



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 7, 2017 – 12:37 PM EST

PDB ID : 5GKO  
Title : Crystal structure of tripartite-type ABC transporter, MacB from *Acinetobacter baumannii*  
Authors : Murakami, S.; Okada, U.; Yamashita, E.  
Deposited on : unknown  
Resolution : 3.39 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](mailto: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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
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-20030345

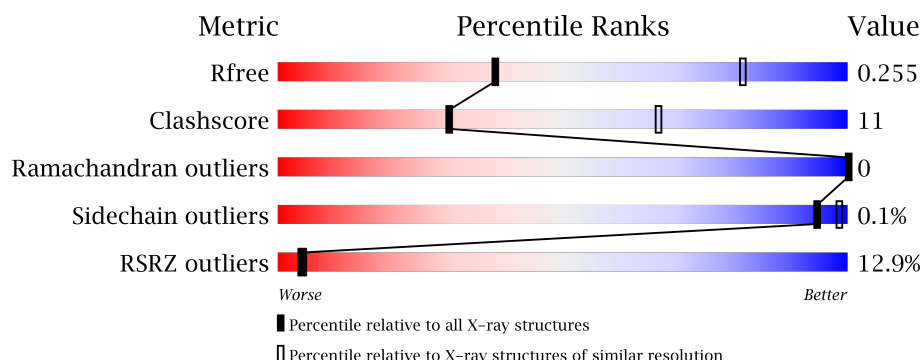
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.39 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	1679 (3.50-3.30)
Clashscore	112137	1832 (3.50-3.30)
Ramachandran outliers	110173	1789 (3.50-3.30)
Sidechain outliers	110143	1789 (3.50-3.30)
RSRZ outliers	101464	1709 (3.50-3.30)

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  $\geq 3$ , 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	671	<div> <div>11%</div> <div> <div></div> <div>71%</div> <div>25%</div> <div>.</div> </div> </div>
1	B	671	<div> <div>13%</div> <div> <div></div> <div>73%</div> <div>24%</div> <div>.</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 9770 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 Macrolide export ATP-binding/permease protein MacB.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	650	Total	C	N	O	S	Se	0	0	0
			4885	3056	854	952	3	20			
1	B	650	Total	C	N	O	S	Se	0	0	0
			4885	3056	854	952	3	20			

