



# Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 08:19 am BST

PDB ID : 5UCS  
Title : Class II fructose-1,6-bisphosphate aldolase E149A variant of *Helicobacter pylori*  
Authors : Jacques, B.; Sygusch, J.  
Deposited on : 2016-12-22  
Resolution : 1.41 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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	:	2.11
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)	:	2.11

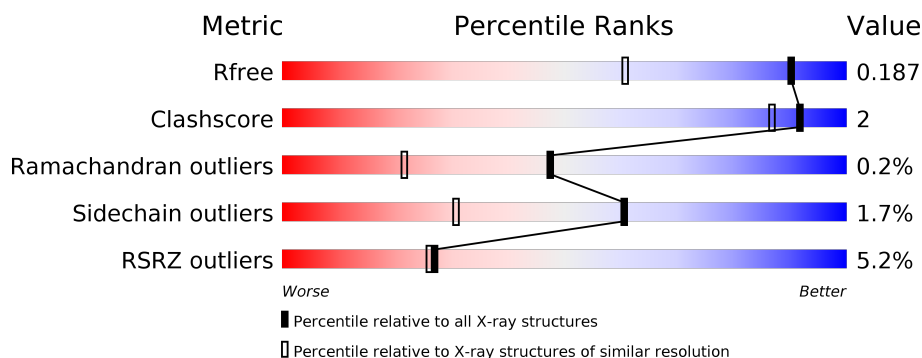
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.41 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	1714 (1.40-1.40)
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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 on 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	307	<div> <div>6%</div> <div> <div></div> <div>89%</div> <div>5%</div> <div>6%</div> </div> </div>
1	B	307	<div> <div>4%</div> <div> <div></div> <div>91%</div> <div>•</div> <div>6%</div> </div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9715 atoms, of which 4590 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 Fructose-bisphosphate aldolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	290	Total	C	H	N	O	S	0	7	0
			4565	1447	2292	389	425	12			
1	B	290	Total	C	H	N	O	S	0	7	0
			4571	1448	2298	389	424	12			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	48	ALA	THR	conflict	UNP P56109
A	67	ILE	THR	conflict	UNP P56109
A	149	ALA	GLU	engineered mutation	UNP P56109
B	48	ALA	THR	conflict	UNP P56109
B	67	ILE	THR	conflict	UNP P56109
B	149	ALA	GLU	engineered mutation	UNP P56109

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

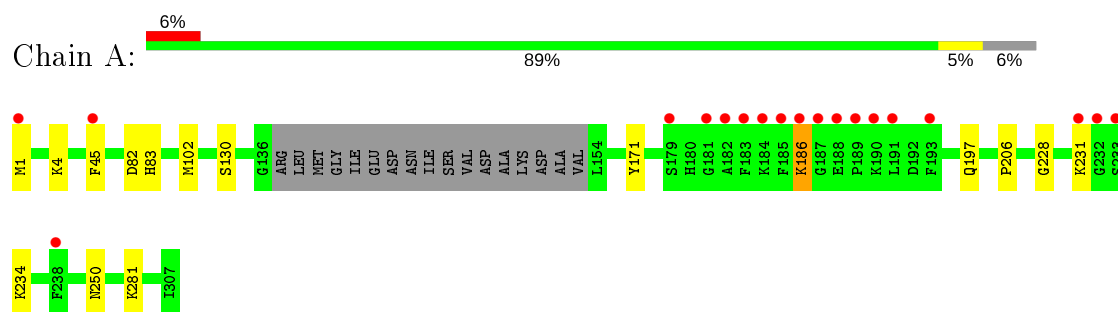
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	270	Total 270	O 270	0	1
4	B	305	Total 305	O 305	0	1

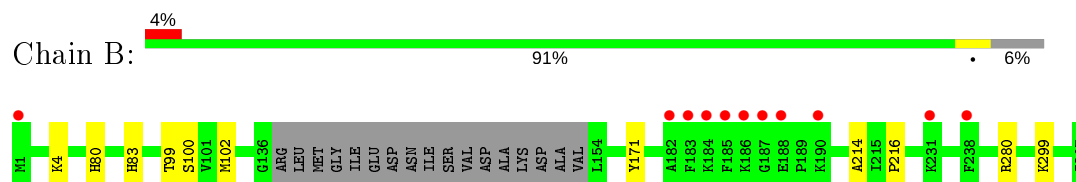
### 3 Residue-property plots [i](#)

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: Fructose-bisphosphate aldolase



