



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 5, 2017 – 07:14 AM EST

PDB ID : 4DDX  
Title : Thermotoga maritima reverse gyrase, primitive monoclinic form  
Authors : Rudolph, M.G.; Klostermeier, D.  
Deposited on : unknown  
Resolution : 4.17 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](mailto: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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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-20030345

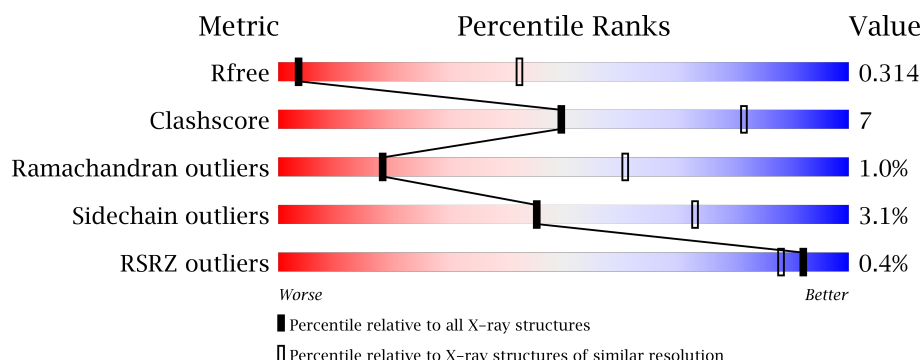
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 4.17 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))
$R_{free}$	100719	1167 (4.76-3.60)
Clashscore	112137	1024 (4.70-3.66)
Ramachandran outliers	110173	1022 (4.72-3.62)
Sidechain outliers	110143	1007 (4.72-3.62)
RSRZ outliers	101464	1179 (4.76-3.60)

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.

Mol	Chain	Length	Quality of chain
1	A	1104	<div> <div></div> <div> <div></div> <div>77%</div> <div>21%</div> <div>.</div> </div> </div>
1	B	1104	<div> <div></div> <div> <div></div> <div>77%</div> <div>21%</div> <div>.</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 18064 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 Reverse gyrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1102	Total	C	N	O	S	0	0	0
			9030	5759	1563	1682	26			
1	B	1102	Total	C	N	O	S	0	0	0
			9030	5759	1563	1682	26			

