



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 07:12 PM GMT

PDB ID : 3M6Y  
Title : Structure of 4-hydroxy-2-oxoglutaratealdolase from bacillus cereus at 1.45 a resolution.  
Authors : Filippova, E.V.; Minasov, G.; Shuvalova, L.; Kiryukhina, O.; Clancy, S.; Joachimiak, A.; Anderson, F.W.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2010-03-16  
Resolution : 1.45 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

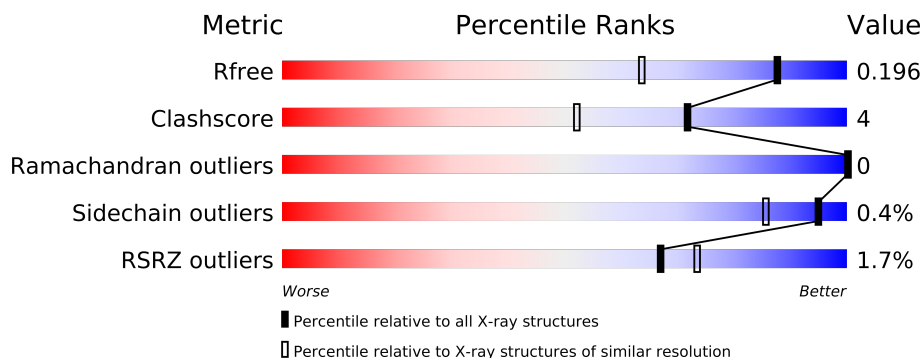
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.45 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))
$R_{free}$	66092	2476 (1.50-1.42)
Clashscore	79885	2834 (1.50-1.42)
Ramachandran outliers	78287	2769 (1.50-1.42)
Sidechain outliers	78261	2767 (1.50-1.42)
RSRZ outliers	66119	2477 (1.50-1.42)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	275	
1	B	275	
1	C	275	
1	D	275	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8851 atoms, of which 0 are hydrogen and 0 are deuterium.

