



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 09:55 am GMT

PDB ID : 1XV2  
Title : Crystal Structure of a Protein of Unknown Function Similar to Alpha-acetolactate Decarboxylase from Staphylococcus aureus  
Authors : Duke, N.E.C.; Zhang, R.; Quartey, P.; Collart, F.; Holzle, D.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2004-10-26  
Resolution : 2.00 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

<http://wwpdb.org/validation/2016/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)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

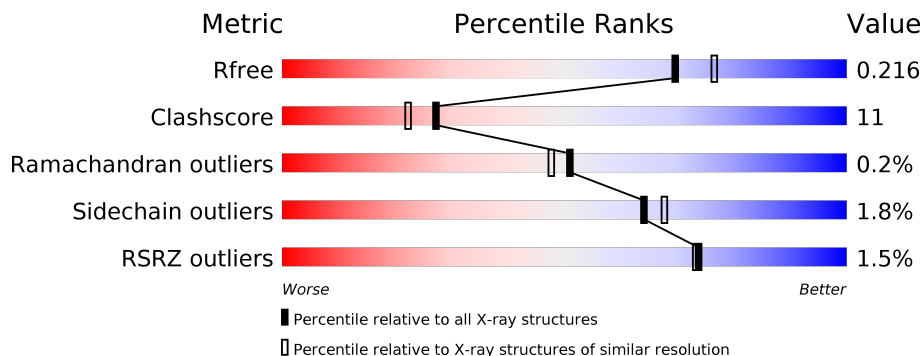
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	237	<div> <div>%</div> <div> <div></div> <div>76%</div> <div>20%</div> <div>..</div> </div> </div>
1	B	237	<div> <div>4%</div> <div> <div></div> <div>72%</div> <div>25%</div> <div>..</div> </div> </div>
1	C	237	<div> <div></div> <div> <div></div> <div>81%</div> <div>17%</div> <div>..</div> </div> </div>
1	D	237	<div> <div></div> <div> <div></div> <div>82%</div> <div>15%</div> <div>..</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8068 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 hypothetical protein, similar to alpha-acetolactate decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	232	Total	C	N	O	Se	0	0	0
			1850	1177	313	355	5			
1	B	232	Total	C	N	O	Se	0	0	0
			1850	1177	313	355	5			
1	C	233	Total	C	N	O	Se	0	0	0
			1857	1181	314	357	5			
1	D	232	Total	C	N	O	Se	0	0	0
			1850	1177	313	355	5			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	CLONING ARTIFACT	UNP Q7A3A7
A	-1	ASN	-	CLONING ARTIFACT	UNP Q7A3A7
A	0	ALA	-	CLONING ARTIFACT	UNP Q7A3A7
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
A	15	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
A	103	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
A	121	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
A	125	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
A	126	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
B	-2	SER	-	CLONING ARTIFACT	UNP Q7A3A7
B	-1	ASN	-	CLONING ARTIFACT	UNP Q7A3A7
B	0	ALA	-	CLONING ARTIFACT	UNP Q7A3A7
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
B	15	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
B	103	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
B	121	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
B	125	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
B	126	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
C	-2	SER	-	CLONING ARTIFACT	UNP Q7A3A7
C	-1	ASN	-	CLONING ARTIFACT	UNP Q7A3A7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	ALA	-	CLONING ARTIFACT	UNP Q7A3A7
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
C	15	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
C	103	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
C	121	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
C	125	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
C	126	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
D	-2	SER	-	CLONING ARTIFACT	UNP Q7A3A7
D	-1	ASN	-	CLONING ARTIFACT	UNP Q7A3A7
D	0	ALA	-	CLONING ARTIFACT	UNP Q7A3A7
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
D	15	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
D	103	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
D	121	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
D	125	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7
D	126	MSE	MET	MODIFIED RESIDUE	UNP Q7A3A7

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Zn 1 1	0	0
2	A	1	Total Zn 1 1	0	0
2	D	1	Total Zn 1 1	0	0
2	C	1	Total Zn 1 1	0	0

