

# Full wwPDB X-ray Structure Validation Report (i)

#### Jan 24, 2023 – 09:16 PM EST

PDB ID : 4LBA

Title: Crystal structure of a conjugative transposon lipoprotein (BACEGG\_03088)

from Bacteroides eggerthii DSM 20697 at 1.70 A resolution

Authors : Joint Center for Structural Genomics (JCSG)

Deposited on : 2013-06-20

Resolution : 1.70 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

https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

 $Mol Probity \quad : \quad 4.02b\text{--}467$ 

Mogul: 1.8.5 (274361), CSD as541be (2020)

Xtriage (Phenix) : 1.13 EDS : 2.31.2

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$ 

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

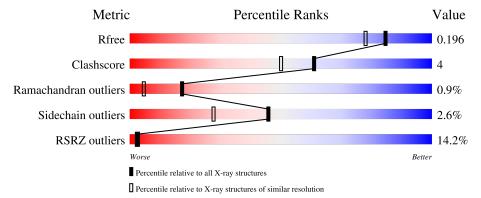
Validation Pipeline (wwPDB-VP) : 2.31.2

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X- $RAY\ DIFFRACTION$ 

The reported resolution of this entry is 1.70 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 $(\# \mathrm{Entries})$	$\begin{array}{c} {\rm Similar\ resolution} \\ (\#{\rm Entries},{\rm resolution\ range}(\mathring{\rm A})) \end{array}$
$R_{free}$	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=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	137	11% 71% 7% 23%				
1	В	137	69%	10%	20%		



# 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2170 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.

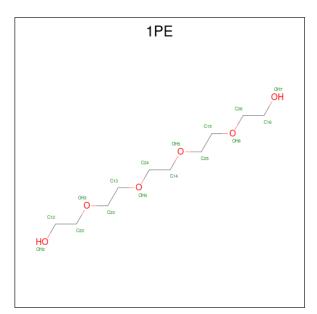
• Molecule 1 is a protein called conjugative transposon lipoprotein.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace			
1	A	106	Total 902	C 587	N 137		S 1	Se 4	0	6	0
1	В	109	Total 962	C 623	N 147	O 188	S 1	Se 3	0	13	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Α	0	GLY	-	expression tag	UNP B7AKX5
В	0	GLY	-	expression tag	UNP B7AKX5

• Molecule 2 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C<sub>10</sub>H<sub>22</sub>O<sub>6</sub>).



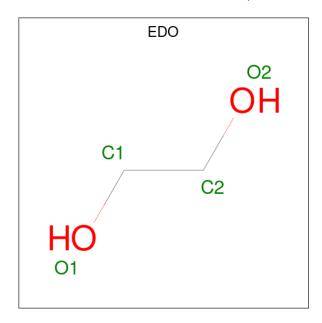
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 13 8 5	0	0



 $Continued\ from\ previous\ page...$ 

Mo	Chain	Residues	Atoms		ZeroOcc	AltConf	
2	В	1	Total 16	C 10	O 6	0	0

 $\bullet$  Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $\mathrm{C_2H_6O_2}).$ 



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	В	1	Total C O 4 2 2	0	0
3	В	1	Total C O 4 2 2	0	0
3	В	1	Total C O 8 4 4	0	1

• Molecule 4 is water.

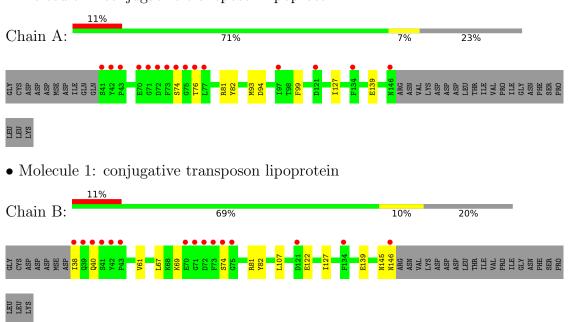
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	116	Total O 117 117	0	2
4	В	129	Total O 132 132	0	3