- Molecule 1: Fructose-bisphosphate aldolase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	39.06Å 82.83Å 91.09Å 90.00° 99.64° 90.00°	Depositor
Resolution (Å)	29.94 – 1.41 29.94 – 1.41	Depositor EDS
% Data completeness (in resolution range)	89.4 (29.94-1.41) 89.3 (29.94-1.41)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.09 (at 1.41Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575)	Depositor
R, $R_{free}$	0.164 , 0.188 0.164 , 0.187	Depositor DCC
$R_{free}$ test set	5506 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.6	Xtriage
Anisotropy	0.524	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 49.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.55$ , $\langle L^2 \rangle = 0.39$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	9715	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 45.73 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2666e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

Bond lengths and bond angles in the following residue types are not validated in this section: NA, 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	0.40	0/2343	0.57	0/3149
1	B	0.44	0/2341	0.58	0/3145
All	All	0.42	0/4684	0.58	0/6294

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	2273	2292	2264	9	0
1	B	2273	2298	2273	6	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	270	0	0	4	1
4	B	305	0	0	3	1
All	All	5125	4590	4537	15	2

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 2.

All (15) 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:280:ARG:NH1	4:B:501:HOH:O	2.10	0.83
1:B:299:LYS:NZ	4:B:505:HOH:O	2.25	0.66
1:A:281:LYS:NZ	4:A:503:HOH:O	2.29	0.65
1:B:83[B]:HIS:NE2	4:B:502[B]:HOH:O	2.20	0.63
1:A:83[B]:HIS:NE2	4:A:501[B]:HOH:O	2.13	0.60
1:A:197:GLN:NE2	4:A:504:HOH:O	2.31	0.51
1:A:186:LYS:N	1:A:186:LYS:HD2	2.28	0.49
1:A:4:LYS:CE	4:A:517:HOH:O	2.64	0.44
1:B:214:ALA:O	1:B:216:PRO:HD3	2.18	0.43
1:A:4:LYS:HG2	1:A:130:SER:OG	2.17	0.43
1:A:82:ASP:OD1	1:A:83[B]:HIS:HD2	2.02	0.42
1:B:4:LYS:HA	1:B:99:THR:HB	2.02	0.41
1:A:206:PRO:HA	1:A:250:ASN:OD1	2.20	0.41
1:B:80:HIS:HA	1:B:100:SER:O	2.22	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:B:580:HOH:O	4:B:639:HOH:O[2_556]	1.92	0.28
4:A:673:HOH:O	4:A:685:HOH:O[2_555]	2.19	0.01

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	293/307 (95%)	289 (99%)	3 (1%)	1 (0%)	41	18
1	B	293/307 (95%)	287 (98%)	6 (2%)	0	100	100
All	All	586/614 (95%)	576 (98%)	9 (2%)	1 (0%)	47	21



All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	231	LYS

### 5.3.2 Protein sidechains ⓘ

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	246/253 (97%)	240 (98%)	6 (2%)	49	16
1	B	245/253 (97%)	243 (99%)	2 (1%)	81	62
All	All	491/506 (97%)	483 (98%)	8 (2%)	60	33

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	45	PHE
1	A	102	MET
1	A	171	TYR
1	A	186	LYS
1	A	234	LYS
1	B	102	MET
1	B	171	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 [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

### 6.1 Protein, DNA and RNA chains

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	290/307 (94%)	0.01	19 (6%) 18 16	13, 21, 60, 136	0
1	B	290/307 (94%)	-0.14	11 (3%) 40 40	12, 18, 52, 122	0
All	All	580/614 (94%)	-0.07	30 (5%) 27 26	12, 20, 57, 136	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	186	LYS	8.3
1	A	185	PHE	6.3
1	A	186	LYS	6.2
1	A	187	GLY	6.2
1	B	185	PHE	6.1
1	A	1	MET	4.0
1	A	183	PHE	3.9
1	A	238	PHE	3.9
1	A	182	ALA	3.8
1	B	238	PHE	3.7
1	A	181	GLY	3.7
1	A	231	LYS	3.6
1	B	188	GLU	3.6
1	B	183	PHE	3.5
1	A	232	GLY	3.3
1	A	191	LEU	3.2
1	B	187	GLY	3.2
1	B	182	ALA	3.1
1	B	184	LYS	2.9
1	A	189	PRO	2.8
1	A	45	PHE	2.8
1	A	184	LYS	2.7
1	A	190	LYS	2.6
1	A	193	PHE	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	190	LYS	2.3
1	A	179	SER	2.3
1	A	233	SER	2.2
1	B	231	LYS	2.2
1	A	188	GLU	2.2
1	B	1	MET	2.2

## 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. 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
2	NA	A	401	1/1	0.98	0.10	25,25,25,25	0
2	NA	B	401	1/1	0.99	0.10	23,23,23,23	0
3	ZN	A	402	1/1	1.00	0.01	22,22,22,22	0
3	ZN	B	402	1/1	1.00	0.03	19,19,19,19	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.