



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 11:28 PM UTC

PDB ID : 4F20 / pdb\_00004f20  
Title : Crystal structures reveal the multi-ligand binding mechanism of the Staphylococcus aureus ClfB  
Authors : Yang, M.J.; Xiang, H.; Wang, J.W.; Liu, B.; Chen, Y.G.; Liu, L.; Deng, X.M.; Feng, Y.  
Deposited on : 2012-05-07  
Resolution : 2.50 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 : 4-5-2 with Phenix2.0  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

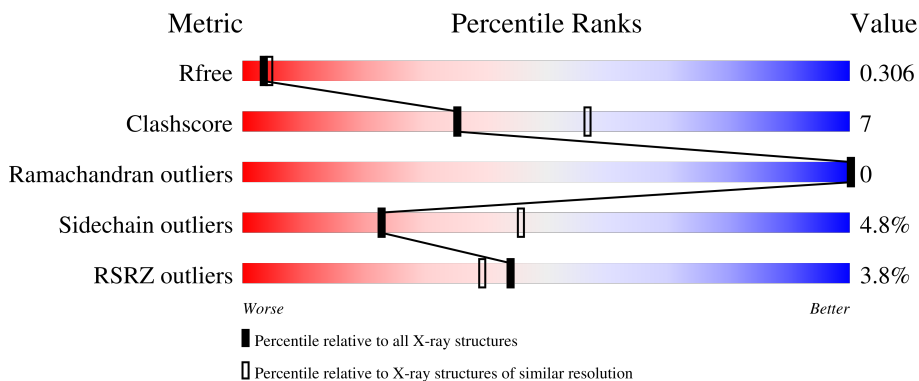
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	180053	5829 (2.50-2.50)
Clashscore	190562	6492 (2.50-2.50)
Ramachandran outliers	187476	6378 (2.50-2.50)
Sidechain outliers	187428	6380 (2.50-2.50)
RSRZ outliers	180081	5833 (2.50-2.50)

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  $\leq 5\%$ . 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	363	 3% 70% 19% 9%
2	Q	13	 15% 62% 23% 15%

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 2635 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 Clumping factor B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	331	2560	1601	421	535	3	0	0	0

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	180	MET	-	expression tag	UNP Q6GDH2
A	181	ARG	-	expression tag	UNP Q6GDH2
A	182	GLY	-	expression tag	UNP Q6GDH2
A	183	SER	-	expression tag	UNP Q6GDH2
A	184	HIS	-	expression tag	UNP Q6GDH2
A	185	HIS	-	expression tag	UNP Q6GDH2
A	186	HIS	-	expression tag	UNP Q6GDH2
A	187	HIS	-	expression tag	UNP Q6GDH2
A	188	HIS	-	expression tag	UNP Q6GDH2
A	189	HIS	-	expression tag	UNP Q6GDH2
A	190	GLU	-	expression tag	UNP Q6GDH2
A	191	ASN	-	expression tag	UNP Q6GDH2
A	192	LEU	-	expression tag	UNP Q6GDH2
A	193	TYR	-	expression tag	UNP Q6GDH2
A	194	PHE	-	expression tag	UNP Q6GDH2
A	195	GLN	-	expression tag	UNP Q6GDH2
A	196	GLY	-	expression tag	UNP Q6GDH2

- Molecule 2 is a protein called peptide from Dermokine.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	Q	11	60	30	12	18	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	315	GLY	-	expression tag	UNP Q6E0U4

- Molecule 3 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total Mg 2 2	0	0

