



wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 08:36 PM GMT

PDB ID : 3IFZ
Title : crystal structure of the first part of the Mycobacterium tuberculosis DNA gyrase reaction core: the breakage and reunion domain at 2.7 Å resolution
Authors : Piton, J.; Aubry, A.; Delarue, M.; Mayer, C.
Deposited on : 2009-07-27
Resolution : 2.70 Å (reported)

This is a wwPDB 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/ValidationPDFNotes.html>

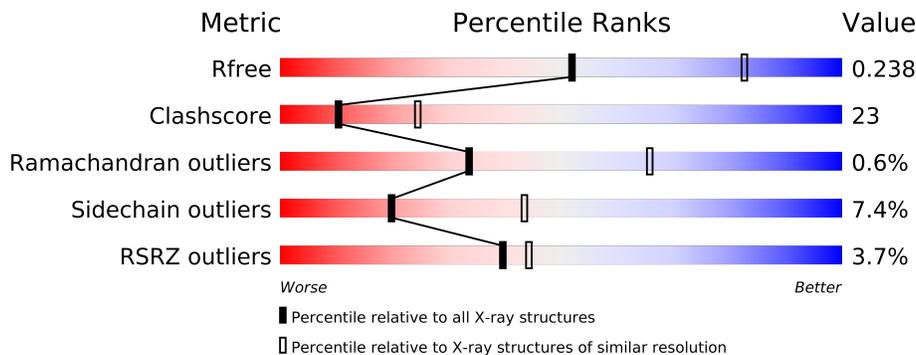
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.70 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}	66092	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	508	
1	B	508	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	MPD	A	509	-	X
2	MPD	B	509	-	X
3	NA	A	510	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7774 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

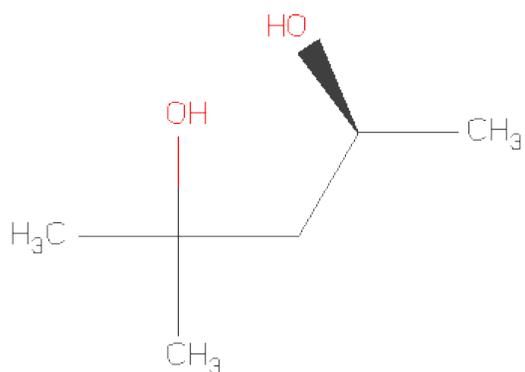
- Molecule 1 is a protein called DNA gyrase subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	487	Total	C	N	O	S	0	0	0
			3834	2386	701	734	13			
1	B	473	Total	C	N	O	S	0	0	0
			3700	2303	682	703	12			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	502	GLU	-	EXPRESSION TAG	UNP Q07702
A	503	HIS	-	EXPRESSION TAG	UNP Q07702
A	504	HIS	-	EXPRESSION TAG	UNP Q07702
A	505	HIS	-	EXPRESSION TAG	UNP Q07702
A	506	HIS	-	EXPRESSION TAG	UNP Q07702
A	507	HIS	-	EXPRESSION TAG	UNP Q07702
A	508	HIS	-	EXPRESSION TAG	UNP Q07702
B	502	GLU	-	EXPRESSION TAG	UNP Q07702
B	503	HIS	-	EXPRESSION TAG	UNP Q07702
B	504	HIS	-	EXPRESSION TAG	UNP Q07702
B	505	HIS	-	EXPRESSION TAG	UNP Q07702
B	506	HIS	-	EXPRESSION TAG	UNP Q07702
B	507	HIS	-	EXPRESSION TAG	UNP Q07702
B	508	HIS	-	EXPRESSION TAG	UNP Q07702

- Molecule 2 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: C₆H₁₄O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 8 6 2	0	0
2	B	1	Total C O 8 6 2	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Na 1 1	0	0
3	A	1	Total Na 1 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	104	Total O 104 104	0	0
4	B	118	Total O 118 118	0	0

