



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 06:07 AM GMT

PDB ID : 2AAL  
Title : Crystal Structures of the Wild-type, Mutant-P1A and Inactivated Malonate Semialdehyde Decarboxylase: A Structural Basis for the Decarboxylase and Hydratase Activities  
Authors : Almrud, J.J.; Poelarends, G.J.; Johnson Jr., W.H.; Serrano, H.; Hackert, M.L.; Whitman, C.P.  
Deposited on : 2005-07-13  
Resolution : 1.65 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>

---

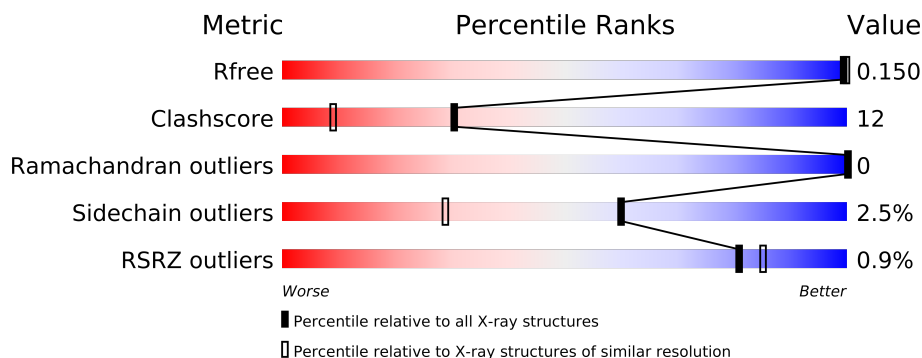
The following versions of software and data (see [references](#)) were used in the production of this report:

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 1.65 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	1404 (1.68-1.64)
Clashscore	79885	1001 (1.66-1.66)
Ramachandran outliers	78287	1581 (1.68-1.64)
Sidechain outliers	78261	1580 (1.68-1.64)
RSRZ outliers	66119	1404 (1.68-1.64)

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	131	
1	B	131	
1	C	131	
1	D	131	
1	E	131	
1	F	131	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7105 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 Malonate Semialdehyde Decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	130	Total	C	N	O	S	0	5	0
			1040	651	182	203	4			
1	B	131	Total	C	N	O	S	0	1	0
			1009	634	173	198	4			
1	C	130	Total	C	N	O	S	0	5	0
			1035	650	177	204	4			
1	D	130	Total	C	N	O	S	0	6	0
			1041	654	177	205	5			
1	E	130	Total	C	N	O	S	0	3	0
			1020	643	174	199	4			
1	F	130	Total	C	N	O	S	0	1	0
			1004	631	172	196	5			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MLI	-	PROPIONATION	GB 10637971
A	7	LEU	ILE	CONFLICT	GB 10637971
A	130	GLY	-	CLONING ARTIFACT	GB 10637971
B	0	MLI	-	PROPIONATION	GB 10637971
B	7	LEU	ILE	CONFLICT	GB 10637971
B	130	GLY	-	CLONING ARTIFACT	GB 10637971
C	0	MLI	-	PROPIONATION	GB 10637971
C	7	LEU	ILE	CONFLICT	GB 10637971
C	130	GLY	-	CLONING ARTIFACT	GB 10637971
D	0	MLI	-	PROPIONATION	GB 10637971
D	7	LEU	ILE	CONFLICT	GB 10637971
D	130	GLY	-	CLONING ARTIFACT	GB 10637971
E	0	MLI	-	PROPIONATION	GB 10637971
E	7	LEU	ILE	CONFLICT	GB 10637971
E	130	GLY	-	CLONING ARTIFACT	GB 10637971
F	0	MLI	-	PROPIONATION	GB 10637971
F	7	LEU	ILE	CONFLICT	GB 10637971

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	130	GLY	-	CLONING ARTIFACT	GB 10637971

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	148	Total O 148 148	0	0
2	B	135	Total O 135 135	0	0
2	C	158	Total O 158 158	0	0
2	D	168	Total O 168 168	0	0
2	E	171	Total O 171 171	0	0
2	F	176	Total O 176 176	0	0

### 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.

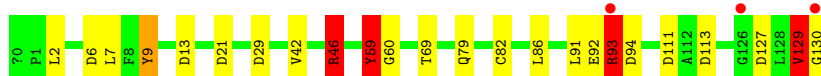
- Molecule 1: Malonate Semialdehyde Decarboxylase

Chain A: 



- Molecule 1: Malonate Semialdehyde Decarboxylase

Chain B: 



- Molecule 1: Malonate Semialdehyde Decarboxylase

Chain C: 



- Molecule 1: Malonate Semialdehyde Decarboxylase

Chain D: 



- Molecule 1: Malonate Semialdehyde Decarboxylase

Chain E: 



- Molecule 1: Malonate Semialdehyde Decarboxylase

Chain F: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.89Å 51.89Å 219.08Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	72.55 – 1.65 73.03 – 1.60	Depositor EDS
% Data completeness (in resolution range)	99.0 (72.55-1.65) 98.8 (73.03-1.60)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.50 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.184 , 0.239 0.155 , 0.150	Depositor DCC
$R_{free}$ test set	4294 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	9.1	Xtriage
Anisotropy	0.027	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 16.2	EDS
Estimated twinning fraction	0.069 for -h,-k,l 0.237 for h,-h-k,-l 0.074 for -k,-h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	0 of 86018 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7105	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	11.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.87% 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: MLI

