WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 22, 2024 – 10:54 PM EDT

PDB ID	:	6J73
Title	:	Crystal structure of IniA from Mycobacterium smegmatis
Authors	:	Wang, M.F.; Guo, X.Y.; Hu, J.J.; Li, J.; Rao, Z.H.
Deposited on	:	2019-01-16
Resolution	:	3.21 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
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
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)	:	2.37.1

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
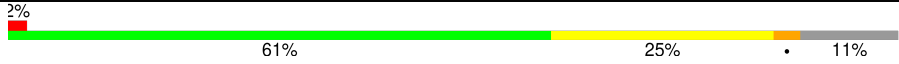*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.21 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.
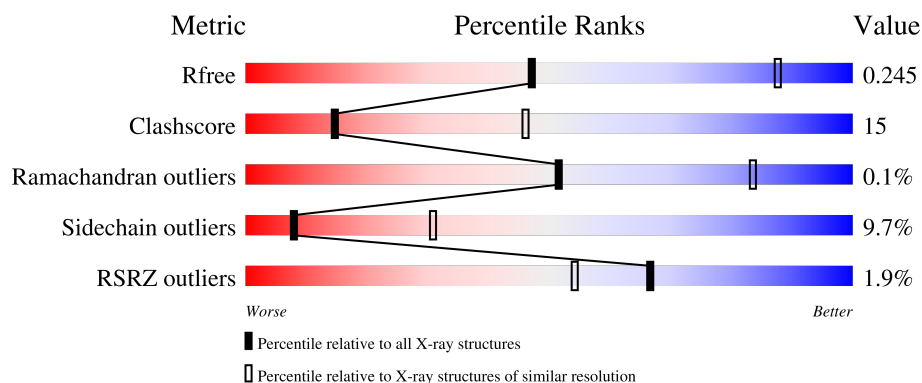
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1335 (3.24-3.20)
Clashscore	141614	1460 (3.24-3.20)
Ramachandran outliers	138981	1437 (3.24-3.20)
Sidechain outliers	138945	1436 (3.24-3.20)
RSRZ outliers	127900	1291 (3.24-3.20)

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	606	2% 61% 25% • 11%
1	B	606	% 65% 21% • 11%

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 8375 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 Isoniazid inducible gene protein IniA.

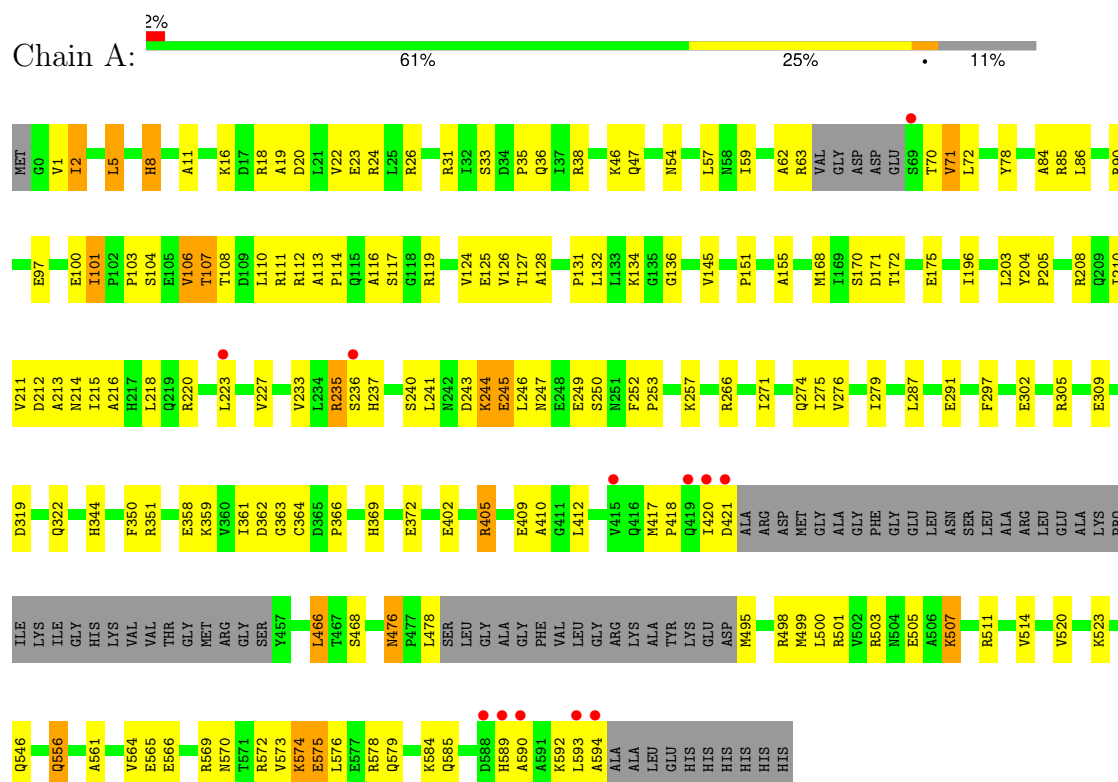| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	539	Total 4179	C 2604	N 767	O 797	S 11	0	1	0
1	B	541	Total 4196	C 2616	N 769	O 800	S 11	0	1	0

