



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 18, 2024 – 09:13 AM EST

PDB ID : 4F6L  
Title : Crystal structure of Aureusimine biosynthetic cluster reductase domain  
Authors : Mok, M.; Junop, M.  
Deposited on : 2012-05-14  
Resolution : 3.86 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  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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

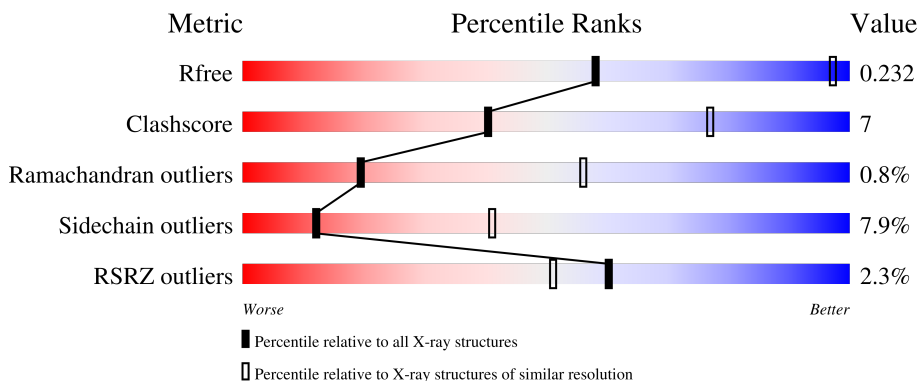
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.86 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	1048 (4.10-3.62)
Clashscore	141614	1015 (4.08-3.64)
Ramachandran outliers	138981	1069 (4.10-3.62)
Sidechain outliers	138945	1062 (4.10-3.62)
RSRZ outliers	127900	1206 (4.12-3.60)

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 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	508	 3% 56% 16% 26%
1	B	508	 59% 13% 27%

## 2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 5714 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 AusA reductase domain protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	370	2842	1808	479	537	18	0	0	0
1	A	378	2872	1828	482	546	16	0	0	0

There are 90 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1895	MET	-	expression tag	UNP Q99X42
B	1896	GLY	-	expression tag	UNP Q99X42
B	1897	SER	-	expression tag	UNP Q99X42
B	1898	SER	-	expression tag	UNP Q99X42
B	1899	HIS	-	expression tag	UNP Q99X42
B	1900	HIS	-	expression tag	UNP Q99X42
B	1901	HIS	-	expression tag	UNP Q99X42
B	1902	HIS	-	expression tag	UNP Q99X42
B	1903	HIS	-	expression tag	UNP Q99X42
B	1904	HIS	-	expression tag	UNP Q99X42
B	1905	SER	-	expression tag	UNP Q99X42
B	1906	SER	-	expression tag	UNP Q99X42
B	1907	GLY	-	expression tag	UNP Q99X42
B	1908	LEU	-	expression tag	UNP Q99X42
B	1909	VAL	-	expression tag	UNP Q99X42
B	1910	PRO	-	expression tag	UNP Q99X42
B	1911	ARG	-	expression tag	UNP Q99X42
B	1912	GLY	-	expression tag	UNP Q99X42
B	1913	SER	-	expression tag	UNP Q99X42
B	1914	HIS	-	expression tag	UNP Q99X42
B	1915	MET	-	expression tag	UNP Q99X42
B	1916	ALA	-	expression tag	UNP Q99X42
B	1917	SER	-	expression tag	UNP Q99X42
B	1918	MET	-	expression tag	UNP Q99X42
B	1919	THR	-	expression tag	UNP Q99X42

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1920	GLY	-	expression tag	UNP Q99X42
B	1921	GLY	-	expression tag	UNP Q99X42
B	1922	GLN	-	expression tag	UNP Q99X42
B	1923	GLN	-	expression tag	UNP Q99X42
B	1924	MET	-	expression tag	UNP Q99X42
B	1925	GLY	-	expression tag	UNP Q99X42
B	1926	ARG	-	expression tag	UNP Q99X42
B	1927	ASP	-	expression tag	UNP Q99X42
B	1928	PRO	-	expression tag	UNP Q99X42
B	2392	ALA	-	expression tag	UNP Q99X42
B	2393	ALA	-	expression tag	UNP Q99X42
B	2394	ALA	-	expression tag	UNP Q99X42
B	2395	LEU	-	expression tag	UNP Q99X42
B	2396	GLU	-	expression tag	UNP Q99X42
B	2397	HIS	-	expression tag	UNP Q99X42
B	2398	HIS	-	expression tag	UNP Q99X42
B	2399	HIS	-	expression tag	UNP Q99X42
B	2400	HIS	-	expression tag	UNP Q99X42
B	2401	HIS	-	expression tag	UNP Q99X42
B	2402	HIS	-	expression tag	UNP Q99X42
A	1895	MET	-	expression tag	UNP Q99X42
A	1896	GLY	-	expression tag	UNP Q99X42
A	1897	SER	-	expression tag	UNP Q99X42
A	1898	SER	-	expression tag	UNP Q99X42
A	1899	HIS	-	expression tag	UNP Q99X42
A	1900	HIS	-	expression tag	UNP Q99X42
A	1901	HIS	-	expression tag	UNP Q99X42
A	1902	HIS	-	expression tag	UNP Q99X42
A	1903	HIS	-	expression tag	UNP Q99X42
A	1904	HIS	-	expression tag	UNP Q99X42
A	1905	SER	-	expression tag	UNP Q99X42
A	1906	SER	-	expression tag	UNP Q99X42
A	1907	GLY	-	expression tag	UNP Q99X42
A	1908	LEU	-	expression tag	UNP Q99X42
A	1909	VAL	-	expression tag	UNP Q99X42
A	1910	PRO	-	expression tag	UNP Q99X42
A	1911	ARG	-	expression tag	UNP Q99X42
A	1912	GLY	-	expression tag	UNP Q99X42
A	1913	SER	-	expression tag	UNP Q99X42
A	1914	HIS	-	expression tag	UNP Q99X42
A	1915	MET	-	expression tag	UNP Q99X42
A	1916	ALA	-	expression tag	UNP Q99X42

