



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 13, 2017 – 09:55 AM EST

PDB ID : 3WPA  
Title : Acinetobacter sp. Tol 5 AtaA C-terminal stalk\_FL fused to GCN4 adaptors (CstalkFL)  
Authors : Koiwai, K.; Hartmann, M.D.; Yoshimoto, S.; Nur 'Izzah, N.; Suzuki, A.; Linke, D.; Lupas, A.N.; Hori, K.  
Deposited on : unknown  
Resolution : 1.79 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

<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

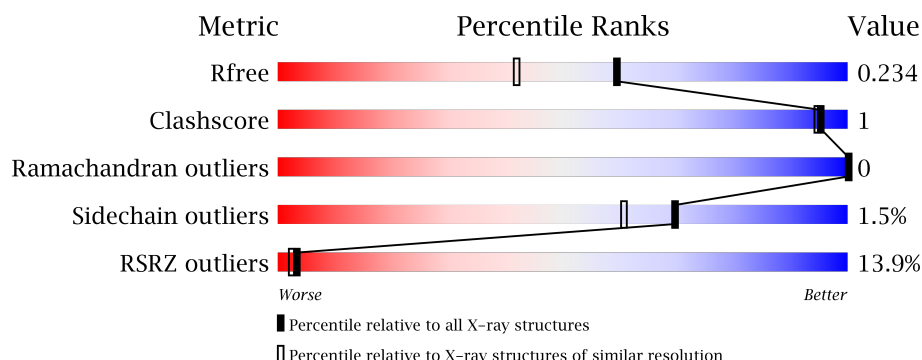
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.79 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	4827 (1.80-1.80)
Clashscore	112137	5742 (1.80-1.80)
Ramachandran outliers	110173	5676 (1.80-1.80)
Sidechain outliers	110143	5675 (1.80-1.80)
RSRZ outliers	101464	4906 (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 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	427	<div> <div>11%</div> <div>78%</div> <div>19%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2889 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
1	A	346	Total	C	N	O	0	6	0
			2511	1500	457	554			

There are 35 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3141	MET	-	EXPRESSION TAG	UNP K7ZP88
A	3142	LYS	-	EXPRESSION TAG	UNP K7ZP88
A	3143	GLN	-	EXPRESSION TAG	UNP K7ZP88
A	3144	ILE	-	EXPRESSION TAG	UNP K7ZP88
A	3145	GLU	-	EXPRESSION TAG	UNP K7ZP88
A	3146	ASP	-	EXPRESSION TAG	UNP K7ZP88
A	3147	LYS	-	EXPRESSION TAG	UNP K7ZP88
A	3148	ILE	-	EXPRESSION TAG	UNP K7ZP88
A	3149	GLU	-	EXPRESSION TAG	UNP K7ZP88
A	3150	GLU	-	EXPRESSION TAG	UNP K7ZP88
A	3151	ILE	-	EXPRESSION TAG	UNP K7ZP88
A	3152	LEU	-	EXPRESSION TAG	UNP K7ZP88
A	3153	SER	-	EXPRESSION TAG	UNP K7ZP88
A	3154	LYS	-	EXPRESSION TAG	UNP K7ZP88
A	3155	ILE	-	EXPRESSION TAG	UNP K7ZP88
A	3156	TYR	-	EXPRESSION TAG	UNP K7ZP88
A	3157	HIS	-	EXPRESSION TAG	UNP K7ZP88
A	3158	ILE	-	EXPRESSION TAG	UNP K7ZP88
A	3159	GLU	-	EXPRESSION TAG	UNP K7ZP88
A	3160	ASN	-	EXPRESSION TAG	UNP K7ZP88
A	3161	GLU	-	EXPRESSION TAG	UNP K7ZP88
A	3162	ILE	-	EXPRESSION TAG	UNP K7ZP88
A	3163	ALA	-	EXPRESSION TAG	UNP K7ZP88
A	3164	ARG	-	EXPRESSION TAG	UNP K7ZP88
A	3165	ILE	-	EXPRESSION TAG	UNP K7ZP88
A	3166	LYS	-	EXPRESSION TAG	UNP K7ZP88
A	3167	LYS	-	EXPRESSION TAG	UNP K7ZP88

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	3168	LEU	-	EXPRESSION TAG	UNP K7ZP88
A	3169	ILE	-	EXPRESSION TAG	UNP K7ZP88
A	3562	HIS	-	EXPRESSION TAG	UNP K7ZP88
A	3563	HIS	-	EXPRESSION TAG	UNP K7ZP88
A	3564	HIS	-	EXPRESSION TAG	UNP K7ZP88
A	3565	HIS	-	EXPRESSION TAG	UNP K7ZP88
A	3566	HIS	-	EXPRESSION TAG	UNP K7ZP88
A	3567	HIS	-	EXPRESSION TAG	UNP K7ZP88