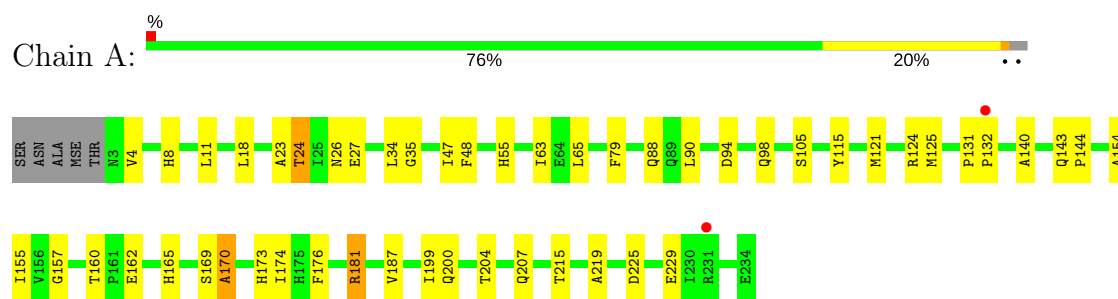
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	107	Total O 107 107	0	0
3	B	82	Total O 82 82	0	0
3	C	246	Total O 246 246	0	0
3	D	222	Total O 222 222	0	0

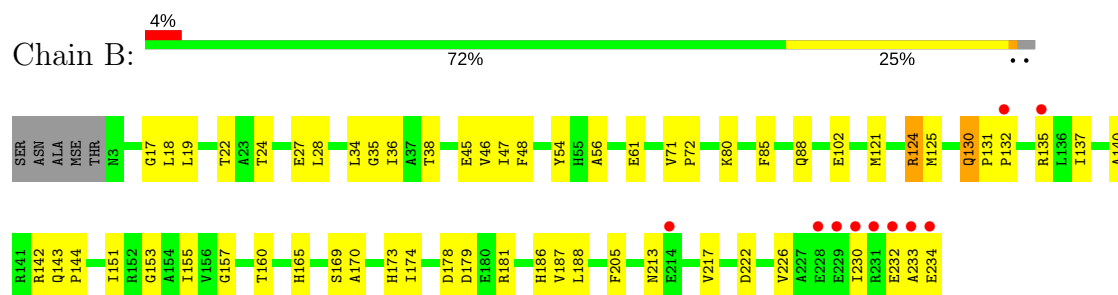
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

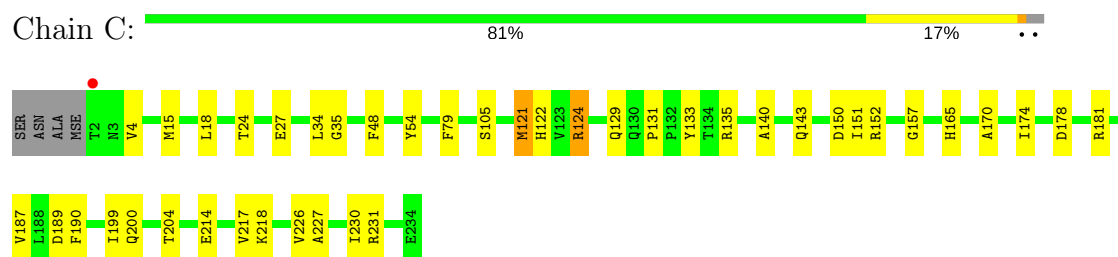
- Molecule 1: hypothetical protein, similar to alpha-acetolactate decarboxylase



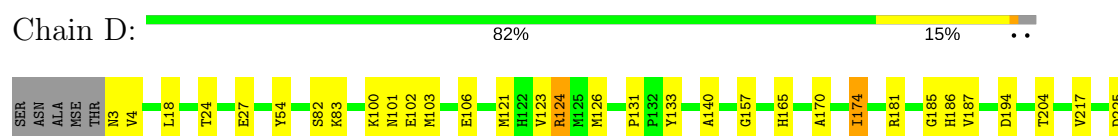
- Molecule 1: hypothetical protein, similar to alpha-acetolactate decarboxylase



