



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 16, 2018 – 01:13 PM EST

PDB ID : 4JBF
Title : Crystal structure of peptidoglycan glycosyltransferase from *Atopobium parvulum* DSM 20469.
Authors : Filippova, E.V.; Wawrzak, Z.; Minasov, G.; Shuvalova, L.; Kiryukhina, O.; Babnigg, G.; Rubin, E.; Sacchettini, J.; Joachimiak, A.; Anderson, W.F.; Midwest Center for Structural Genomics (MCSG); Structures of Mtb Proteins Conferring Susceptibility to Known Mtb Inhibitors (MTBI)
Deposited on : 2013-02-19
Resolution : 1.92 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030736
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20030736

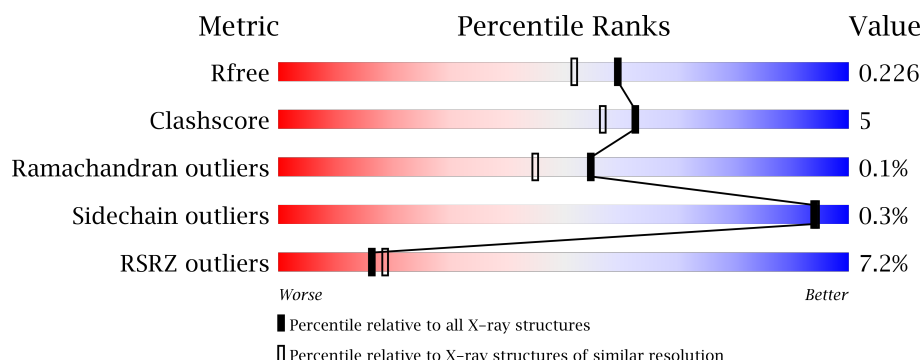
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.92 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}	100719	6276 (1.94-1.90)
Clashscore	112137	7025 (1.94-1.90)
Ramachandran outliers	110173	6947 (1.94-1.90)
Sidechain outliers	110143	6948 (1.94-1.90)
RSRZ outliers	101464	6332 (1.94-1.90)

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	482	
1	B	482	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6378 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 Peptidoglycan glycosyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	410	Total	C	N	O	S	Se	0	1	0
			2948	1833	499	601	2	13			
1	B	410	Total	C	N	O	S	Se	0	2	0
			2953	1832	501	605	2	13			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	473	MSE	-	EXPRESSION TAG	UNP C8W8H7
A	474	HIS	-	EXPRESSION TAG	UNP C8W8H7
A	475	HIS	-	EXPRESSION TAG	UNP C8W8H7
A	476	HIS	-	EXPRESSION TAG	UNP C8W8H7
A	477	HIS	-	EXPRESSION TAG	UNP C8W8H7
A	478	HIS	-	EXPRESSION TAG	UNP C8W8H7
A	479	HIS	-	EXPRESSION TAG	UNP C8W8H7
A	480	SER	-	EXPRESSION TAG	UNP C8W8H7
A	481	SER	-	EXPRESSION TAG	UNP C8W8H7
A	482	GLY	-	EXPRESSION TAG	UNP C8W8H7
A	483	VAL	-	EXPRESSION TAG	UNP C8W8H7
A	484	ASP	-	EXPRESSION TAG	UNP C8W8H7
A	485	LEU	-	EXPRESSION TAG	UNP C8W8H7
A	486	TRP	-	EXPRESSION TAG	UNP C8W8H7
A	487	SER	-	EXPRESSION TAG	UNP C8W8H7
A	488	HIS	-	EXPRESSION TAG	UNP C8W8H7
A	489	PRO	-	EXPRESSION TAG	UNP C8W8H7
A	490	GLN	-	EXPRESSION TAG	UNP C8W8H7
A	491	PHE	-	EXPRESSION TAG	UNP C8W8H7
A	492	GLU	-	EXPRESSION TAG	UNP C8W8H7
A	493	LYS	-	EXPRESSION TAG	UNP C8W8H7
A	494	GLY	-	EXPRESSION TAG	UNP C8W8H7
A	495	THR	-	EXPRESSION TAG	UNP C8W8H7
A	496	GLU	-	EXPRESSION TAG	UNP C8W8H7
A	497	ASN	-	EXPRESSION TAG	UNP C8W8H7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	498	LEU	-	EXPRESSION TAG	UNP C8W8H7
A	499	TYR	-	EXPRESSION TAG	UNP C8W8H7
A	500	PHE	-	EXPRESSION TAG	UNP C8W8H7
A	501	GLN	-	EXPRESSION TAG	UNP C8W8H7
A	502	SER	-	EXPRESSION TAG	UNP C8W8H7
A	503	ASN	-	EXPRESSION TAG	UNP C8W8H7
A	504	ALA	-	EXPRESSION TAG	UNP C8W8H7
B	473	MSE	-	EXPRESSION TAG	UNP C8W8H7
B	474	HIS	-	EXPRESSION TAG	UNP C8W8H7
B	475	HIS	-	EXPRESSION TAG	UNP C8W8H7
B	476	HIS	-	EXPRESSION TAG	UNP C8W8H7
B	477	HIS	-	EXPRESSION TAG	UNP C8W8H7
B	478	HIS	-	EXPRESSION TAG	UNP C8W8H7
B	479	HIS	-	EXPRESSION TAG	UNP C8W8H7
B	480	SER	-	EXPRESSION TAG	UNP C8W8H7
B	481	SER	-	EXPRESSION TAG	UNP C8W8H7
B	482	GLY	-	EXPRESSION TAG	UNP C8W8H7
B	483	VAL	-	EXPRESSION TAG	UNP C8W8H7
B	484	ASP	-	EXPRESSION TAG	UNP C8W8H7
B	485	LEU	-	EXPRESSION TAG	UNP C8W8H7
B	486	TRP	-	EXPRESSION TAG	UNP C8W8H7
B	487	SER	-	EXPRESSION TAG	UNP C8W8H7
B	488	HIS	-	EXPRESSION TAG	UNP C8W8H7
B	489	PRO	-	EXPRESSION TAG	UNP C8W8H7
B	490	GLN	-	EXPRESSION TAG	UNP C8W8H7
B	491	PHE	-	EXPRESSION TAG	UNP C8W8H7
B	492	GLU	-	EXPRESSION TAG	UNP C8W8H7
B	493	LYS	-	EXPRESSION TAG	UNP C8W8H7
B	494	GLY	-	EXPRESSION TAG	UNP C8W8H7
B	495	THR	-	EXPRESSION TAG	UNP C8W8H7
B	496	GLU	-	EXPRESSION TAG	UNP C8W8H7
B	497	ASN	-	EXPRESSION TAG	UNP C8W8H7
B	498	LEU	-	EXPRESSION TAG	UNP C8W8H7
B	499	TYR	-	EXPRESSION TAG	UNP C8W8H7
B	500	PHE	-	EXPRESSION TAG	UNP C8W8H7
B	501	GLN	-	EXPRESSION TAG	UNP C8W8H7
B	502	SER	-	EXPRESSION TAG	UNP C8W8H7
B	503	ASN	-	EXPRESSION TAG	UNP C8W8H7
B	504	ALA	-	EXPRESSION TAG	UNP C8W8H7

