP Q5SI81
B	392	HIS	-	EXPRESSION TAG	UNP Q5SI81
B	393	HIS	-	EXPRESSION TAG	UNP Q5SI81

- Molecule 2 is S-ADENOSYLMETHIONINE (three-letter code: SAM) (formula: $C_{15}H_{22}N_6O_5S$).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			27	15	6	5	1		
2	B	1	Total	C	N	O	S	0	0
			27	15	6	5	1		

• Molecule 3 is water.
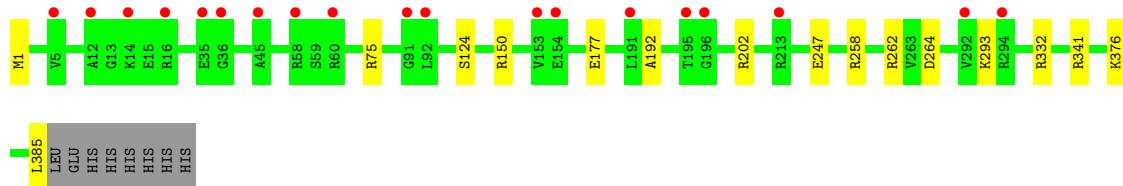
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	380	Total	O	0	0
			380	380		
3	B	308	Total	O	0	0
			308	308		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.
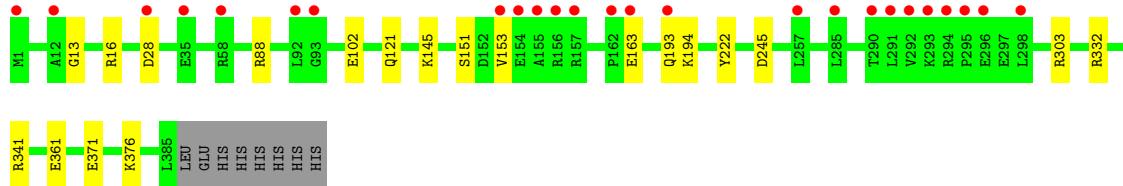
• Molecule 1: Putative uncharacterized protein TTHA1493

Chain A:

• Molecule 1: Putative uncharacterized protein TTHA1493

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	63.80Å 46.00Å 129.36Å 90.00° 96.78° 90.00°	Depositor
Resolution (Å)	27.37 – 1.70 27.37 – 1.70	Depositor EDS
% Data completeness (in resolution range)	98.5 (27.37-1.70) 93.4 (27.37-1.70)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.15 (at 1.70Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, R_{free}	0.179 , 0.208 0.173 , 0.202	Depositor DCC
R_{free} test set	3877 reflections (5.01%)	DCC
Wilson B-factor (Å^2)	19.3	Xtriage
Anisotropy	0.485	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 41.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Outliers	1 of 81675 reflections (0.001%)	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6937	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 49.48 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.4201e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	0.29	0/3185	0.49	0/4316
1	B	0.27	0/3168	0.47	0/4293
All	All	0.28	0/6353	0.48	0/8609

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3103	0	0	14	0
1	B	3092	0	0	14	0
2	A	27	0	22	0	0
2	B	27	0	22	1	0
3	A	380	0	0	12	1
3	B	308	0	0	9	1
All	All	6937	0	44	28	2

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

All (28) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:121:GLN:NE2	3:B:679:HOH:O	1.97	0.97
1:B:371:GLU:OE2	3:B:569:HOH:O	1.95	0.82
1:B:163:GLU:OE2	3:B:789:HOH:O	1.97	0.81
1:A:124:SER:OG	3:A:585:HOH:O	2.02	0.77
1:A:385:LEU:O	3:A:841:HOH:O	2.05	0.75
1:A:150:ARG:NE	3:A:589:HOH:O	2.21	0.72
1:A:341:ARG:NH2	3:A:656:HOH:O	2.26	0.67
1:B:303:ARG:NH2	3:B:726:HOH:O	2.32	0.62
1:A:247:GLU:OE1	3:A:818:HOH:O	2.17	0.59
1:B:193:GLN:NE2	3:B:587:HOH:O	2.35	0.58
1:B:16:ARG:NH1	3:B:730:HOH:O	2.38	0.57
1:B:194:LYS:NZ	1:B:245:ASP:OD2	2.38	0.56
1:A:177:GLU:OE2	1:A:258:ARG:NH2	2.41	0.53
1:B:13:GLY:O	1:B:16:ARG:NH1	2.46	0.48
1:A:262:ARG:NE	3:A:718:HOH:O	2.45	0.48
1:A:1:MET:N	3:A:561:HOH:O	2.47	0.48
1:A:293:LYS:NZ	3:A:590:HOH:O	2.48	0.47
1:A:264:ASP:OD1	3:A:718:HOH:O	2.20	0.46
1:B:332:ARG:NH2	3:B:699:HOH:O	2.49	0.46
1:A:202:ARG:NH1	3:A:646:HOH:O	2.49	0.45
1:B:341:ARG:NH1	3:B:581:HOH:O	2.49	0.44
1:A:150:ARG:NH1	1:A:192:ALA:O	2.50	0.44
1:A:75:ARG:NH1	3:A:836:HOH:O	2.52	0.42
1:B:222:TYR:O	2:B:401:SAM:HG2	2.20	0.42
1:B:88:ARG:NH2	1:B:361:GLU:OE2	2.52	0.42
1:B:341:ARG:NH2	3:B:547:HOH:O	2.53	0.41
1:A:332:ARG:NE	3:A:598:HOH:O	2.55	0.40

