wwPDB X-ray Structure Validation Summary Report (i)

May 22, 2020 – 05:49 pm BST

PDB ID	:	3WP8
Title	:	Acinetobacter sp. Tol 5 AtaA C-terminal Ylhead fused to GCN4 adaptors (Chead)
Authors	:	Koiwai, K.; Hartmann, M.D.; Yoshimoto, S.; Nur 'Izzah, N.; Suzuki, A.; Linke, D.; Lupas, A.N.; Hori, K.
Deposited on	:	2014-01-10
Resolution	:	1.97 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

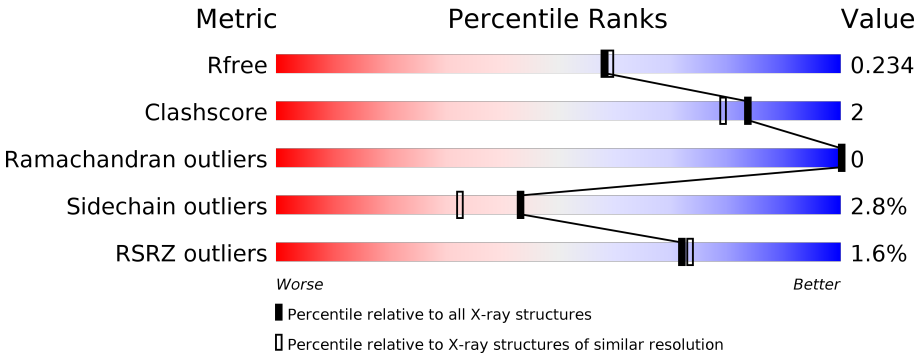
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.97 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 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	330	2% 88% 8% .

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2474 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	317	Total	C	N	O	S	0	0	0
			2293	1402	403	486	2			

There are 67 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2876	MET	-	expression tag	UNP K7ZP88
A	2877	LYS	-	expression tag	UNP K7ZP88
A	2878	GLN	-	expression tag	UNP K7ZP88
A	2879	ILE	-	expression tag	UNP K7ZP88
A	2880	GLU	-	expression tag	UNP K7ZP88
A	2881	ASP	-	expression tag	UNP K7ZP88
A	2882	LYS	-	expression tag	UNP K7ZP88
A	2883	ILE	-	expression tag	UNP K7ZP88
A	2884	GLU	-	expression tag	UNP K7ZP88
A	2885	GLU	-	expression tag	UNP K7ZP88
A	2886	ILE	-	expression tag	UNP K7ZP88
A	2887	LEU	-	expression tag	UNP K7ZP88
A	2888	SER	-	expression tag	UNP K7ZP88
A	2889	LYS	-	expression tag	UNP K7ZP88
A	2890	ILE	-	expression tag	UNP K7ZP88
A	2891	TYR	-	expression tag	UNP K7ZP88
A	2892	HIS	-	expression tag	UNP K7ZP88
A	2893	ILE	-	expression tag	UNP K7ZP88
A	2894	GLU	-	expression tag	UNP K7ZP88
A	2895	ASN	-	expression tag	UNP K7ZP88
A	2896	GLU	-	expression tag	UNP K7ZP88
A	2897	ILE	-	expression tag	UNP K7ZP88
A	2898	ALA	-	expression tag	UNP K7ZP88
A	2899	ARG	-	expression tag	UNP K7ZP88
A	2900	ILE	-	expression tag	UNP K7ZP88
A	2901	LYS	-	expression tag	UNP K7ZP88
A	2902	LYS	-	expression tag	UNP K7ZP88

