wwPDB X-ray Structure Validation Summary Report (i)

Nov 1, 2023 – 05:23 AM EDT

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 A resolution
Authors	:	Piton, J.; Aubry, A.; Delarue, M.; Mayer, C.
Deposited on	:	2009-07-27
Resolution	:	2.70 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

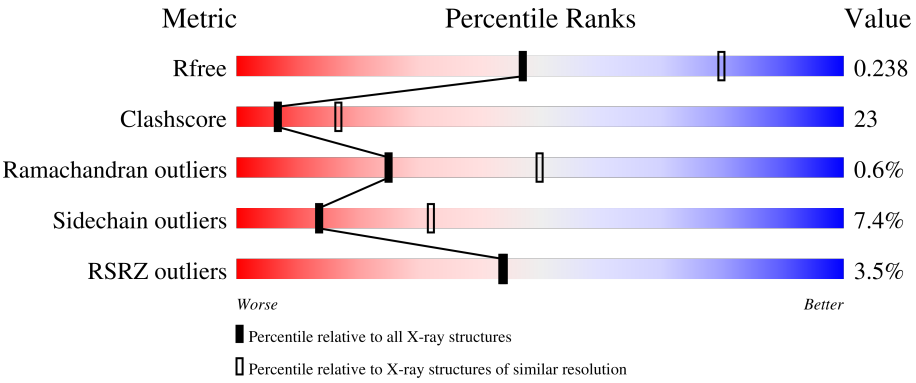
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 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}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (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 of 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 <=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	508	4% 59% 33%
1	B	508	2% 60% 30% 7%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	MPD	A	509	-	-	X	-
2	MPD	B	509	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7774 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.

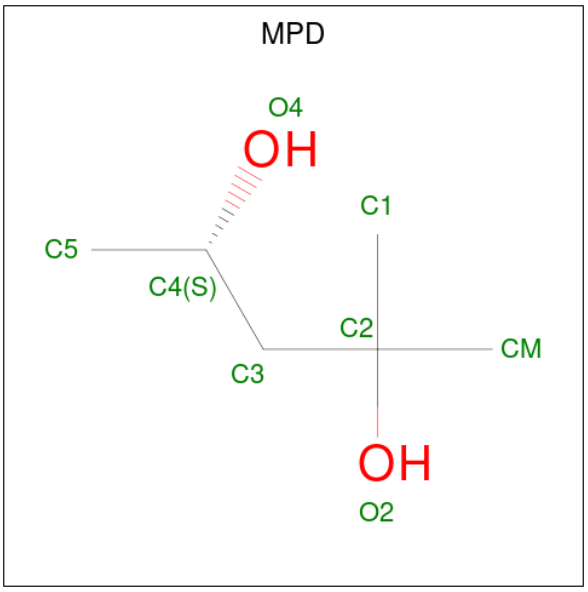
- Molecule 1 is a protein called DNA gyrase subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
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: $C_6H_{14}O_2$).



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	A	1	Total Na 1 1	0	0
3	B	1	Total Na 1 1	0	0

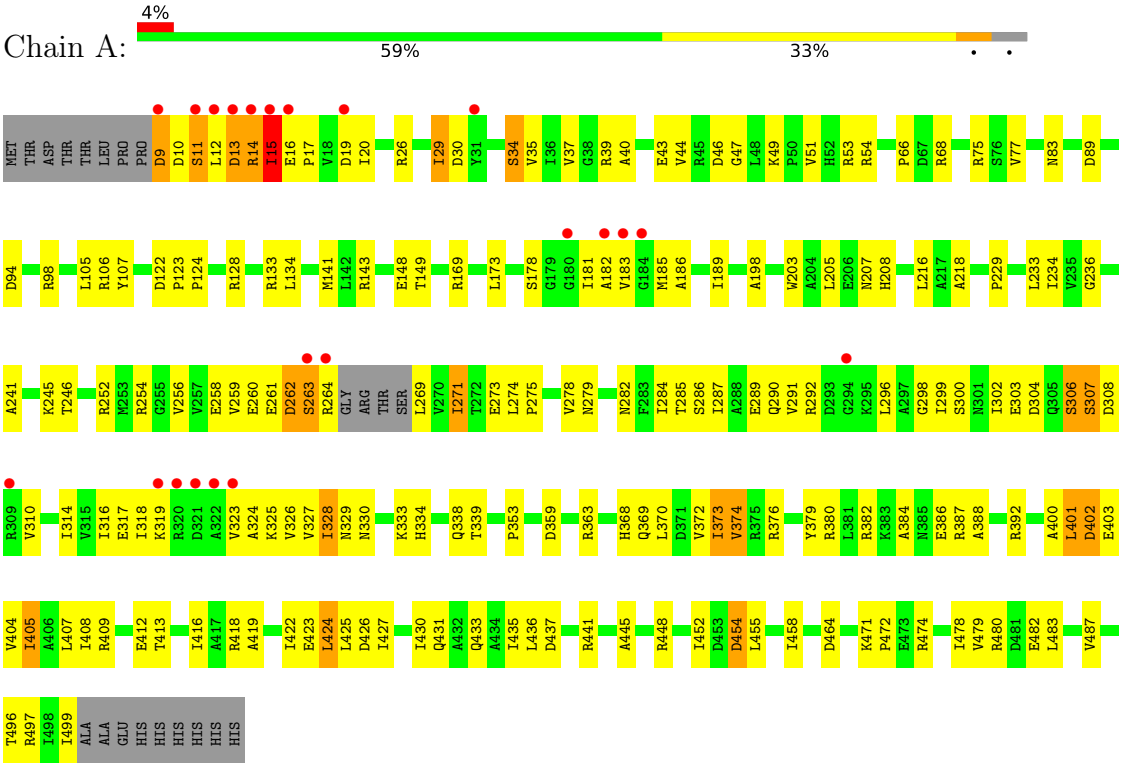
- Molecule 4 is water.

