



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:57 AM GMT

PDB ID : 2YW4  
Title : Crystal Structure Analysis of the 4-hydroxy-2-oxoglutarate aldolase/2-dehydro-3-deoxyphosphogluconate aldolase from TTHB1  
Authors : Kawano, Y.; Hashimoto, K.; Kamiya, N.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)  
Deposited on : 2007-04-19  
Resolution : 2.53 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  
<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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

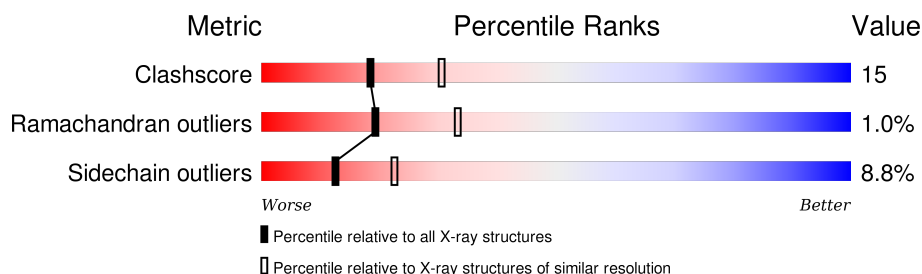
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.53 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))
Clashscore	102246	4968 (2.54-2.50)
Ramachandran outliers	100387	4873 (2.54-2.50)
Sidechain outliers	100360	4875 (2.54-2.50)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	207	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1545 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 4-hydroxy-2-oxoglutarate aldolase/2-deydro-3-deoxyphosphogluconate aldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	200	Total	C	N	O	S	0	0	0
			1483	958	257	267	1			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	62	Total	O	0	0
			62	62		

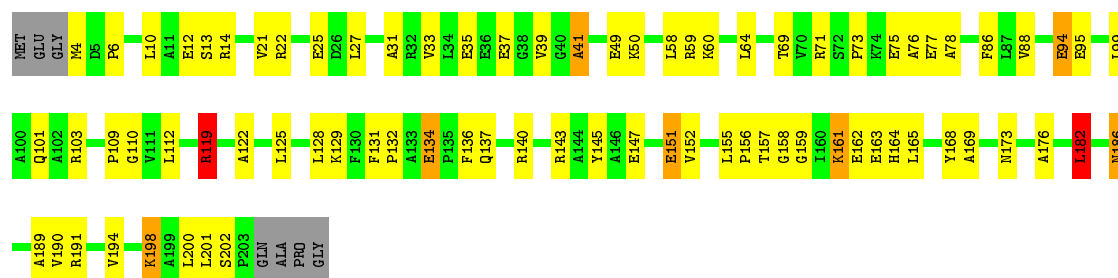
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

- Molecule 1: 4-hydroxy-2-oxoglutarate aldolase/2-deydro-3-deoxyphosphogluconate aldolase

Chain A: 



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.69 Å 66.69 Å 241.30 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.80 – 2.53	Depositor
% Data completeness (in resolution range)	100.0 (30.80-2.53)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.164 , 0.241	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1545	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

## 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	1.60	18/1508 (1.2%)	1.33	10/2049 (0.5%)

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	94	GLU	CD-OE2	8.45	1.34	1.25
1	A	78	ALA	CA-CB	8.40	1.70	1.52
1	A	147	GLU	CG-CD	8.12	1.64	1.51
1	A	95	GLU	CG-CD	7.80	1.63	1.51
1	A	86	PHE	CE2-CZ	6.76	1.50	1.37
1	A	134	GLU	CG-CD	6.70	1.61	1.51
1	A	151	GLU	CG-CD	6.70	1.61	1.51
1	A	94	GLU	CG-CD	6.64	1.61	1.51
1	A	86	PHE	CG-CD2	6.12	1.48	1.38
1	A	88	VAL	CB-CG2	5.85	1.65	1.52
1	A	41	ALA	CA-CB	-5.69	1.40	1.52
1	A	77	GLU	CD-OE2	5.39	1.31	1.25
1	A	75	GLU	CB-CG	-5.33	1.42	1.52
1	A	145	TYR	CE1-CZ	-5.21	1.31	1.38
1	A	21	VAL	CB-CG2	-5.16	1.42	1.52
1	A	147	GLU	CB-CG	5.04	1.61	1.52
1	A	168	TYR	CD1-CE1	5.03	1.46	1.39
1	A	163	GLU	CD-OE2	5.03	1.31	1.25

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	119	ARG	NE-CZ-NH1	10.89	125.75	120.30
1	A	119	ARG	NE-CZ-NH2	-10.32	115.14	120.30
1	A	112	LEU	CB-CG-CD1	-8.09	97.24	111.00
1	A	59	ARG	NE-CZ-NH2	6.22	123.41	120.30
1	A	151	GLU	CA-CB-CG	5.52	125.53	113.40
1	A	119	ARG	CD-NE-CZ	5.26	130.96	123.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	151	GLU	OE1-CD-OE2	-5.26	116.99	123.30
1	A	140	ARG	NE-CZ-NH2	5.17	122.89	120.30
1	A	59	ARG	NE-CZ-NH1	-5.13	117.73	120.30
1	A	71	ARG	NE-CZ-NH1	-5.10	117.75	120.30

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	1483	0	1559	47	0
2	A	62	0	0	5	0
All	All	1545	0	1559	47	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 15.

