



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 2, 2024 – 09:28 am GMT

PDB ID : 4USX  
Title : The Structure of the C-terminal YadA-like domain of BPSL2063 from Burkholderia pseudomallei  
Authors : Perletti, L.; Gourlay, L.J.; Peano, C.; Pietrelli, A.; DeBellis, G.; Deantonio, C.; Santoro, C.; Sblattero, D.; Bolognesi, M.  
Deposited on : 2014-07-16  
Resolution : 1.80 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](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 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) : 2.36

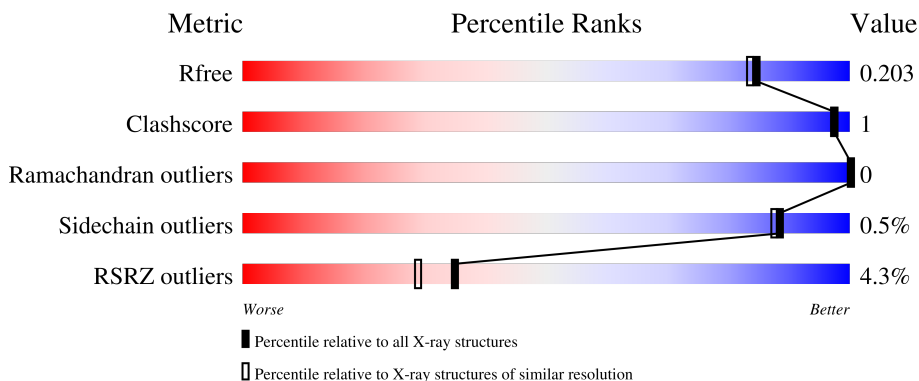
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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))
$R_{free}$	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	371	 2% 55% 44%
1	B	371	 4% 56% 43%
1	C	371	 2% 56% 42%

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 4746 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 TRIMERIC AUTOTRANSPORTER ADHESIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	208	1415	860	247	307	1	0	1	0
1	B	211	1431	870	250	310	1	0	0	0
1	C	214	1445	875	254	315	1	0	1	0

