



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2017 – 08:01 AM EDT

PDB ID : 3ETX
Title : Crystal structure of bacterial adhesin FadA L14A mutant
Authors : Nithianantham, S.; Xu, M.; Wu, N.; Shoham, M.; Han, Y.W.
Deposited on : unknown
Resolution : 3.00 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

<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

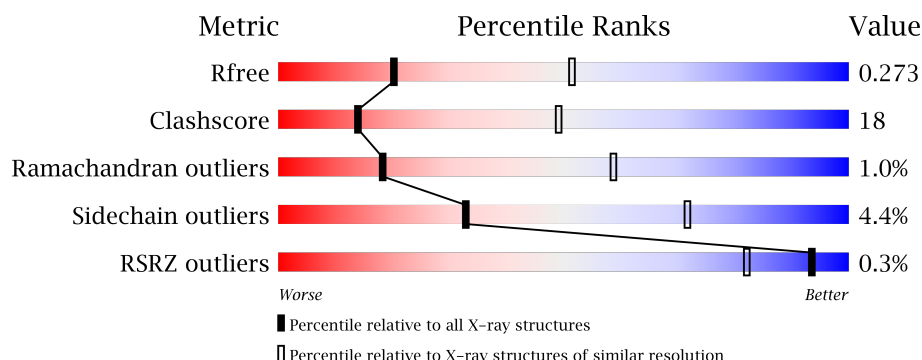
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	1692 (3.00-3.00)
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)
RSRZ outliers	101464	1716 (3.00-3.00)

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. 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	119	
1	B	119	
1	C	119	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2657 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 Adhesin A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	106	Total	C	N	O	S	0	0	0
			849	518	154	176	1			
1	B	107	Total	C	N	O	S	0	0	0
			857	522	156	178	1			
1	C	106	Total	C	N	O	S	0	0	0
			849	518	154	176	1			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	14	ALA	LEU	ENGINEERED	UNP Q5I6B0
A	112	LEU	-	EXPRESSION TAG	UNP Q5I6B0
A	113	GLU	-	EXPRESSION TAG	UNP Q5I6B0
A	114	HIS	-	EXPRESSION TAG	UNP Q5I6B0
A	115	HIS	-	EXPRESSION TAG	UNP Q5I6B0
A	116	HIS	-	EXPRESSION TAG	UNP Q5I6B0
A	117	HIS	-	EXPRESSION TAG	UNP Q5I6B0
A	118	HIS	-	EXPRESSION TAG	UNP Q5I6B0
A	119	HIS	-	EXPRESSION TAG	UNP Q5I6B0
B	14	ALA	LEU	ENGINEERED	UNP Q5I6B0
B	112	LEU	-	EXPRESSION TAG	UNP Q5I6B0
B	113	GLU	-	EXPRESSION TAG	UNP Q5I6B0
B	114	HIS	-	EXPRESSION TAG	UNP Q5I6B0
B	115	HIS	-	EXPRESSION TAG	UNP Q5I6B0
B	116	HIS	-	EXPRESSION TAG	UNP Q5I6B0
B	117	HIS	-	EXPRESSION TAG	UNP Q5I6B0
B	118	HIS	-	EXPRESSION TAG	UNP Q5I6B0
B	119	HIS	-	EXPRESSION TAG	UNP Q5I6B0
C	14	ALA	LEU	ENGINEERED	UNP Q5I6B0
C	112	LEU	-	EXPRESSION TAG	UNP Q5I6B0
C	113	GLU	-	EXPRESSION TAG	UNP Q5I6B0
C	114	HIS	-	EXPRESSION TAG	UNP Q5I6B0
C	115	HIS	-	EXPRESSION TAG	UNP Q5I6B0

Continued on next page...

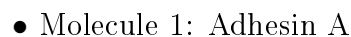
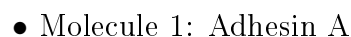
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	116	HIS	-	EXPRESSION TAG	UNP Q5I6B0
C	117	HIS	-	EXPRESSION TAG	UNP Q5I6B0
C	118	HIS	-	EXPRESSION TAG	UNP Q5I6B0
C	119	HIS	-	EXPRESSION TAG	UNP Q5I6B0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	30	Total O 30 30	0	0
2	B	37	Total O 37 37	0	0
2	C	35	Total O 35 35	0	0

- Molecule 1: Adhesin A



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	59.32Å 104.81Å 59.85Å 90.00° 105.10° 90.00°	Depositor
Resolution (Å)	38.80 – 3.00 38.82 – 3.00	Depositor EDS
% Data completeness (in resolution range)	96.0 (38.80-3.00) 96.2 (38.82-3.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.14 (at 3.01Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.238 , 0.280 0.234 , 0.273	Depositor DCC
R_{free} test set	1389 reflections (10.18%)	DCC
Wilson B-factor (Å ²)	68.7	Xtriage
Anisotropy	0.159	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 64.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.028 for l,-k,h	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	2657	wwPDB-VP
Average B, all atoms (Å ²)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.58% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

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	0.53	0/856	0.64	1/1148 (0.1%)
1	B	0.55	0/864	0.55	0/1159
1	C	0.48	0/856	0.55	0/1148
All	All	0.52	0/2576	0.58	1/3455 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	107	LEU	CA-CB-CG	5.26	127.40	115.30

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	849	0	813	28	0
1	B	857	0	819	36	0
1	C	849	0	813	25	0
2	A	30	0	0	10	0
2	B	37	0	0	14	0
2	C	35	0	0	7	0
All	All	2657	0	2445	88	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 18.

The worst 5 of 88 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:79:LYS:HE3	2:C:139:HOH:O	1.78	0.82
1:A:40:ARG:HA	1:A:98:ILE:HD11	1.67	0.76
1:B:71:SER:HB3	2:B:124:HOH:O	1.87	0.75
1:B:41:GLN:HG3	2:B:146:HOH:O	1.86	0.75
1:B:54:SER:HA	1:B:84:LEU:HD21	1.70	0.74

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	104/119 (87%)	96 (92%)	7 (7%)	1 (1%)	18	59
1	B	105/119 (88%)	98 (93%)	6 (6%)	1 (1%)	18	59
1	C	104/119 (87%)	94 (90%)	9 (9%)	1 (1%)	18	59
All	All	313/357 (88%)	288 (92%)	22 (7%)	3 (1%)	18	59

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	64	ALA
1	C	62	ALA
1	B	69	TYR

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	83/94 (88%)	78 (94%)	5 (6%)	22	60
1	B	84/94 (89%)	81 (96%)	3 (4%)	40	77
1	C	83/94 (88%)	80 (96%)	3 (4%)	40	77
All	All	250/282 (89%)	239 (96%)	11 (4%)	33	72

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	108	ARG
1	B	26	GLU
1	C	26	GLU
1	A	52	GLU
1	B	67	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	46	ASN
1	B	51	ASN
1	C	30	ASN
1	B	23	ASN
1	B	30	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 [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, 95th 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	106/119 (89%)	-0.10	1 (0%) 84 61	33, 63, 90, 118	0
1	B	107/119 (89%)	0.03	0 100 100	39, 72, 94, 112	0
1	C	106/119 (89%)	-0.07	0 100 100	38, 70, 95, 115	0
All	All	319/357 (89%)	-0.05	1 (0%) 93 82	33, 69, 95, 118	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	5	ALA	2.7

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.