



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 19, 2016 – 08:57 PM EDT

PDB ID : 5D0O  
Title : BamABCDE complex, outer membrane beta barrel assembly machinery entire complex  
Authors : Gu, Y.; Paterson, N.; Zeng, Y.; Dong, H.; Wang, W.; Dong, C.  
Deposited on : 2015-08-03  
Resolution : 2.90 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.

---

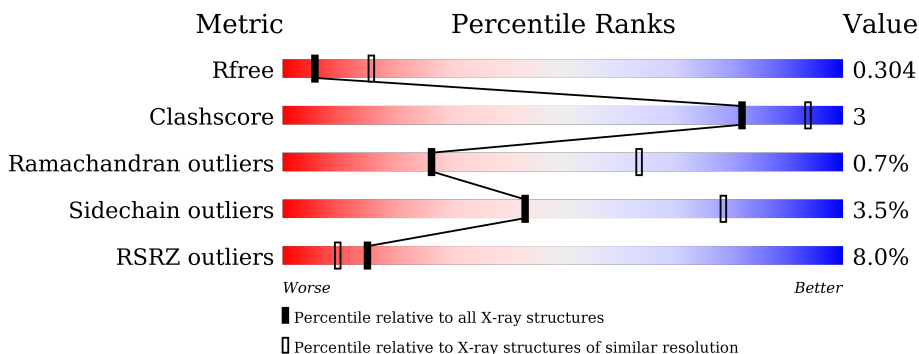
The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027107  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0122  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027107

**i**

## X-RAY DIFFRACTION

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.



<b>Metric</b>	<b>Whole archive (#Entries)</b>	<b>Similar resolution (#Entries, resolution range(Å))</b>
R <sub>free</sub>	91344	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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  $\geq 3$ , 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	810	<p>6% 88% 8% . .</p>
2	B	392	<p>12% 80% 9% . 10%</p>
3	C	344	<p>3% 14% . 84%</p>
4	D	245	<p>2% 78% 7% 15%</p>
5	E	123	<p>7% 46% 19% . 33%</p>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 22815 atoms, of which 11201 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 Outer membrane protein assembly factor BamA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	786	Total	C	H	N	O	S	0	0	0
			12142	3919	5928	1047	1232	16			

- Molecule 2 is a protein called Outer membrane protein assembly factor BamB.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	354	Total	C	H	N	O	S	0	0	0
			5280	1675	2615	457	527	6			

- Molecule 3 is a protein called Outer membrane protein assembly factor BamC.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	56	Total	C	H	N	O	S	0	0	0
			817	258	409	71	78	1			

- Molecule 4 is a protein called Outer membrane protein assembly factor BamD.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	D	208	Total	C	H	N	O	S	0	0	0
			3306	1057	1624	296	322	7			

- Molecule 5 is a protein called Outer membrane protein assembly factor BamE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
5	E	83	Total	C	H	N	O	S	0	0	0
			1270	405	625	112	126	2			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	114	GLY	-	expression tag	UNP P0A937
E	115	GLY	-	expression tag	UNP P0A937

*Continued on next page...*

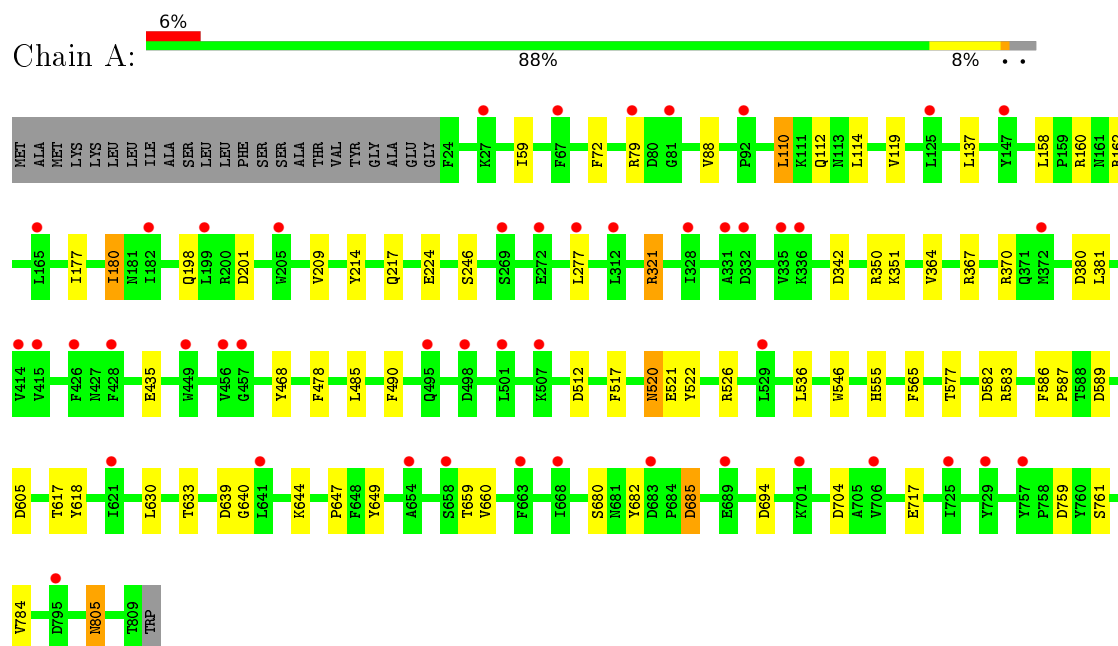
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	116	HIS	-	expression tag	UNP P0A937
E	117	HIS	-	expression tag	UNP P0A937
E	118	HIS	-	expression tag	UNP P0A937
E	119	HIS	-	expression tag	UNP P0A937
E	120	HIS	-	expression tag	UNP P0A937
E	121	HIS	-	expression tag	UNP P0A937
E	122	HIS	-	expression tag	UNP P0A937
E	123	HIS	-	expression tag	UNP P0A937