There are 24 discrepancies between the modelled and reference sequences:

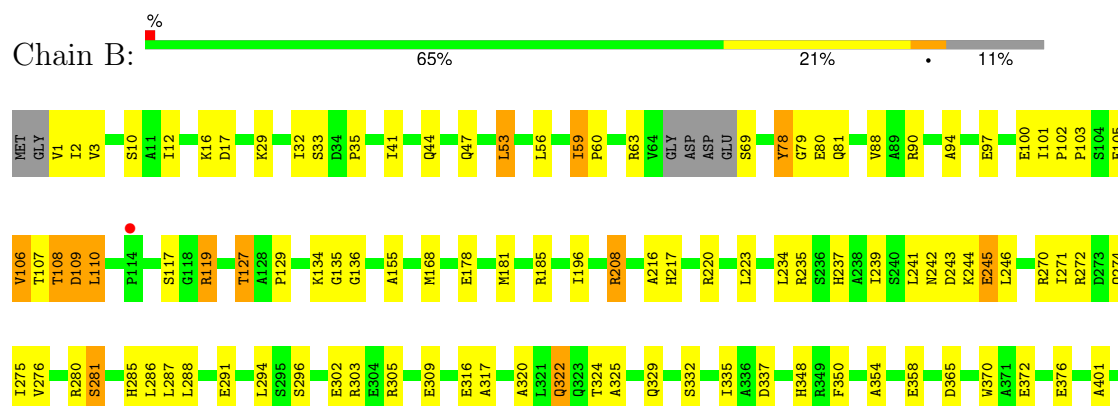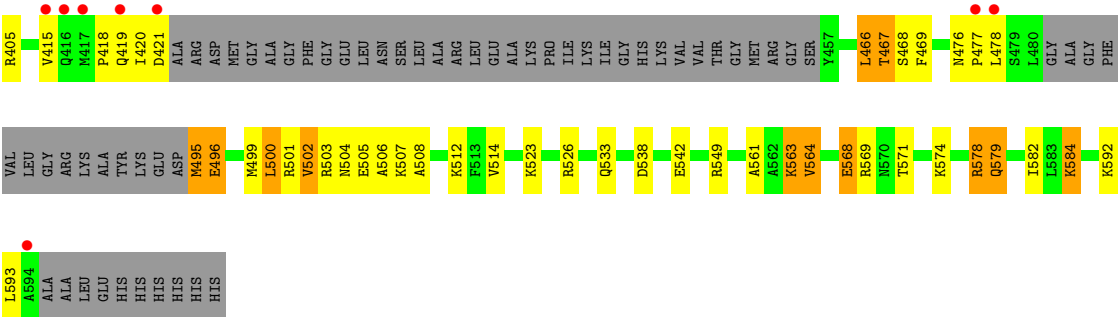| Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	initiating methionine	UNP I7FE16
A	0	GLY	-	expression tag	UNP I7FE16
A	595	ALA	-	expression tag	UNP I7FE16
A	596	ALA	-	expression tag	UNP I7FE16
A	597	LEU	-	expression tag	UNP I7FE16
A	598	GLU	-	expression tag	UNP I7FE16
A	599	HIS	-	expression tag	UNP I7FE16
A	600	HIS	-	expression tag	UNP I7FE16
A	601	HIS	-	expression tag	UNP I7FE16
A	602	HIS	-	expression tag	UNP I7FE16
A	603	HIS	-	expression tag	UNP I7FE16
A	604	HIS	-	expression tag	UNP I7FE16
B	-1	MET	-	initiating methionine	UNP I7FE16
B	0	GLY	-	expression tag	UNP I7FE16
B	595	ALA	-	expression tag	UNP I7FE16
B	596	ALA	-	expression tag	UNP I7FE16
B	597	LEU	-	expression tag	UNP I7FE16
B	598	GLU	-	expression tag	UNP I7FE16
B	599	HIS	-	expression tag	UNP I7FE16
B	600	HIS	-	expression tag	UNP I7FE16
B	601	HIS	-	expression tag	UNP I7FE16
B	602	HIS	-	expression tag	UNP I7FE16
B	603	HIS	-	expression tag	UNP I7FE16
B	604	HIS	-	expression tag	UNP I7FE16