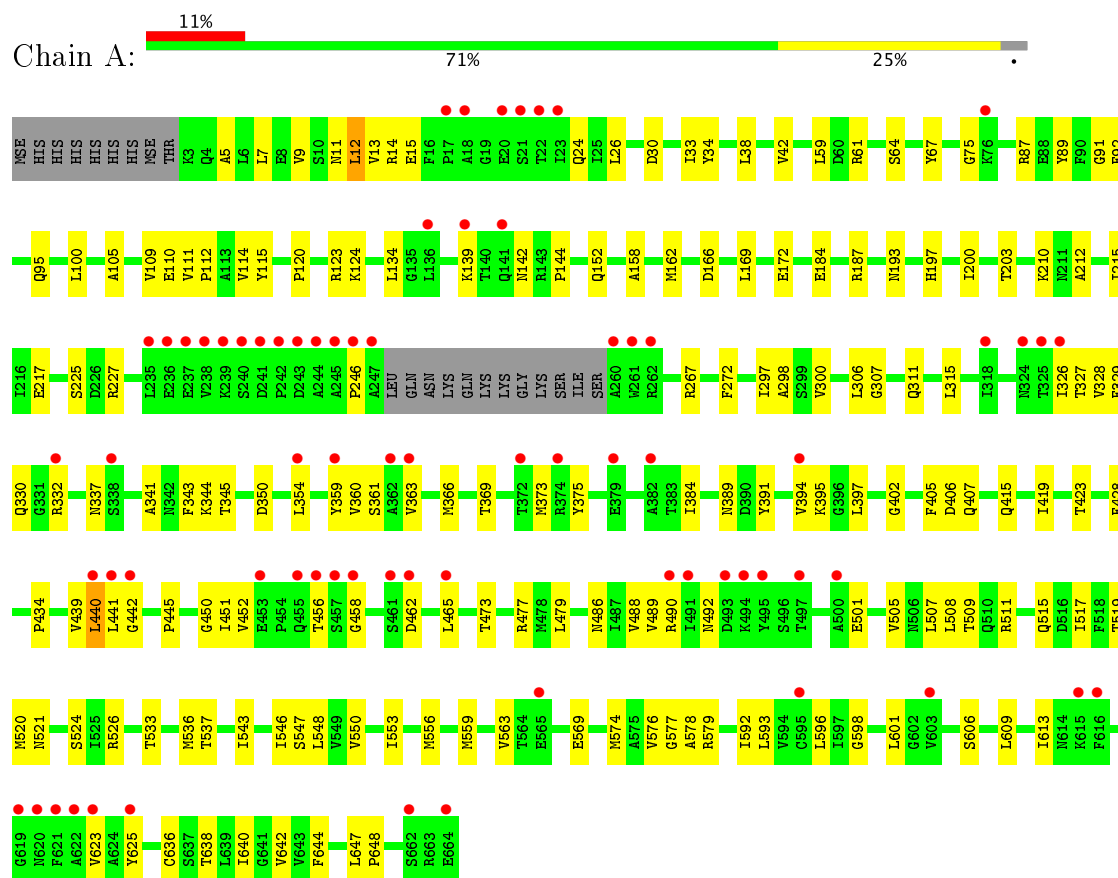
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MSE	-	expression tag	UNP A0A0D8G707
A	-5	HIS	-	expression tag	UNP A0A0D8G707
A	-4	HIS	-	expression tag	UNP A0A0D8G707
A	-3	HIS	-	expression tag	UNP A0A0D8G707
A	-2	HIS	-	expression tag	UNP A0A0D8G707
A	-1	HIS	-	expression tag	UNP A0A0D8G707
A	0	HIS	-	expression tag	UNP A0A0D8G707
B	-6	MSE	-	expression tag	UNP A0A0D8G707
B	-5	HIS	-	expression tag	UNP A0A0D8G707
B	-4	HIS	-	expression tag	UNP A0A0D8G707
B	-3	HIS	-	expression tag	UNP A0A0D8G707
B	-2	HIS	-	expression tag	UNP A0A0D8G707
B	-1	HIS	-	expression tag	UNP A0A0D8G707
B	0	HIS	-	expression tag	UNP A0A0D8G707

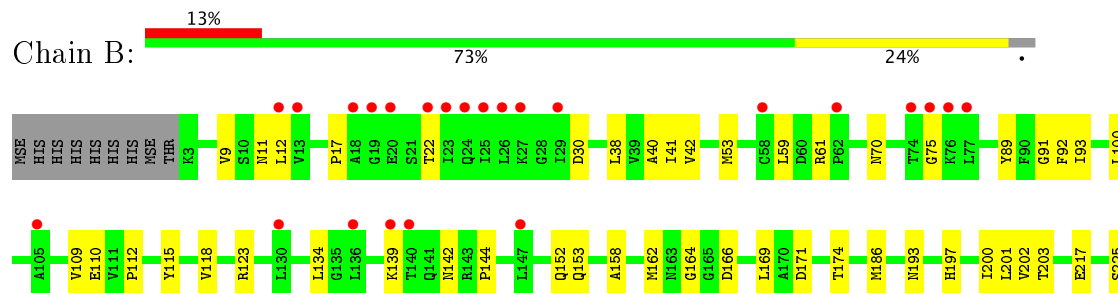
### 3 Residue-property plots [i](#)