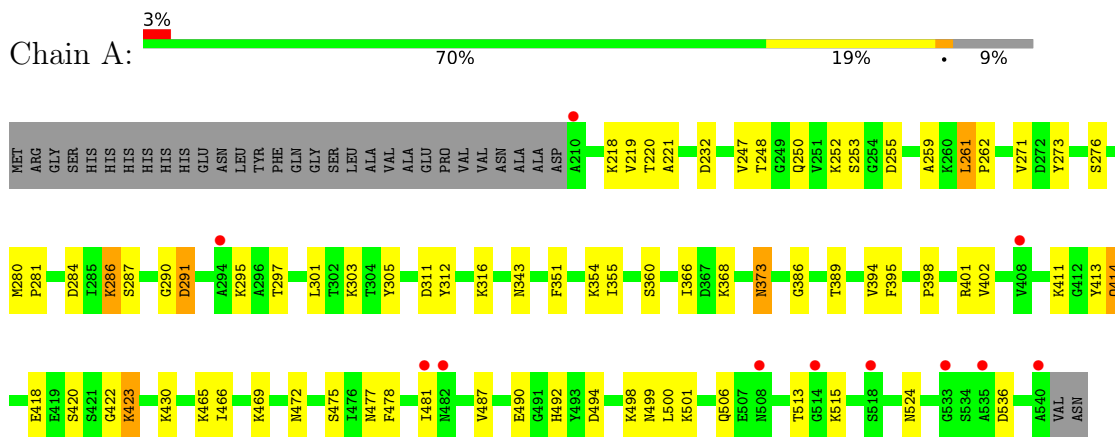
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	12	Total O 12 12	0	0
4	Q	1	Total O 1 1	0	0

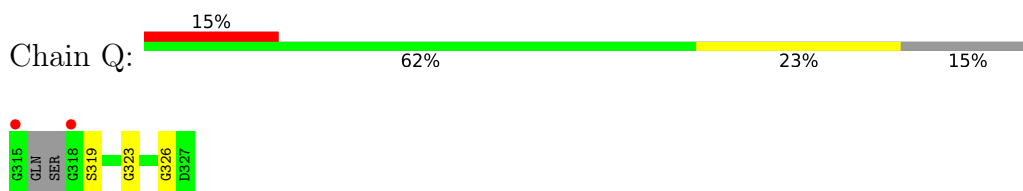
### 3 Residue-property plots [i](#)

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 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: Clumping factor B



- Molecule 2: peptide from Dermokine



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.09Å 70.09Å 175.76Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.37 – 2.50 34.37 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.5 (34.37-2.50) 97.9 (34.37-2.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.56 (at 2.51Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, $R_{free}$	0.224 , 0.270 (Not available) , 0.306	Depositor DCC
$R_{free}$ test set	896 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	64.2	Xtrriage
Anisotropy	0.655	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 64.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.036 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2635	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	112.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.29% 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: MG

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	0.30	0/2608	0.73	2/3538 (0.1%)
2	Q	0.27	0/58	0.58	0/73
All	All	0.29	0/2666	0.73	2/3611 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	386	GLY	N-CA-C	-5.26	108.08	114.92
1	A	414	GLN	CB-CA-C	-5.02	110.40	117.23

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	2560	0	2447	38	0
2	Q	60	0	43	3	0
3	A	2	0	0	0	0
4	A	12	0	0	0	0
4	Q	1	0	0	0	0
All	All	2635	0	2490	38	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 7.

The worst 5 of 38 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:423:LYS:NZ	1:A:494:ASP:OD2	2.31	0.64
1:A:420:SER:HB2	1:A:501:LYS:HB2	1.81	0.61
1:A:287:SER:OG	1:A:291:ASP:OD2	2.17	0.59
1:A:413:TYR:HA	1:A:420:SER:HB3	1.83	0.59
1:A:411:LYS:NZ	1:A:414:GLN:O	2.30	0.57

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	329/363 (91%)	301 (92%)	28 (8%)	0	100	100
2	Q	8/13 (62%)	8 (100%)	0	0	100	100
All	All	337/376 (90%)	309 (92%)	28 (8%)	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	283/309 (92%)	269 (95%)	14 (5%)	22	45
2	Q	6/8 (75%)	6 (100%)	0	100	100
All	All	289/317 (91%)	275 (95%)	14 (5%)	23	46

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	301	LEU
1	A	354	LYS
1	A	536	ASP
1	A	402	VAL
1	A	423	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	405	ASN
1	A	472	ASN

### 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 oligosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 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	331/363 (91%)	0.53	11 (3%) 49 45	78, 106, 162, 219	0
2	Q	11/13 (84%)	1.06	2 (18%) 3 2	80, 92, 131, 181	0
All	All	342/376 (90%)	0.55	13 (3%) 44 39	78, 106, 162, 219	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	540	ALA	4.4
1	A	210	ALA	3.7
1	A	482	ASN	3.4
2	Q	318	GLY	3.2
1	A	481	ILE	3.0

### 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 oligosaccharides 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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MG	A	601	1/1	0.72	0.28	134,134,134,134	0
3	MG	A	602	1/1	0.74	0.26	120,120,120,120	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.