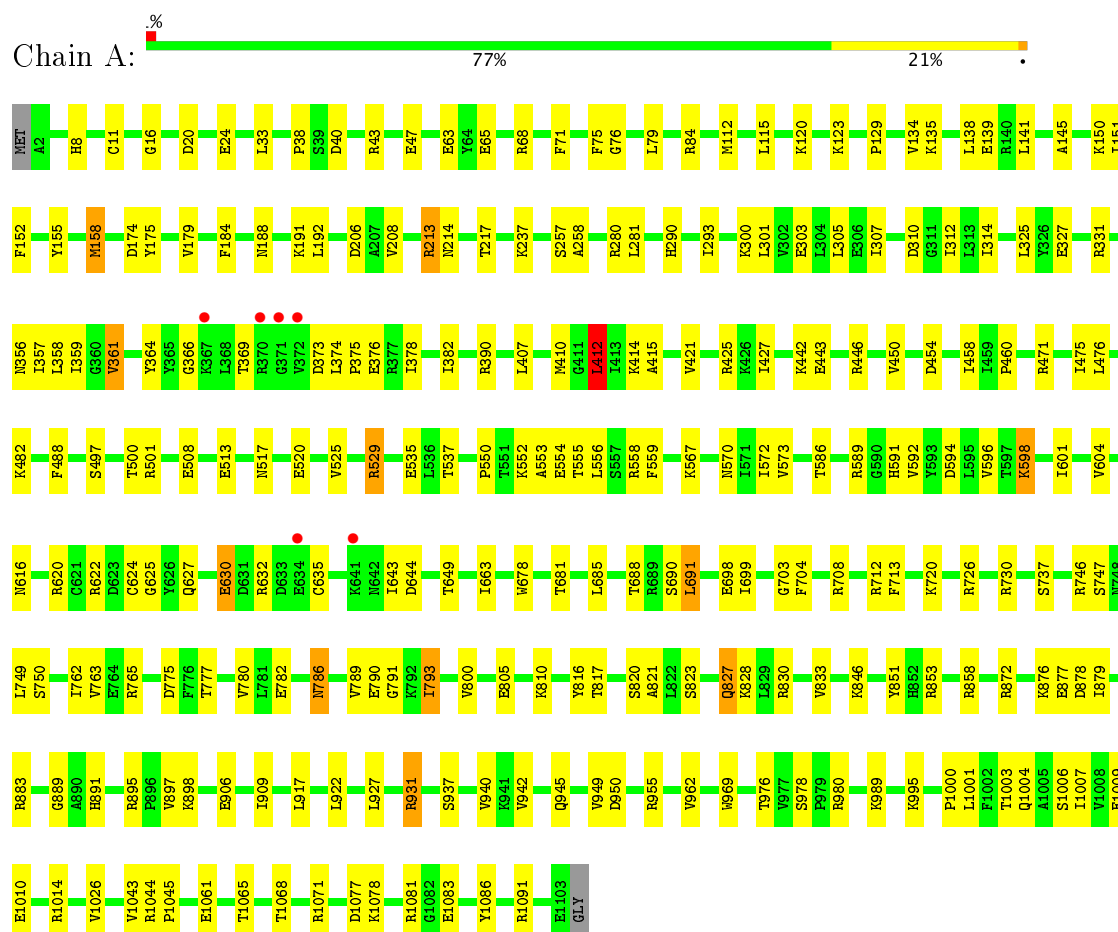
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Zn	0	0
			2	2		
2	A	2	Total	Zn	0	0
			2	2		

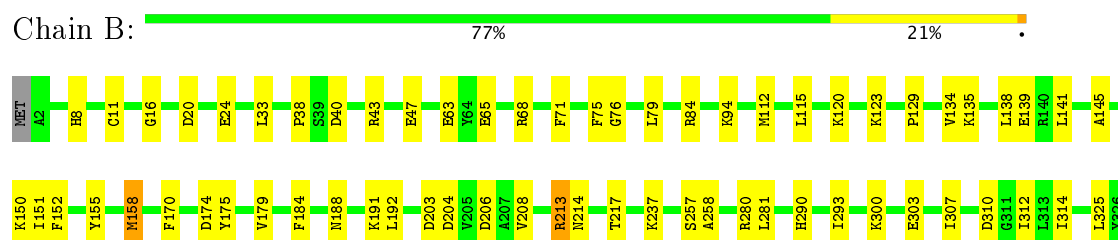
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Reverse gyrase



#### • Molecule 1: Reverse gyrase



T1021	R883	L749	V604	K482	E327
V1026	G889	S750	N616	K482	R331
R1044	A890	I762	R620	F488	N356
P1045	H891	V763	C621	S457	I357
E1061	R895	E764	R622	T500	L358
P896	P896	R765	D623	R501	L359
V897	V897	D775	C624	E508	G360
K898	K898	F776	G625	E508	V361
T1065	T1065	T777	G626	E513	Y364
T1068	E906	V780	D627	E513	Y365
E907	E907	L781	E630	N517	G366
R908	R908	E782	D631	E520	T369
I909	I909	N786	R632	E520	G371
L917	L917	V789	C635	V525	V372
L922	L922	E790	I643	R529	D373
G1082	G1082	G791	D644	E535	L374
E1083	E1083	K792	T649	T537	P375
Y1086	Y1086	I793	I663	L536	E376
R1091	R1091	V800	I663	T537	R377
E1103	E1103	E805	N678	P550	L378
GLY	GLY	D679	D679	P550	T382
		V680	V680	K552	R390
		K810	T681	A553	L407
		Y816	L685	E554	M410
		T817	T688	L556	G411
		S820	R689	S557	L412
		A821	S690	P559	I413
		L822	L691	R566	K414
		S823	E698	K567	A415
		Y977	I699	N570	Q416
		P979	G703	I571	V421
		S978	F704	I572	R425
		R980	R830	V573	K426
		K989	V833	T586	I427
		K995	R846	R589	E443
		P1000	L1001	G590	R446
		F1002	Y851	H591	D454
		T1003	H852	V592	I458
		Q1004	R853	Y593	L459
		A1005	R858	D594	P460
		S1006	R872	V596	R471
		E1009	R876	K598	I475
		E1010	E877	T601	L476
		R1014	D878		
		R1018	I879		