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

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:B:765:HOH:O	3:B:767:HOH:O[2_757]	1.88	0.32
3:A:634:HOH:O	3:A:695:HOH:O[2_756]	2.19	0.01

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	387/393 (98%)	378 (98%)	9 (2%)	0	100	100
1	B	385/393 (98%)	375 (97%)	10 (3%)	0	100	100
All	All	772/786 (98%)	753 (98%)	19 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2   Protein sidechains ⓘ

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	321/325 (99%)	320 (100%)	1 (0%)	96	93
1	B	319/325 (98%)	314 (98%)	5 (2%)	75	58
All	All	640/650 (98%)	634 (99%)	6 (1%)	87	79

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	376	LYS
1	B	28	ASP
1	B	145	LYS
1	B	151	SER
1	B	153	VAL
1	B	376	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3   RNA ⓘ

There are no RNA chains in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

2 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	SAM	A	401	-	26,29,29	1.00	2 (7%)	38,42,42	2.52	8 (21%)
2	SAM	B	401	-	26,29,29	0.99	2 (7%)	38,42,42	2.45	9 (23%)

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	SAM	A	401	-	-	0/13/33/33	0/1/3/3
2	SAM	B	401	-	-	0/13/33/33	0/1/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	401	SAM	C2-N3	3.27	1.38	1.32
2	A	401	SAM	C2-N3	2.93	1.38	1.32
2	B	401	SAM	C2-N1	2.08	1.38	1.33
2	A	401	SAM	C2-N1	2.05	1.38	1.33

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	401	SAM	N3-C2-N1	-10.99	119.52	128.71
2	A	401	SAM	N3-C2-N1	-10.59	119.86	128.71
2	A	401	SAM	O4'-C1'-N9	-7.56	101.41	108.44
2	B	401	SAM	O4'-C1'-N9	-6.08	102.79	108.44
2	B	401	SAM	C1'-N9-C4	-3.33	120.88	126.64
2	A	401	SAM	C1'-N9-C4	-2.95	121.54	126.64
2	A	401	SAM	O4'-C1'-C2'	-2.69	102.65	106.77
2	B	401	SAM	C2'-C1'-N9	2.56	119.85	113.27
2	B	401	SAM	C8-N9-C1'	2.50	131.30	126.38
2	B	401	SAM	N3-C4-N9	2.49	129.92	125.43
2	A	401	SAM	N3-C4-N9	2.48	129.92	125.43
2	A	401	SAM	C4-C5-N7	-2.47	107.41	109.52
2	B	401	SAM	C4-C5-N7	-2.28	107.57	109.52
2	A	401	SAM	N7-C8-N9	-2.26	107.96	114.36
2	B	401	SAM	CG-CB-CA	2.25	115.53	112.22
2	B	401	SAM	N7-C8-N9	-2.09	108.44	114.36
2	A	401	SAM	C2'-C1'-N9	2.01	118.42	113.27

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7　Other polymers　ⓘ

There are no such residues in this entry.

5.8　Polymer linkage issues

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, 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	385/393 (97%)	0.08	19 (4%)	28	33	11, 22, 49, 86	0
1	B	385/393 (97%)	0.27	25 (6%)	18	23	13, 27, 53, 80	0
All	All	770/786 (97%)	0.17	44 (5%)	23	27	11, 24, 51, 86	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	153	VAL	7.6
1	A	153	VAL	7.2
1	B	257	LEU	5.0
1	A	35	GLU	5.0
1	B	291	LEU	4.4
1	B	292	VAL	4.2
1	B	155	ALA	4.0
1	B	1	MET	4.0
1	B	298	LEU	3.7
1	A	196	GLY	3.5
1	B	295	PRO	3.5
1	B	290	THR	3.4
1	A	60	ARG	3.1
1	B	58	ARG	3.0
1	B	163	GLU	3.0
1	B	154	GLU	2.9
1	B	296	GLU	2.9
1	B	12	ALA	2.8
1	B	35	GLU	2.7
1	B	93	GLY	2.7
1	A	58	ARG	2.6
1	A	213	ARG	2.6
1	B	285	LEU	2.5
1	B	157	ARG	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	193	GLN	2.4
1	A	92	LEU	2.3
1	A	191	LEU	2.3
1	A	294	ARG	2.3
1	B	92	LEU	2.3
1	A	14	LYS	2.2
1	B	294	ARG	2.2
1	A	16	ARG	2.2
1	A	5	VAL	2.2
1	A	36	GLY	2.2
1	A	91	GLY	2.2
1	B	28	ASP	2.1
1	A	292	VAL	2.1
1	B	293	LYS	2.1
1	A	195	THR	2.1
1	A	45	ALA	2.1
1	A	12	ALA	2.0
1	A	154	GLU	2.0
1	B	156	ARG	2.0
1	B	162	PRO	2.0

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	SAM	A	401	27/27	0.10	-0.06	12,16,21,24	0
2	SAM	B	401	27/27	0.10	-0.29	15,23,30,38	0

6.5 Other polymers (i)

There are no such residues in this entry.