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-oxoglutaratealdolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	252	Total	C	N	O	S	Se	0	6	0
			1937	1230	325	375	1	6			
1	B	250	Total	C	N	O	S	Se	0	3	0
			1910	1209	322	374	1	4			
1	C	248	Total	C	N	O	S	Se	0	1	0
			1882	1193	319	365	1	4			
1	D	248	Total	C	N	O	S	Se	0	1	0
			1886	1195	319	367	1	4			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-23	MSE	-	expression tag	UNP Q812M4
A	-22	HIS	-	expression tag	UNP Q812M4
A	-21	HIS	-	expression tag	UNP Q812M4
A	-20	HIS	-	expression tag	UNP Q812M4
A	-19	HIS	-	expression tag	UNP Q812M4
A	-18	HIS	-	expression tag	UNP Q812M4
A	-17	HIS	-	expression tag	UNP Q812M4
A	-16	SER	-	expression tag	UNP Q812M4
A	-15	SER	-	expression tag	UNP Q812M4
A	-14	GLY	-	expression tag	UNP Q812M4
A	-13	VAL	-	expression tag	UNP Q812M4
A	-12	ASP	-	expression tag	UNP Q812M4
A	-11	LEU	-	expression tag	UNP Q812M4
A	-10	GLY	-	expression tag	UNP Q812M4
A	-9	THR	-	expression tag	UNP Q812M4
A	-8	GLU	-	expression tag	UNP Q812M4
A	-7	ASN	-	expression tag	UNP Q812M4
A	-6	LEU	-	expression tag	UNP Q812M4
A	-5	TYR	-	expression tag	UNP Q812M4
A	-4	PHE	-	expression tag	UNP Q812M4
A	-3	GLN	-	expression tag	UNP Q812M4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q812M4
A	-1	ASN	-	expression tag	UNP Q812M4
A	0	ALA	-	expression tag	UNP Q812M4
B	-23	MSE	-	expression tag	UNP Q812M4
B	-22	HIS	-	expression tag	UNP Q812M4
B	-21	HIS	-	expression tag	UNP Q812M4
B	-20	HIS	-	expression tag	UNP Q812M4
B	-19	HIS	-	expression tag	UNP Q812M4
B	-18	HIS	-	expression tag	UNP Q812M4
B	-17	HIS	-	expression tag	UNP Q812M4
B	-16	SER	-	expression tag	UNP Q812M4
B	-15	SER	-	expression tag	UNP Q812M4
B	-14	GLY	-	expression tag	UNP Q812M4
B	-13	VAL	-	expression tag	UNP Q812M4
B	-12	ASP	-	expression tag	UNP Q812M4
B	-11	LEU	-	expression tag	UNP Q812M4
B	-10	GLY	-	expression tag	UNP Q812M4
B	-9	THR	-	expression tag	UNP Q812M4
B	-8	GLU	-	expression tag	UNP Q812M4
B	-7	ASN	-	expression tag	UNP Q812M4
B	-6	LEU	-	expression tag	UNP Q812M4
B	-5	TYR	-	expression tag	UNP Q812M4
B	-4	PHE	-	expression tag	UNP Q812M4
B	-3	GLN	-	expression tag	UNP Q812M4
B	-2	SER	-	expression tag	UNP Q812M4
B	-1	ASN	-	expression tag	UNP Q812M4
B	0	ALA	-	expression tag	UNP Q812M4
C	-23	MSE	-	expression tag	UNP Q812M4
C	-22	HIS	-	expression tag	UNP Q812M4
C	-21	HIS	-	expression tag	UNP Q812M4
C	-20	HIS	-	expression tag	UNP Q812M4
C	-19	HIS	-	expression tag	UNP Q812M4
C	-18	HIS	-	expression tag	UNP Q812M4
C	-17	HIS	-	expression tag	UNP Q812M4
C	-16	SER	-	expression tag	UNP Q812M4
C	-15	SER	-	expression tag	UNP Q812M4
C	-14	GLY	-	expression tag	UNP Q812M4
C	-13	VAL	-	expression tag	UNP Q812M4
C	-12	ASP	-	expression tag	UNP Q812M4
C	-11	LEU	-	expression tag	UNP Q812M4
C	-10	GLY	-	expression tag	UNP Q812M4
C	-9	THR	-	expression tag	UNP Q812M4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-8	GLU	-	expression tag	UNP Q812M4
C	-7	ASN	-	expression tag	UNP Q812M4
C	-6	LEU	-	expression tag	UNP Q812M4
C	-5	TYR	-	expression tag	UNP Q812M4
C	-4	PHE	-	expression tag	UNP Q812M4
C	-3	GLN	-	expression tag	UNP Q812M4
C	-2	SER	-	expression tag	UNP Q812M4
C	-1	ASN	-	expression tag	UNP Q812M4
C	0	ALA	-	expression tag	UNP Q812M4
D	-23	MSE	-	expression tag	UNP Q812M4
D	-22	HIS	-	expression tag	UNP Q812M4
D	-21	HIS	-	expression tag	UNP Q812M4
D	-20	HIS	-	expression tag	UNP Q812M4
D	-19	HIS	-	expression tag	UNP Q812M4
D	-18	HIS	-	expression tag	UNP Q812M4
D	-17	HIS	-	expression tag	UNP Q812M4
D	-16	SER	-	expression tag	UNP Q812M4
D	-15	SER	-	expression tag	UNP Q812M4
D	-14	GLY	-	expression tag	UNP Q812M4
D	-13	VAL	-	expression tag	UNP Q812M4
D	-12	ASP	-	expression tag	UNP Q812M4
D	-11	LEU	-	expression tag	UNP Q812M4
D	-10	GLY	-	expression tag	UNP Q812M4
D	-9	THR	-	expression tag	UNP Q812M4
D	-8	GLU	-	expression tag	UNP Q812M4
D	-7	ASN	-	expression tag	UNP Q812M4
D	-6	LEU	-	expression tag	UNP Q812M4
D	-5	TYR	-	expression tag	UNP Q812M4
D	-4	PHE	-	expression tag	UNP Q812M4
D	-3	GLN	-	expression tag	UNP Q812M4
D	-2	SER	-	expression tag	UNP Q812M4
D	-1	ASN	-	expression tag	UNP Q812M4
D	0	ALA	-	expression tag	UNP Q812M4

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total 3 Ca 3	0	1
2	A	2	Total 3 Ca 3	0	1

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	389	Total 399	O 399	3	12
4	B	326	Total 328	O 328	0	3
4	C	274	Total 276	O 276	0	3
4	D	216	Total 221	O 221	3	7



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.68Å 106.19Å 88.20Å 90.00° 90.95° 90.00°	Depositor
Resolution (Å)	88.05 – 1.45 28.33 – 1.45	Depositor EDS
% Data completeness (in resolution range)	99.7 (88.05-1.45) 99.7 (28.33-1.45)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.47 (at 1.45Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.152 , 0.180 0.169 , 0.196	Depositor DCC
$R_{free}$ test set	7916 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	12.7	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 34.3	EDS
Estimated twinning fraction	0.025 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 157810 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8851	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.04% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.