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	125.81Å 104.90Å 126.18Å 90.00° 91.72° 90.00°	Depositor
Resolution (Å)	48.43 – 4.17 48.43 – 4.17	Depositor EDS
% Data completeness (in resolution range)	96.8 (48.43-4.17) 87.1 (48.43-4.17)	Depositor EDS
$R_{merge}$	0.28	Depositor
$R_{sym}$	0.28	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.49 (at 4.14Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_881)	Depositor
R, $R_{free}$	0.244 , 0.305 0.251 , 0.314	Depositor DCC
$R_{free}$ test set	1107 reflections (5.10%)	DCC
Wilson B-factor (Å <sup>2</sup> )	84.8	Xtriage
Anisotropy	0.684	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 95.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.007 for l,k,-h 0.023 for h,-k,-l 0.239 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	18064	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	141.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% 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

### 5.1 Standard geometry

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

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	0.21	0/9191	0.37	0/12356
1	B	0.21	0/9191	0.37	0/12356
All	All	0.21	0/18382	0.37	0/24712

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	9030	0	9186	122	0
1	B	9030	0	9186	124	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
All	All	18064	0	18372	244	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 7.

The worst 5 of 244 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:B:407:LEU:HA	1:B:410:MET:HB2	1.69	0.74
1:A:407:LEU:HA	1:A:410:MET:HB2	1.69	0.73
1:A:897:VAL:HG23	1:A:898:LYS:HG3	1.72	0.72
1:A:375:PRO:HG2	1:A:476:LEU:HD11	1.71	0.72
1:B:375:PRO:HG2	1:B:476:LEU:HD11	1.71	0.71

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	1100/1104 (100%)	979 (89%)	110 (10%)	11 (1%)	18	61
1	B	1100/1104 (100%)	979 (89%)	110 (10%)	11 (1%)	18	61
All	All	2200/2208 (100%)	1958 (89%)	220 (10%)	22 (1%)	18	61

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	129	PRO
1	B	129	PRO
1	A	174	ASP
1	A	412	LEU
1	A	535	GLU

### 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	992/993 (100%)	961 (97%)	31 (3%)	45	73
1	B	992/993 (100%)	961 (97%)	31 (3%)	45	73
All	All	1984/1986 (100%)	1922 (97%)	62 (3%)	45	73

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	978	SER
1	B	138	LEU
1	B	931	ARG
1	B	8	HIS
1	B	206	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	728	GLN
1	B	728	GLN

### 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 ⓘ

Of 4 ligands modelled in this entry, 4 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 ⓘ

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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	1102/1104 (99%)	-0.57	6 (0%) 90 86	23, 135, 216, 337	0
1	B	1102/1104 (99%)	-0.59	2 (0%) 94 93	41, 135, 210, 367	0
All	All	2204/2208 (99%)	-0.58	8 (0%) 92 88	23, 135, 212, 367	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	371	GLY	4.0
1	B	371	GLY	4.0
1	B	372	VAL	3.4
1	A	634	GLU	2.2
1	A	372	VAL	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. 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	ZN	A	1201	1/1	0.99	0.04	-1.94	203,203,203,203	0
2	ZN	A	1202	1/1	0.98	0.05	-2.14	125,125,125,125	0
2	ZN	B	1202	1/1	0.99	0.04	-2.21	101,101,101,101	0
2	ZN	B	1201	1/1	0.99	0.03	-2.86	189,189,189,189	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.