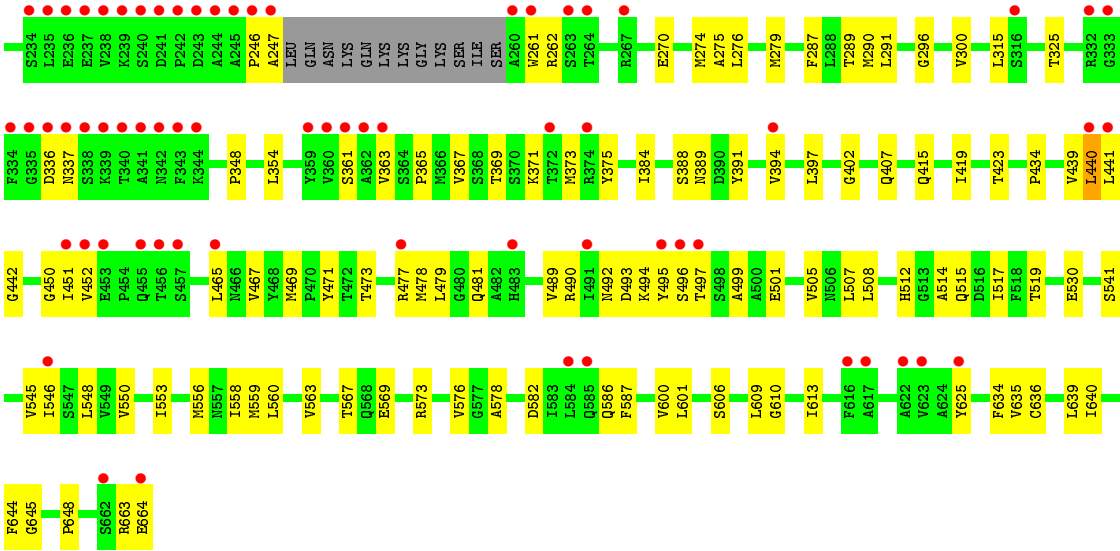
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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 ( $RSRZ > 2$ ). 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: Macrolide export ATP-binding/permease protein MacB



- Molecule 1: Macrolide export ATP-binding/permease protein MacB





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	229.15Å 229.15Å 155.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.96 – 3.39 45.23 – 3.39	Depositor EDS
% Data completeness (in resolution range)	99.2 (43.96-3.39) 99.2 (45.23-3.39)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.64 (at 3.40Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.223 , 0.255 0.220 , 0.255	Depositor DCC
$R_{free}$ test set	2900 reflections (5.08%)	DCC
Wilson B-factor (Å <sup>2</sup> )	135.9	Xtriage
Anisotropy	0.481	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 123.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9770	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	161.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 39.96 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 2.9358e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical 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](#)

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	0.28	0/4925	0.54	2/6643 (0.0%)
1	B	0.28	0/4925	0.54	1/6643 (0.0%)
All	All	0.28	0/9850	0.54	3/13286 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	12	LEU	CA-CB-CG	6.47	130.19	115.30
1	A	440	LEU	CA-CB-CG	6.06	129.23	115.30
1	B	440	LEU	CA-CB-CG	5.91	128.89	115.30

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	4885	0	4973	109	0
1	B	4885	0	4973	105	0
All	All	9770	0	9946	208	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 11.

The worst 5 of 208 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:546:ILE:O	1:B:550:VAL:HB	1.57	1.05
1:A:547:SER:HB3	1:B:548:LEU:HD12	1.60	0.83
1:A:109:VAL:HG12	1:A:162:MSE:HE2	1.62	0.79
1:A:394:VAL:HG23	1:A:490:ARG:HB2	1.65	0.78
1:B:109:VAL:HG12	1:B:162:MSE:HE2	1.63	0.78

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	646/671 (96%)	618 (96%)	28 (4%)	0	100	100
1	B	646/671 (96%)	613 (95%)	33 (5%)	0	100	100
All	All	1292/1342 (96%)	1231 (95%)	61 (5%)	0	100	100

There are no Ramachandran outliers to report.

### 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	538/536 (100%)	537 (100%)	1 (0%)	94	98
1	B	538/536 (100%)	538 (100%)	0	100	100
All	All	1076/1072 (100%)	1075 (100%)	1 (0%)	94	98



All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	405	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	330	GLN
1	B	512	HIS
1	B	153	GLN
1	A	193	ASN
1	B	193	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 [i](#)

There are no ligands in this entry.

## 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	630/671 (93%)	0.70	72 (11%) 6 6	98, 144, 239, 333	0
1	B	630/671 (93%)	0.84	90 (14%) 3 3	109, 156, 255, 334	0
All	All	1260/1342 (93%)	0.77	162 (12%) 4 4	98, 150, 249, 334	0

The worst 5 of 162 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	243	ASP	12.2
1	A	261	TRP	10.3
1	A	243	ASP	10.3
1	A	244	ALA	8.5
1	A	239	LYS	8.5

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.