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

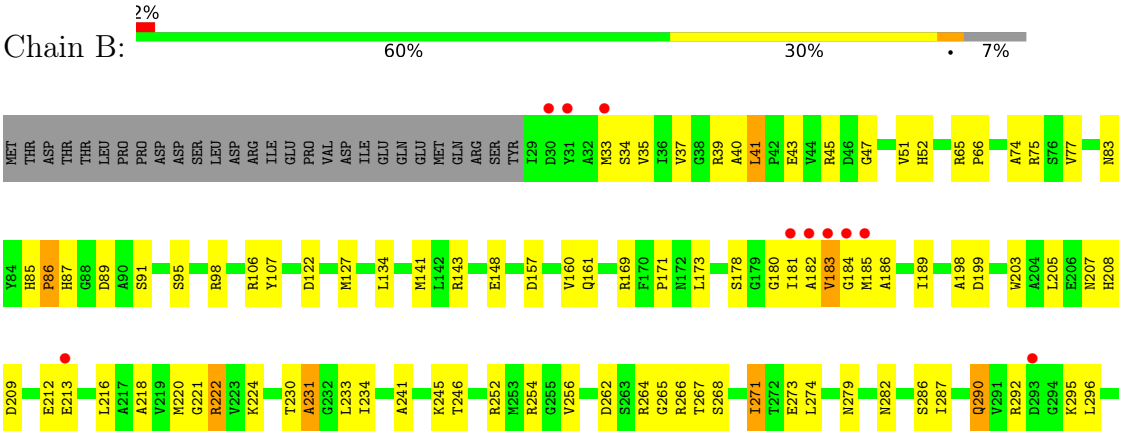
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: DNA gyrase subunit A



• Molecule 1: DNA gyrase subunit A




I302 Q305 S306 S307 D308 G311 L312 R313 R320 V323 A324 K325 V326 V327 I328 L331 Y332 K333 H334 Q338 T339 N344 I348 R357 L358 D359 R363 H368 Q369 I373 R376 T377 T378 Y379 R380 L381 A388 H389 I390 L391 R392 L401 D402 E403 V404
I405 A406 L407 I408 R409 E412 T413 I416 A417 R418 I422 E423 L424 L425 D426 I427 D428 E429 I430 Q431 A432 Q433 L436 L440 R441 A445 L446 E447 R448 I452 D453 D454 L455 I458 I462 I468 K471 P472 E473 R474 V479 R480 D481 E482 E485 I486 V487
D488 R489 T496 R497 A501 GLU HIS HIS HIS HIS HIS HIS

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
$< I/\sigma(I) >$ [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 (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	51.2	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 56.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.009 for -1/2*h+1/2*k+l,1/2*h-1/2*k+l,1/2*h+1/2*k 0.015 for -1/2*h-1/2*k+l,-1/2*h-1/2*k-l,1/2*h-1/2*k	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: MPD, NA

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 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	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

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 23.

The worst 5 of 352 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)
2:B:509:MPD:H12	2:B:509:MPD:C5	1.61	1.22
2:B:509:MPD:H53	2:B:509:MPD:C1	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 (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	483/508 (95%)	449 (93%)	32 (7%)	2 (0%)	34 60
1	B	471/508 (93%)	445 (94%)	22 (5%)	4 (1%)	19 43
All	All	954/1016 (94%)	894 (94%)	54 (6%)	6 (1%)	25 50

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 (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	409/427 (96%)	377 (92%)	32 (8%)	12 29
1	B	391/427 (92%)	364 (93%)	27 (7%)	15 35
All	All	800/854 (94%)	741 (93%)	59 (7%)	13 32

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

Mol	Chain	Res	Type
1	A	403	GLU
1	B	402	ASP
1	B	43	GLU
1	B	401	LEU
1	B	308	ASP

Sometimes 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	238	GLN
1	B	338	GLN
1	B	329	ASN
1	B	344	ASN
1	A	369	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 monosaccharides 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	0.56	0	9,10,10	0.88	0
2	MPD	B	509	-	7,7,7	0.68	0	9,10,10	0.59	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	-	-	1/5/5/5	-
2	MPD	B	509	-	-	2/5/5/5	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	509	MPD	C2-C3-C4-C5
2	B	509	MPD	C2-C3-C4-C5
2	B	509	MPD	C2-C3-C4-O4

There are no ring outliers.

2 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	509	MPD	9	0
2	B	509	MPD	9	0

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

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	487/508 (95%)	-0.05	22 (4%)	33	31	29, 56, 109, 144	0
1	B	473/508 (93%)	-0.27	12 (2%)	57	59	26, 50, 80, 121	0
All	All	960/1016 (94%)	-0.16	34 (3%)	44	44	26, 53, 103, 144	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	182	ALA	4.8
1	A	322	ALA	4.7
1	A	9	ASP	4.6
1	B	182	ALA	4.5
1	A	319	LYS	4.5

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 monosaccharides 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. 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	B-factors(Å^2)	Q<0.9
2	MPD	A	509	8/8	0.82	0.38	96,97,99,99	0
2	MPD	B	509	8/8	0.93	0.26	63,65,67,67	0
3	NA	A	510	1/1	0.96	0.31	54,54,54,54	0
3	NA	B	510	1/1	0.96	0.21	77,77,77,77	0

6.5 Other polymers (i)

There are no such residues in this entry.