



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 11, 2016 – 03:02 PM EDT

PDB ID : 5ERX  
Title : Crystal Structure of APO MenD from M. tuberculosis - I222  
Authors : Johnston, J.M.; Jirgis, E.N.M.; Bashiri, G.; Bulloch, E.M.M.; Baker, E.N.  
Deposited on : 2015-11-16  
Resolution : 1.73 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 : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027790  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027790

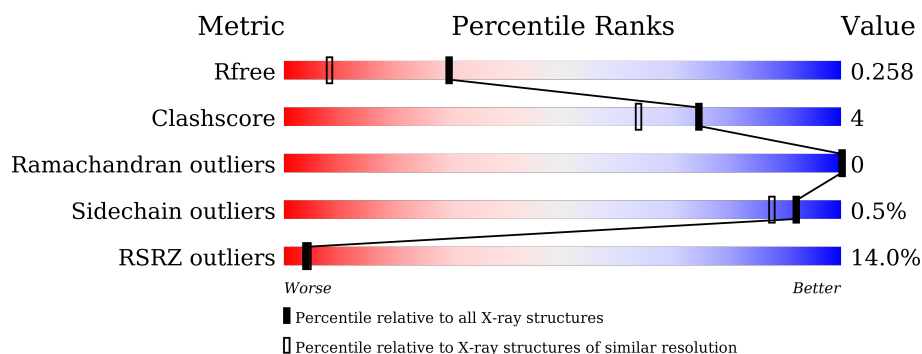
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.73 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))
$R_{free}$	91344	3998 (1.74-1.70)
Clashscore	102246	4425 (1.74-1.70)
Ramachandran outliers	100387	4360 (1.74-1.70)
Sidechain outliers	100360	4360 (1.74-1.70)
RSRZ outliers	91569	4010 (1.74-1.70)

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.

Mol	Chain	Length	Quality of chain
1	A	574	<div> <div>12%</div> <div> <div></div> <div>80%</div> <div>8%</div> <div>11%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4013 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 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	508	Total	C	N	O	S	0	5	0
			3731	2331	697	693	10			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP P9WK11
A	-18	GLY	-	expression tag	UNP P9WK11
A	-17	SER	-	expression tag	UNP P9WK11
A	-16	SER	-	expression tag	UNP P9WK11
A	-15	HIS	-	expression tag	UNP P9WK11
A	-14	HIS	-	expression tag	UNP P9WK11
A	-13	HIS	-	expression tag	UNP P9WK11
A	-12	HIS	-	expression tag	UNP P9WK11
A	-11	HIS	-	expression tag	UNP P9WK11
A	-10	HIS	-	expression tag	UNP P9WK11
A	-9	SER	-	expression tag	UNP P9WK11
A	-8	SER	-	expression tag	UNP P9WK11
A	-7	GLY	-	expression tag	UNP P9WK11
A	-6	LEU	-	expression tag	UNP P9WK11
A	-5	VAL	-	expression tag	UNP P9WK11
A	-4	PRO	-	expression tag	UNP P9WK11
A	-3	ARG	-	expression tag	UNP P9WK11
A	-2	GLY	-	expression tag	UNP P9WK11
A	-1	SER	-	expression tag	UNP P9WK11
A	0	HIS	-	expression tag	UNP P9WK11

