



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 23, 2024 – 04:27 PM EDT

PDB ID : 5J83  
Title : Crystal structure of L-arabinonate dehydratase in apo-form  
Authors : Rahman, M.M.; Rouvinen, J.; Hakulinen, N.  
Deposited on : 2016-04-07  
Resolution : 3.00 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

<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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

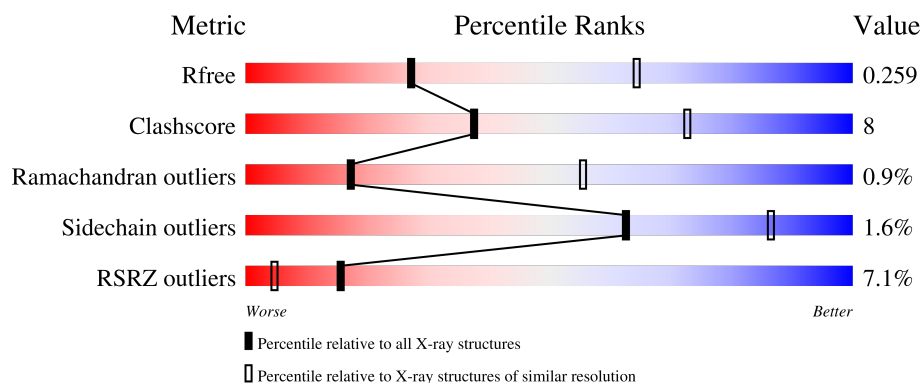
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	588	<div> <div>9%</div> <div>76%</div> <div>21%</div> <div>..</div> </div>
1	B	588	<div> <div>5%</div> <div>80%</div> <div>17%</div> <div>..</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8698 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 Dihydroxyacid dehydratase/phosphogluconate dehydratase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	575	Total	C	N	O	S	0	0	0
			4349	2724	766	825	34			
1	B	575	Total	C	N	O	S	0	0	0
			4349	2724	766	825	34			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	MET	-	initiating methionine	UNP I9XDU6
A	-7	ASP	-	expression tag	UNP I9XDU6
A	-6	TRP	-	expression tag	UNP I9XDU6
A	-5	SER	-	expression tag	UNP I9XDU6
A	-4	HIS	-	expression tag	UNP I9XDU6
A	-3	PRO	-	expression tag	UNP I9XDU6
A	-2	GLN	-	expression tag	UNP I9XDU6
A	-1	PHE	-	expression tag	UNP I9XDU6
A	0	GLU	-	expression tag	UNP I9XDU6
A	1	LYS	-	expression tag	UNP I9XDU6
B	-8	MET	-	initiating methionine	UNP I9XDU6
B	-7	ASP	-	expression tag	UNP I9XDU6
B	-6	TRP	-	expression tag	UNP I9XDU6
B	-5	SER	-	expression tag	UNP I9XDU6
B	-4	HIS	-	expression tag	UNP I9XDU6
B	-3	PRO	-	expression tag	UNP I9XDU6
B	-2	GLN	-	expression tag	UNP I9XDU6
B	-1	PHE	-	expression tag	UNP I9XDU6
B	0	GLU	-	expression tag	UNP I9XDU6
B	1	LYS	-	expression tag	UNP I9XDU6



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	161.13Å 161.13Å 114.53Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.91 – 3.00 47.91 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (47.91-3.00) 99.6 (47.91-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.92 (at 3.01Å)	Xtriage
Refinement program	PHENIX (1.11 _ 2567: ???)	Depositor
R, $R_{free}$	0.208 , 0.259 0.208 , 0.259	Depositor DCC
$R_{free}$ test set	1689 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	68.0	Xtriage
Anisotropy	0.314	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 81.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.059 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	8698	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.31% 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

Bond lengths and bond angles in the following residue types are not validated in this section: KCX

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	0.24	0/4422	0.43	0/5988
1	B	0.24	0/4422	0.43	0/5988
All	All	0.24	0/8844	0.43	0/11976

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	4349	0	4320	82	1
1	B	4349	0	4320	61	1
All	All	8698	0	8640	134	1

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 8.

The worst 5 of 134 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:60:ASN:ND2	1:A:127:CYS:SG	2.58	0.76
1:A:19:GLY:O	1:A:28:ARG:NH2	2.24	0.70
1:A:129:KCX:OQ1	1:A:478:ARG:NH1	2.27	0.68
1:B:19:GLY:O	1:B:28:ARG:NH2	2.28	0.67
1:B:461:PRO:HB2	1:B:465:LYS:HE2	1.75	0.66

All (1) 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)
1:A:195:ARG:NH2	1:B:325:TYR:OH[4_455]	2.11	0.09

## 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	572/588 (97%)	529 (92%)	38 (7%)	5 (1%)	17	55
1	B	572/588 (97%)	539 (94%)	28 (5%)	5 (1%)	17	55
All	All	1144/1176 (97%)	1068 (93%)	66 (6%)	10 (1%)	17	55

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	491	THR
1	A	491	THR
1	B	126	GLY
1	A	126	GLY
1	B	422	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	460/473 (97%)	452 (98%)	8 (2%)	60	85
1	B	460/473 (97%)	453 (98%)	7 (2%)	65	87
All	All	920/946 (97%)	905 (98%)	15 (2%)	62	86

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	538	ASN
1	B	420	ASP
1	B	39	ASP
1	B	490	HIS
1	B	195	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	572	ASN
1	B	270	ASN
1	B	278	ASN
1	A	223	ASN
1	A	60	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The



Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	KCX	A	129	1	9,11,12	0.81	0	5,12,14	1.66	1 (20%)
1	KCX	B	129	1	9,11,12	0.86	0	5,12,14	1.54	1 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	KCX	A	129	1	-	3/9/10/12	-
1	KCX	B	129	1	-	2/9/10/12	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	129	KCX	OQ1-CX-NZ	-3.14	120.10	124.96
1	B	129	KCX	OQ1-CX-NZ	-3.12	120.12	124.96

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	129	KCX	N-CA-CB-CG
1	A	129	KCX	C-CA-CB-CG
1	B	129	KCX	C-CA-CB-CG
1	A	129	KCX	CA-CB-CG-CD
1	B	129	KCX	CA-CB-CG-CD

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	129	KCX	2	0
1	B	129	KCX	2	0

## 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 [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	574/588 (97%)	0.38	50 (8%)	10 3	34, 68, 103, 122	0
1	B	574/588 (97%)	0.06	32 (5%)	24 8	21, 56, 98, 118	0
All	All	1148/1176 (97%)	0.22	82 (7%)	16 5	21, 63, 100, 122	0

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	522	ASP	4.7
1	B	526	GLU	4.2
1	A	528	LEU	4.0
1	B	187	LEU	3.9
1	A	503	VAL	3.9

### 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q < 0.9
1	KCX	A	129	12/13	0.90	0.30	50,61,85,88	0
1	KCX	B	129	12/13	0.90	0.25	48,66,76,100	0

### 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.