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Cl 1 1	0	0

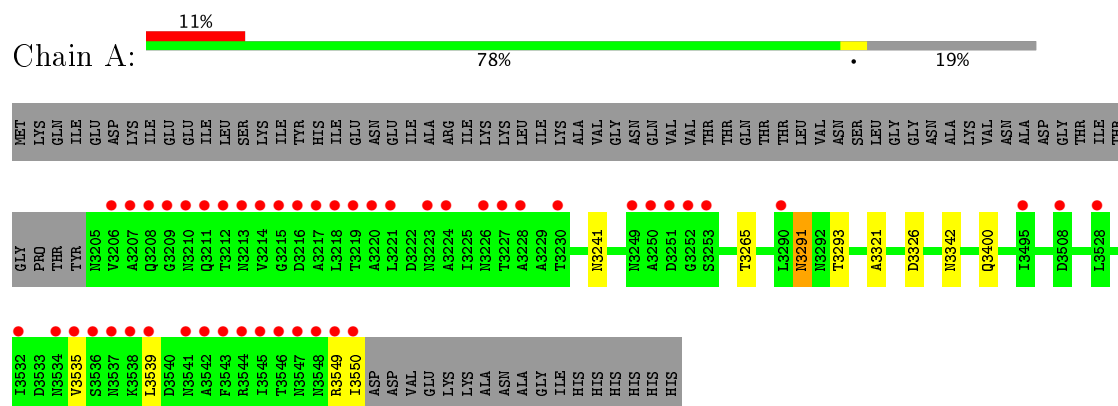
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	377	Total O 377 377	0	0

### 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: Trimeric autotransporter adhesin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	46.66 Å 46.66 Å 406.11 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	33.16 – 1.79 33.16 – 1.79	Depositor EDS
% Data completeness (in resolution range)	99.3 (33.16-1.79) 99.4 (33.16-1.79)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.30 (at 1.79 Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.203 , 0.227 0.209 , 0.234	Depositor DCC
$R_{free}$ test set	2559 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	26.0	Xtriage
Anisotropy	0.189	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 52.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.048 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2889	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.27% 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:  
CL

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	0.34	0/2523	0.56	0/3439

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	2511	0	2438	7	0
2	A	1	0	0	0	0
3	A	377	0	0	1	0
All	All	2889	0	2438	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:A:3241:ASN:HD21	1:A:3265:THR:H	1.45	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:3400[A]:GLN:NE2	3:A:4016:HOH:O	2.38	0.57
1:A:3535:VAL:O	1:A:3539:LEU:N	2.46	0.48
1:A:3291:ASN:C	1:A:3291:ASN:HD22	2.20	0.44
1:A:3321:ALA:HB1	1:A:3326:ASP:HB2	2.00	0.43
1:A:3241:ASN:ND2	1:A:3265:THR:H	2.16	0.42
1:A:3291:ASN:ND2	1:A:3293:THR:H	2.19	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	350/427 (82%)	350 (100%)	0	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	274/337 (81%)	270 (98%)	4 (2%)	70	61

All (4) residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	A	3291	ASN
1	A	3342	ASN
1	A	3549	ARG
1	A	3550	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	3238	ASN
1	A	3241	ASN
1	A	3291	ASN
1	A	3399	GLN
1	A	3425	GLN
1	A	3494	ASN
1	A	3513	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](#)

Of 1 ligands modelled in this entry, 1 is 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	346/427 (81%)	0.88	48 (13%) 3 2	13, 29, 120, 145	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3550	ILE	12.8
1	A	3546	THR	11.9
1	A	3543	PHE	11.2
1	A	3547	ASN	11.0
1	A	3544	ARG	10.3
1	A	3212	THR	9.6
1	A	3545	ILE	9.4
1	A	3250	ALA	8.9
1	A	3218	LEU	8.1
1	A	3209	GLY	7.6
1	A	3548	ASN	7.5
1	A	3206	VAL	6.8
1	A	3537	ASN	6.6
1	A	3219	THR	6.6
1	A	3214	VAL	6.4
1	A	3539	LEU	6.4
1	A	3216	ASP	6.1
1	A	3549	ARG	6.1
1	A	3538	LYS	6.1
1	A	3210	ASN	6.0
1	A	3226	ASN	5.8
1	A	3213	ASN	5.3
1	A	3211	GLN	5.2
1	A	3541	ASN	5.0
1	A	3208	GLN	4.8
1	A	3217	ALA	4.8
1	A	3535	VAL	4.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	3221	LEU	4.6
1	A	3215	GLY	4.6
1	A	3542	ALA	4.3
1	A	3220	ALA	4.1
1	A	3534	ASN	3.9
1	A	3508	ASP	3.8
1	A	3252	GLY	3.1
1	A	3536	SER	3.1
1	A	3223	ASN	3.1
1	A	3251	ASP	3.0
1	A	3528	LEU	2.7
1	A	3532	ILE	2.5
1	A	3249	ASN	2.5
1	A	3227	THR	2.3
1	A	3228	ALA	2.3
1	A	3230	THR	2.3
1	A	3224	ALA	2.2
1	A	3495	ILE	2.2
1	A	3253	SER	2.1
1	A	3207	ALA	2.1
1	A	3290	LEU	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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	CL	A	3601	1/1	0.99	0.08	-	26,26,26,26	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.