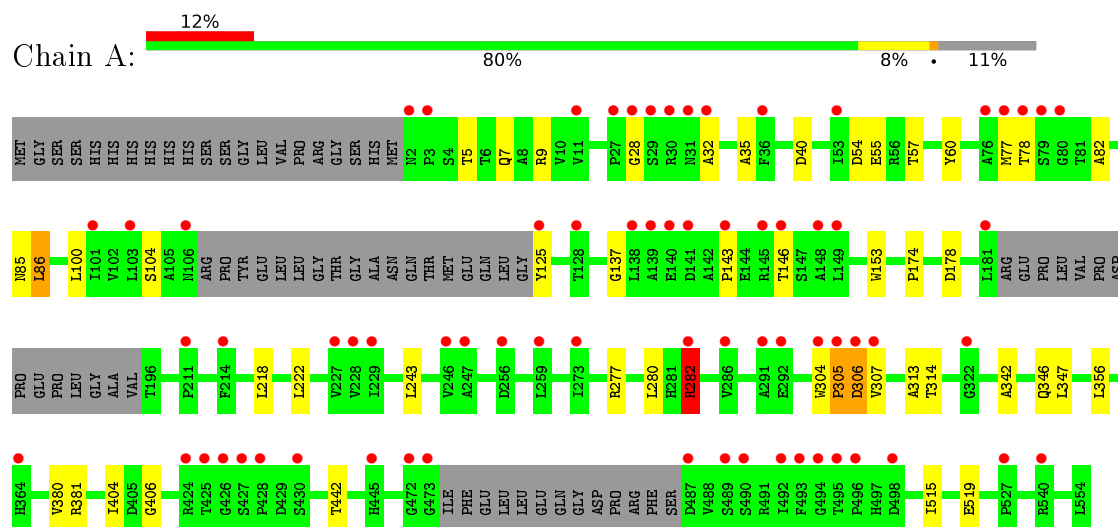
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	282	Total	O	0	0
			282	282		

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

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: 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.06 Å   105.50 Å   120.19 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	19.66 – 1.73 19.66 – 1.73	Depositor EDS
% Data completeness (in resolution range)	94.4 (19.66-1.73) 94.5 (19.66-1.73)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.51 (at 1.73 Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.229   ,   0.259 0.228   ,   0.258	Depositor DCC
$R_{free}$ test set	3205 reflections (5.10%)	DCC
Wilson B-factor (Å <sup>2</sup> )	27.4	Xtriage
Anisotropy	0.448	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 46.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4013	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.46% 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

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	0.30	0/3820	0.49	3/5231 (0.1%)

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	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	306	ASP	CB-CG-OD2	5.86	123.57	118.30
1	A	282	ARG	NE-CZ-NH1	-5.66	117.47	120.30
1	A	86	LEU	CB-CG-CD1	-5.16	102.22	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	305	PRO	Peptide

### 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	3731	0	3767	30	0
2	A	282	0	0	0	0
All	All	4013	0	3767	30	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 4.

All (30) 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:32:ALA:HB3	1:A:35:ALA:HB2	1.61	0.81
1:A:280:LEU:HD11	1:A:380:VAL:HG22	1.81	0.63
1:A:32:ALA:HA	1:A:78:THR:HG22	1.81	0.61
1:A:5:THR:O	1:A:9:ARG:HG3	2.02	0.59
1:A:143:PRO:HA	1:A:146:THR:HG23	1.84	0.58
1:A:304:TRP:CD2	1:A:305:PRO:HD2	2.39	0.57
1:A:515:ILE:HG23	1:A:519:GLU:HB2	1.87	0.56
1:A:306:ASP:HB2	1:A:307:VAL:HG13	1.86	0.56
1:A:77[B]:MET:HB3	1:A:82:ALA:HB1	1.89	0.53
1:A:55:GLU:HB2	1:A:85:ASN:HB3	1.91	0.51
1:A:218:LEU:HD11	1:A:313:ALA:HB1	1.92	0.51
1:A:305:PRO:HG3	1:A:314:THR:HG22	1.95	0.49
1:A:7:GLN:NE2	1:A:137:GLY:O	2.41	0.49
1:A:347:LEU:HD21	1:A:356:LEU:HD23	1.94	0.48
1:A:305:PRO:HG3	1:A:314:THR:CG2	2.44	0.47
1:A:277:ARG:CZ	1:A:306:ASP:HB3	2.43	0.47
1:A:404:ILE:HB	1:A:442:THR:HG23	1.96	0.47
1:A:77[B]:MET:SD	1:A:86:LEU:HD11	2.56	0.46
1:A:60:TYR:CG	1:A:406:GLY:HA3	2.51	0.45
1:A:54:ASP:HB3	1:A:57:THR:HB	2.01	0.43
1:A:280:LEU:HD22	1:A:381:ARG:HG2	2.00	0.43
1:A:7:GLN:HA	1:A:153:TRP:CZ3	2.54	0.42
1:A:9:ARG:HH12	1:A:40:ASP:CG	2.23	0.42
1:A:28:GLY:HA3	1:A:78:THR:OG1	2.19	0.42
1:A:104:SER:OG	1:A:178:ASP:OD1	2.29	0.42
1:A:100:LEU:O	1:A:174:PRO:HA	2.19	0.41
1:A:222:LEU:HD22	1:A:243:LEU:HD11	2.03	0.41
1:A:342:ALA:O	1:A:346:GLN:HG2	2.21	0.40
1:A:282:ARG:HA	1:A:282:ARG:HE	1.87	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	505/574 (88%)	496 (98%)	9 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	389/445 (87%)	387 (100%)	2 (0%)	92	87

