



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2017 – 01:28 PM EDT

PDB ID : 3F72  
Title : Crystal Structure of the Staphylococcus aureus pl258 CadC Metal Binding Site 2 Mutant  
Authors : Kandegedara, A.; Thiyagarajan, S.; Kondapalli, K.C.; Stemmler, T.L.; Rosen, B.P.  
Deposited on : unknown  
Resolution : 2.31 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](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

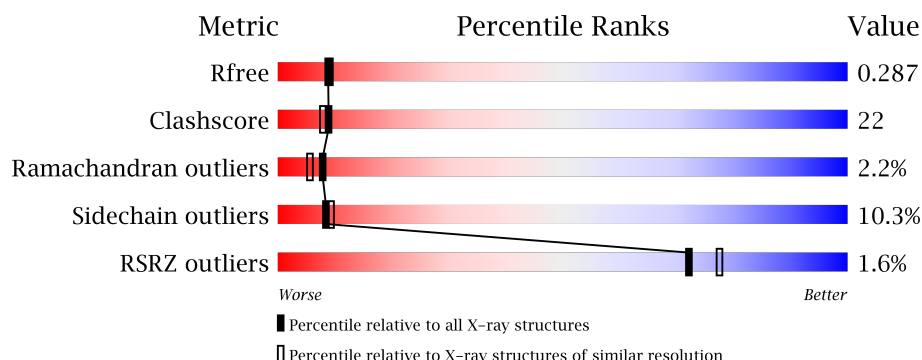
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.31 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	4787 (2.34-2.30)
Clashscore	112137	5439 (2.34-2.30)
Ramachandran outliers	110173	5386 (2.34-2.30)
Sidechain outliers	110143	5385 (2.34-2.30)
RSRZ outliers	101464	4814 (2.34-2.30)

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	122	<div> <div>60%</div> <div>25%</div> <div>12%</div> </div>
1	B	122	<div> <div>56%</div> <div>23%</div> <div>10%</div> <div>11%</div> </div>
1	C	122	<div> <div>2%</div> <div>46%</div> <div>30%</div> <div>8%</div> <div>16%</div> </div>
1	D	122	<div> <div>2%</div> <div>42%</div> <div>36%</div> <div>18%</div> </div>
1	E	122	<div> <div>4%</div> <div>34%</div> <div>39%</div> <div>8%</div> <div>17%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	122	 A horizontal bar chart showing the quality of chain F. The bar is divided into four segments: green (49%), yellow (26%), orange (8%), and grey (16%).

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4843 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 Cadmium efflux system accessory protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	107	Total	C	N	O	S	0	0	0
			821	517	145	154	5			
1	B	108	Total	C	N	O	S	0	0	0
			832	521	147	159	5			
1	C	103	Total	C	N	O	S	0	0	0
			797	500	142	150	5			
1	D	100	Total	C	N	O	S	0	0	0
			761	479	130	147	5			
1	E	101	Total	C	N	O	S	0	0	0
			737	459	126	147	5			
1	F	102	Total	C	N	O	S	0	0	0
			748	465	130	148	5			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	11	GLY	CYS	ENGINEERED	UNP P20047
A	101	GLY	ASP	ENGINEERED	UNP P20047
A	103	ALA	HIS	ENGINEERED	UNP P20047
B	11	GLY	CYS	ENGINEERED	UNP P20047
B	101	GLY	ASP	ENGINEERED	UNP P20047
B	103	ALA	HIS	ENGINEERED	UNP P20047
C	11	GLY	CYS	ENGINEERED	UNP P20047
C	101	GLY	ASP	ENGINEERED	UNP P20047
C	103	ALA	HIS	ENGINEERED	UNP P20047
D	11	GLY	CYS	ENGINEERED	UNP P20047
D	101	GLY	ASP	ENGINEERED	UNP P20047
D	103	ALA	HIS	ENGINEERED	UNP P20047
E	11	GLY	CYS	ENGINEERED	UNP P20047
E	101	GLY	ASP	ENGINEERED	UNP P20047
E	103	ALA	HIS	ENGINEERED	UNP P20047
F	11	GLY	CYS	ENGINEERED	UNP P20047
F	101	GLY	ASP	ENGINEERED	UNP P20047