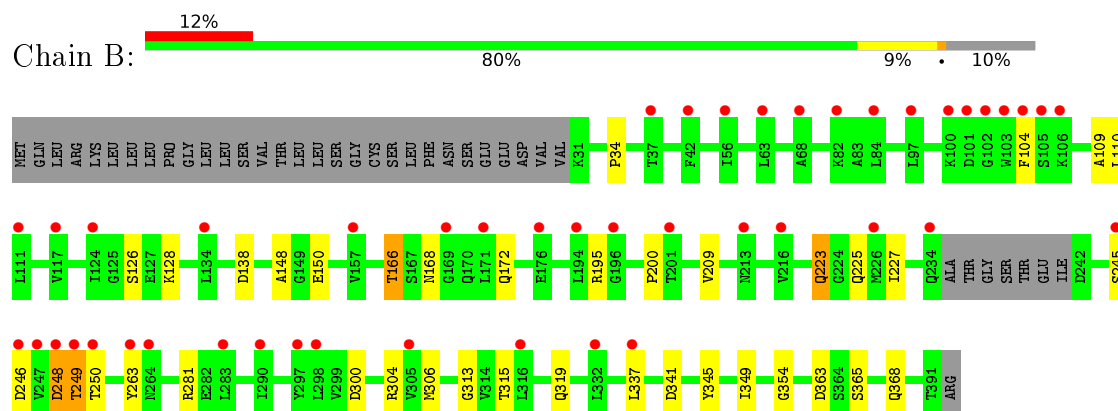
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Outer membrane protein assembly factor BamA



- Molecule 2: Outer membrane protein assembly factor BamB



- Molecule 3: Outer membrane protein assembly factor BamC





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	116.69Å 116.69Å 435.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.65 – 2.90 49.65 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.65-2.90) 100.0 (49.65-2.75)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	0.02	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.16 (at 2.77Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.276 , 0.304 0.276 , 0.304	Depositor DCC
$R_{free}$ test set	3222 reflections (4.77%)	DCC
Wilson B-factor (Å <sup>2</sup> )	84.2	Xtriage
Anisotropy	0.560	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 64.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	4 of 79346 reflections (0.005%)	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	22815	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	128.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% of the height of the origin peak. No significant pseudotranslation is detected.*

<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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.24	0/6357	0.44	0/8624
2	B	0.23	0/2714	0.47	0/3700
3	C	0.22	0/417	0.44	0/569
4	D	0.24	0/1719	0.41	0/2336
5	E	0.25	0/659	0.50	0/899
All	All	0.24	0/11866	0.44	0/16128

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 [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	6214	5928	5928	36	0
2	B	2665	2615	2615	19	0
3	C	408	409	409	3	0
4	D	1682	1624	1624	10	0
5	E	645	625	625	7	0
All	All	11614	11201	11201	72	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 3.

The worst 5 of 72 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:160:ARG:NH1	4:D:61:ARG:O	2.11	0.83
1:A:364:VAL:O	1:A:367:ARG:NH1	2.15	0.79
1:A:217:GLN:OE1	1:A:217:GLN:N	2.25	0.69
1:A:582:ASP:OD2	1:A:589:ASP:N	2.30	0.65
1:A:605:ASP:O	1:A:644:LYS:NZ	2.32	0.62

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	784/810 (97%)	736 (94%)	45 (6%)	3 (0%)	39	74
2	B	350/392 (89%)	318 (91%)	28 (8%)	4 (1%)	17	51
3	C	54/344 (16%)	46 (85%)	6 (11%)	2 (4%)	4	17
4	D	204/245 (83%)	196 (96%)	8 (4%)	0	100	100
5	E	81/123 (66%)	74 (91%)	6 (7%)	1 (1%)	16	48
All	All	1473/1914 (77%)	1370 (93%)	93 (6%)	10 (1%)	26	63

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	249	THR
3	C	38	ASP
5	E	41	ASN
3	C	77	GLY
1	A	277	LEU

### 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	670/688 (97%)	651 (97%)	19 (3%)	51	84
2	B	287/321 (89%)	281 (98%)	6 (2%)	61	88
3	C	41/276 (15%)	41 (100%)	0	100	100
4	D	176/204 (86%)	174 (99%)	2 (1%)	80	95
5	E	72/103 (70%)	56 (78%)	16 (22%)	1	3
All	All	1246/1592 (78%)	1203 (96%)	43 (4%)	43	78

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	104	PHE
2	B	368	GLN
5	E	96	ASN
2	B	166	THR
2	B	223	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	HIS

### 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 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	786/810 (97%)	0.44	47 (5%) 25 18	92, 112, 150, 182	0
2	B	354/392 (90%)	0.81	46 (12%) 5 3	94, 113, 154, 214	0
3	C	56/344 (16%)	1.09	12 (21%) 1 0	90, 119, 181, 188	0
4	D	208/245 (84%)	0.39	6 (2%) 55 49	94, 112, 162, 213	0
5	E	83/123 (67%)	0.70	8 (9%) 10 6	99, 122, 156, 189	0
All	All	1487/1914 (77%)	0.56	119 (8%) 15 10	90, 113, 156, 214	0

The worst 5 of 119 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	30	ARG	8.8
2	B	104	PHE	7.7
5	E	83	HIS	6.9
2	B	103	TRP	6.3
3	C	31	TYR	5.9

### 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.