



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 20, 2016 – 12:03 PM EDT

PDB ID : 5L8X
Title : X-RAY STRUCTURE OF APO METHANOCALDOCOCUS JANNASCHII
METHYLTRANSFERASE SUBUNIT A AT 1.85 ANGSTROM
Authors : Wagner, T.; Ermler, U.; Shima, S.
Deposited on : 2016-06-08
Resolution : 1.85 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

<http://wwpdb.org/validation/2016/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 : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027939
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027939

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MLT	B	201	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2688 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 Tetrahydromethanopterin S-methyltransferase subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	172	Total	C	N	O	S	0	0	0
			1287	808	219	251	9			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	expression tag	UNP Q58261
A	0	HIS	-	expression tag	UNP Q58261

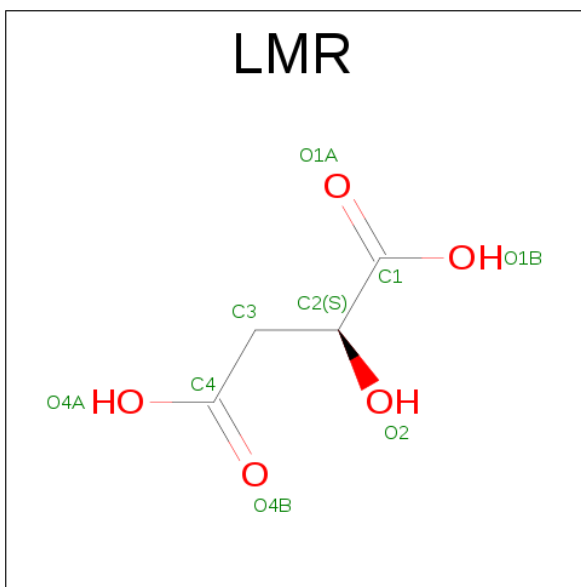
- Molecule 2 is a protein called Tetrahydromethanopterin S-methyltransferase subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	172	Total	C	N	O	S	0	0	0
			1282	805	219	249	9			

There are 3 discrepancies between the modelled and reference sequences:

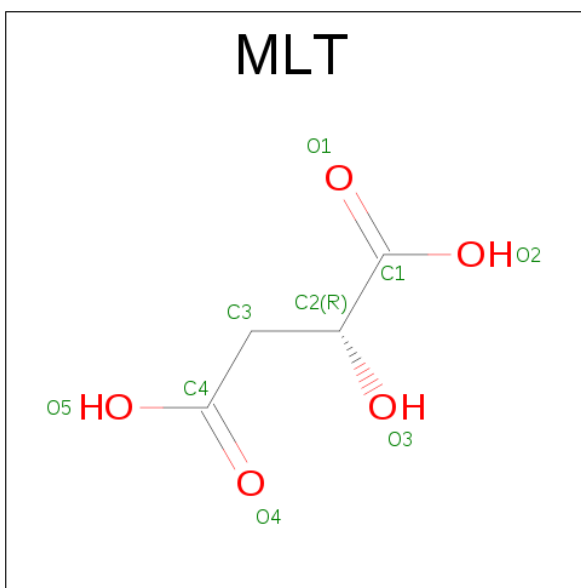
Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	GLY	-	expression tag	UNP Q58261
B	-1	SER	-	expression tag	UNP Q58261
B	0	HIS	-	expression tag	UNP Q58261

- Molecule 3 is (2S)-2-hydroxybutanedioic acid (three-letter code: LMR) (formula: C₄H₆O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			9	4	5		

- Molecule 4 is D-MALATE (three-letter code: MLT) (formula: $C_4H_6O_5$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			9	4	5		

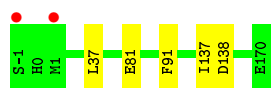
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	52	Total 52	O 52	0	0
5	B	49	Total 49	O 49	0	0

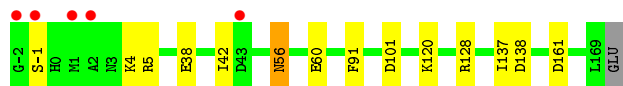
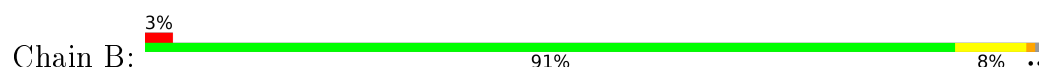
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Tetrahydromethanopterin S-methyltransferase subunit A