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	103	ALA	HIS	ENGINEERED	UNP P20047

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

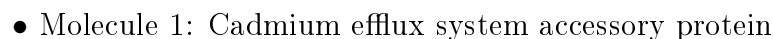
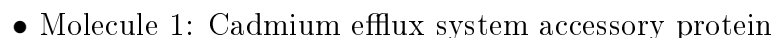
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Na 2 2	0	0
2	A	2	Total Na 2 2	0	0
2	F	1	Total Na 1 1	0	0

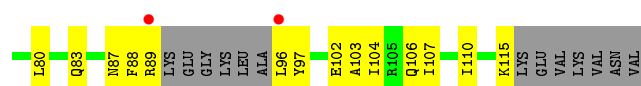
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	26	Total O 26 26	0	0
3	B	27	Total O 27 27	0	0
3	C	25	Total O 25 25	0	0
3	D	24	Total O 24 24	0	0
3	E	24	Total O 24 24	0	0
3	F	16	Total O 16 16	0	0

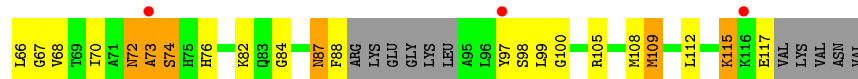
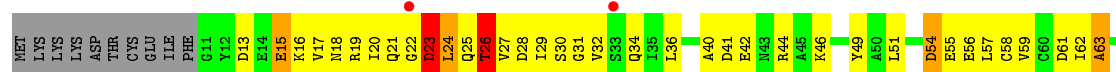


- Molecule 1: Cadmium efflux system accessory protein





- Molecule 1: Cadmium efflux system accessory protein



- Molecule 1: Cadmium efflux system accessory protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	89.47Å 89.47Å 148.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.30 – 2.31 34.30 – 2.31	Depositor EDS
% Data completeness (in resolution range)	97.3 (34.30-2.31) 97.3 (34.30-2.31)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.39 (at 2.31Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.237 , 0.288 0.235 , 0.287	Depositor DCC
$R_{free}$ test set	2532 reflections (5.36%)	DCC
Wilson B-factor (Å <sup>2</sup> )	49.6	Xtriage
Anisotropy	0.031	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 41.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.366 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4843	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.68% 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: NA

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	1.46	5/828 (0.6%)	1.18	1/1114 (0.1%)
1	B	1.60	9/840 (1.1%)	1.38	6/1130 (0.5%)
1	C	1.59	7/803 (0.9%)	1.37	7/1079 (0.6%)
1	D	1.36	3/768 (0.4%)	1.18	3/1037 (0.3%)
1	E	1.32	2/742 (0.3%)	1.17	3/1006 (0.3%)
1	F	1.22	2/754 (0.3%)	1.23	4/1021 (0.4%)
All	All	1.44	28/4735 (0.6%)	1.26	24/6387 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1
1	E	0	2
All	All	0	3

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	81	TYR	CD2-CE2	7.92	1.51	1.39
1	C	81	TYR	CD1-CE1	7.09	1.50	1.39
1	A	42	GLU	CG-CD	7.03	1.62	1.51
1	B	63	ALA	CA-CB	6.48	1.66	1.52
1	A	86	VAL	CB-CG1	-6.25	1.39	1.52

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	44	ARG	NE-CZ-NH1	11.39	125.99	120.30
1	C	28	ASP	CB-CG-OD2	-8.62	110.54	118.30
1	F	66	LEU	CA-CB-CG	8.14	134.02	115.30
1	F	80	LEU	CA-CB-CG	8.05	133.81	115.30
1	F	16	LYS	CD-CE-NZ	6.95	127.67	111.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	91	GLU	Peptide
1	E	26	THR	Peptide
1	E	87	ASN	Peptide

## 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	821	0	834	31	0
1	B	832	0	838	38	0
1	C	797	0	820	47	0
1	D	761	0	754	43	0
1	E	737	0	701	54	0
1	F	748	0	707	33	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	F	1	0	0	0	0
3	A	26	0	0	6	0
3	B	27	0	0	0	0
3	C	25	0	0	1	0
3	D	24	0	0	3	0
3	E	24	0	0	1	0
3	F	16	0	0	3	0
All	All	4843	0	4654	210	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 22.

