



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 05:45 am BST

PDB ID : 4Q37
Title : Crystal structure of the hypothetical protein TM0182 *Thermotoga maritima*, N-terminal domain.
Authors : Hocker, B.; Farias-Rico, J.A.
Deposited on : 2014-04-11
Resolution : 3.19 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
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

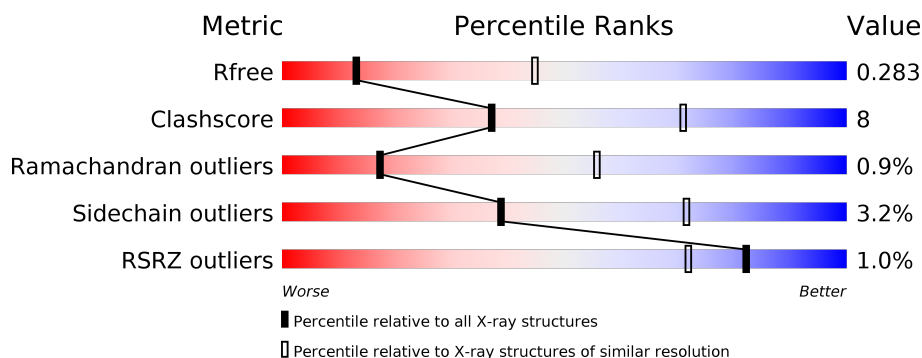
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.19 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 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 $\leq 5\%$. 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.

Mol	Chain	Length	Quality of chain
1	A	128	<div> <div></div> <div>74% 16% .. 6%</div> </div>
1	B	128	<div> <div>%</div> <div>78% 15% • 6%</div> </div>
1	C	128	<div> <div></div> <div>78% 13% • 6%</div> </div>
1	D	128	<div> <div>%</div> <div>75% 18% • 6%</div> </div>
1	E	128	<div> <div>4%</div> <div>65% 15% .. 19%</div> </div>
1	F	128	<div> <div></div> <div>81% 13% 6%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5272 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 Radical SAM protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	120	Total	C	N	O	S	0	0	0
			912	590	149	169	4			
1	B	120	Total	C	N	O	S	0	0	0
			895	585	143	162	5			
1	C	120	Total	C	N	O	S	0	0	0
			905	587	143	169	6			
1	D	120	Total	C	N	O	S	0	0	0
			905	585	146	169	5			
1	E	104	Total	C	N	O	S	0	0	0
			747	487	122	133	5			
1	F	120	Total	C	N	O	S	0	0	0
			889	579	139	166	5			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	121	LEU	-	EXPRESSION TAG	UNP Q9WY26
A	122	GLU	-	EXPRESSION TAG	UNP Q9WY26
A	123	HIS	-	EXPRESSION TAG	UNP Q9WY26
A	124	HIS	-	EXPRESSION TAG	UNP Q9WY26
A	125	HIS	-	EXPRESSION TAG	UNP Q9WY26
A	126	HIS	-	EXPRESSION TAG	UNP Q9WY26
A	127	HIS	-	EXPRESSION TAG	UNP Q9WY26
A	128	HIS	-	EXPRESSION TAG	UNP Q9WY26
B	121	LEU	-	EXPRESSION TAG	UNP Q9WY26
B	122	GLU	-	EXPRESSION TAG	UNP Q9WY26
B	123	HIS	-	EXPRESSION TAG	UNP Q9WY26
B	124	HIS	-	EXPRESSION TAG	UNP Q9WY26
B	125	HIS	-	EXPRESSION TAG	UNP Q9WY26
B	126	HIS	-	EXPRESSION TAG	UNP Q9WY26
B	127	HIS	-	EXPRESSION TAG	UNP Q9WY26
B	128	HIS	-	EXPRESSION TAG	UNP Q9WY26
C	121	LEU	-	EXPRESSION TAG	UNP Q9WY26

