wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 03:46 PM GMT

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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

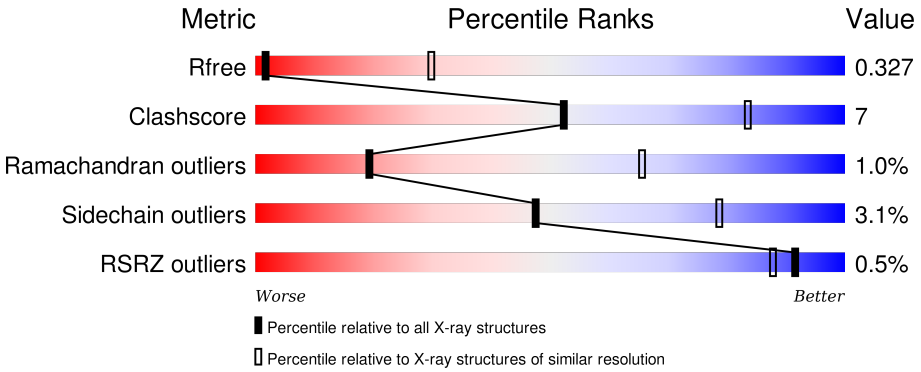
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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}	91344	1030 (4.76-3.60)
Clashscore	102246	1130 (4.76-3.60)
Ramachandran outliers	100387	1076 (4.76-3.60)
Sidechain outliers	100360	1061 (4.76-3.60)
RSRZ outliers	91569	1034 (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 >=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 <=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	% 77% 21%
1	B	1104	77% 21%

2 Entry composition (i)

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).

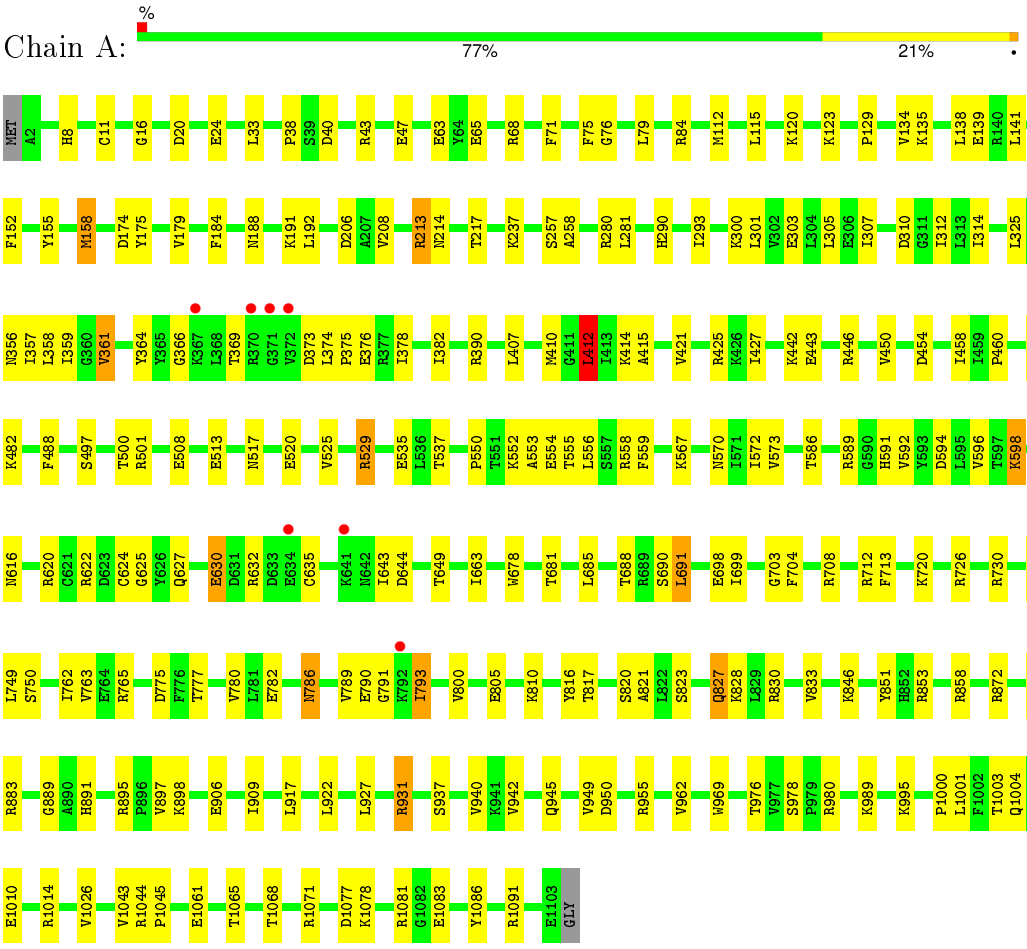
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Zn	0	0
			2	2		
2	A	2	Total	Zn	0	0
			2	2		

