



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:24 PM GMT

PDB ID : 4I9G
Title : Crystal structure of glycerol phosphate phosphatase Rv1692 from Mycobacterium tuberculosis in complex with magnesium
Authors : Biswas, T.; Larrouy-Maumus, G.; de Carvalho, L.P.; Tsodikov, O.V.
Deposited on : 2012-12-05
Resolution : 3.25 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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

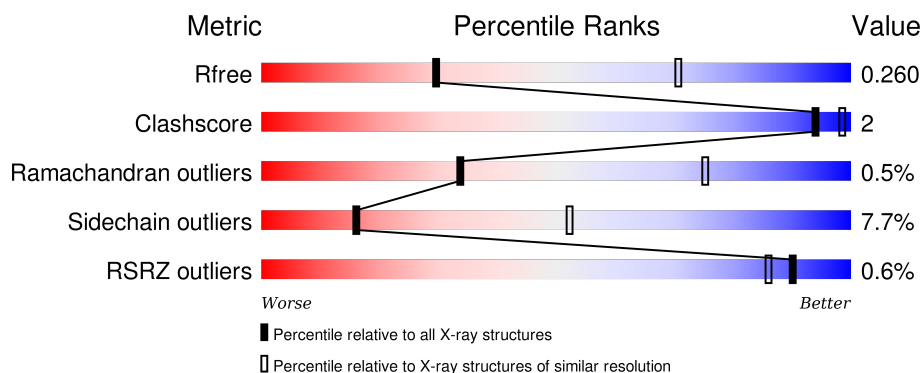
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.25 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	1624 (3.32-3.20)
Clashscore	102246	1806 (3.32-3.20)
Ramachandran outliers	100387	1773 (3.32-3.20)
Sidechain outliers	100360	1771 (3.32-3.20)
RSRZ outliers	91569	1632 (3.32-3.20)

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 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 $\leq 5\%$. 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	361	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 83%, green 8%, grey 9%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 83% 8% 9% </div> </div>
1	B	361	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 81%, green 8%, grey 10%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 81% 8% 10% </div> </div>

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	MG	A	400	-	-	-	X
2	MG	B	400	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4787 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 Glycerol 3-phosphate phosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	329	Total	C	N	O	S	0	0	0
			2395	1487	440	458	10			
1	B	326	Total	C	N	O	S	0	0	0
			2388	1481	441	456	10			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	EXPRESSION TAG	UNP O33194
A	-6	ALA	-	EXPRESSION TAG	UNP O33194
A	-5	HIS	-	EXPRESSION TAG	UNP O33194
A	-4	HIS	-	EXPRESSION TAG	UNP O33194
A	-3	HIS	-	EXPRESSION TAG	UNP O33194
A	-2	HIS	-	EXPRESSION TAG	UNP O33194
A	-1	HIS	-	EXPRESSION TAG	UNP O33194
A	0	HIS	-	EXPRESSION TAG	UNP O33194
B	-7	MET	-	EXPRESSION TAG	UNP O33194
B	-6	ALA	-	EXPRESSION TAG	UNP O33194
B	-5	HIS	-	EXPRESSION TAG	UNP O33194
B	-4	HIS	-	EXPRESSION TAG	UNP O33194
B	-3	HIS	-	EXPRESSION TAG	UNP O33194
B	-2	HIS	-	EXPRESSION TAG	UNP O33194
B	-1	HIS	-	EXPRESSION TAG	UNP O33194
B	0	HIS	-	EXPRESSION TAG	UNP O33194

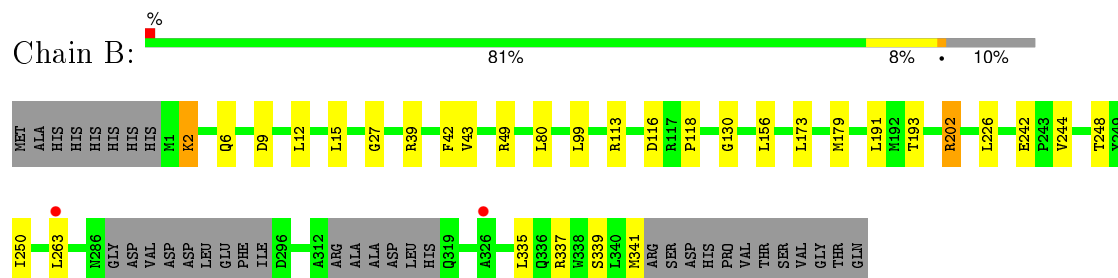
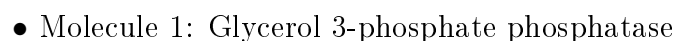
- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		

- Molecule 3 is water.