There are 105 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	641	MET	-	expression tag	UNP Q63TA4
A	642	ALA	-	expression tag	UNP Q63TA4
A	643	SER	-	expression tag	UNP Q63TA4
A	644	MET	-	expression tag	UNP Q63TA4
A	645	THR	-	expression tag	UNP Q63TA4
A	646	GLY	-	expression tag	UNP Q63TA4
A	647	GLY	-	expression tag	UNP Q63TA4
A	648	GLN	-	expression tag	UNP Q63TA4
A	649	GLN	-	expression tag	UNP Q63TA4
A	650	MET	-	expression tag	UNP Q63TA4
A	651	GLY	-	expression tag	UNP Q63TA4
A	652	ARG	-	expression tag	UNP Q63TA4
A	653	GLY	-	expression tag	UNP Q63TA4
A	654	SER	-	expression tag	UNP Q63TA4
A	655	MET	-	expression tag	UNP Q63TA4
A	656	ALA	-	expression tag	UNP Q63TA4
A	993	ASN	-	expression tag	UNP Q63TA4
A	994	SER	-	expression tag	UNP Q63TA4
A	995	SER	-	expression tag	UNP Q63TA4
A	996	SER	-	expression tag	UNP Q63TA4
A	997	VAL	-	expression tag	UNP Q63TA4
A	998	ASP	-	expression tag	UNP Q63TA4
A	999	LYS	-	expression tag	UNP Q63TA4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1000	LEU	-	expression tag	UNP Q63TA4
A	1001	ALA	-	expression tag	UNP Q63TA4
A	1002	ALA	-	expression tag	UNP Q63TA4
A	1003	ALA	-	expression tag	UNP Q63TA4
A	1004	LEU	-	expression tag	UNP Q63TA4
A	1005	GLU	-	expression tag	UNP Q63TA4
A	1006	HIS	-	expression tag	UNP Q63TA4
A	1007	HIS	-	expression tag	UNP Q63TA4
A	1008	HIS	-	expression tag	UNP Q63TA4
A	1009	HIS	-	expression tag	UNP Q63TA4
A	1010	HIS	-	expression tag	UNP Q63TA4
A	1011	HIS	-	expression tag	UNP Q63TA4
B	641	MET	-	expression tag	UNP Q63TA4
B	642	ALA	-	expression tag	UNP Q63TA4
B	643	SER	-	expression tag	UNP Q63TA4
B	644	MET	-	expression tag	UNP Q63TA4
B	645	THR	-	expression tag	UNP Q63TA4
B	646	GLY	-	expression tag	UNP Q63TA4
B	647	GLY	-	expression tag	UNP Q63TA4
B	648	GLN	-	expression tag	UNP Q63TA4
B	649	GLN	-	expression tag	UNP Q63TA4
B	650	MET	-	expression tag	UNP Q63TA4
B	651	GLY	-	expression tag	UNP Q63TA4
B	652	ARG	-	expression tag	UNP Q63TA4
B	653	GLY	-	expression tag	UNP Q63TA4
B	654	SER	-	expression tag	UNP Q63TA4
B	655	MET	-	expression tag	UNP Q63TA4
B	656	ALA	-	expression tag	UNP Q63TA4
B	993	ASN	-	expression tag	UNP Q63TA4
B	994	SER	-	expression tag	UNP Q63TA4
B	995	SER	-	expression tag	UNP Q63TA4
B	996	SER	-	expression tag	UNP Q63TA4
B	997	VAL	-	expression tag	UNP Q63TA4
B	998	ASP	-	expression tag	UNP Q63TA4
B	999	LYS	-	expression tag	UNP Q63TA4
B	1000	LEU	-	expression tag	UNP Q63TA4
B	1001	ALA	-	expression tag	UNP Q63TA4
B	1002	ALA	-	expression tag	UNP Q63TA4
B	1003	ALA	-	expression tag	UNP Q63TA4
B	1004	LEU	-	expression tag	UNP Q63TA4
B	1005	GLU	-	expression tag	UNP Q63TA4
B	1006	HIS	-	expression tag	UNP Q63TA4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1007	HIS	-	expression tag	UNP Q63TA4
B	1008	HIS	-	expression tag	UNP Q63TA4
B	1009	HIS	-	expression tag	UNP Q63TA4
B	1010	HIS	-	expression tag	UNP Q63TA4
B	1011	HIS	-	expression tag	UNP Q63TA4
C	641	MET	-	expression tag	UNP Q63TA4
C	642	ALA	-	expression tag	UNP Q63TA4
C	643	SER	-	expression tag	UNP Q63TA4
C	644	MET	-	expression tag	UNP Q63TA4
C	645	THR	-	expression tag	UNP Q63TA4
C	646	GLY	-	expression tag	UNP Q63TA4
C	647	GLY	-	expression tag	UNP Q63TA4
C	648	GLN	-	expression tag	UNP Q63TA4
C	649	GLN	-	expression tag	UNP Q63TA4
C	650	MET	-	expression tag	UNP Q63TA4
C	651	GLY	-	expression tag	UNP Q63TA4
C	652	ARG	-	expression tag	UNP Q63TA4
C	653	GLY	-	expression tag	UNP Q63TA4
C	654	SER	-	expression tag	UNP Q63TA4
C	655	MET	-	expression tag	UNP Q63TA4
C	656	ALA	-	expression tag	UNP Q63TA4
C	993	ASN	-	expression tag	UNP Q63TA4
C	994	SER	-	expression tag	UNP Q63TA4
C	995	SER	-	expression tag	UNP Q63TA4
C	996	SER	-	expression tag	UNP Q63TA4
C	997	VAL	-	expression tag	UNP Q63TA4
C	998	ASP	-	expression tag	UNP Q63TA4
C	999	LYS	-	expression tag	UNP Q63TA4
C	1000	LEU	-	expression tag	UNP Q63TA4
C	1001	ALA	-	expression tag	UNP Q63TA4
C	1002	ALA	-	expression tag	UNP Q63TA4
C	1003	ALA	-	expression tag	UNP Q63TA4
C	1004	LEU	-	expression tag	UNP Q63TA4
C	1005	GLU	-	expression tag	UNP Q63TA4
C	1006	HIS	-	expression tag	UNP Q63TA4
C	1007	HIS	-	expression tag	UNP Q63TA4
C	1008	HIS	-	expression tag	UNP Q63TA4
C	1009	HIS	-	expression tag	UNP Q63TA4
C	1010	HIS	-	expression tag	UNP Q63TA4
C	1011	HIS	-	expression tag	UNP Q63TA4

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Mg 1 1	0	0
2	B	4	Total Mg 4 4	0	0
2	C	2	Total Mg 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	184	Total O 184 184	0	0
3	B	150	Total O 150 150	0	0
3	C	114	Total O 114 114	0	0



SER ASP GLY ARG SER ASN VAL VAL SER GLY PRO GLY GLN VAL THR ASN VAL ALA GLY THR GLN ALA THR ASP ALA VAL ASN LEU GLY MET ASN GLY ALA SER SER SER VAL LYS LEU ALA ALA LEU HIS HIS HIS HIS



