



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 04:38 am BST

PDB ID : 4DDX
Title : Thermotoga maritima reverse gyrase, primitive monoclinic form
Authors : Rudolph, M.G.; Klostermeier, D.
Deposited on : 2012-01-19
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

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

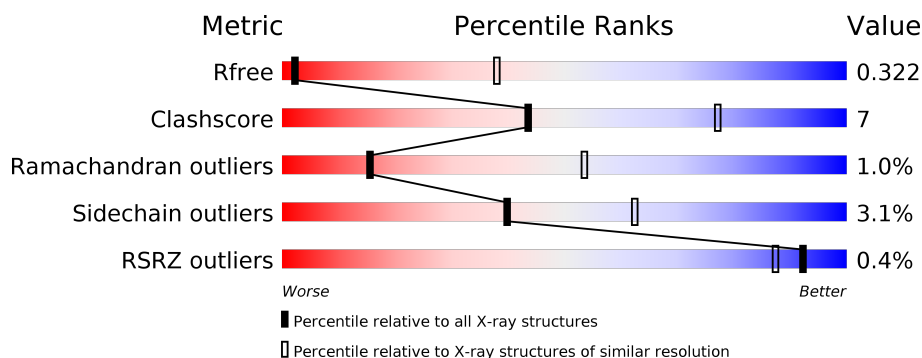
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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}	130704	1034 (4.60-3.76)
Clashscore	141614	1030 (4.54-3.80)
Ramachandran outliers	138981	1006 (4.58-3.78)
Sidechain outliers	138945	1037 (4.60-3.76)
RSRZ outliers	127900	1056 (4.66-3.70)

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 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	1104	<div> <div style="width: 100%; height: 10px; background-color: red;"></div> <div style="display: flex; justify-content: space-between; align-items: center;"> % <div style="width: 100%; height: 10px; background-color: green;"></div> </div> <div style="display: flex; justify-content: space-between; align-items: center;"> 77% 21% </div> </div>
1	B	1104	<div> <div style="width: 100%; height: 10px; background-color: green;"></div> <div style="display: flex; justify-content: space-between; align-items: center;"> 77% 21% </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).

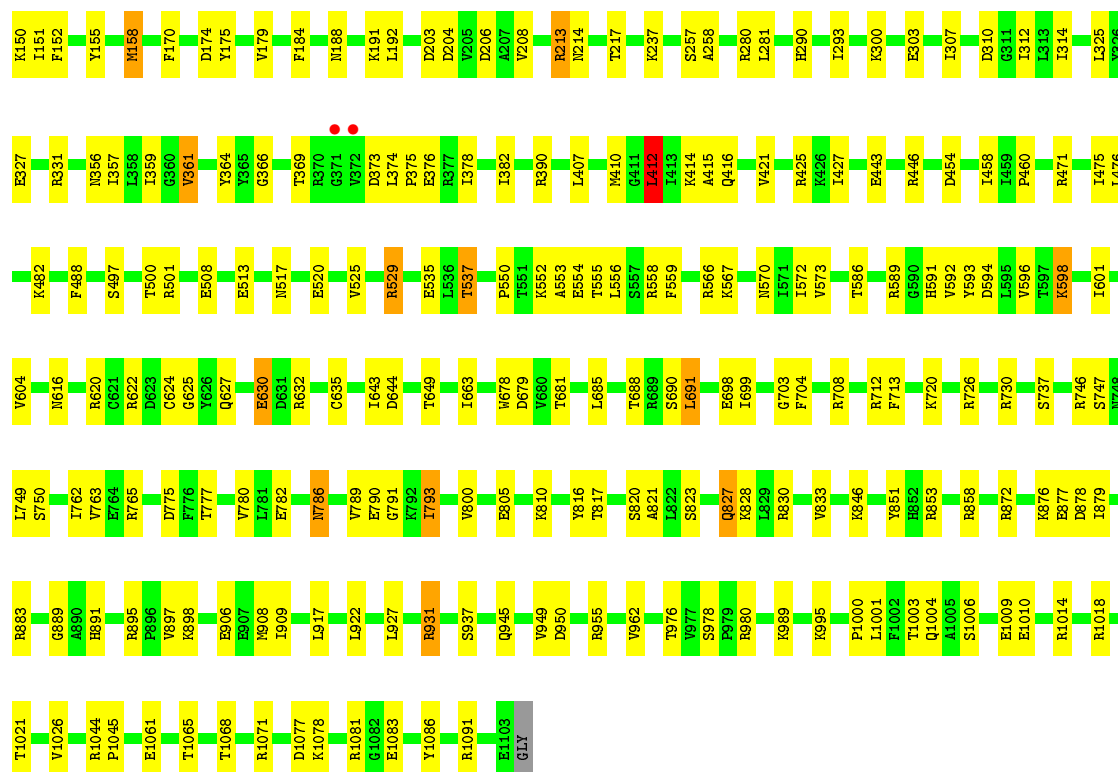
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Zn	0	0
			2	2		
2	A	2	Total	Zn	0	0
			2	2		

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 ($\text{RSRZ} > 2$). 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.