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	122	GLU	-	EXPRESSION TAG	UNP Q9WY26
C	123	HIS	-	EXPRESSION TAG	UNP Q9WY26
C	124	HIS	-	EXPRESSION TAG	UNP Q9WY26
C	125	HIS	-	EXPRESSION TAG	UNP Q9WY26
C	126	HIS	-	EXPRESSION TAG	UNP Q9WY26
C	127	HIS	-	EXPRESSION TAG	UNP Q9WY26
C	128	HIS	-	EXPRESSION TAG	UNP Q9WY26
D	121	LEU	-	EXPRESSION TAG	UNP Q9WY26
D	122	GLU	-	EXPRESSION TAG	UNP Q9WY26
D	123	HIS	-	EXPRESSION TAG	UNP Q9WY26
D	124	HIS	-	EXPRESSION TAG	UNP Q9WY26
D	125	HIS	-	EXPRESSION TAG	UNP Q9WY26
D	126	HIS	-	EXPRESSION TAG	UNP Q9WY26
D	127	HIS	-	EXPRESSION TAG	UNP Q9WY26
D	128	HIS	-	EXPRESSION TAG	UNP Q9WY26
E	121	LEU	-	EXPRESSION TAG	UNP Q9WY26
E	122	GLU	-	EXPRESSION TAG	UNP Q9WY26
E	123	HIS	-	EXPRESSION TAG	UNP Q9WY26
E	124	HIS	-	EXPRESSION TAG	UNP Q9WY26
E	125	HIS	-	EXPRESSION TAG	UNP Q9WY26
E	126	HIS	-	EXPRESSION TAG	UNP Q9WY26
E	127	HIS	-	EXPRESSION TAG	UNP Q9WY26
E	128	HIS	-	EXPRESSION TAG	UNP Q9WY26
F	121	LEU	-	EXPRESSION TAG	UNP Q9WY26
F	122	GLU	-	EXPRESSION TAG	UNP Q9WY26
F	123	HIS	-	EXPRESSION TAG	UNP Q9WY26
F	124	HIS	-	EXPRESSION TAG	UNP Q9WY26
F	125	HIS	-	EXPRESSION TAG	UNP Q9WY26
F	126	HIS	-	EXPRESSION TAG	UNP Q9WY26
F	127	HIS	-	EXPRESSION TAG	UNP Q9WY26
F	128	HIS	-	EXPRESSION TAG	UNP Q9WY26

- Molecule 2 is PLATINUM (II) ION (three-letter code: PT) (formula: Pt).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	F	1	Total Pt 1 1	0	0
2	E	1	Total Pt 1 1	0	0


- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	O 1	0	0
3	B	5	Total 5	O 5	0	0
3	C	4	Total 4	O 4	0	0
3	D	5	Total 5	O 5	0	0
3	F	2	Total 2	O 2	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 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: Radical SAM protein

Chain A: 




- Molecule 1: Radical SAM protein

Chain B: 




- Molecule 1: Radical SAM protein

Chain C: 



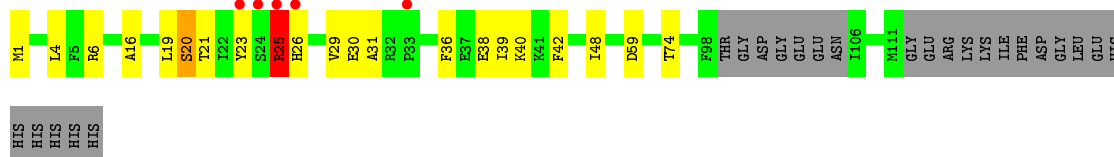
- Molecule 1: Radical SAM protein

Chain D: 

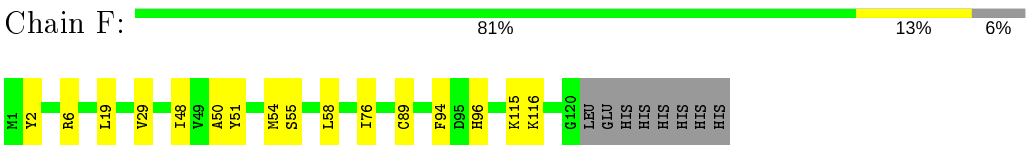


- Molecule 1: Radical SAM protein

Chain E: 



● Molecule 1: Radical SAM protein



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	102.05Å 145.30Å 141.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.60 – 3.19 39.60 – 3.19	Depositor EDS
% Data completeness (in resolution range)	99.1 (39.60-3.19) 99.4 (39.60-3.19)	Depositor EDS
R_{merge}	0.01	Depositor
R_{sym}	0.01	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.40 (at 3.18Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.235 , 0.280 0.240 , 0.283	Depositor DCC
R_{free} test set	892 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	69.6	Xtriage
Anisotropy	0.312	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 70.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	5272	wwPDB-VP
Average B, all atoms (Å ²)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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: PT