- Molecule 1: hypothetical protein, similar to alpha-acetolactate decarboxylase



- Molecule 1: hypothetical protein, similar to alpha-acetolactate decarboxylase



V226	
A227	
E228	
E229	
I230	
R231	
E232	
A233	
E234	



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.43Å 160.00Å 68.06Å 90.00° 107.55° 90.00°	Depositor
Resolution (Å)	30.24 – 2.00 30.24 – 1.66	Depositor EDS
% Data completeness (in resolution range)	94.3 (30.24-2.00) 55.2 (30.24-1.66)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.07 (at 1.66Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.191 , 0.236 0.193 , 0.216	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	12.5	Xtriage
Anisotropy	0.714	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 44.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	8068	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.81% 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: ZN

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	0.31	0/1887	0.58	0/2540
1	B	0.31	0/1887	0.56	0/2540
1	C	0.35	0/1894	0.62	0/2550
1	D	0.37	0/1887	0.62	0/2540
All	All	0.34	0/7555	0.60	0/10170

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	1850	0	1781	38	0
1	B	1850	0	1781	46	0
1	C	1857	0	1788	42	0
1	D	1850	0	1781	42	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	107	0	0	1	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	82	0	0	1	0
3	C	246	0	0	3	0
3	D	222	0	0	2	0
All	All	8068	0	7131	165	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 11.

The worst 5 of 165 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:C:121:MSE:HE3	1:C:190:PHE:HB3	1.20	1.17
1:C:121:MSE:HE2	1:C:122:HIS:H	1.10	1.08
1:D:174:ILE:HG12	1:D:187:VAL:CG1	1.87	1.04
1:D:186:HIS:HE2	1:D:234:GLU:HG3	1.27	0.99
1:C:121:MSE:HE1	1:C:187:VAL:HG13	1.46	0.97

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	230/237 (97%)	221 (96%)	8 (4%)	1 (0%)	38	33
1	B	230/237 (97%)	221 (96%)	8 (4%)	1 (0%)	38	33
1	C	231/237 (98%)	226 (98%)	5 (2%)	0	100	100
1	D	230/237 (97%)	223 (97%)	7 (3%)	0	100	100
All	All	921/948 (97%)	891 (97%)	28 (3%)	2 (0%)	51	48

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	142	ARG
1	A	170	ALA

### 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	196/194 (101%)	193 (98%)	3 (2%)	70	74
1	B	196/194 (101%)	193 (98%)	3 (2%)	70	74
1	C	197/194 (102%)	194 (98%)	3 (2%)	70	74
1	D	196/194 (101%)	191 (97%)	5 (3%)	51	52
All	All	785/776 (101%)	771 (98%)	14 (2%)	64	68

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

Mol	Chain	Res	Type
1	C	121	MSE
1	C	124	ARG
1	D	124	ARG
1	B	130	GLN
1	D	121	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	200	GLN
1	C	88	GLN
1	D	200	GLN
1	C	3	ASN
1	C	26	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 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	227/237 (95%)	-0.40	2 (0%) 84 83	12, 21, 33, 39	0
1	B	227/237 (95%)	-0.07	10 (4%) 35 35	10, 24, 48, 59	0
1	C	228/237 (96%)	-0.83	1 (0%) 92 92	3, 10, 20, 36	0
1	D	227/237 (95%)	-0.73	1 (0%) 92 92	3, 10, 20, 40	0
All	All	909/948 (95%)	-0.51	14 (1%) 74 73	3, 16, 36, 59	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	233	ALA	5.0
1	A	132	PRO	3.6
1	B	230	ILE	3.4
1	C	2	THR	3.1
1	B	214	GLU	3.1

### 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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	ZN	B	802	1/1	0.98	0.06	-1.11	29,29,29,29	0
2	ZN	A	801	1/1	1.00	0.04	-1.27	17,17,17,17	0
2	ZN	C	803	1/1	1.00	0.04	-2.83	10,10,10,10	0
2	ZN	D	804	1/1	1.00	0.05	-	12,12,12,12	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.