# 3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: conjugative transposon lipoprotein





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 64	Depositor
Cell constants	126.92Å 126.92Å 39.35Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $120.00^{\circ}$	Depositor
Resolution (Å)	25.22 - 1.70	Depositor
Resolution (A)	25.22 - 1.70	EDS
% Data completeness	99.8 (25.22-1.70)	Depositor
(in resolution range)	99.8 (25.22-1.70)	EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$< I/\sigma(I) > 1$	1.99 (at 1.70Å)	Xtriage
Refinement program	REFMAC	Depositor
D.D.	0.165 , 0.187	Depositor
$R, R_{free}$	0.177 , $0.196$	DCC
$R_{free}$ test set	2013 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.2	Xtriage
Anisotropy	0.122	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.34 , 48.4	EDS
L-test for twinning <sup>2</sup>	$< L >=0.50, < L^2>=0.33$	Xtriage
Estimated twinning fraction	0.032 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	2170	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.19% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical values of <|L|>,  $<L^2>$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: EDO, 1PE

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		
IVIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	A	0.59	0/939	0.77	1/1266 (0.1%)	
1	В	0.61	0/1011	0.83	1/1364 (0.1%)	
All	All	0.60	0/1950	0.81	2/2630 (0.1%)	

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
1	A	81	ARG	NE-CZ-NH1	5.65	123.12	120.30
1	В	81	ARG	NE-CZ-NH1	5.65	123.13	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	902	0	877	6	0
1	В	962	0	933	9	0
2	A	13	0	17	0	0
2	В	16	0	22	0	0
3	A	12	0	18	0	0
3	В	16	0	24	3	0
4	A	117	0	0	0	0



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	В	132	0	0	2	0
All	All	2170	0	1891	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 4.

All (17) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic	Clash
Atom-1	Atom-2	${f distance}({ m \AA})$	overlap (Å)
1:B:122[B]:GLU:HG2	1:B:122[B]:GLU:O	1.49	1.04
1:B:122[B]:GLU:O	1:B:122[B]:GLU:CG	2.35	0.65
1:A:127[B]:ILE:CG2	1:A:139:GLU:HG3	2.26	0.64
1:B:127[B]:ILE:CG2	1:B:139:GLU:HG3	2.27	0.64
1:A:93[B]:MSE:HG3	1:A:99:PHE:CE1	2.36	0.60
3:B:203:EDO:H21	4:B:385:HOH:O	2.02	0.59
1:A:127[B]:ILE:CG2	1:A:139:GLU:CG	2.88	0.51
1:B:67[B]:LEU:HD11	1:B:107[B]:LEU:HD23	1.93	0.51
1:B:127[B]:ILE:CG2	1:B:139:GLU:CG	2.88	0.51
1:B:38:ILE:N	1:B:139:GLU:H	2.09	0.51
1:B:127[B]:ILE:HD13	1:B:139:GLU:OE2	2.13	0.47
1:A:127[B]:ILE:HG21	1:A:139:GLU:CG	2.45	0.47
1:B:61:VAL:HG13	3:B:203:EDO:H11	1.96	0.47
1:A:93[A]:MSE:CG	1:A:94:ASP:H	2.30	0.45
3:B:203:EDO:C2	4:B:385:HOH:O	2.66	0.42
1:B:145:ASN:O	1:B:146:ASN:HB2	2.21	0.41
1:A:127[B]:ILE:HD13	1:A:139:GLU:OE2	2.20	0.40

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	110/137 (80%)	107 (97%)	2 (2%)	1 (1%)	17 5
1	В	121/137 (88%)	118 (98%)	2 (2%)	1 (1%)	19 6
All	All	231/274 (84%)	225 (97%)	4 (2%)	2 (1%)	17 5

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	74	SER
1	В	74	SER

#### 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	102/122~(84%)	100 (98%)	2 (2%)	55 38		
1	В	109/122 (89%)	105 (96%)	4 (4%)	34 15		
All	All	211/244 (86%)	205 (97%)	6 (3%)	46 25		

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	76	THR
1	A	82	TYR
1	В	40[A]	GLN
1	В	40[B]	GLN
1	В	69	LYS
1	В	82	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 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 monosaccharides in this entry.