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	0/934	0.50	0/1269
1	B	0.26	0/918	0.53	0/1250
1	C	0.28	0/928	0.45	0/1263
1	D	0.25	0/928	0.46	0/1263
1	E	0.28	0/767	0.52	1/1048 (0.1%)
1	F	0.25	0/912	0.46	0/1244
All	All	0.27	0/5387	0.49	1/7337 (0.0%)

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	A	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	25	ARG	NE-CZ-NH1	5.11	122.85	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	38	GLU	Peptide
1	A	39	ILE	Peptide

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	912	0	850	21	0
1	B	895	0	824	12	0
1	C	905	0	825	12	0
1	D	905	0	820	16	0
1	E	747	0	644	19	0
1	F	889	0	799	11	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	1	0	0	0	0
3	B	5	0	0	0	0
3	C	4	0	0	0	0
3	D	5	0	0	0	0
3	F	2	0	0	0	0
All	All	5272	0	4762	78	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 8.

The worst 5 of 78 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:39:ILE:HG22	1:A:40:LYS:HA	1.63	0.79
1:E:25:ARG:HG2	1:E:25:ARG:HH11	1.50	0.75
1:E:16:ALA:HA	1:E:19:LEU:HB3	1.75	0.68
1:A:39:ILE:HA	1:A:41:LYS:H	1.61	0.66
1:E:21:THR:O	1:E:25:ARG:N	2.28	0.66

There are no symmetry-related clashes.

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	118/128 (92%)	111 (94%)	5 (4%)	2 (2%)	9	42
1	B	118/128 (92%)	109 (92%)	9 (8%)	0	100	100
1	C	118/128 (92%)	109 (92%)	8 (7%)	1 (1%)	19	58
1	D	118/128 (92%)	107 (91%)	9 (8%)	2 (2%)	9	42
1	E	100/128 (78%)	90 (90%)	9 (9%)	1 (1%)	15	54
1	F	118/128 (92%)	112 (95%)	6 (5%)	0	100	100
All	All	690/768 (90%)	638 (92%)	46 (7%)	6 (1%)	17	56

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	39	ILE
1	C	39	ILE
1	E	39	ILE
1	D	39	ILE
1	D	40	LYS

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	90/113 (80%)	85 (94%)	5 (6%)	21	57
1	B	85/113 (75%)	83 (98%)	2 (2%)	49	77
1	C	88/113 (78%)	85 (97%)	3 (3%)	37	70
1	D	87/113 (77%)	85 (98%)	2 (2%)	50	78
1	E	64/113 (57%)	62 (97%)	2 (3%)	40	72
1	F	84/113 (74%)	82 (98%)	2 (2%)	49	77
All	All	498/678 (74%)	482 (97%)	16 (3%)	39	71

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	11	ASN
1	C	27	LEU
1	E	20	SER
1	B	58	LEU
1	E	25	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	105	ASN

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 2 ligands modelled in this entry, 2 are monoatomic - 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 [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	120/128 (93%)	-0.18	0 100 100	29, 51, 82, 116	0
1	B	120/128 (93%)	0.01	1 (0%) 86 78	30, 64, 100, 139	0
1	C	120/128 (93%)	-0.06	0 100 100	33, 59, 92, 138	0
1	D	120/128 (93%)	0.10	1 (0%) 86 78	42, 72, 115, 140	0
1	E	104/128 (81%)	0.33	5 (4%) 30 18	39, 85, 134, 293	0
1	F	120/128 (93%)	0.07	0 100 100	38, 74, 110, 153	0
All	All	704/768 (91%)	0.04	7 (0%) 82 72	29, 67, 111, 293	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	26	HIS	3.7
1	E	25	ARG	3.5
1	E	23	TYR	2.5
1	E	24	SER	2.4
1	E	33	PRO	2.2

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. 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	B-factors(\AA^2)	Q<0.9
2	PT	E	201	1/1	0.97	0.12	78,78,78,78	1
2	PT	F	201	1/1	0.99	0.13	70,70,70,70	1

6.5 Other polymers [i](#)

There are no such residues in this entry.