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	125	TYR
1	A	282	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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

### 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	508/574 (88%)	0.73	71 (13%) <b>4</b> <b>4</b>	21, 37, 72, 92	0

All (71) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	36	PHE	7.9
1	A	494	GLY	6.7
1	A	495	THR	6.6
1	A	125	TYR	6.6
1	A	496	PRO	6.1
1	A	428	PRO	5.8
1	A	181	LEU	5.8
1	A	426	GLY	5.8
1	A	29	SER	5.6
1	A	291	ALA	5.3
1	A	138	LEU	5.1
1	A	106	ASN	4.9
1	A	527	PRO	4.5
1	A	211	PRO	4.5
1	A	493	PHE	4.4
1	A	31	ASN	4.4
1	A	214[A]	PHE	4.3
1	A	292	GLU	4.3
1	A	306	ASP	4.3
1	A	141	ASP	4.3
1	A	139	ALA	4.2
1	A	145[A]	ARG	4.0
1	A	143	PRO	3.9
1	A	28	GLY	3.8
1	A	140	GLU	3.8
1	A	32	ALA	3.8
1	A	490	SER	3.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	487	ASP	3.5
1	A	307	VAL	3.5
1	A	2	ASN	3.5
1	A	256	ASP	3.4
1	A	498	ASP	3.4
1	A	282	ARG	3.3
1	A	228	VAL	3.3
1	A	80	GLY	3.1
1	A	11	VAL	3.1
1	A	101	ILE	3.1
1	A	473	GLY	3.1
1	A	425	THR	3.0
1	A	305	PRO	2.7
1	A	79	SER	2.7
1	A	427	SER	2.7
1	A	259	LEU	2.7
1	A	424	ARG	2.6
1	A	30	ARG	2.6
1	A	27	PRO	2.6
1	A	146	THR	2.5
1	A	286	VAL	2.5
1	A	76	ALA	2.4
1	A	472	GLY	2.4
1	A	227	VAL	2.3
1	A	103	LEU	2.3
1	A	229	ILE	2.3
1	A	489	SER	2.3
1	A	273	ILE	2.3
1	A	445[A]	HIS	2.3
1	A	77[A]	MET	2.3
1	A	3	PRO	2.3
1	A	246	VAL	2.3
1	A	492	ILE	2.3
1	A	430	SER	2.3
1	A	78	THR	2.2
1	A	53	ILE	2.2
1	A	304	TRP	2.2
1	A	322	GLY	2.2
1	A	364	HIS	2.1
1	A	247	ALA	2.1
1	A	128	THR	2.1
1	A	540	ARG	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	148	ALA	2.1
1	A	149	LEU	2.1

## 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 carbohydrates 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.