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	1.32	6/1052 (0.6%)	1.47	16/1423 (1.1%)
1	B	1.42	13/1021 (1.3%)	1.58	19/1382 (1.4%)
1	C	1.52	10/1047 (1.0%)	1.57	18/1417 (1.3%)
1	D	1.56	13/1054 (1.2%)	1.34	19/1426 (1.3%)
1	E	1.64	14/1032 (1.4%)	1.72	23/1397 (1.6%)
1	F	1.33	8/1016 (0.8%)	1.58	11/1375 (0.8%)
All	All	1.47	64/6222 (1.0%)	1.55	106/8420 (1.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	2
1	C	0	3
1	F	0	1
All	All	0	6

The worst 5 of 64 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	39	TYR	CE1-CZ	-16.42	1.17	1.38
1	D	59	TYR	CE1-CZ	-15.58	1.18	1.38
1	C	129	VAL	CA-CB	-14.02	1.25	1.54
1	C	129	VAL	CB-CG1	-13.67	1.24	1.52
1	D	59	TYR	CG-CD2	-12.92	1.22	1.39

The worst 5 of 106 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	61	ARG	NE-CZ-NH1	33.53	137.07	120.30
1	F	3	LEU	O-C-N	-31.10	72.95	122.70
1	C	129	VAL	CB-CA-C	-24.75	64.37	111.40
1	B	93	ARG	NE-CZ-NH2	-21.31	109.65	120.30
1	B	46	ARG	NE-CZ-NH1	-19.91	110.35	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	59	TYR	Peptide
1	B	93	ARG	Sidechain
1	C	128	LEU	Mainchain,Peptide
1	C	59	TYR	Peptide
1	F	3	LEU	Mainchain

## 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	1040	0	0	17	2
1	B	1009	0	0	16	3
1	C	1035	0	0	10	2
1	D	1041	0	0	15	0
1	E	1020	0	0	15	0
1	F	1004	0	0	7	0
2	A	148	0	0	9	4
2	B	135	0	0	9	1
2	C	158	0	0	7	6
2	D	168	0	0	7	3
2	E	171	0	0	12	4
2	F	176	0	0	5	3
All	All	7105	0	0	73	14

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 12.



The worst 5 of 73 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:46:ARG:NH2	1:B:9:TYR:OH	1.73	1.21
1:D:82[A]:CYS:SG	2:D:293:HOH:O	2.04	1.16
1:B:82:CYS:SG	2:B:261:HOH:O	2.04	1.15
1:B:46:ARG:CG	1:B:46:ARG:NH1	2.00	1.14
1:E:82:CYS:SG	2:E:303:HOH:O	2.07	1.10

The worst 5 of 14 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	Distance(Å)	Clash(Å)
2:B:240:HOH:O	2:D:279:HOH:O[1_444]	1.18	1.02
1:B:93:ARG:NH2	1:C:93:ARG:CB[1_455]	1.67	0.53
1:A:93:ARG:O	1:B:93:ARG:NH1[1_665]	1.86	0.34
2:A:269:HOH:O	2:C:172:HOH:O[1_455]	1.89	0.31
1:C:46:ARG:NH2	2:D:216:HOH:O[1_544]	1.95	0.25

## 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	132/131 (101%)	128 (97%)	4 (3%)	0	100	100
1	B	129/131 (98%)	126 (98%)	3 (2%)	0	100	100
1	C	132/131 (101%)	128 (97%)	4 (3%)	0	100	100
1	D	133/131 (102%)	130 (98%)	3 (2%)	0	100	100
1	E	130/131 (99%)	125 (96%)	5 (4%)	0	100	100
1	F	128/131 (98%)	125 (98%)	3 (2%)	0	100	100
All	All	784/786 (100%)	762 (97%)	22 (3%)	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	111/106 (105%)	108 (97%)	3 (3%)	57	25
1	B	107/106 (101%)	105 (98%)	2 (2%)	69	43
1	C	111/106 (105%)	109 (98%)	2 (2%)	71	46
1	D	112/106 (106%)	109 (97%)	3 (3%)	57	25
1	E	109/106 (103%)	104 (95%)	5 (5%)	37	9
1	F	107/106 (101%)	105 (98%)	2 (2%)	69	43
All	All	657/636 (103%)	640 (97%)	17 (3%)	60	27

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	1	PRO
1	D	7	LEU
1	E	61	ARG
1	C	129	VAL
1	E	106	LEU

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 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 ⓘ

There are no ligands in this entry.

## 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	130/131 (99%)	-0.41	1 (0%) 83 87	6, 9, 15, 19	0
1	B	131/131 (100%)	-0.33	3 (2%) 57 58	6, 9, 15, 25	0
1	C	130/131 (99%)	-0.38	1 (0%) 83 87	6, 9, 15, 19	0
1	D	130/131 (99%)	-0.42	0 100 100	6, 9, 15, 19	0
1	E	130/131 (99%)	-0.42	1 (0%) 83 87	6, 9, 15, 19	0
1	F	130/131 (99%)	-0.43	1 (0%) 83 87	6, 10, 15, 19	0
All	All	781/786 (99%)	-0.40	7 (0%) 81 85	6, 9, 15, 25	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	129	VAL	5.0
1	B	130	GLY	3.8
1	F	129	VAL	2.5
1	B	93	ARG	2.4
1	B	126	GLY	2.4

### 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

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.