### 5.6 Ligand geometry (i)

9 ligands are 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	Trunc	Chain	Res	Link	Во	Bond lengths			Bond angles		
IVIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2	
3	EDO	В	204[A]	-	3,3,3	0.45	0	2,2,2	0.39	0	
3	EDO	A	203	-	3,3,3	0.54	0	2,2,2	0.29	0	
3	EDO	A	204	-	3,3,3	0.53	0	2,2,2	0.19	0	
2	1PE	A	201	-	12,12,15	0.53	0	11,11,14	0.53	0	
3	EDO	A	202	-	3,3,3	0.54	0	2,2,2	0.50	0	
3	EDO	В	202	-	3,3,3	0.32	0	2,2,2	0.52	0	
3	EDO	В	204[B]	-	3,3,3	0.46	0	2,2,2	0.30	0	
2	1PE	В	201	-	15,15,15	0.55	0	14,14,14	0.36	0	
3	EDO	В	203	-	3,3,3	0.97	0	2,2,2	0.73	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
3	EDO	В	204[A]	-	=	1/1/1/1	-
3	EDO	A	203	-	-	1/1/1/1	-
3	EDO	A	204	-	-	1/1/1/1	-
2	1PE	A	201	-	-	6/10/10/13	-
3	EDO	A	202	-	-	0/1/1/1	-
3	EDO	В	202	-	=	0/1/1/1	-



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	В	204[B]	-	-	1/1/1/1	-
2	1PE	В	201	-	-	5/13/13/13	_
3	EDO	В	203	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (16) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	201	1PE	ОН4-С13-С23-ОН3
2	В	201	1PE	ОН7-С16-С26-ОН6
2	В	201	1PE	OH5-C14-C24-OH4
2	В	201	1PE	OH2-C12-C22-OH3
3	A	203	EDO	O1-C1-C2-O2
3	A	204	EDO	O1-C1-C2-O2
3	В	204[A]	EDO	O1-C1-C2-O2
3	В	204[B]	EDO	O1-C1-C2-O2
2	A	201	1PE	OH2-C12-C22-OH3
2	A	201	1PE	C23-C13-OH4-C24
2	A	201	1PE	C14-C24-OH4-C13
3	В	203	EDO	O1-C1-C2-O2
2	В	201	1PE	C12-C22-OH3-C23
2	A	201	1PE	C13-C23-OH3-C22
2	A	201	1PE	OH5-C14-C24-OH4
2	В	201	1PE	OH4-C13-C23-OH3

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	В	203	EDO	3	0

# 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,  $95^{th}$  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 $>$	$\# \mathrm{RSRZ} {>} 2$		$OWAB(A^2)$	Q<0.9	
1	A	104/137~(75%)	0.62	15 (14%)	2	2	20, 34, 93, 120	0
1	В	107/137 (78%)	0.63	15 (14%)	2	3	16, 27, 75, 119	0
All	All	211/274 (77%)	0.63	30 (14%)	2	3	16, 30, 85, 120	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	74	SER	11.0
1	В	74	SER	9.4
1	В	75	GLY	8.3
1	В	71	GLY	7.7
1	A	71	GLY	6.8
1	В	73	PHE	6.0
1	В	38	ILE	6.0
1	В	39	GLN	5.9
1	A	73	PHE	5.8
1	A	146	ASN	5.8
1	В	72	ASP	5.7
1	A	75	GLY	5.5
1	A	72	ASP	4.8
1	A	76	THR	4.4
1	A	77	LEU	4.2
1	A	70	GLU	4.2
1	A	134	PHE	4.0
1	A	121	ASP	3.8
1	A	42	TYR	3.6
1	A	41	SER	3.4
1	В	146	ASN	3.4
1	В	40[A]	GLN	3.4
1	В	41	SER	3.3
1	В	70	GLU	2.8



Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	В	134	PHE	2.8
1	В	43	PRO	2.5
1	A	43	PRO	2.4
1	В	42	TYR	2.2
1	В	121[A]	ASP	2.1
1	A	97	ILE	2.0

### 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 monosaccharides 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. The B-factors column lists the minimum, median,  $95^{th}$  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	$\operatorname{B-factors}(\mathring{\mathrm{A}}^2)$	Q<0.9
3	EDO	В	204[A]	4/4	0.64	0.28	44,52,54,56	4
3	EDO	В	204[B]	4/4	0.64	0.28	49,52,56,59	4
2	1PE	A	201	13/16	0.80	0.30	64,76,98,99	0
2	1PE	В	201	16/16	0.83	0.26	57,68,88,95	0
3	EDO	A	202	4/4	0.84	0.23	42,58,59,65	0
3	EDO	A	203	4/4	0.86	0.33	44,48,58,61	0
3	EDO	A	204	4/4	0.89	0.10	51,55,56,56	0
3	EDO	В	202	4/4	0.91	0.13	54,58,58,59	0
3	EDO	В	203	4/4	0.92	0.17	31,31,35,41	0

### 6.5 Other polymers (i)

There are no such residues in this entry.

