



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 17, 2024 – 02:23 AM EST

PDB ID : 4F6C  
Title : Crystal structure of Aureusimine biosynthetic cluster reductase domain  
Authors : Mok, M.; Junop, M.  
Deposited on : 2012-05-14  
Resolution : 2.81 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  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

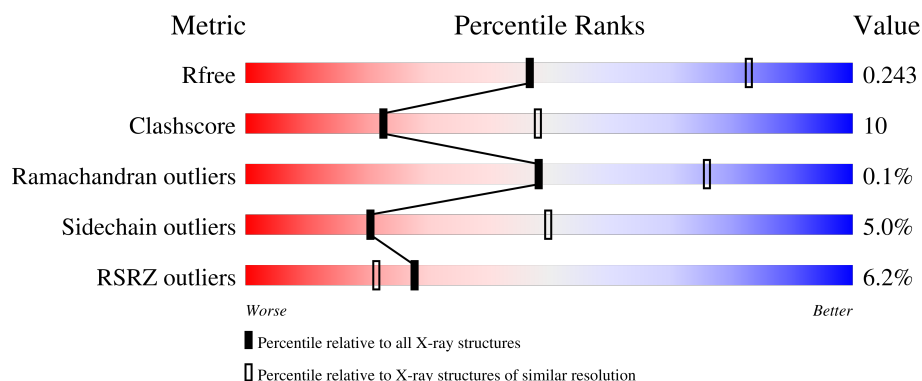
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.81 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))
$R_{free}$	164625	4293 (2.84-2.80)
Clashscore	180529	4801 (2.84-2.80)
Ramachandran outliers	177936	4739 (2.84-2.80)
Sidechain outliers	177891	4741 (2.84-2.80)
RSRZ outliers	164620	4295 (2.84-2.80)

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	427	
1	B	427	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5545 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
1	A	368	Total	C	N	O	S	Se	0	0	0
			2740	1748	461	514	3	14			
1	B	368	Total	C	N	O	S	Se	0	0	0
			2801	1790	473	520	3	15			

There are 90 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1976	MSE	-	expression tag	UNP Q99X42
A	1977	GLY	-	expression tag	UNP Q99X42
A	1978	SER	-	expression tag	UNP Q99X42
A	1979	SER	-	expression tag	UNP Q99X42
A	1980	HIS	-	expression tag	UNP Q99X42
A	1981	HIS	-	expression tag	UNP Q99X42
A	1982	HIS	-	expression tag	UNP Q99X42
A	1983	HIS	-	expression tag	UNP Q99X42
A	1984	HIS	-	expression tag	UNP Q99X42
A	1985	HIS	-	expression tag	UNP Q99X42
A	1986	SER	-	expression tag	UNP Q99X42
A	1987	SER	-	expression tag	UNP Q99X42
A	1988	GLY	-	expression tag	UNP Q99X42
A	1989	LEU	-	expression tag	UNP Q99X42
A	1990	VAL	-	expression tag	UNP Q99X42
A	1991	PRO	-	expression tag	UNP Q99X42
A	1992	ARG	-	expression tag	UNP Q99X42
A	1993	GLY	-	expression tag	UNP Q99X42
A	1994	SER	-	expression tag	UNP Q99X42
A	1995	HIS	-	expression tag	UNP Q99X42
A	1996	MSE	-	expression tag	UNP Q99X42
A	1997	ALA	-	expression tag	UNP Q99X42
A	1998	SER	-	expression tag	UNP Q99X42
A	1999	MSE	-	expression tag	UNP Q99X42
A	2000	THR	-	expression tag	UNP Q99X42

