



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 3, 2024 – 03:51 am BST

PDB ID : 8RP5  
Title : Alpha-Methylacyl-CoA racemase from Mycobacterium tuberculosis (E241A mutant)  
Authors : Mojanaga, O.O.; Acharya, K.R.; Lloyd, M.D.  
Deposited on : 2024-01-12  
Resolution : 1.85 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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

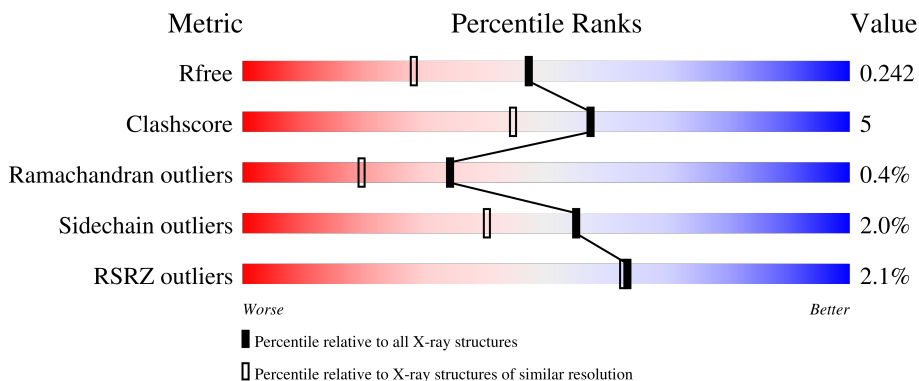
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	365	 2% 87% 8% . .
1	B	365	 2% 83% 13% . .

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11146 atoms, of which 5328 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 Alpha-methylacyl-CoA racemase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	354	5326	1687	2639	481	503	16	56	3	0
1	B	354	5323	1686	2639	481	501	16	56	2	0

There are 12 discrepancies between the modelled and reference sequences:

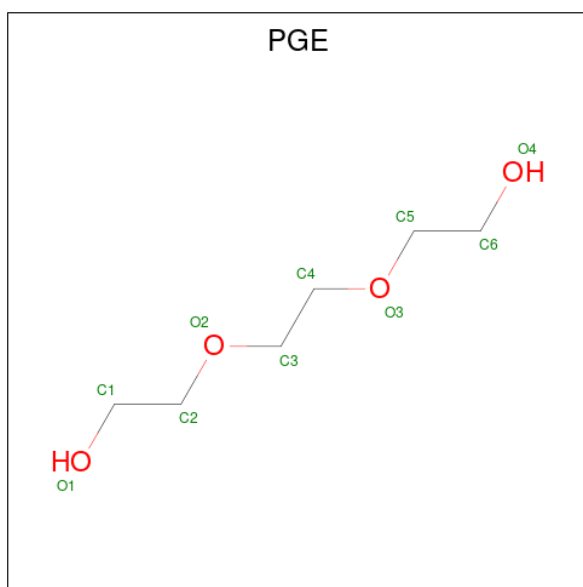
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP O06543
A	241	ALA	GLU	engineered mutation	UNP O06543
A	361	GLY	-	expression tag	UNP O06543
A	362	SER	-	expression tag	UNP O06543
A	363	GLY	-	expression tag	UNP O06543
A	364	CYS	-	expression tag	UNP O06543
B	0	MET	-	initiating methionine	UNP O06543
B	241	ALA	GLU	engineered mutation	UNP O06543
B	361	GLY	-	expression tag	UNP O06543
B	362	SER	-	expression tag	UNP O06543
B	363	GLY	-	expression tag	UNP O06543
B	364	CYS	-	expression tag	UNP O06543

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	A	1	10	2	6	2	1	0
2	A	1	10	2	6	2	1	0
2	A	1	10	2	6	2	1	0
2	B	1	10	2	6	2	1	0
2	B	1	10	2	6	2	1	0
2	B	1	10	2	6	2	1	0

- Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
3	A	1	24	6	14	4	1	0

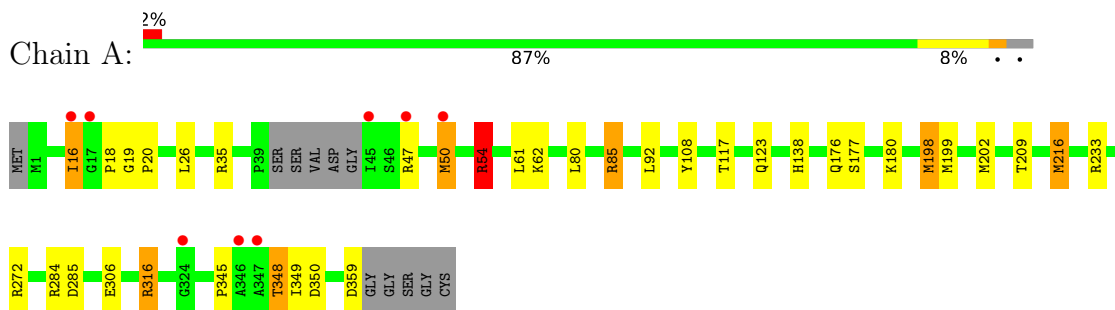
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	191	Total	O	0	0
			191	191		
4	B	222	Total	O	0	0
			222	222		

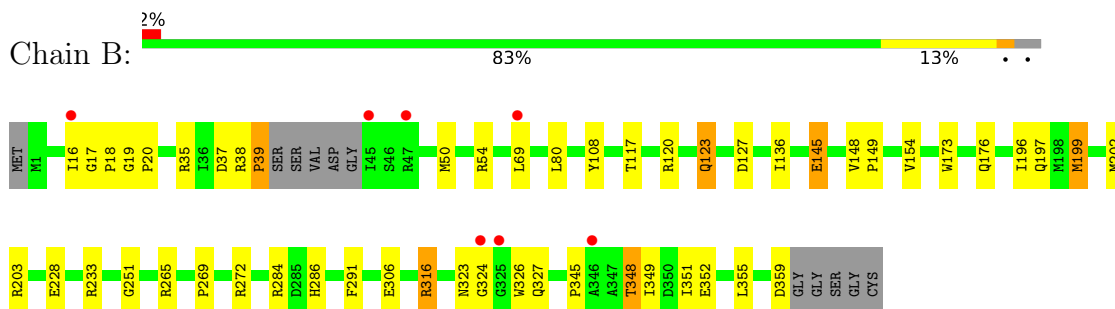
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Alpha-methylacyl-CoA racemase



- Molecule 1: Alpha-methylacyl-CoA racemase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	180.62Å 78.99Å 58.92Å 90.00° 92.19° 90.00°	Depositor
Resolution (Å)	90.24 – 1.85 90.24 – 1.85	Depositor EDS
% Data completeness (in resolution range)	100.0 (90.24-1.85) 99.8 (90.24-1.85)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.21 (at 1.84Å)	Xtrriage
Refinement program	REFMAC 5.8.0419, REFMAC 5.8.0419	Depositor
R, $R_{free}$	0.193 , 0.236 0.203 , 0.242	Depositor DCC
$R_{free}$ test set	3583 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.1	Xtrriage
Anisotropy	0.451	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 48.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.160 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11146	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.00% 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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.76	0/2763	1.24	20/3758 (0.5%)
1	B	0.80	2/2754 (0.1%)	1.30	22/3746 (0.6%)
All	All	0.78	2/5517 (0.0%)	1.27	42/7504 (0.6%)

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	A	0	3
1	B	0	1
All	All	0	4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	145	GLU	CD-OE1	6.67	1.32	1.25
1	B	228	GLU	CD-OE2	-5.09	1.20	1.25

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	47	ARG	NE-CZ-NH2	11.42	126.01	120.30
1	B	54	ARG	NE-CZ-NH2	-9.82	115.39	120.30
1	A	54	ARG	NE-CZ-NH2	-9.75	115.42	120.30
1	A	216	MET	CG-SD-CE	9.73	115.76	100.20
1	B	54	ARG	CB-CA-C	-9.58	91.24	110.40

There are no chirality outliers.