All (47) 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:161:LYS:H	1:A:164:HIS:HD2	0.95	0.94
1:A:161:LYS:H	1:A:164:HIS:CD2	1.87	0.89
1:A:101:GLN:NE2	2:A:269:HOH:O	2.11	0.83
1:A:186:ASN:H	1:A:189:ALA:HB3	1.45	0.80
1:A:49:GLU:CD	1:A:49:GLU:H	1.85	0.80
1:A:110:GLY:HA2	1:A:129:LYS:HB3	1.63	0.80
1:A:161:LYS:N	1:A:164:HIS:HD2	1.80	0.72
1:A:169:ALA:HB2	1:A:201:LEU:HD23	1.73	0.70
1:A:136:PHE:O	1:A:137:GLN:HB2	1.95	0.65
1:A:143:ARG:HD3	2:A:240:HOH:O	1.99	0.63
1:A:119:ARG:HD3	2:A:234:HOH:O	1.99	0.62
1:A:134:GLU:HG3	1:A:159:GLY:O	1.99	0.61
1:A:182:LEU:O	1:A:190:VAL:HG13	2.02	0.60
1:A:31:ALA:O	1:A:35:GLU:HG3	2.02	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:99:LEU:O	1:A:103:ARG:HG2	2.03	0.57
1:A:162:GLU:HG2	1:A:200:LEU:CD1	2.35	0.56
1:A:165:LEU:N	1:A:165:LEU:HD23	2.21	0.56
1:A:39:VAL:O	1:A:39:VAL:HG12	2.04	0.56
1:A:129:LYS:HE3	1:A:131:PHE:HB2	1.87	0.56
1:A:13:SER:O	1:A:14:ARG:HB2	2.07	0.55
1:A:33:VAL:HG22	1:A:191:ARG:NH2	2.23	0.54
1:A:194:VAL:O	1:A:198:LYS:HB2	2.08	0.53
1:A:129:LYS:HE3	1:A:131:PHE:CB	2.39	0.53
1:A:49:GLU:N	1:A:49:GLU:OE1	2.42	0.51
1:A:156:PRO:HD2	1:A:176:ALA:O	2.11	0.51
1:A:25:GLU:H	1:A:25:GLU:CD	2.14	0.49
1:A:39:VAL:HG21	1:A:194:VAL:HG13	1.93	0.49
1:A:94:GLU:HG3	1:A:119:ARG:NH1	2.31	0.46
1:A:27:LEU:HB3	1:A:58:LEU:HD21	1.97	0.46
1:A:157:THR:HG22	1:A:158:GLY:N	2.32	0.45
1:A:143:ARG:NH1	2:A:223:HOH:O	2.48	0.45
1:A:162:GLU:O	1:A:165:LEU:HG	2.16	0.45
1:A:110:GLY:CA	1:A:129:LYS:HB3	2.39	0.45
1:A:60:LYS:CE	1:A:60:LYS:HA	2.47	0.45
1:A:37:GLU:OE2	1:A:191:ARG:HD3	2.17	0.44
1:A:73:PRO:O	1:A:76:ALA:HB3	2.16	0.44
1:A:109:PRO:HG3	1:A:125:LEU:HD13	2.00	0.43
1:A:125:LEU:O	1:A:128:LEU:HD21	2.18	0.43
1:A:122:ALA:HB3	2:A:216:HOH:O	2.19	0.42
1:A:13:SER:O	1:A:14:ARG:CB	2.62	0.42
1:A:132:PRO:HD2	1:A:136:PHE:CE2	2.55	0.42
1:A:186:ASN:O	1:A:189:ALA:HB3	2.20	0.41
1:A:60:LYS:HE3	1:A:60:LYS:HA	2.02	0.41
1:A:22:ARG:HA	1:A:22:ARG:HD3	1.89	0.41
1:A:173:ASN:OD1	1:A:173:ASN:N	2.54	0.41
1:A:128:LEU:HD12	1:A:152:VAL:HG11	2.03	0.40
1:A:41:ALA:HA	1:A:64:LEU:O	2.21	0.40

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	198/207 (96%)	186 (94%)	10 (5%)	2 (1%)	19	33

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	182	LEU
1	A	202	SER

### 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	147/153 (96%)	134 (91%)	13 (9%)	12	22

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	MET
1	A	6	PRO
1	A	10	LEU
1	A	12	GLU
1	A	50	LYS
1	A	69	THR
1	A	119	ARG
1	A	151	GLU
1	A	155	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	161	LYS
1	A	182	LEU
1	A	186	ASN
1	A	198	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	101	GLN
1	A	164	HIS

### 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](#)

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](#)

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.