



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 03:13 am BST

PDB ID : 4MEE
Title : Crystal structure of the transport unit of the autotransporter AIDA-I from Escherichia coli
Authors : Gawarzewski, I.; Tschapek, B.; Hoepfner, A.; Smits, S.H.; Jose, J.; Schmitt, L.
Deposited on : 2013-08-26
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

<https://www.wwpdb.org/validation/2017/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)	:	1.13
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

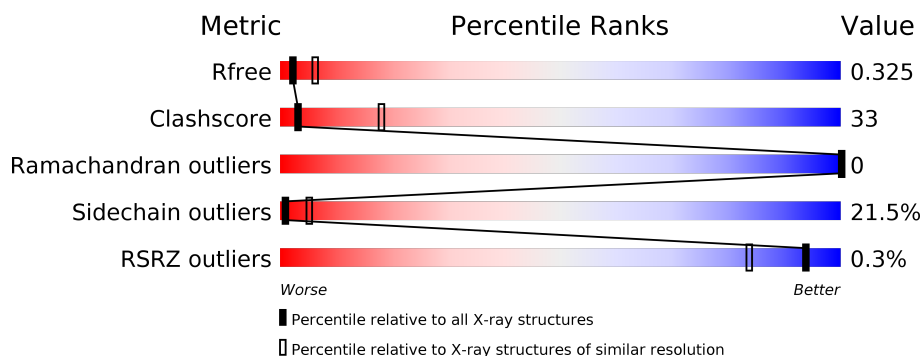
1 Overall quality at a glance

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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (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 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	469	<div> <div></div> <div>37%</div> <div>22%</div> <div>5%</div> <div>35%</div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2234 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 Diffuse adherence adhesin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	0	0	0
			2234	1389	391	447	7			

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	818	HIS	-	EXPRESSION TAG	UNP D7PPP4
A	819	HIS	-	EXPRESSION TAG	UNP D7PPP4
A	820	HIS	-	EXPRESSION TAG	UNP D7PPP4
A	821	HIS	-	EXPRESSION TAG	UNP D7PPP4
A	822	HIS	-	EXPRESSION TAG	UNP D7PPP4
A	823	HIS	-	EXPRESSION TAG	UNP D7PPP4
A	824	SER	-	EXPRESSION TAG	UNP D7PPP4
A	825	ARG	-	EXPRESSION TAG	UNP D7PPP4
A	826	ARG	-	EXPRESSION TAG	UNP D7PPP4
A	827	SER	-	EXPRESSION TAG	UNP D7PPP4
A	828	PRO	-	EXPRESSION TAG	UNP D7PPP4
A	829	GLU	-	EXPRESSION TAG	UNP D7PPP4
A	830	TYR	-	EXPRESSION TAG	UNP D7PPP4
A	831	PHE	-	EXPRESSION TAG	UNP D7PPP4
A	832	LYS	-	EXPRESSION TAG	UNP D7PPP4
A	833	GLY	-	EXPRESSION TAG	UNP D7PPP4
A	834	PRO	-	EXPRESSION TAG	UNP D7PPP4
A	835	PRO	-	EXPRESSION TAG	UNP D7PPP4
A	836	SER	-	EXPRESSION TAG	UNP D7PPP4
A	837	PRO	-	EXPRESSION TAG	UNP D7PPP4
A	838	ARG	-	EXPRESSION TAG	UNP D7PPP4
A	839	SER	-	EXPRESSION TAG	UNP D7PPP4
A	975	THR	ALA	CONFLICT	UNP D7PPP4
A	1008	ARG	LYS	CONFLICT	UNP D7PPP4
A	1212	SER	ARG	CONFLICT	UNP D7PPP4

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 ($\text{RSRZ} > 2$). 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.

- [illegible]

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	40.33Å 85.85Å 134.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.67 – 3.00 39.64 – 3.00	Depositor EDS
% Data completeness (in resolution range)	97.8 (39.67-3.00) 97.9 (39.64-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.56 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.248 , 0.306 0.262 , 0.325	Depositor DCC
R_{free} test set	492 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	62.0	Xtriage
Anisotropy	0.094	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 35.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	2234	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 7.27% 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	0.75	0/2276	0.79	3/3089 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1125	LEU	CA-CB-CG	7.70	133.01	115.30
1	A	1258	LEU	CA-CB-CG	6.79	130.92	115.30
1	A	1275	ALA	N-CA-C	-5.66	95.71	111.00

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	2234	0	2046	143	0
All	All	2234	0	2046	143	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 33.

The worst 5 of 143 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:1117:GLY:C	1:A:1118:LEU:HD23	1.42	1.35
1:A:1106:TYR:HE2	1:A:1108:TRP:NE1	1.34	1.25
1:A:1063:LYS:NZ	1:A:1065:ILE:CD1	2.01	1.22
1:A:1106:TYR:CE2	1:A:1108:TRP:NE1	2.09	1.19
1:A:1020:ASP:OD2	1:A:1023:ASN:OD1	1.61	1.19

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	295/469 (63%)	290 (98%)	5 (2%)	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	219/381 (58%)	172 (78%)	47 (22%)	1	5

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1059	ASN
1	A	1087	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1258	LEU
1	A	1065	ILE
1	A	1090	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1032	ASN
1	A	1075	ASN
1	A	1107	ASN
1	A	1023	ASN
1	A	1105	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 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	303/469 (64%)	-0.32	1 (0%)	94 84	30, 41, 50, 60	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1070	ASN	2.6

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.