3 Residue-property plots (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.

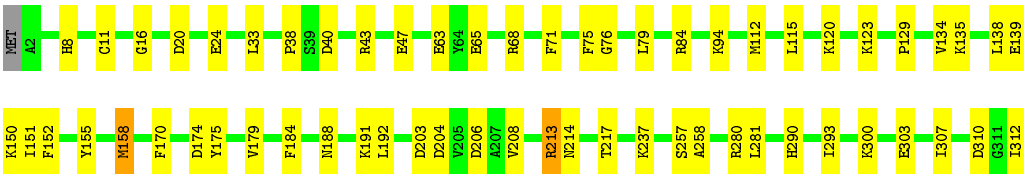
- Molecule 1: Reverse gyrase

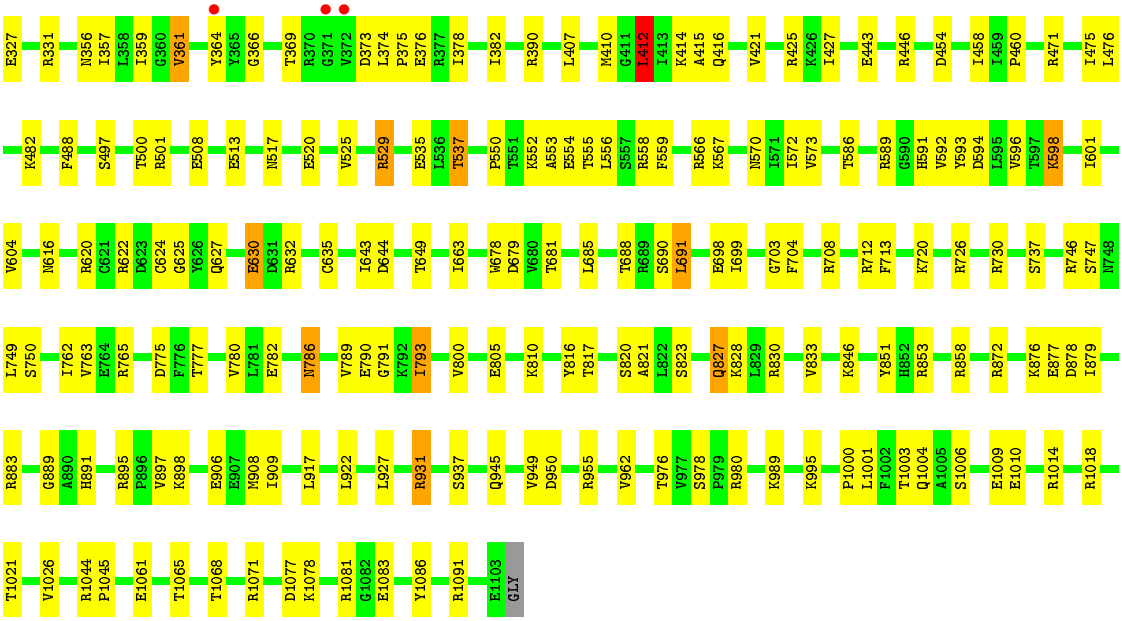
Chain A:



- Molecule 1: Reverse gyrase

Chain B:




E327 R331 N356 I357 L358 I359 G360 V361 Y364 Y365 G366 T369 R370 G371 V372 D373 L374 P375 E376 R377 I378 I382 R390 L407 M410 G411 L412 T413 K414 A415 Q416 V421 R425 K426 I427 E443 R446 D454 I458 I459 P460 R471 I475 L476
K482 F488 S497 T500 R501 E508 E513 N517 E520 V525 R529 E535 L536 T537 P550 T551 K552 A553 E554 T555 L556 S557 R558 F559 R566 K567 N570 I571 I572 V573 T586 R589 G590 H591 V592 Y593 D594 L595 V596 T597 K598 I601
V604 N616 R620 C621 R622 D623 C624 G625 Y626 Q627 E630 D631 R632 C635 I643 D644 T649 I663 W678 D679 V680 T681 L685 T688 R689 S690 L691 E698 I699 G703 F704 R708 R712 F713 K720 R726 R730 S737 R746 S747 N748
L749 S750 I762 V763 E764 R765 D775 F776 T777 V780 L781 E782 N786 V789 E790 G791 K792 I793 V800 E805 K810 Y816 T817 S820 A821 L822 S823 Q827 K828 L829 R830 V833 K846 Y851 H852 R853 R858 R872 K876 E877 D878 I879
R883 G889 A890 H891 R895 P896 V897 K898 E906 E907 M908 I909 L917 L922 L927 R931 S937 Q945 V949 D950 R955 V962 T976 V977 S978 P979 R980 K989 K995 P1000 L1001 F1002 T1003 Q1004 A1005 S1006 E1009 E1010 R1014 R1018
T1021 V1026 R1044 P1045 E1061 T1065 T1068 R1071 D1077 K1078 R1081 G1082 E1083 Y1086 R1091 E1103 GLY

4 Data and refinement statistics (i)

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
$< I/\sigma(I) >$ [1]	1.49 (at 4.14Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_881)	Depositor
R, R_{free}	0.244 , 0.305 0.275 , 0.327	Depositor DCC
R_{free} test set	1110 reflections (5.11%)	DCC
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
Estimated twinning fraction	0.007 for l,k,-h 0.023 for h,-k,-l 0.239 for l,-k,h	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Outliers	0 of 23987 reflections	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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%)	19	65
1	B	1100/1104 (100%)	979 (89%)	110 (10%)	11 (1%)	19	65
All	All	2200/2208 (100%)	1958 (89%)	220 (10%)	22 (1%)	19	65

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%)	47 79
1	B	992/993 (100%)	961 (97%)	31 (3%)	47 79
All	All	1984/1986 (100%)	1922 (97%)	62 (3%)	47 79

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 (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 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

6.1 Protein, DNA and RNA chains

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^{th} 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%)	-0.55	7 (0%)	90	86	23, 135, 216, 337	0
1	B	1102/1104 (99%)	-0.57	3 (0%)	94	92	41, 135, 210, 367	0
All	All	2204/2208 (99%)	-0.56	10 (0%)	91	88	23, 135, 212, 367	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	371	GLY	4.3
1	B	371	GLY	4.2
1	B	372	VAL	3.6
1	A	372	VAL	2.4
1	A	634	GLU	2.3

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

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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^{th} 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(Å^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.