- Molecule 2: Tetrahydromethanopterin S-methyltransferase subunit A



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	64.69Å 37.03Å 65.16Å 90.00° 91.98° 90.00°	Depositor
Resolution (Å)	45.11 – 1.85 45.11 – 1.85	Depositor EDS
% Data completeness (in resolution range)	95.1 (45.11-1.85) 95.1 (45.11-1.85)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.04 (at 1.86Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.187 , 0.237 0.182 , 0.235	Depositor DCC
R_{free} test set	1279 reflections (5.03%)	DCC
Wilson B-factor (Å ²)	21.1	Xtriage
Anisotropy	0.419	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 46.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.032 for l,k,-h 0.035 for h,-k,-l 0.027 for l,-k,h	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	2688	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.23% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: LMR, MLT

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.44	0/1310	0.58	0/1776
2	B	0.44	0/1305	0.56	0/1769
All	All	0.44	0/2615	0.57	0/3545

There are no bond length outliers.

There are no bond angle outliers.

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	1287	0	1268	2	0
2	B	1282	0	1265	8	0
3	A	9	0	4	0	0
4	B	9	0	4	1	0
5	A	52	0	0	0	0
5	B	49	0	0	3	0
All	All	2688	0	2541	11	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 2.

All (11) 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:B:101:ASP:OD2	2:B:120:LYS:NZ	2.24	0.70
2:B:56:ASN:O	2:B:60:GLU:HG3	2.08	0.54
2:B:5:ARG:NH1	2:B:161:ASP:OD1	2.47	0.48
2:B:4:LYS:NZ	5:B:303:HOH:O	2.36	0.47
4:B:201:MLT:H31	5:B:307:HOH:O	2.15	0.46
2:B:137:ILE:O	2:B:138:ASP:HB2	2.17	0.45
2:B:128:ARG:HD3	5:B:344:HOH:O	2.17	0.43
1:A:137:ILE:O	1:A:138:ASP:HB2	2.20	0.42
1:A:81:GLU:HG2	1:A:137:ILE:CG2	2.50	0.41
2:B:38:GLU:O	2:B:42:ILE:HG12	2.21	0.41
2:B:4:LYS:HA	2:B:4:LYS:HD3	1.88	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	
1	A	170/172 (99%)	166 (98%)	4 (2%)	0	100	100
2	B	170/173 (98%)	165 (97%)	4 (2%)	1 (1%)	30	13
All	All	340/345 (99%)	331 (97%)	8 (2%)	1 (0%)	46	29

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	-1	SER

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	138/138 (100%)	136 (99%)	2 (1%)	74	63
2	B	137/138 (99%)	135 (98%)	2 (2%)	72	60
All	All	275/276 (100%)	271 (98%)	4 (2%)	72	60

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	LEU
1	A	91	PHE
2	B	56	ASN
2	B	91	PHE

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 molecules 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	$\# Z > 2$
3	LMR	A	201	-	1,8,8	0.04	0	2,10,10	4.83	1 (50%)
4	MLT	B	201	-	1,8,8	0.56	0	2,10,10	2.67	1 (50%)

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
3	LMR	A	201	-	-	0/2/8/8	0/0/0/0
4	MLT	B	201	-	-	0/2/8/8	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	A	201	LMR	C4-C3-C2	-6.84	107.08	114.33
4	B	201	MLT	C4-C3-C2	-3.66	110.44	114.33

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	201	MLT	1	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

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

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

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	172/172 (100%)	-0.17	2 (1%) 81 81	10, 21, 40, 59	0
2	B	172/173 (99%)	0.13	5 (2%) 55 52	10, 24, 44, 58	0
All	All	344/345 (99%)	-0.02	7 (2%) 68 67	10, 22, 42, 59	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	-2	GLY	4.0
2	B	2	ALA	3.4
2	B	-1	SER	3.0
2	B	43	ASP	2.6
1	A	1	MET	2.3
1	A	-1	SER	2.1
2	B	1	MET	2.1

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	MLT	B	201	9/9	0.87	0.20	2.86	22,42,60,63	0
3	LMR	A	201	9/9	0.94	0.11	0.39	18,21,31,36	0

6.5 Other polymers ⓘ

There are no such residues in this entry.