- Molecule 2 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C₈H₁₈O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	C	O	0	0
			7	4	3		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	218	Total	O	0	4
			222	222		
3	B	247	Total	O	0	1
			248	248		

These plots are drawn for all protein, RNA and DNA 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 ($\text{RSRZ} > 2$). 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.

Chain A:

6% 78% 7% 15%

MSE HIS HIS HIS HIS HIS HIS HIS SER SER GLY VAL ASP LEU TRP SER HIS PRO GLN PHE GLU LYS GLY THR GLU ASN LEU TYR PHE GLN SER ASN ILE ASP ALA PRO ARG LEU GLN LEU LEU PRO THR ASN ASN HIS THR ILE ALA LYS S523 A524 Y525 Y526 E541 S542 V543 I545 G546

D546 D547 T549 Y550 V551 R552 M580 F582 T583 L584 T585 G593 R594 I603 M604 T605 T606 G607 S632 K646 G658 T659 I661 I662 G663 G664 G665 G666 G667 T673 G681 K685 G715 L742 M784 W787 E788

W791 G795 L796 A819 S838 E840 GLY A842 V843 V844 R865 K886 D895 V896 GLU M898 G899 N900 V917 N925 G926 N928 Y932 N946 L950 GLY ALA ALA SER

Chain B:

77% 8% 15%

Label	Value
W737	0.000
W791	0.000
L796	0.000
M812	0.000
A819	0.000
P839	0.000
E840	0.000
M878	0.000
R881	0.000
D885	0.000
V896	0.000
E897	0.000
N898	0.000
G899	0.000
N900	0.000
D912	0.000
V917	0.000
V920	0.000
N925	0.000
G926	0.000
E927	0.000
N928	0.000
R939	0.000
Q943	0.000
N946	0.000
L950	0.000
GLY	0.000
ALA	0.000
ALA	0.000
SER	0.000
LVS	0.000
Q545	0.000
D546	0.000
D547	0.000
GLY	0.000
THR	0.000
THR	0.000
V551	0.000
R552	0.000
S578	0.000
N581	0.000
E582	0.000
T583	0.000
L584	0.000
T585	0.000
GLY	0.000
ALA	0.000
GLU	0.000
ASP	0.000
HIS	0.000
SER	0.000
ASP	0.000
E593	0.000
R594	0.000
M600	0.000
A601	0.000
G602	0.000
I603	0.000
ASN	0.000
T605	0.000
T606	0.000
G607	0.000
S608	0.000
I660	0.000
ILE	0.000
GLU	0.000
SER	0.000
GLY	0.000
GLY	0.000
S667	0.000
I712	0.000
G715	0.000
T716	0.000
H718	0.000
N722	0.000
R759	0.000
Y763	0.000
G764	0.000
S523	0.000
A524	0.000
Y525	0.000
V526	0.000
Q527	0.000
R528	0.000
E541	0.000
S542	0.000
V543	0.000
MSE	0.000
HIS	0.000
HIS	0.000
HIS	0.000
HIS	0.000
HIS	0.000
HIS	0.000
SER	0.000
GLY	0.000
VAL	0.000
ASP	0.000
TRP	0.000
SER	0.000
HIS	0.000
PRO	0.000
GLN	0.000
PHE	0.000
GLU	0.000
LVS	0.000
GLY	0.000
THR	0.000
GLU	0.000
ASN	0.000
LEU	0.000
TYR	0.000
PHE	0.000
GLN	0.000
SER	0.000
ASN	0.000
ALA	0.000
ILE	0.000
ASP	0.000
ALA	0.000
PRO	0.000
ARG	0.000
LEU	0.000
GLN	0.000
ILE	0.000
LEU	0.000
PRO	0.000
THR	0.000
THR	0.000
HIS	0.000
ALA	0.000
ALA	0.000
LVS	0.000

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	66.93Å 70.17Å 114.82Å 90.00° 97.35° 90.00°	Depositor
Resolution (Å)	29.48 – 1.92 29.48 – 1.92	Depositor EDS
% Data completeness (in resolution range)	98.5 (29.48-1.92) 98.4 (29.48-1.92)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.43 (at 1.92Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.195 , 0.237 0.192 , 0.226	Depositor DCC
R_{free} test set	4008 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	27.4	Xtriage
Anisotropy	0.413	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 56.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	6378	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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: PG4

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	0.74	0/2982	0.88	1/4044 (0.0%)
1	B	0.81	1/2985 (0.0%)	0.91	3/4047 (0.1%)
All	All	0.78	1/5967 (0.0%)	0.89	4/8091 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	812	MSE	N-CA	5.59	1.57	1.46

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	759[B]	ARG	NE-CZ-NH1	5.96	123.28	120.30
1	B	759[C]	ARG	NE-CZ-NH1	5.96	123.28	120.30
1	B	920	VAL	CB-CA-C	-5.59	100.77	111.40
1	A	865	ARG	NE-CZ-NH1	5.59	123.09	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	2948	0	2863	22	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2953	0	2874	36	0
2	B	7	0	9	0	0
3	A	222	0	0	4	0
3	B	248	0	0	3	0
All	All	6378	0	5746	58	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 5.

The worst 5 of 58 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:759[C]:ARG:HG2	1:B:759[C]:ARG:HH21	1.27	0.98
1:B:878:MSE:CE	1:B:881:ARG:HH11	1.90	0.84
1:B:759[C]:ARG:HH21	1:B:759[C]:ARG:CG	1.90	0.84
1:B:878:MSE:HE2	1:B:881:ARG:HH11	1.45	0.80
1:A:545:GLN:HA	3:A:1146:HOH:O	1.83	0.77

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	395/482 (82%)	385 (98%)	10 (2%)	0	100	100
1	B	400/482 (83%)	384 (96%)	15 (4%)	1 (0%)	44	33
All	All	795/964 (82%)	769 (97%)	25 (3%)	1 (0%)	55	45

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	925	ASN

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	311/356 (87%)	310 (100%)	1 (0%)	94	94
1	B	311/356 (87%)	310 (100%)	1 (0%)	94	94
All	All	622/712 (87%)	620 (100%)	2 (0%)	94	94

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	900	ASN
1	B	900	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	527	GLN
1	B	925	ASN
1	B	722	ASN
1	A	722	ASN
1	B	570	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

There are no carbohydrates in this entry.

5.6 Ligand geometry

1 ligand is modelled in this entry.

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	PG4	B	1001	-	6,6,12	0.51	0	5,5,11	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	PG4	B	1001	-	-	0/4/4/10	0/0/0/0

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	397/482 (82%)	0.22	28 (7%) 17 19	21, 37, 81, 125	0
1	B	397/482 (82%)	0.13	29 (7%) 16 18	21, 34, 77, 97	0
All	All	794/964 (82%)	0.17	57 (7%) 16 19	21, 35, 79, 125	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	603	ILE	5.8
1	B	525	TYR	5.6
1	A	898	ASN	5.5
1	A	551	VAL	5.1
1	B	898	ASN	5.1

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	PG4	B	1001	7/13	0.86	0.13	1.27	42,44,49,50	0

6.5 Other polymers ⓘ

There are no such residues in this entry.