The worst 5 of 210 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:E:24:LEU:N	1:E:25:GLN:HA	1.36	1.23
1:E:23:ASP:HB3	1:E:25:GLN:CA	1.70	1.20
1:B:47:ILE:HD11	1:B:66:LEU:HD11	1.25	1.18
1:E:20:ILE:HD11	1:F:54:ASP:OD1	1.51	1.11
1:E:23:ASP:HB3	1:E:25:GLN:CB	1.80	1.10

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	103/122 (84%)	96 (93%)	7 (7%)	0	100	100
1	B	106/122 (87%)	100 (94%)	6 (6%)	0	100	100
1	C	99/122 (81%)	91 (92%)	7 (7%)	1 (1%)	18	19
1	D	96/122 (79%)	88 (92%)	7 (7%)	1 (1%)	18	19
1	E	97/122 (80%)	83 (86%)	7 (7%)	7 (7%)	1	0
1	F	98/122 (80%)	85 (87%)	9 (9%)	4 (4%)	3	1
All	All	599/732 (82%)	543 (91%)	43 (7%)	13 (2%)	8	5

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	23	ASP
1	E	24	LEU
1	E	73	ALA
1	E	16	LYS
1	F	65	ILE

### 5.3.2 Protein sidechains ⓘ

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	86/103 (84%)	82 (95%)	4 (5%)	30	41
1	B	87/103 (84%)	77 (88%)	10 (12%)	6	7
1	C	85/103 (82%)	78 (92%)	7 (8%)	13	16
1	D	80/103 (78%)	72 (90%)	8 (10%)	9	10
1	E	73/103 (71%)	62 (85%)	11 (15%)	3	3
1	F	74/103 (72%)	64 (86%)	10 (14%)	4	4
All	All	485/618 (78%)	435 (90%)	50 (10%)	8	9

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	18	ASN
1	D	56	GLU
1	F	49	TYR
1	D	20	ILE
1	D	29	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	43	ASN
1	F	53	GLN
1	E	76	HIS
1	B	72	ASN
1	F	43	ASN

### 5.3.3 RNA ⓘ

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 5 ligands modelled in this entry, 5 are 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 [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, 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	107/122 (87%)	0.26	1 (0%) 84 87	34, 48, 73, 81	0
1	B	108/122 (88%)	0.26	0 100 100	29, 42, 68, 78	1 (0%)
1	C	103/122 (84%)	0.38	2 (1%) 67 73	36, 53, 71, 75	0
1	D	100/122 (81%)	0.33	2 (2%) 65 72	37, 55, 76, 82	0
1	E	101/122 (82%)	0.50	5 (4%) 30 37	42, 60, 83, 91	0
1	F	102/122 (83%)	0.37	0 100 100	42, 66, 86, 91	0
All	All	621/732 (84%)	0.35	10 (1%) 72 77	29, 54, 79, 91	1 (0%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	22	GLY	4.8
1	D	96	LEU	3.3
1	C	12	TYR	2.7
1	E	116	LYS	2.6
1	D	89	ARG	2.5

### 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	NA	A	123	1/1	0.88	0.12	-1.62	41,41,41,41	0
2	NA	B	123	1/1	0.88	0.17	-	55,55,55,55	0
2	NA	B	124	1/1	0.94	0.34	-	55,55,55,55	0
2	NA	A	124	1/1	0.96	0.35	-	58,58,58,58	0
2	NA	F	123	1/1	0.93	0.26	-	69,69,69,69	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.