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	2903	LEU	-	expression tag	UNP K7ZP88
A	2904	ILE	-	expression tag	UNP K7ZP88
A	3061	GLY	PRO	engineered mutation	UNP K7ZP88
A	3169	MET	-	expression tag	UNP K7ZP88
A	3170	LYS	-	expression tag	UNP K7ZP88
A	3171	GLN	-	expression tag	UNP K7ZP88
A	3172	ILE	-	expression tag	UNP K7ZP88
A	3173	GLU	-	expression tag	UNP K7ZP88
A	3174	ASP	-	expression tag	UNP K7ZP88
A	3175	LYS	-	expression tag	UNP K7ZP88
A	3176	ILE	-	expression tag	UNP K7ZP88
A	3177	GLU	-	expression tag	UNP K7ZP88
A	3178	GLU	-	expression tag	UNP K7ZP88
A	3179	ILE	-	expression tag	UNP K7ZP88
A	3180	LEU	-	expression tag	UNP K7ZP88
A	3181	SER	-	expression tag	UNP K7ZP88
A	3182	LYS	-	expression tag	UNP K7ZP88
A	3183	ILE	-	expression tag	UNP K7ZP88
A	3184	TYR	-	expression tag	UNP K7ZP88
A	3185	HIS	-	expression tag	UNP K7ZP88
A	3186	ILE	-	expression tag	UNP K7ZP88
A	3187	GLU	-	expression tag	UNP K7ZP88
A	3188	ASN	-	expression tag	UNP K7ZP88
A	3189	GLU	-	expression tag	UNP K7ZP88
A	3190	ILE	-	expression tag	UNP K7ZP88
A	3191	ALA	-	expression tag	UNP K7ZP88
A	3192	ARG	-	expression tag	UNP K7ZP88
A	3193	ILE	-	expression tag	UNP K7ZP88
A	3194	LYS	-	expression tag	UNP K7ZP88
A	3195	LYS	-	expression tag	UNP K7ZP88
A	3196	LEU	-	expression tag	UNP K7ZP88
A	3197	ILE	-	expression tag	UNP K7ZP88
A	3198	LYS	-	expression tag	UNP K7ZP88
A	3199	LEU	-	expression tag	UNP K7ZP88
A	3200	HIS	-	expression tag	UNP K7ZP88
A	3201	HIS	-	expression tag	UNP K7ZP88
A	3202	HIS	-	expression tag	UNP K7ZP88
A	3203	HIS	-	expression tag	UNP K7ZP88
A	3204	HIS	-	expression tag	UNP K7ZP88
A	3205	HIS	-	expression tag	UNP K7ZP88

- Molecule 2 is water.

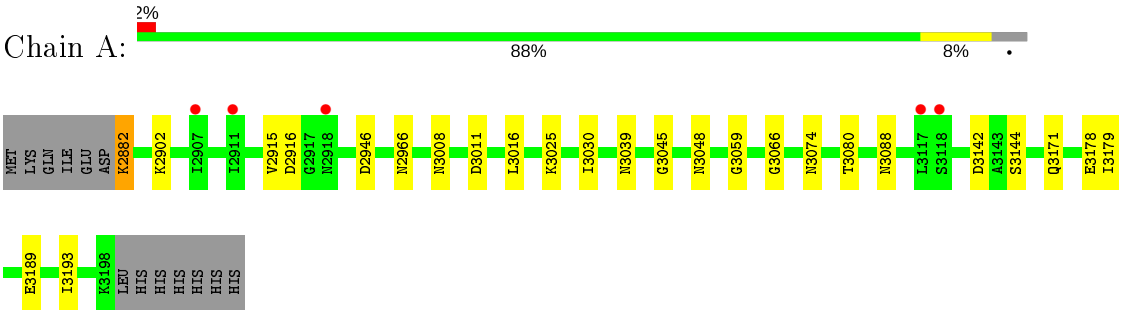
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	181	Total	O	0	0
			181	181		

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

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, α, β, γ	54.41Å 54.41Å 210.73Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.24 – 1.97 19.24 – 1.97	Depositor EDS
% Data completeness (in resolution range)	97.8 (19.24-1.97) 98.0 (19.24-1.97)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	8.15 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.175 , 0.230 0.182 , 0.234	Depositor DCC
R_{free} test set	1241 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å^2)	34.8	Xtriage
Anisotropy	0.237	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 31.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.212 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2474	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.03% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] 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.

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	0.88	0/2308	0.92	2/3124 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	2946	ASP	CB-CG-OD1	6.50	124.15	118.30
1	A	3011	ASP	CB-CG-OD1	5.27	123.04	118.30

There are no chirality outliers.

There are no planarity outliers.

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	2293	0	2298	11	0
2	A	181	0	0	4	0
All	All	2474	0	2298	11	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 2.

The worst 5 of 11 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:3074:ASN:H	1:A:3088:ASN:HD22	1.41	0.69
1:A:3025:LYS:H	1:A:3039:ASN:HD22	1.44	0.66
1:A:3048:ASN:HB2	2:A:3323:HOH:O	2.04	0.57
1:A:3142:ASP:OD1	1:A:3144:SER:HB2	2.14	0.48
1:A:2916:ASP:HB2	2:A:3315:HOH:O	2.14	0.47

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	315/330 (96%)	305 (97%)	10 (3%)	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	249/262 (95%)	242 (97%)	7 (3%)	43 32

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3008	ASN
1	A	3179	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	3171	GLN
1	A	2915	VAL
1	A	3178	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2963	GLN
1	A	2966	ASN
1	A	3023	ASN
1	A	3039	ASN
1	A	3088	ASN

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

There are no ligands in this entry.

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^{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>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	317/330 (96%)	-0.23	5 (1%) 72 73	25, 37, 68, 88	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3118	SER	2.9
1	A	3117	LEU	2.7
1	A	2907	ILE	2.6
1	A	2911	ILE	2.1
1	A	2918	ASN	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.