



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 13, 2024 – 08:55 AM EDT

PDB ID : 4EM1  
Title : staphylococcus aureus MarR native  
Authors : Chang, Y.M.; Chen, C.K.-M.; Wang, A.H.-J.  
Deposited on : 2012-04-11  
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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtrriage (Phenix) : 1.20.1  
EDS : 2.36.2  
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) : 2.36.2

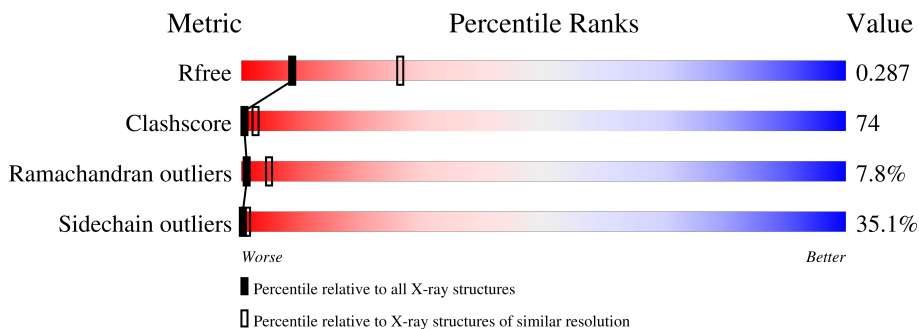
# 1 Overall quality at a glance

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)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	178	 18% 34% 23% 5% 20%

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 1286 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 Uncharacterized HTH-type transcriptional regulator SAR2349.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	143	1183	759	202	218	4	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	THR	-	expression tag	UNP Q6GEG9
A	8	ALA	-	expression tag	UNP Q6GEG9
A	9	ALA	-	expression tag	UNP Q6GEG9
A	10	ALA	-	expression tag	UNP Q6GEG9
A	11	LYS	-	expression tag	UNP Q6GEG9
A	12	PHE	-	expression tag	UNP Q6GEG9
A	13	GLU	-	expression tag	UNP Q6GEG9
A	14	ARG	-	expression tag	UNP Q6GEG9
A	15	GLN	-	expression tag	UNP Q6GEG9
A	16	HIS	-	expression tag	UNP Q6GEG9
A	17	MET	-	expression tag	UNP Q6GEG9
A	18	ASP	-	expression tag	UNP Q6GEG9
A	19	SER	-	expression tag	UNP Q6GEG9
A	20	PRO	-	expression tag	UNP Q6GEG9
A	21	ASP	-	expression tag	UNP Q6GEG9
A	22	LEU	-	expression tag	UNP Q6GEG9
A	23	GLY	-	expression tag	UNP Q6GEG9
A	24	THR	-	expression tag	UNP Q6GEG9
A	25	ASP	-	expression tag	UNP Q6GEG9
A	26	ASP	-	expression tag	UNP Q6GEG9
A	27	ASP	-	expression tag	UNP Q6GEG9
A	28	ASP	-	expression tag	UNP Q6GEG9
A	29	LYS	-	expression tag	UNP Q6GEG9
A	30	ALA	-	expression tag	UNP Q6GEG9
A	31	MET	-	expression tag	UNP Q6GEG9
A	32	ALA	-	expression tag	UNP Q6GEG9
A	33	ASP	-	expression tag	UNP Q6GEG9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	ILE	-	expression tag	UNP Q6GEG9
A	35	GLY	-	expression tag	UNP Q6GEG9
A	36	SER	-	expression tag	UNP Q6GEG9
A	37	ASP	-	expression tag	UNP Q6GEG9
A	38	PHE	-	expression tag	UNP Q6GEG9

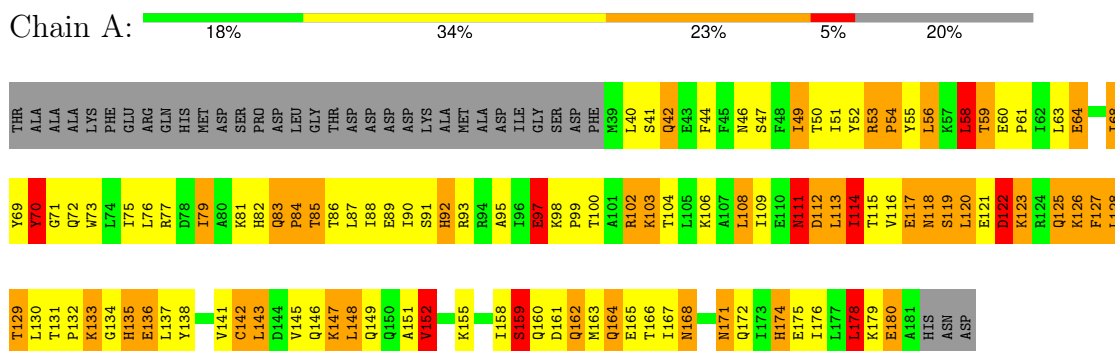
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	103	Total O 103 103	0	0

### 3 Residue-property plots [i](#)

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. 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. 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: Uncharacterized HTH-type transcriptional regulator SAR2349



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	36.12Å 75.22Å 109.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	22.00 – 3.00 32.56 – 2.43	Depositor EDS
% Data completeness (in resolution range)	95.8 (22.00-3.00) 95.3 (32.56-2.43)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.00 (at 2.42Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.269 , 0.293 0.269 , 0.287	Depositor DCC
$R_{free}$ test set	135 reflections (2.39%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.3	Xtrriage
Anisotropy	0.091	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 127.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	1286	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.40% 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](#)

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	1.23	7/1203 (0.6%)	1.40	15/1626 (0.9%)

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	178	LEU	CB-CG	13.15	1.90	1.52
1	A	108	LEU	CB-CG	11.17	1.84	1.52
1	A	152	VAL	CA-CB	6.16	1.67	1.54
1	A	152	VAL	CB-CG2	5.59	1.64	1.52
1	A	159	SER	CB-OG	5.42	1.49	1.42

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	108	LEU	CA-CB-CG	-12.06	87.55	115.30
1	A	152	VAL	N-CA-CB	-10.63	88.11	111.50
1	A	152	VAL	CB-CA-C	10.40	131.15	111.40
1	A	120	LEU	CA-CB-CG	8.95	135.89	115.30
1	A	178	LEU	CB-CG-CD2	-8.68	96.24	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	1183	0	1225	179	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	103	0	0	5	0
All	All	1286	0	1225	179	1

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 74.

The worst 5 of 179 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:108:LEU:CB	1:A:108:LEU:CG	1.85	1.52
1:A:178:LEU:CB	1:A:178:LEU:CG	1.90	1.48
1:A:149:GLN:O	1:A:152:VAL:HG12	1.20	1.27
1:A:176:ILE:O	1:A:180:GLU:HB2	1.26	1.23
1:A:112:ASP:HB3	1:A:133:LYS:HD3	1.25	1.10

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:69:TYR:OH	1:A:160:GLN:O[1_455]	1.77	0.43

## 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	141/178 (79%)	112 (79%)	18 (13%)	11 (8%)	<b>1</b> <b>4</b>

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	ASP

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	41	SER
1	A	85	THR
1	A	97	GLU
1	A	111	ASN

### 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	134/162 (83%)	87 (65%)	47 (35%)	<b>0</b> <b>1</b>

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

Mol	Chain	Res	Type
1	A	125	GLN
1	A	143	LEU
1	A	126	LYS
1	A	133	LYS
1	A	148	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	168	ASN
1	A	171	ASN
1	A	172	GLN
1	A	149	GLN
1	A	160	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 monosaccharides 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

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.