*Continued on next page...*

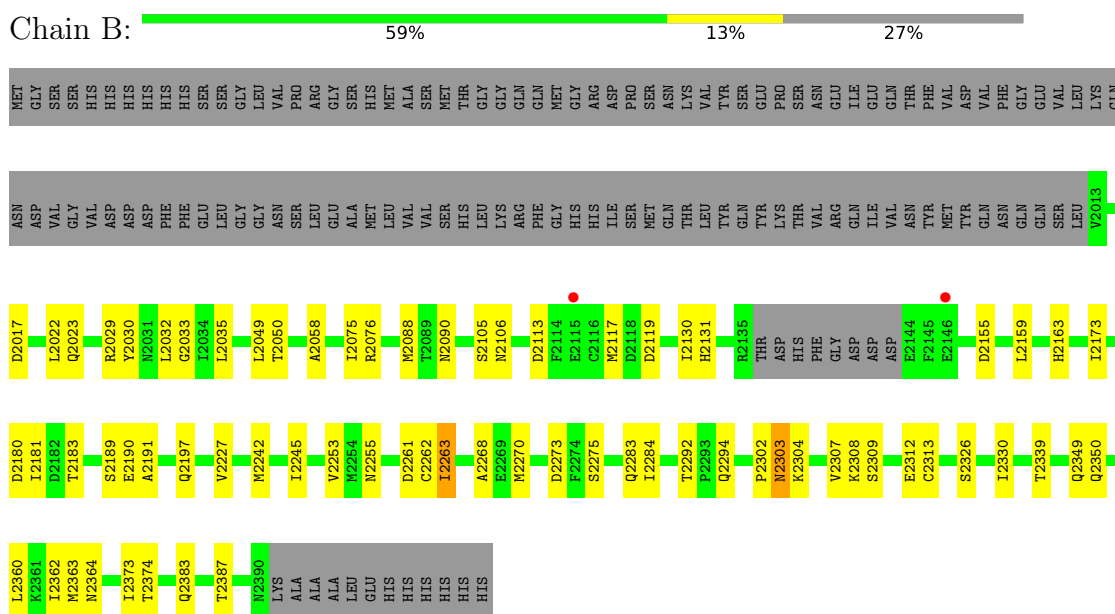
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1917	SER	-	expression tag	UNP Q99X42
A	1918	MET	-	expression tag	UNP Q99X42
A	1919	THR	-	expression tag	UNP Q99X42
A	1920	GLY	-	expression tag	UNP Q99X42
A	1921	GLY	-	expression tag	UNP Q99X42
A	1922	GLN	-	expression tag	UNP Q99X42
A	1923	GLN	-	expression tag	UNP Q99X42
A	1924	MET	-	expression tag	UNP Q99X42
A	1925	GLY	-	expression tag	UNP Q99X42
A	1926	ARG	-	expression tag	UNP Q99X42
A	1927	ASP	-	expression tag	UNP Q99X42
A	1928	PRO	-	expression tag	UNP Q99X42
A	2392	ALA	-	expression tag	UNP Q99X42
A	2393	ALA	-	expression tag	UNP Q99X42
A	2394	ALA	-	expression tag	UNP Q99X42
A	2395	LEU	-	expression tag	UNP Q99X42
A	2396	GLU	-	expression tag	UNP Q99X42
A	2397	HIS	-	expression tag	UNP Q99X42
A	2398	HIS	-	expression tag	UNP Q99X42
A	2399	HIS	-	expression tag	UNP Q99X42
A	2400	HIS	-	expression tag	UNP Q99X42
A	2401	HIS	-	expression tag	UNP Q99X42
A	2402	HIS	-	expression tag	UNP Q99X42

