W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2020 – 09:36 AM BST

PDB ID : 6Y4R
Title : Cytoplasmic domain of TssL from Acinetobacter baumannii
Authors : Ruiz, F.M.; Romero, A.
Deposited on : 2020-02-22
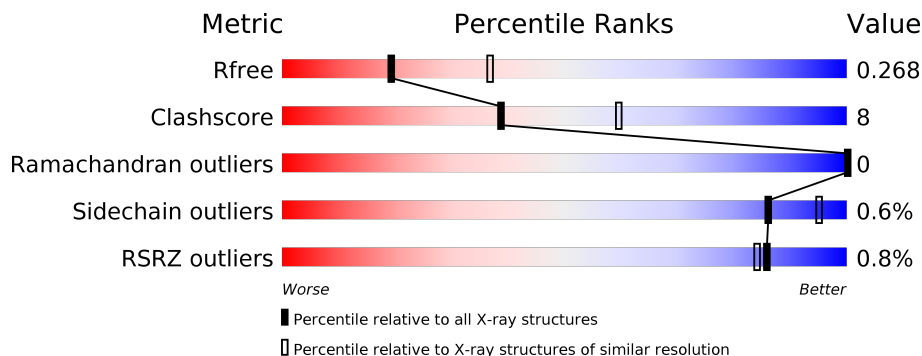Resolution : 2.59 Å(reported)

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 2.59 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	Percentile Ranks	Value
Rfree		0.268
Clashscore		8
Ramachandran outliers		0
Sidechain outliers		0.6%
RSRZ outliers		0.8%

Worse　　　　　　*Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	161	% 81% 19% •
1	B	161	% 85% 15%
1	C	161	% 76% 21% ••
1	D	161	75% 24% ••
1	E	161	3% 75% 24% ••
1	F	161	80% 20% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 8196 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 DotU family type IV/VI secretion system protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	160	Total	C	N	O	S	0	0	0
			1325	861	222	240	2			
1	B	161	Total	C	N	O	S	0	0	0
			1329	863	223	241	2			
1	C	160	Total	C	N	O	S	0	0	0
			1325	861	222	240	2			
1	D	160	Total	C	N	O	S	0	0	0
			1325	861	222	240	2			
1	E	159	Total	C	N	O	S	0	0	0
			1320	858	221	239	2			
1	F	160	Total	C	N	O	S	0	0	0
			1324	860	222	240	2			

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	43	Total	O	0	0
			43	43		
2	B	47	Total	O	0	0
			47	47		
2	C	43	Total	O	0	0
			43	43		
2	D	39	Total	O	0	0
			39	39		
2	E	30	Total	O	0	0
			30	30		
2	F	46	Total	O	0	0
			46	46		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: DotU family type IV/VI secretion system protein
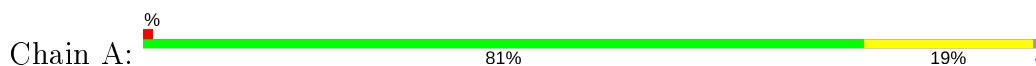
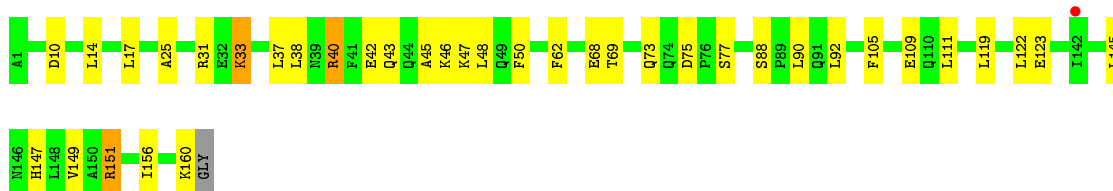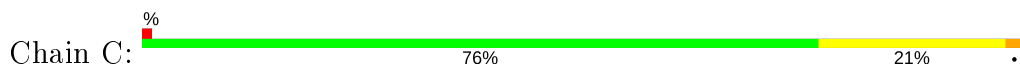• Molecule 1: DotU family type IV/VI secretion system protein
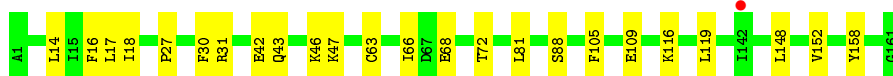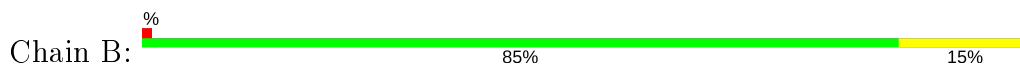
• Molecule 1: DotU family type IV/VI secretion system protein

