



# Full wwPDB X-ray Structure Validation Report ⓘ

May 28, 2020 – 09:13 pm BST

PDB ID : 1ZGH  
Title : Methionyl-tRNA formyltransferase from Clostridium thermocellum  
Authors : Yang, H.; Kataeva, I.; Xu, H.; Zhao, M.; Chang, J.; Liu, Z.; Chen, L.; Tempel, W.; Habel, J.; Zhou, W.; Lee, D.; Lin, D.; Chang, S.; Arendall III, W.B.; Richardson, J.S.; Richardson, D.C.; Rose, J.P.; Wang, B.; Southeast Collaboratory for Structural Genomics (SECSG)  
Deposited on : 2005-04-21  
Resolution : 2.05 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

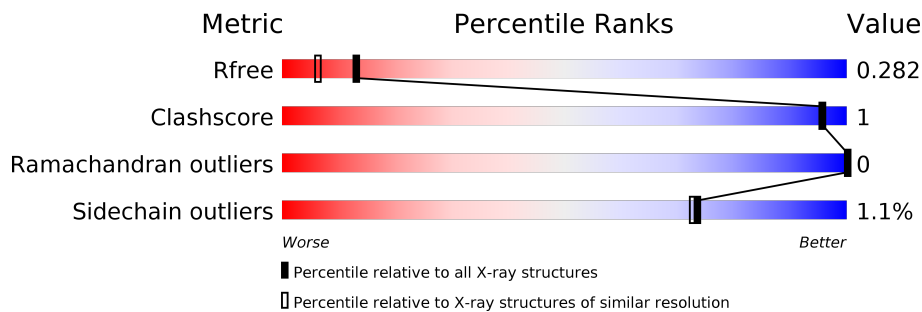
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


The reported resolution of this entry is 2.05 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)

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 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	260	 83% 13%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 1873 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 Methionyl-tRNA formyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	227	1809	1176	293	333	1	6	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-29	MSE	-	CLONING ARTIFACT	GB 48859327
A	-28	GLY	-	CLONING ARTIFACT	GB 48859327
A	-27	SER	-	CLONING ARTIFACT	GB 48859327
A	-26	SER	-	CLONING ARTIFACT	GB 48859327
A	-25	HIS	-	EXPRESSION TAG	GB 48859327
A	-24	HIS	-	EXPRESSION TAG	GB 48859327
A	-23	HIS	-	EXPRESSION TAG	GB 48859327
A	-22	HIS	-	EXPRESSION TAG	GB 48859327
A	-21	HIS	-	EXPRESSION TAG	GB 48859327
A	-20	HIS	-	EXPRESSION TAG	GB 48859327
A	-19	SER	-	CLONING ARTIFACT	GB 48859327
A	-18	SER	-	CLONING ARTIFACT	GB 48859327
A	-17	GLY	-	CLONING ARTIFACT	GB 48859327
A	-16	LEU	-	CLONING ARTIFACT	GB 48859327
A	-15	VAL	-	CLONING ARTIFACT	GB 48859327
A	-14	PRO	-	CLONING ARTIFACT	GB 48859327
A	-13	ARG	-	CLONING ARTIFACT	GB 48859327
A	-12	GLY	-	CLONING ARTIFACT	GB 48859327
A	-11	SER	-	CLONING ARTIFACT	GB 48859327
A	-10	GLN	-	CLONING ARTIFACT	GB 48859327
A	-9	SER	-	CLONING ARTIFACT	GB 48859327
A	-8	THR	-	CLONING ARTIFACT	GB 48859327
A	-7	SER	-	CLONING ARTIFACT	GB 48859327
A	-6	LEU	-	CLONING ARTIFACT	GB 48859327
A	-5	TYR	-	CLONING ARTIFACT	GB 48859327
A	-4	LYS	-	CLONING ARTIFACT	GB 48859327
A	-3	LYS	-	CLONING ARTIFACT	GB 48859327

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	ALA	-	CLONING ARTIFACT	GB 48859327
A	-1	GLY	-	CLONING ARTIFACT	GB 48859327
A	0	LEU	-	CLONING ARTIFACT	GB 48859327
A	1	MSE	MET	MODIFIED RESIDUE	GB 48859327
A	75	MSE	MET	MODIFIED RESIDUE	GB 48859327
A	131	MSE	MET	MODIFIED RESIDUE	GB 48859327
A	141	MSE	MET	MODIFIED RESIDUE	GB 48859327
A	187	MSE	MET	MODIFIED RESIDUE	GB 48859327
A	212	MSE	MET	MODIFIED RESIDUE	GB 48859327

- Molecule 2 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	17	Total X 17 17	0	0

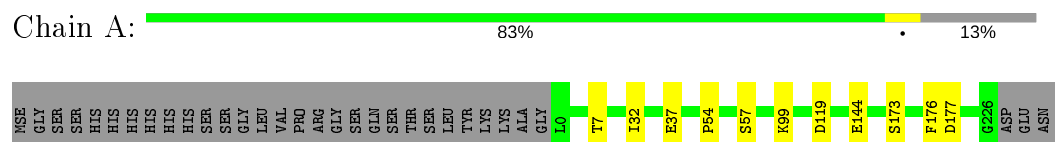
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	47	Total O 47 47	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 the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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: Methionyl-tRNA formyltransferase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.40Å 85.40Å 104.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.19 – 2.05 30.81 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.9 (30.19-2.05) 99.9 (30.81-2.05)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.34 (at 2.05Å)	Xtrriage
Refinement program	REFMAC refmac_5.2.0005	Depositor
R, $R_{free}$	0.209 , 0.237 0.261 , 0.282	Depositor DCC
$R_{free}$ test set	1223 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.7	Xtrriage
Anisotropy	0.202	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 67.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	1873	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.16% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.71	0/1844	0.71	1/2483 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	177	ASP	CB-CG-OD1	5.45	123.21	118.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	1809	0	1779	4	0
2	A	17	0	0	0	0
3	A	47	0	0	0	0
All	All	1873	0	1779	4	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 1.

All (4) 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:A:32:ILE:HG23	1:A:37:GLU:HB2	1.93	0.51
1:A:99:LYS:NZ	1:A:119:ASP:OD1	2.44	0.50
1:A:173:SER:HB3	1:A:176:PHE:CE1	2.50	0.46
1:A:7:THR:HB	1:A:54:PRO:HD2	2.00	0.43

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	225/260 (86%)	219 (97%)	6 (3%)	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	189/225 (84%)	187 (99%)	2 (1%)	73 73

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	57	SER
1	A	144	GLU



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

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 17 ligands modelled in this entry, 17 are unknown - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.