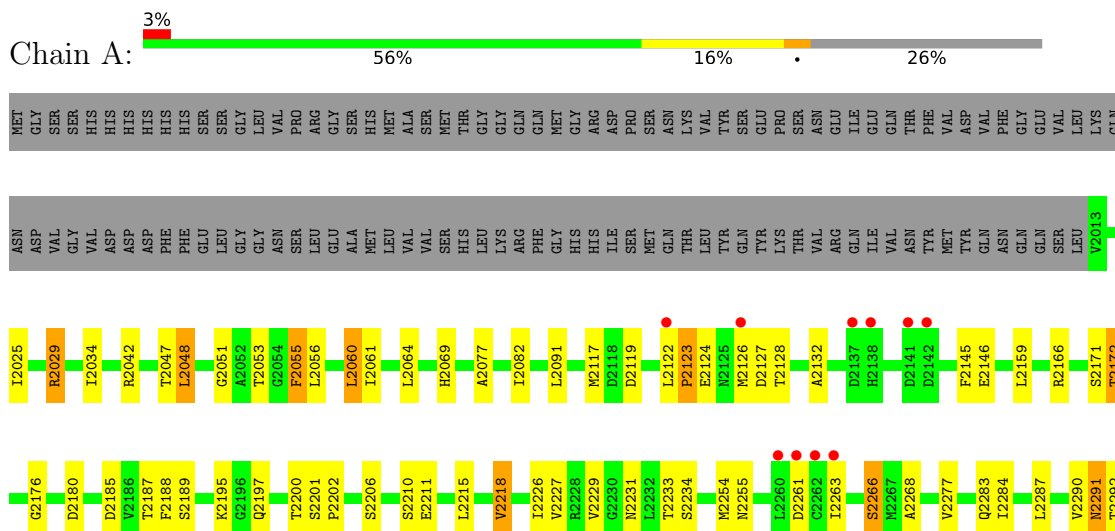
### 3 Residue-property plots

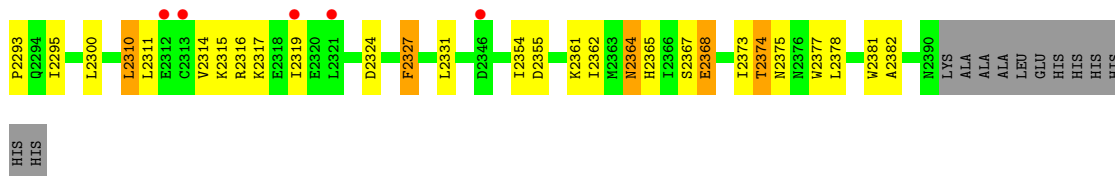
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: AusA reductase domain protein



- Molecule 1: AusA reductase domain protein





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.81Å 107.78Å 127.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.63 – 3.86 49.63 – 3.56	Depositor EDS
% Data completeness (in resolution range)	98.6 (49.63-3.86) 94.8 (49.63-3.56)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.69 (at 3.57Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, $R_{free}$	0.200 , 0.246 0.185 , 0.232	Depositor DCC
$R_{free}$ test set	1759 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	113.6	Xtrriage
Anisotropy	0.169	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 122.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.025 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5714	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	140.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% 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	0.27	0/2927	0.47	0/3994
1	B	0.27	0/2895	0.46	0/3944
All	All	0.27	0/5822	0.46	0/7938

There are no bond length outliers.

There are no bond angle outliers.

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	2872	0	2720	47	0
1	B	2842	0	2724	33	0
All	All	5714	0	5444	80	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 80 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:B:2117:MET:HG2	1:B:2155:ASP:HB3	1.67	0.77
1:A:2146:GLU:HG2	1:A:2202:PRO:HB2	1.69	0.75

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2047:THR:HG22	1:A:2128:THR:HB	1.72	0.70
1:A:2311:LEU:HA	1:A:2314:VAL:HG12	1.74	0.69
1:A:2029:ARG:NH2	1:A:2293:PRO:O	2.26	0.69

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	376/508 (74%)	342 (91%)	30 (8%)	4 (1%)	14	50
1	B	366/508 (72%)	343 (94%)	21 (6%)	2 (0%)	29	66
All	All	742/1016 (73%)	685 (92%)	51 (7%)	6 (1%)	19	56

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2124	GLU
1	B	2119	ASP
1	A	2069	HIS
1	A	2123	PRO
1	B	2035	LEU

### 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	298/456 (65%)	266 (89%)	32 (11%)	6	29
1	B	300/456 (66%)	285 (95%)	15 (5%)	24	53
All	All	598/912 (66%)	551 (92%)	47 (8%)	12	41

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

Mol	Chain	Res	Type
1	A	2206	SER
1	A	2283	GLN
1	A	2210	SER
1	A	2234	SER
1	A	2290	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2364	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 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

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	378/508 (74%)	0.20	15 (3%) 38 31	89, 143, 228, 250	0
1	B	370/508 (72%)	-0.02	2 (0%) 91 86	77, 129, 182, 220	0
All	All	748/1016 (73%)	0.09	17 (2%) 60 51	77, 135, 216, 250	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2263	ILE	4.9
1	A	2319	ILE	3.9
1	A	2262	CYS	3.8
1	A	2260	LEU	3.8
1	B	2146	GLU	2.9

### 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.