• Molecule 1: DotU family type IV/VI secretion system protein
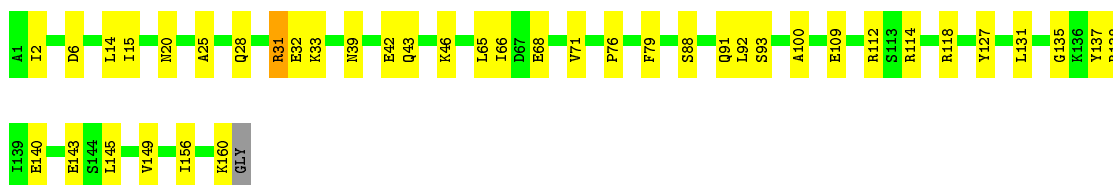
• Molecule 1: DotU family type IV/VI secretion system protein

Chain E:

3%

75%   24%

ALA I2 N3 L4 I5 L14 I15 I18 F30 R31 E32 K33 L38 E42 Q43 K46 K47 L48 Q49 D57 F62 C63 A64 L65 I66 D67 E68 V71 T72 Q73 W85 S88 F105 E106 E109 Q110 R118 V124 C128 Q134 Y137 R138

I139 E140 S141 I142 E143 H147 L148 R151 V152 K160 GLY

● Molecule 1: DotU family type IV/VI secretion system protein

Chain F:

80%   20%

ALA I2 N3 L4 I5 D6 L14 I15 R31 E32 K33 F41 E42 K46 K47 K59 C63 A64 L65 I66 D67 E68 V71 T72 W85 Q91 L94 A100 L108 E109 Q110 L111 R112 L129 L145 V149 I156 K160 G161
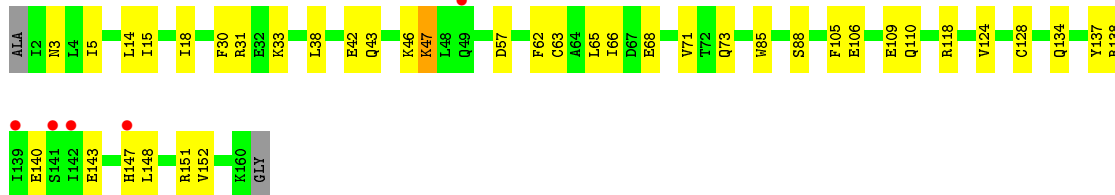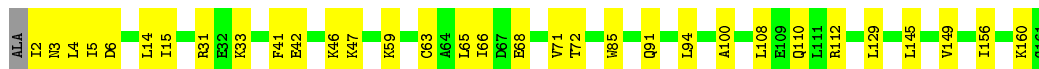
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	96.84Å 78.83Å 96.97Å 90.00° 119.77° 90.00°	Depositor
Resolution (Å)	48.62 – 2.59 48.62 – 2.59	Depositor EDS
% Data completeness (in resolution range)	95.6 (48.62-2.59) 95.6 (48.62-2.59)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.37 (at 2.58Å)	Xtriage
Refinement program	PHENIX v1.10	Depositor
R, R_{free}	0.209 , 0.269 0.207 , 0.268	Depositor DCC
R_{free} test set	1812 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å²)	28.2	Xtriage
Anisotropy	1.065	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.35 , 54.9	EDS
L-test for twinning [2]	$<|L|> = 0.49$, $<L^2> = 0.32$	Xtriage
Estimated twinning fraction	0.014 for -h-l,k,h 0.014 for l,k,-h-l 0.000 for h,-k,-h-l 0.000 for -h-l,-k,l 0.000 for l,-k,h	Xtriage
Reported twinning fraction	0.030 for l,k,-h-l	Depositor
Outliers	3 of 37812 reflections (0.008%)	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8196	wwPDB-VP
Average B, all atoms (Å²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 45.90 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2372e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 (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	0.49	0/1354	0.70	2/1829 (0.1%)
1	B	0.44	0/1358	0.69	0/1834
1	C	0.49	0/1354	0.76	3/1829 (0.2%)
1	D	0.43	0/1354	0.68	1/1829 (0.1%)
1	E	0.47	0/1349	0.72	3/1822 (0.2%)
1	F	0.45	0/1353	0.71	1/1827 (0.1%)
All	All	0.46	0/8122	0.71	10/10970 (0.1%)

