



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 1, 2021 – 02:46 PM EDT

PDB ID : 2ZGU  
Title : Crystal structure Agrocybe aegerita lectin AAL mutant I144G  
Authors : Li, D.F.; Yang, N.; Wang, D.C.  
Deposited on : 2008-01-23  
Resolution : 2.40 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 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 : 2.23.2  
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) : 2.23.2

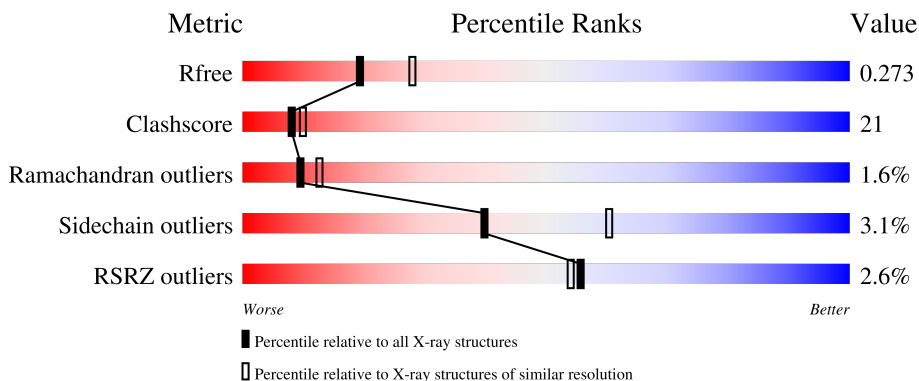
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	167	<div> <div>2%</div> <div> <div></div> <div>62%</div> <div>30%</div> <div>• 5%</div> </div> </div>
1	B	167	<div> <div>2%</div> <div> <div></div> <div>59%</div> <div>29%</div> <div>•• 10%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2451 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 Anti-tumor lectin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	159	Total	C	N	O	S	0	0	0
			1202	759	199	242	2			
1	B	151	Total	C	N	O	S	0	0	0
			1153	732	188	231	2			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q6WY08
A	132	SER	LEU	SEE REMARK 999	UNP Q6WY08
A	144	GLY	ILE	engineered mutation	UNP Q6WY08
A	159	LEU	-	expression tag	UNP Q6WY08
A	160	GLU	-	expression tag	UNP Q6WY08
A	161	HIS	-	expression tag	UNP Q6WY08
A	162	HIS	-	expression tag	UNP Q6WY08
A	163	HIS	-	expression tag	UNP Q6WY08
A	164	HIS	-	expression tag	UNP Q6WY08
A	165	HIS	-	expression tag	UNP Q6WY08
A	166	HIS	-	expression tag	UNP Q6WY08
B	0	MET	-	initiating methionine	UNP Q6WY08
B	132	SER	LEU	SEE REMARK 999	UNP Q6WY08
B	144	GLY	ILE	engineered mutation	UNP Q6WY08
B	159	LEU	-	expression tag	UNP Q6WY08
B	160	GLU	-	expression tag	UNP Q6WY08
B	161	HIS	-	expression tag	UNP Q6WY08
B	162	HIS	-	expression tag	UNP Q6WY08
B	163	HIS	-	expression tag	UNP Q6WY08
B	164	HIS	-	expression tag	UNP Q6WY08
B	165	HIS	-	expression tag	UNP Q6WY08
B	166	HIS	-	expression tag	UNP Q6WY08

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	43	Total 43	O 43	0	0
2	B	53	Total 53	O 53	0	0



- Molecule 1: Anti-tumor lectin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.17 Å 42.17 Å 135.12 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	24.80 – 2.40 24.80 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (24.80-2.40) 100.0 (24.80-2.40)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.30 (at 2.39 Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.221 , 0.272 0.221 , 0.273	Depositor DCC
$R_{free}$ test set	1083 reflections (10.31%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.8	Xtriage
Anisotropy	0.618	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 20.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.046 for -h,-k,l 0.489 for h,-h-k,-l 0.055 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2451	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.82% 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	0.39	0/1223	0.67	1/1669 (0.1%)
1	B	0.39	0/1172	0.68	1/1597 (0.1%)
All	All	0.39	0/2395	0.67	2/3266 (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	138	ALA	N-CA-C	-6.44	93.61	111.00
1	B	138	ALA	N-CA-C	-5.91	95.06	111.00

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	1202	0	1172	52	0
1	B	1153	0	1130	49	0
2	A	43	0	0	0	0
2	B	53	0	0	1	0
All	All	2451	0	2302	99	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 21.

The worst 5 of 99 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:42:PRO:HB3	1:A:64:LEU:HB2	1.39	1.01
1:A:43:ASN:HB3	1:A:63:ARG:NH2	1.82	0.95
1:B:63:ARG:HB3	1:B:65:GLN:HE22	1.31	0.94
1:B:139:THR:C	1:B:140:GLU:HG2	1.98	0.82
1:A:110:LYS:HG2	1:A:124:THR:HA	1.64	0.80

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	157/167 (94%)	137 (87%)	17 (11%)	3 (2%)	8	10
1	B	147/167 (88%)	134 (91%)	11 (8%)	2 (1%)	11	15
All	All	304/334 (91%)	271 (89%)	28 (9%)	5 (2%)	9	13

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	141	GLU
1	A	39	ALA
1	A	139	THR
1	B	140	GLU
1	B	142	THR

### 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	131/139 (94%)	126 (96%)	5 (4%)	33	51
1	B	127/139 (91%)	124 (98%)	3 (2%)	49	68
All	All	258/278 (93%)	250 (97%)	8 (3%)	40	60

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	140	GLU
1	B	57	LEU
1	A	97	ASP
1	A	77	ASP
1	B	33	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	126	GLN
1	B	92	GLN
1	A	92	GLN
1	A	91	ASN
1	B	65	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

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 [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	159/167 (95%)	-0.46	4 (2%) 57 55	15, 28, 64, 79	0
1	B	151/167 (90%)	-0.45	4 (2%) 56 54	16, 29, 52, 67	0
All	All	310/334 (92%)	-0.46	8 (2%) 56 54	15, 29, 59, 79	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	139	THR	6.5
1	A	40	GLY	3.1
1	B	94	ALA	2.8
1	A	139	THR	2.6
1	B	65	GLN	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 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.