3 Residue-property plots ⓘ

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: Isoniazid inducible gene protein IniA

• Molecule 1: Isoniazid inducible gene protein IniA

R405 V415 Q416 M417 P418 Q419 I420 D421 ALA ARG ASP MET GLY ALA GLY PHE GLY GLU LEU ASN SER LEU ALA ARG LEU GLU ALA LYS PRO ILE LYS ILE GLY HIS LYS VAL VAL THR GLY MET ARG GLY SER Y457 L466 T467 S468 F469 N476 P477 L478 S479 L480 GLY ALA GLY PHE

VAL LEU GLY ARG LYS ALA TYR LYS GLU ASP M495 E496 M499 L500 R501 V502 R503 N504 E505 A506 K507 A508 K512 F513 V514 K523 R526 Q533 D538 E542 R549 A561 A562 K563 V564 E568 R569 N570 T571 K574 R578 Q579 I582 L583 K584 K592

L593 A594 ALA ALA LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	91.76Å 91.76Å 137.35Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.73 – 3.21 39.73 – 3.21	Depositor EDS
% Data completeness (in resolution range)	99.8 (39.73-3.21) 99.9 (39.73-3.21)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.67 (at 3.18Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.219 , 0.258 0.217 , 0.245	Depositor DCC
R_{free} test set	973 reflections (4.61%)	wwPDB-VP
Wilson B-factor (Å^2)	54.3	Xtriage
Anisotropy	0.329	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 57.7	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.033 for -h,-k,l 0.049 for h,-h-k,-l 0.028 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	8375	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.32% 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 ⓘ

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	0.55	0/4241	0.75	0/5746
1	B	0.58	0/4258	0.75	0/5770
All	All	0.56	0/8499	0.75	0/11516

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	4179	0	4221	135	0
1	B	4196	0	4243	125	0
All	All	8375	0	8464	257	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 15.

The worst 5 of 257 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:288:LEU:HD13	1:B:415:VAL:O	1.29	1.30
1:B:41:ILE:CD1	1:B:53:LEU:HD22	1.68	1.24

Continued on next page...

WORLDWIDE wwPDB PROTEIN DATA BANK

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:103:PRO:O	1:B:106:VAL:HG23	1.42	1.20
1:B:41:ILE:HD12	1:B:53:LEU:HD22	1.19	1.18
1:B:103:PRO:O	1:B:106:VAL:CG2	1.96	1.12

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	532/606 (88%)	511 (96%)	20 (4%)	1 (0%)	47	79
1	B	534/606 (88%)	516 (97%)	18 (3%)	0	100	100
All	All	1066/1212 (88%)	1027 (96%)	38 (4%)	1 (0%)	51	83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	245	GLU

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	443/491 (90%)	403 (91%)	40 (9%)	9	34
1	B	446/491 (91%)	400 (90%)	46 (10%)	7	28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	889/982 (90%)	803 (90%)	86 (10%)	8 31

5 of 86 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	241	LEU
1	B	469	PHE
1	B	245	GLU
1	B	419	GLN
1	B	500	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	369	HIS
1	A	546	GLN
1	B	36	GLN
1	B	217	HIS
1	B	242	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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	539/606 (88%)	0.05	12 (2%)	62	49	21, 48, 94, 118	0
1	B	541/606 (89%)	0.07	9 (1%)	70	58	23, 53, 96, 132	0
All	All	1080/1212 (89%)	0.06	21 (1%)	66	54	21, 50, 95, 132	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	421	ASP	7.0
1	A	420	ILE	4.1
1	A	415	VAL	3.4
1	B	417	MET	3.3
1	B	477	PRO	3.3

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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.