- Chain A: 77% 21% 2%

Residue	Identity (%)	Label
E1010	77%	
R1014	77%	
V1026	77%	
V1043	77%	
R1044	77%	
P1045	77%	
E1061	77%	
T1065	77%	
T1068	77%	
R1071	77%	
D1077	77%	
K1078	77%	
R1081	77%	
G1082	77%	
E1083	77%	
Y1086	77%	
R1091	77%	
E1103	77%	
GLY	77%	
R883	77%	
G889	77%	
A890	77%	
B891	77%	
R895	77%	
P896	77%	
B897	77%	
K898	77%	
E906	77%	
T909	77%	
L917	77%	
L922	77%	
L927	77%	
R931	77%	
S937	77%	
V940	77%	
V942	77%	
Q945	77%	
V949	77%	
D950	77%	
R955	77%	
V962	77%	
N969	77%	
T976	77%	
V977	77%	
S978	77%	
P979	77%	
R980	77%	
K989	77%	
K995	77%	
P1000	77%	
L1001	77%	
F1002	77%	
T1003	77%	
Q1004	77%	
A1005	77%	
S1006	77%	
I1007	77%	
E1008	77%	
E1009	77%	
L749	77%	
S750	77%	
I762	77%	
V763	77%	
E764	77%	
K765	77%	
D775	77%	
F776	77%	
I777	77%	
V780	77%	
L781	77%	
E782	77%	
N786	77%	
V789	77%	
E790	77%	
G791	77%	
K792	77%	
I793	77%	
W800	77%	
E805	77%	
K810	77%	
Y816	77%	
T817	77%	
S820	77%	
A821	77%	
L822	77%	
S823	77%	
Q827	77%	
K828	77%	
L829	77%	
R830	77%	
W833	77%	
K846	77%	
Y851	77%	
R853	77%	
R858	77%	
R872	77%	
K876	77%	
B877	77%	
D878	77%	
L879	77%	
M616	77%	
R620	77%	
G621	77%	
D622	77%	
D623	77%	
C624	77%	
G625	77%	
G626	77%	
D627	77%	
E630	77%	
D631	77%	
R632	77%	
D633	77%	
E634	77%	
C635	77%	
K641	77%	
I642	77%	
T643	77%	
D644	77%	
T649	77%	
T663	77%	
V678	77%	
T681	77%	
L685	77%	
T688	77%	
R689	77%	
S690	77%	
L691	77%	
E698	77%	
K699	77%	
L704	77%	
G703	77%	
R708	77%	
R712	77%	
F713	77%	
K720	77%	
R726	77%	
L730	77%	
S737	77%	
R746	77%	
S747	77%	
E749	77%	
K482	77%	
F488	77%	
S497	77%	
T500	77%	
R501	77%	
E508	77%	
E513	77%	
M517	77%	
E520	77%	
V525	77%	
R529	77%	
E535	77%	
L536	77%	
T537	77%	
P550	77%	
K552	77%	
A553	77%	
E554	77%	
T555	77%	
L556		

- Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	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
$\langle I/\sigma(I) \rangle$ ¹	1.49 (at 4.14Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_881)	Depositor
R, R_{free}	0.244 , 0.305 0.257 , 0.322	Depositor DCC
R_{free} test set	1221 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	84.8	Xtriage
Anisotropy	0.684	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 95.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	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 ²)	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.*

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: 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%)	15	53
1	B	1100/1104 (100%)	979 (89%)	110 (10%)	11 (1%)	15	53
All	All	2200/2208 (100%)	1958 (89%)	220 (10%)	22 (1%)	15	53

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%)	40	62
1	B	992/993 (100%)	961 (97%)	31 (3%)	40	62
All	All	1984/1986 (100%)	1922 (97%)	62 (3%)	40	62

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, 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	1102/1104 (99%)	-0.57	7 (0%) 89 84	23, 135, 216, 337	0
1	B	1102/1104 (99%)	-0.59	2 (0%) 95 93	41, 135, 210, 367	0
All	All	2204/2208 (99%)	-0.58	9 (0%) 92 87	23, 135, 212, 367	0

The worst 5 of 9 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. 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	ZN	A	1202	1/1	0.98	0.05	125,125,125,125	0
2	ZN	A	1201	1/1	0.99	0.04	203,203,203,203	0
2	ZN	B	1202	1/1	0.99	0.04	101,101,101,101	0
2	ZN	B	1201	1/1	0.99	0.03	189,189,189,189	0

6.5 Other polymers [i](#)

There are no such residues in this entry.