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	2001	GLY	-	expression tag	UNP Q99X42
A	2002	GLY	-	expression tag	UNP Q99X42
A	2003	GLN	-	expression tag	UNP Q99X42
A	2004	GLN	-	expression tag	UNP Q99X42
A	2005	MSE	-	expression tag	UNP Q99X42
A	2006	GLY	-	expression tag	UNP Q99X42
A	2007	ARG	-	expression tag	UNP Q99X42
A	2008	ASP	-	expression tag	UNP Q99X42
A	2009	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
B	1976	MSE	-	expression tag	UNP Q99X42
B	1977	GLY	-	expression tag	UNP Q99X42
B	1978	SER	-	expression tag	UNP Q99X42
B	1979	SER	-	expression tag	UNP Q99X42
B	1980	HIS	-	expression tag	UNP Q99X42
B	1981	HIS	-	expression tag	UNP Q99X42
B	1982	HIS	-	expression tag	UNP Q99X42
B	1983	HIS	-	expression tag	UNP Q99X42
B	1984	HIS	-	expression tag	UNP Q99X42
B	1985	HIS	-	expression tag	UNP Q99X42
B	1986	SER	-	expression tag	UNP Q99X42
B	1987	SER	-	expression tag	UNP Q99X42
B	1988	GLY	-	expression tag	UNP Q99X42
B	1989	LEU	-	expression tag	UNP Q99X42
B	1990	VAL	-	expression tag	UNP Q99X42
B	1991	PRO	-	expression tag	UNP Q99X42
B	1992	ARG	-	expression tag	UNP Q99X42
B	1993	GLY	-	expression tag	UNP Q99X42
B	1994	SER	-	expression tag	UNP Q99X42
B	1995	HIS	-	expression tag	UNP Q99X42
B	1996	MSE	-	expression tag	UNP Q99X42
B	1997	ALA	-	expression tag	UNP Q99X42

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1998	SER	-	expression tag	UNP Q99X42
B	1999	MSE	-	expression tag	UNP Q99X42
B	2000	THR	-	expression tag	UNP Q99X42
B	2001	GLY	-	expression tag	UNP Q99X42
B	2002	GLY	-	expression tag	UNP Q99X42
B	2003	GLN	-	expression tag	UNP Q99X42
B	2004	GLN	-	expression tag	UNP Q99X42
B	2005	MSE	-	expression tag	UNP Q99X42
B	2006	GLY	-	expression tag	UNP Q99X42
B	2007	ARG	-	expression tag	UNP Q99X42
B	2008	ASP	-	expression tag	UNP Q99X42
B	2009	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

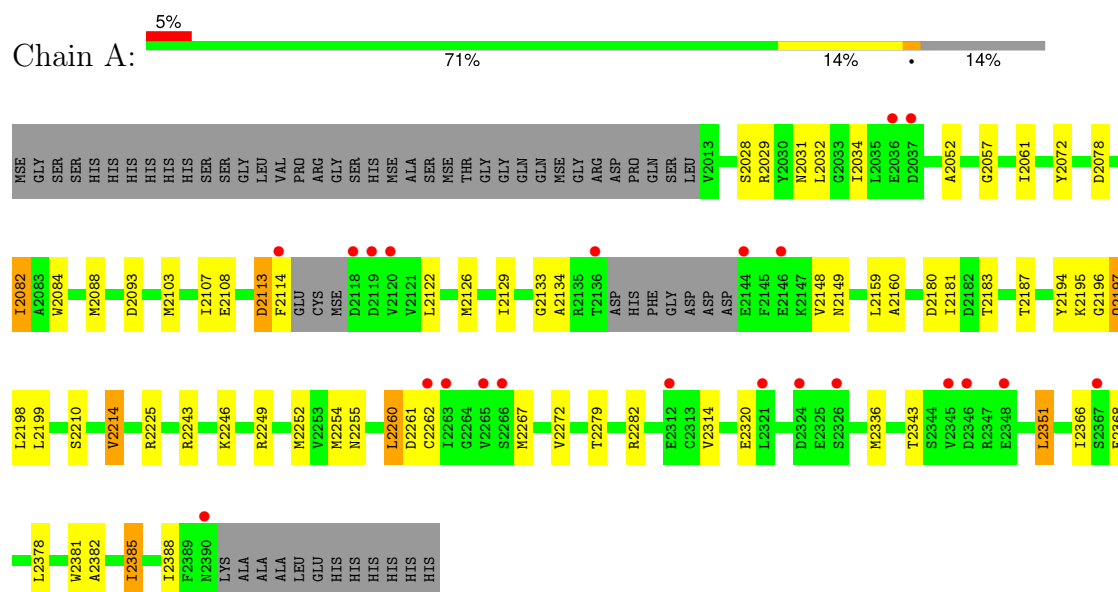
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O 1 1	0	0
2	B	3	Total O 3 3	0	0