All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	233	ARG	Sidechain
1	A	54	ARG	Sidechain
1	A	85	ARG	Sidechain
1	B	233	ARG	Sidechain

## 5.2 Too-close contacts [i](#)

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	2687	2639	2630	28	0
1	B	2684	2639	2632	28	0
2	A	12	18	18	0	0
2	B	12	18	17	0	0
3	A	10	14	14	2	0
4	A	191	0	0	15	0
4	B	222	0	0	9	0
All	All	5818	5328	5311	52	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 5.

The worst 5 of 52 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:286:HIS:HD2	4:B:693:HOH:O	1.69	0.76
1:B:17:GLY:HA3	4:B:558:HOH:O	1.83	0.76
1:A:123:GLN:NE2	4:A:502:HOH:O	2.20	0.74
1:B:345:PRO:HB3	4:B:649:HOH:O	1.90	0.72
1:A:16:ILE:O	1:A:18:PRO:HD2	1.91	0.71

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	353/365 (97%)	332 (94%)	20 (6%)	1 (0%)	41	26
1	B	352/365 (96%)	333 (95%)	17 (5%)	2 (1%)	25	12
All	All	705/730 (97%)	665 (94%)	37 (5%)	3 (0%)	34	19

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	348	THR
1	B	348	THR
1	B	324	GLY

### 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	273/277 (99%)	267 (98%)	6 (2%)	52	36
1	B	272/277 (98%)	267 (98%)	5 (2%)	59	45
All	All	545/554 (98%)	534 (98%)	11 (2%)	55	40

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	145	GLU
1	B	173	TRP
1	B	269	PRO

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	199	MET
1	A	350	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	263	ASN
1	B	286	HIS
1	B	327	GLN
1	B	323	ASN
1	A	176	GLN

### 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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

7 ligands 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	$\# Z  > 2$
2	EDO	B	402	-	3,3,3	0.34	0	2,2,2	0.82	0
2	EDO	B	403	-	3,3,3	0.88	0	2,2,2	0.58	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	EDO	B	401	-	3,3,3	0.43	0	2,2,2	0.23	0
2	EDO	A	401	-	3,3,3	0.38	0	2,2,2	1.02	0
2	EDO	A	403	-	3,3,3	0.35	0	2,2,2	0.30	0
2	EDO	A	402	-	3,3,3	0.20	0	2,2,2	0.23	0
3	PGE	A	404	-	9,9,9	0.51	0	8,8,8	0.65	0

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
2	EDO	B	402	-	-	0/1/1/1	-
2	EDO	B	403	-	-	1/1/1/1	-
2	EDO	B	401	-	-	1/1/1/1	-
2	EDO	A	401	-	-	1/1/1/1	-
2	EDO	A	403	-	-	0/1/1/1	-
2	EDO	A	402	-	-	1/1/1/1	-
3	PGE	A	404	-	-	6/7/7/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	404	PGE	C4-C3-O2-C2
3	A	404	PGE	C6-C5-O3-C4
3	A	404	PGE	O2-C3-C4-O3
2	A	401	EDO	O1-C1-C2-O2
3	A	404	PGE	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	404	PGE	2	0

## 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	354/365 (96%)	0.05	8 (2%) 60 59	17, 29, 59, 86	0
1	B	354/365 (96%)	0.06	7 (1%) 65 64	18, 30, 60, 82	0
All	All	708/730 (96%)	0.06	15 (2%) 63 63	17, 29, 60, 86	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	16	ILE	5.0
1	A	324	GLY	4.1
1	B	324	GLY	4.1
1	A	47	ARG	3.8
1	B	69	LEU	2.9

### 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 monosaccharides 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	EDO	B	402	4/4	0.77	0.25	41,49,51,52	1
3	PGE	A	404	10/10	0.91	0.18	32,43,64,67	1
2	EDO	B	401	4/4	0.94	0.15	34,36,38,41	1
2	EDO	A	401	4/4	0.94	0.28	39,50,57,57	1
2	EDO	B	403	4/4	0.94	0.10	24,26,36,36	1
2	EDO	A	402	4/4	0.94	0.15	35,37,50,55	1
2	EDO	A	403	4/4	0.97	0.11	26,29,32,33	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.