



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:38 PM GMT

PDB ID : 4PK1
Title : Structure of BamB fused to a BamA POTRA domain fragment
Authors : Jansen, K.B.; Sousa, M.C.
Deposited on : 2014-05-13
Resolution : 3.10 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 (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

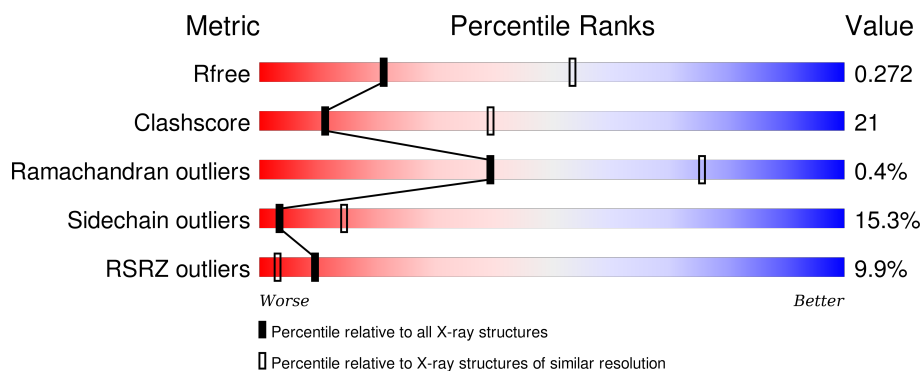
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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	645	<div> <div>8%</div> <div>47%</div> <div>24%</div> <div>• •</div> <div>23%</div> </div>

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 7451 atoms, of which 3689 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 Chimera protein of Outer membrane protein assembly factors BamA and BamB.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	494	Total	C	H	N	O	S	0	0	0
			7451	2357	3689	648	749	8			

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	171	GLY	-	expression tag	UNP P0A940
A	172	ALA	-	expression tag	UNP P0A940
A	173	MET	-	expression tag	UNP P0A940
A	174	GLU	-	expression tag	UNP P0A940
A	1002	PRO	-	linker	UNP P0A940
A	1003	GLY	-	linker	UNP P0A940
A	1004	ALA	-	linker	UNP P0A940
A	1005	SER	-	linker	UNP P0A940
A	1006	GLY	-	linker	UNP P0A940
A	1007	GLY	-	linker	UNP P0A940
A	1008	SER	-	linker	UNP P0A940
A	1009	GLY	-	linker	UNP P0A940
A	1010	GLY	-	linker	UNP P0A940
A	1011	GLY	-	linker	UNP P0A940
A	1012	SER	-	linker	UNP P0A940
A	1013	GLY	-	linker	UNP P0A940
A	1014	GLY	-	linker	UNP P0A940
A	1015	GLY	-	linker	UNP P0A940
A	1016	SER	-	linker	UNP P0A940
A	1017	GLY	-	linker	UNP P0A940
A	1018	GLY	-	linker	UNP P0A940
A	1019	THR	-	linker	UNP P0A940
A	1020	ALA	-	linker	UNP P0A940

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	49.27Å 91.41Å 61.21Å 90.00° 92.85° 90.00°	Depositor
Resolution (Å)	19.89 – 3.10 19.89 – 3.10	Depositor EDS
% Data completeness (in resolution range)	98.9 (19.89-3.10) 99.1 (19.89-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.07 (at 3.09Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.251 , 0.281 0.265 , 0.272	Depositor DCC
R_{free} test set	525 reflections (5.66%)	DCC
Wilson B-factor (Å ²)	68.2	Xtriage
Anisotropy	0.961	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.44 , 63.3	EDS
Estimated twinning fraction	0.062 for h,-k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 9797 reflections	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	7451	wwPDB-VP
Average B, all atoms (Å ²)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.54% 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

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	0.50	0/3822	0.71	8/5192 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	4

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	1194	LEU	CA-CB-CG	8.03	133.76	115.30
1	A	1035	LEU	C-N-CD	5.94	140.87	128.40
1	A	1043	THR	C-N-CD	5.79	140.55	128.40
1	A	1194	LEU	N-CA-C	-5.56	95.98	111.00
1	A	1190	PRO	CA-N-CD	-5.47	103.84	111.50

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1040	ASN	Peptide
1	A	1194	LEU	Peptide
1	A	1195	ARG	Peptide
1	A	1284	GLY	Peptide

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	3762	3689	3690	154	0
All	All	3762	3689	3690	154	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 21.

The worst 5 of 154 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:A:1035:LEU:HD12	1:A:1325:ARG:NH2	1.54	1.22
1:A:1261:LEU:CD2	1:A:1284:GLY:HA2	1.95	0.96
1:A:1158:SER:CB	1:A:1202:THR:HG21	1.96	0.95
1:A:1327:LEU:HD23	1:A:1341:ASP:HB3	1.48	0.94
1:A:1271:ASP:OD1	1:A:1274:SER:OG	1.84	0.94

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	482/645 (75%)	430 (89%)	50 (10%)	2 (0%)	39 75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1188	ASP
1	A	1286	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	406/529 (77%)	344 (85%)	62 (15%)	3 14

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1088	ASP
1	A	1170	GLN
1	A	1364	SER
1	A	1154	ARG
1	A	1185	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1070	ASN
1	A	1186	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, 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	494/645 (76%)	0.57	49 (9%) 9 3	46, 76, 111, 141	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	281	SER	5.7
1	A	1032	MET	4.0
1	A	286	GLN	3.9
1	A	269	SER	3.7
1	A	282	ALA	3.7

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.