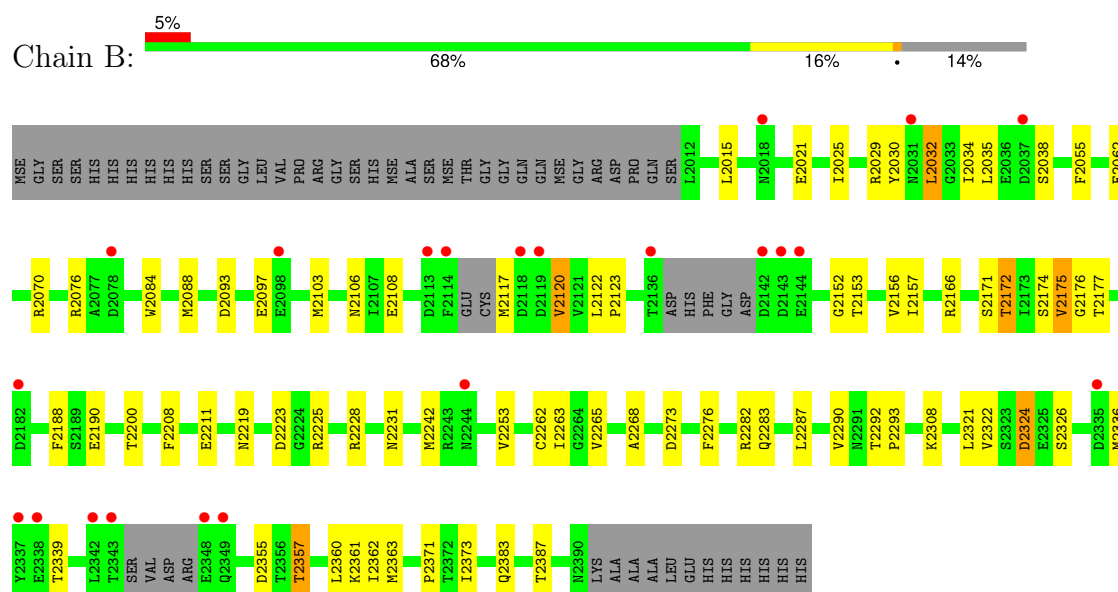
### 3 Residue-property plots [i](#)

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: AusA reductase domain protein



- Molecule 1: AusA reductase domain protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.70Å 106.45Å 124.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.43 – 2.81 44.43 – 2.81	Depositor EDS
% Data completeness (in resolution range)	95.9 (44.43-2.81) 94.8 (44.43-2.81)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.10 (at 2.81Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.206 , 0.243 0.205 , 0.243	Depositor DCC
$R_{free}$ test set	2002 reflections (5.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	75.7	Xtriage
Anisotropy	0.517	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 96.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.015 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5545	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.19% 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

### 5.1 Standard geometry

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	0.41	0/2778	0.56	0/3771
1	B	0.45	0/2837	0.62	0/3841
All	All	0.43	0/5615	0.59	0/7612

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

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	2740	0	2537	51	0
1	B	2801	0	2694	62	0
2	A	1	0	0	1	0
2	B	3	0	0	1	0
All	All	5545	0	5231	111	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 10.

The worst 5 of 111 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:2282:ARG:NH2	1:A:2368:GLU:HG3	1.57	1.17

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:2172:THR:HG21	1:B:2174:SER:OG	1.40	1.16
1:B:2172:THR:HG23	1:B:2174:SER:H	0.94	1.09
1:B:2172:THR:CG2	1:B:2174:SER:H	1.69	1.05
1:B:2172:THR:HG23	1:B:2174:SER:N	1.77	0.99

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	362/427 (85%)	351 (97%)	10 (3%)	1 (0%)	37	65
1	B	360/427 (84%)	336 (93%)	24 (7%)	0	100	100
All	All	722/854 (84%)	687 (95%)	34 (5%)	1 (0%)	48	76

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2133	GLY

### 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	271/362 (75%)	258 (95%)	13 (5%)	21	51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	291/362 (80%)	276 (95%)	15 (5%)	19	48
All	All	562/724 (78%)	534 (95%)	28 (5%)	20	50

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	2032	LEU
1	B	2373	ILE
1	B	2120	VAL
1	B	2336	MSE
1	B	2106	ASN

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

Mol	Chain	Res	Type
1	B	2375	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](#)

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	354/427 (82%)	0.13	22 (6%) 28 21	60, 96, 148, 164	0
1	B	353/427 (82%)	0.07	22 (6%) 28 21	54, 86, 139, 164	0
All	All	707/854 (82%)	0.10	44 (6%) 28 21	54, 90, 144, 164	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2119	ASP	5.8
1	B	2114	PHE	5.7
1	B	2136	THR	4.6
1	B	2142	ASP	4.5
1	B	2118	ASP	4.5

### 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 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.