## 5 Model quality

### 5.1 Standard geometry

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

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	0.79	0/1978	0.81	0/2671
1	B	0.75	1/1943 (0.1%)	0.75	0/2626
1	C	0.68	0/1912	0.73	2/2585 (0.1%)
1	D	0.72	0/1915	0.73	0/2587
All	All	0.74	1/7748 (0.0%)	0.76	2/10469 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	49	GLU	CG-CD	5.06	1.59	1.51

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	7	ARG	NE-CZ-NH2	-5.57	117.52	120.30
1	C	7	ARG	NE-CZ-NH1	5.05	122.83	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1937	0	1972	11	0
1	B	1910	0	1924	17	0
1	C	1882	0	1901	28	0
1	D	1886	0	1904	16	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	399	0	0	2	0
4	B	328	0	0	5	0
4	C	276	0	0	3	0
4	D	221	0	0	3	0
All	All	8851	0	7701	67	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

The worst 5 of 67 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:156[B]:MSE:SE	4:D:575:HOH:O	2.29	1.00
1:B:156[A]:MSE:HE1	1:C:156[A]:MSE:HE2	1.41	0.96
1:B:156[A]:MSE:HE1	1:C:156[A]:MSE:CE	1.95	0.94
1:B:156[A]:MSE:SE	4:B:608:HOH:O	2.37	0.91
1:C:5:GLN:HE22	1:C:214:GLN:HE22	1.14	0.91

There are no symmetry-related clashes.

## 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	256/275 (93%)	250 (98%)	6 (2%)	0	100	100
1	B	251/275 (91%)	247 (98%)	4 (2%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	247/275 (90%)	241 (98%)	6 (2%)	0	100	100
1	D	245/275 (89%)	240 (98%)	5 (2%)	0	100	100
All	All	999/1100 (91%)	978 (98%)	21 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	204/214 (95%)	202 (99%)	2 (1%)	85	64
1	B	200/214 (94%)	199 (100%)	1 (0%)	94	83
1	C	197/214 (92%)	197 (100%)	0	100	100
1	D	198/214 (92%)	198 (100%)	0	100	100
All	All	799/856 (93%)	796 (100%)	3 (0%)	95	85

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	234	GLU
1	A	245	LYS
1	B	211	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	20	ASN
1	C	24	ASN
1	C	159	ASN
1	C	5	GLN
1	C	106	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 12 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.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	252/275 (91%)	-0.29	1 (0%) 90 94	6, 10, 22, 36	0
1	B	250/275 (90%)	-0.27	2 (0%) 83 88	8, 13, 23, 34	0
1	C	248/275 (90%)	0.16	9 (3%) 41 44	8, 16, 30, 43	0
1	D	248/275 (90%)	-0.10	5 (2%) 62 68	6, 13, 27, 40	0
All	All	998/1100 (90%)	-0.13	17 (1%) 67 73	6, 13, 27, 43	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	2	THR	5.4
1	D	109	GLU	4.6
1	C	109	GLU	3.9
1	C	227	GLU	3.8
1	C	72	ALA	3.5

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

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	A	254	1/1	0.09	0.68	21,21,21,21	0
2	CA	B	253[B]	1/1	0.06	-0.94	34,34,34,34	1
3	CL	D	252	1/1	0.05	-0.96	16,16,16,16	0
3	CL	B	255	1/1	0.04	-0.98	17,17,17,17	0
2	CA	B	252	1/1	0.08	-1.08	17,17,17,17	0
2	CA	B	253[A]	1/1	0.06	-1.19	14,14,14,14	1
3	CL	B	254	1/1	0.04	-1.34	18,18,18,18	0
2	CA	A	253[B]	1/1	0.05	-1.61	17,17,17,17	1
3	CL	C	252	1/1	0.03	-1.86	21,21,21,21	0
3	CL	A	255	1/1	0.04	-2.44	24,24,24,24	0
2	CA	A	252	1/1	0.02	-2.44	8,8,8,8	0
2	CA	A	253[A]	1/1	0.05	-2.56	12,12,12,12	1

## 6.5 Other polymers ⓘ

There are no such residues in this entry.