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	163.91Å 109.66Å 102.00Å 90.00° 120.40° 90.00°	Depositor
Resolution (Å)	19.90 – 2.70 19.90 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.2 (19.90-2.70) 99.2 (19.90-2.70)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.46 (at 2.71Å)	Xtriage
Refinement program	BUSTER-TNT 2.5.1	Depositor
R, R_{free}	0.192 , 0.233 0.200 , 0.238	Depositor DCC
R_{free} test set	2125 reflections (5.28%)	DCC
Wilson B-factor (Å ²)	51.2	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 45.6	EDS
Estimated twinning fraction	0.009 for $-1/2*h+1/2*k+1, 1/2*h-1/2*k+1, 1/2*h+1/2*k$ 0.015 for $-1/2*h-1/2*k+1, -1/2*h-1/2*k-1, 1/2*h-1/2*k$	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 42396 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7774	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.89% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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: NA, MPD

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	0.26	0/3892	0.48	0/5271
1	B	0.27	0/3757	0.49	0/5090
All	All	0.26	0/7649	0.48	0/10361

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3834	0	3838	199	0
1	B	3700	0	3715	154	0
2	A	8	0	14	9	0
2	B	8	0	14	9	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	104	0	0	4	0
4	B	118	0	0	0	0
All	All	7774	0	7581	352	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 23.

The worst 5 of 352 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:509:MPD:C5	2:B:509:MPD:H12	1.61	1.22
2:B:509:MPD:C1	2:B:509:MPD:H53	1.57	1.21
1:A:298:GLY:HA2	1:A:319:LYS:HG2	1.32	1.08
1:A:259:VAL:HG12	1:A:269:LEU:HA	1.33	1.07
1:A:318:ILE:HD12	1:A:323:VAL:HA	1.39	1.05

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	483/508 (95%)	449 (93%)	32 (7%)	2 (0%)	43	76
1	B	471/508 (93%)	445 (94%)	22 (5%)	4 (1%)	27	58
All	All	954/1016 (94%)	894 (94%)	54 (6%)	6 (1%)	33	66

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	13	ASP
1	B	171	PRO
1	B	86	PRO
1	B	231	ALA
1	B	183	VAL

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	409/427 (96%)	377 (92%)	32 (8%)	18	40
1	B	391/427 (92%)	364 (93%)	27 (7%)	22	48
All	All	800/854 (94%)	741 (93%)	59 (7%)	20	43

5 of 59 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	401	LEU
1	B	35	VAL
1	B	401	LEU
1	A	402	ASP
1	A	405	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	52	HIS
1	B	85	HIS
1	B	389	HIS
1	B	83	ASN
1	B	208	HIS

5.3.3 RNA [i](#)

There are no RNA chains 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, 2 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	MPD	A	509	-	7,7,7	0.51	0	10,10,10	0.97	1 (10%)
2	MPD	B	509	-	7,7,7	0.61	0	10,10,10	0.76	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	MPD	A	509	-	-	0/5/5/5	0/0/0/0
2	MPD	B	509	-	-	0/5/5/5	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	509	MPD	C2-C3-C4	-2.03	105.61	116.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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, 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	487/508 (95%)	0.01	24 (4%) 28 31	29, 56, 109, 144	0
1	B	473/508 (93%)	-0.22	12 (2%) 54 61	26, 50, 80, 121	0
All	All	960/1016 (94%)	-0.10	36 (3%) 39 43	26, 53, 103, 144	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	14	ARG	5.0
1	B	183	VAL	4.9
1	A	182	ALA	4.8
1	A	322	ALA	4.7
1	A	321	ASP	4.4

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, 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	MPD	A	509	8/8	0.41	13.62	96,97,99,99	0
3	NA	A	510	1/1	0.30	4.99	54,54,54,54	0
2	MPD	B	509	8/8	0.26	2.79	63,65,67,67	0
3	NA	B	510	1/1	0.22	0.37	77,77,77,77	0

6.5 Other polymers

There are no such residues in this entry.