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

- Molecule 1: Glycerol 3-phosphate phosphatase



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	114.37Å 71.20Å 94.62Å 90.00° 105.52° 90.00°	Depositor
Resolution (Å)	40.00 – 3.25 46.97 – 3.24	Depositor EDS
% Data completeness (in resolution range)	99.8 (40.00-3.25) 99.4 (46.97-3.24)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.28 (at 3.25Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.220 , 0.272 0.208 , 0.260	Depositor DCC
R_{free} test set	555 reflections (4.98%)	DCC
Wilson B-factor (Å ²)	89.7	Xtriage
Anisotropy	0.061	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 46.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 11755 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4787	wwPDB-VP
Average B, all atoms (Å ²)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.96% 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, 0.375 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: MG

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	0.27	0/2432	0.52	0/3315
1	B	0.28	0/2425	0.51	0/3304
All	All	0.27	0/4857	0.51	0/6619

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	2395	0	2388	8	0
1	B	2388	0	2389	10	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
All	All	4787	0	4777	17	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 2.

All (17) 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:2:LYS:HB3	1:B:6:GLN:OE1	1.79	0.82
1:B:2:LYS:CE	1:B:2:LYS:HA	2.19	0.72
1:B:2:LYS:HA	1:B:2:LYS:HE2	1.76	0.68
1:B:116:ASP:O	1:B:118:PRO:HD3	1.96	0.65
1:A:116:ASP:O	1:A:118:PRO:HD3	1.97	0.63
1:A:152:VAL:O	1:A:152:VAL:HG12	2.05	0.56
1:A:266:VAL:HG11	1:A:310:TRP:CZ3	2.47	0.49
1:B:9:ASP:OD2	1:B:202:ARG:N	2.43	0.47
1:B:335:LEU:HD13	1:B:341:MET:CE	2.45	0.46
1:B:15:LEU:HD21	1:B:42:PHE:HB3	1.97	0.46
1:A:130:GLY:HA2	1:B:130:GLY:HA2	1.97	0.46
1:A:229:LEU:HD21	1:A:250:ILE:HD11	1.99	0.45
1:B:43:VAL:HG13	1:B:191:LEU:HD21	1.99	0.44
1:A:10:CYS:HB2	1:A:39:ARG:HB3	2.01	0.42
1:A:15:LEU:HD23	1:A:44:THR:HB	2.03	0.41
1:B:335:LEU:HD13	1:B:341:MET:HE1	2.02	0.41
1:A:339:SER:O	1:A:341:MET:N	2.54	0.41

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	323/361 (90%)	313 (97%)	9 (3%)	1 (0%)	46	83
1	B	320/361 (89%)	301 (94%)	17 (5%)	2 (1%)	30	72
All	All	643/722 (89%)	614 (96%)	26 (4%)	3 (0%)	34	75

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	340	LEU
1	B	339	SER
1	B	27	GLY

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	240/272 (88%)	222 (92%)	18 (8%)	17	53
1	B	241/272 (89%)	222 (92%)	19 (8%)	15	51
All	All	481/544 (88%)	444 (92%)	37 (8%)	16	52

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	CYS
1	A	12	LEU
1	A	21	CYS
1	A	39	ARG
1	A	49	ARG
1	A	80	LEU
1	A	99	LEU
1	A	109	ARG
1	A	113	ARG
1	A	156	LEU
1	A	173	LEU
1	A	198	ARG
1	A	226	LEU
1	A	242	GLU
1	A	248	THR
1	A	286	ASN
1	A	322	LEU
1	A	337	ARG
1	B	2	LYS
1	B	12	LEU
1	B	39	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	49	ARG
1	B	80	LEU
1	B	99	LEU
1	B	113	ARG
1	B	156	LEU
1	B	173	LEU
1	B	179	MET
1	B	193	THR
1	B	202	ARG
1	B	226	LEU
1	B	242	GLU
1	B	244	VAL
1	B	248	THR
1	B	250	ILE
1	B	263	LEU
1	B	337	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	101	ASN
1	A	151	ASN
1	A	286	ASN
1	B	101	ASN
1	B	273	GLN
1	B	286	ASN

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

Of 2 ligands modelled in this entry, 2 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	329/361 (91%)	-0.08	2 (0%) 90 86	76, 94, 124, 143	0
1	B	326/361 (90%)	0.02	2 (0%) 90 86	80, 110, 134, 154	0
All	All	655/722 (90%)	-0.03	4 (0%) 90 86	76, 101, 132, 154	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	288	ASP	4.0
1	B	326	ALA	2.4
1	B	263	LEU	2.2
1	A	271	GLY	2.1

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. 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	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	MG	A	400	1/1	0.96	0.35	5.35	76,76,76,76	0
2	MG	B	400	1/1	0.97	0.31	4.53	98,98,98,98	0

6.5 Other polymers [i](#)

There are no such residues in this entry.