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.12Å 58.85Å 74.41Å 90.00° 104.65° 90.00°	Depositor
Resolution (Å)	39.69 – 1.80 39.69 – 1.80	Depositor EDS
% Data completeness (in resolution range)	98.9 (39.69-1.80) 93.9 (39.69-1.80)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.28 (at 1.81Å)	Xtrriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.169 , 0.204 0.169 , 0.203	Depositor DCC
$R_{free}$ test set	2462 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	13.9	Xtrriage
Anisotropy	0.713	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 49.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4746	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.62% 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, 0.333 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: MG

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	0.37	0/1435	0.51	0/1966
1	B	0.37	0/1448	0.50	0/1985
1	C	0.40	0/1465	0.55	1/2006 (0.0%)
All	All	0.38	0/4348	0.52	1/5957 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	795	GLN	CA-CB-CG	-5.38	101.56	113.40

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	1415	0	1370	5	0
1	B	1431	0	1388	2	0
1	C	1445	0	1395	4	0
2	A	1	0	0	0	0
2	B	4	0	0	0	0
2	C	2	0	0	0	0
3	A	184	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	150	0	0	0	0
3	C	114	0	0	0	0
All	All	4746	0	4153	7	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 1.

All (7) 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:C:795:GLN:HG3	1:C:796:ALA:N	2.10	0.67
1:A:847:ILE:HD11	1:C:849:GLY:HA2	1.88	0.55
1:A:841:GLN:HG3	1:C:853:TYR:CE2	2.44	0.53
1:C:800:LEU:HD23	1:C:814:THR:HB	1.90	0.53
1:A:817:ILE:HG21	1:B:827:VAL:HG12	1.91	0.52
1:A:713:ASP:OD1	1:A:715:LYS:HB3	2.19	0.42
1:A:817:ILE:HG12	1:B:827:VAL:HB	2.02	0.41

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	205/371 (55%)	203 (99%)	2 (1%)	0	100	100
1	B	207/371 (56%)	204 (99%)	3 (1%)	0	100	100
1	C	211/371 (57%)	207 (98%)	4 (2%)	0	100	100
All	All	623/1113 (56%)	614 (99%)	9 (1%)	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	147/259 (57%)	147 (100%)	0	100	100
1	B	148/259 (57%)	147 (99%)	1 (1%)	84	81
1	C	149/259 (58%)	148 (99%)	1 (1%)	84	81
All	All	444/777 (57%)	442 (100%)	2 (0%)	88	87

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

Mol	Chain	Res	Type
1	B	749	THR
1	C	795	GLN

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

Mol	Chain	Res	Type
1	C	795	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 [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

Of 7 ligands modelled in this entry, 7 are monoatomic - leaving 0 for Mogul analysis.

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

### 6.1 Protein, DNA and RNA chains

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	208/371 (56%)	-0.32	7 (3%) 45 39	11, 17, 42, 50	0
1	B	211/371 (56%)	-0.24	14 (6%) 18 14	11, 17, 47, 57	0
1	C	214/371 (57%)	-0.25	6 (2%) 53 47	10, 17, 40, 57	0
All	All	633/1113 (56%)	-0.27	27 (4%) 35 29	10, 17, 43, 57	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	874	ALA	5.4
1	B	889	ALA	4.3
1	C	878	GLY	4.0
1	B	883	TYR	3.9
1	B	890	VAL	3.6
1	A	889	ALA	3.6
1	B	876	THR	3.6
1	C	875	ASN	3.4
1	A	883	TYR	3.3
1	C	879	SER	3.3
1	A	678	PRO	3.2
1	B	882	SER	3.1
1	B	881	ALA	3.1
1	A	890	VAL	3.0
1	B	877	LEU	2.9
1	B	886	GLN	2.8
1	B	875	ASN	2.8
1	C	880	GLY	2.7
1	C	873	VAL	2.7
1	B	873	VAL	2.6
1	B	884	ASN	2.6
1	A	875	ASN	2.5
1	B	885	PRO	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	888	GLY	2.3
1	B	677	GLY	2.2
1	A	885	PRO	2.2
1	A	886	GLN	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<sup>th</sup> 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	MG	B	1894	1/1	0.88	0.10	31,31,31,31	0
2	MG	B	1893	1/1	0.95	0.13	28,28,28,28	0
2	MG	B	1892	1/1	0.99	0.10	24,24,24,24	0
2	MG	A	1892	1/1	0.99	0.06	21,21,21,21	0
2	MG	B	1891	1/1	0.99	0.05	20,20,20,20	0
2	MG	C	1893	1/1	0.99	0.06	21,21,21,21	0
2	MG	C	1894	1/1	0.99	0.07	23,23,23,23	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.