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	33	LYS	CA-CB-CG	7.07	128.96	113.40
1	C	33	LYS	CB-CG-CD	-6.66	94.28	111.60
1	A	47	LYS	CB-CA-C	-6.54	97.31	110.40
1	A	47	LYS	CA-CB-CG	6.35	127.36	113.40
1	E	47	LYS	CG-CD-CE	-6.18	93.36	111.90

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	1325	0	1319	20	0
1	B	1329	0	1322	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1325	0	1319	21	0
1	D	1325	0	1319	24	0
1	E	1320	0	1311	27	0
1	F	1324	0	1314	20	0
2	A	43	0	0	2	0
2	B	47	0	0	0	0
2	C	43	0	0	0	0
2	D	39	0	0	0	0
2	E	30	0	0	1	0
2	F	46	0	0	2	0
All	All	8196	0	7904	119	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 8.

The worst 5 of 119 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:105:PHE:O	1:C:109:GLU:HG2	1.78	0.84
1:C:10:ASP:OD2	1:C:40:ARG:NH2	2.12	0.83
1:D:31:ARG:NH2	1:D:88:SER:O	2.12	0.83
1:B:31:ARG:NH2	1:B:88:SER:O	2.16	0.77
1:E:43:GLN:O	1:E:47:LYS:HG2	1.83	0.77

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	158/161 (98%)	155 (98%)	3 (2%)	0	100	100
1	B	159/161 (99%)	158 (99%)	1 (1%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	158/161 (98%)	155 (98%)	3 (2%)	0	100	100
1	D	158/161 (98%)	156 (99%)	2 (1%)	0	100	100
1	E	157/161 (98%)	155 (99%)	2 (1%)	0	100	100
1	F	158/161 (98%)	157 (99%)	1 (1%)	0	100	100
All	All	948/966 (98%)	936 (99%)	12 (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	142/142 (100%)	142 (100%)	0	100	100
1	B	142/142 (100%)	140 (99%)	2 (1%)	67	85
1	C	142/142 (100%)	140 (99%)	2 (1%)	67	85
1	D	142/142 (100%)	141 (99%)	1 (1%)	84	94
1	E	142/142 (100%)	142 (100%)	0	100	100
1	F	142/142 (100%)	142 (100%)	0	100	100
All	All	852/852 (100%)	847 (99%)	5 (1%)	86	95

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	16	PHE
1	B	119	LEU
1	C	40	ARG
1	C	68	GLU
1	D	93	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	73	GLN
1	D	103	GLN
1	E	134	GLN
1	C	82	GLN
1	E	147	HIS

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 monosaccharides 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 (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	160/161 (99%)	0.03	1 (0%)	89	88	23, 36, 61, 92	0
1	B	161/161 (100%)	-0.08	1 (0%)	89	88	23, 34, 56, 84	0
1	C	160/161 (99%)	0.03	1 (0%)	89	88	24, 39, 66, 93	0
1	D	160/161 (99%)	-0.08	0	100	100	21, 33, 53, 65	0
1	E	159/161 (98%)	0.14	5 (3%)	49	42	29, 41, 67, 98	0
1	F	160/161 (99%)	0.08	0	100	100	25, 35, 66, 83	0
All	All	960/966 (99%)	0.02	8 (0%)	86	84	21, 36, 62, 98	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	142	ILE	3.7
1	E	139	ILE	3.6
1	A	139	ILE	3.1
1	E	49	GLN	2.6
1	E	147	HIS	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 monosaccharides 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.