



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:02 PM GMT

PDB ID : 1N8W
Title : Biochemical and Structural Studies of Malate Synthase from Mycobacterium tuberculosis
Authors : Smith, C.V.; Huang, C.C.; Miczak, A.; Russell, D.G.; Sacchettini, J.C.; Honer zu Bentrup, K.; TB Structural Genomics Consortium (TBSGC)
Deposited on : 2002-11-21
Resolution : 2.70 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
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

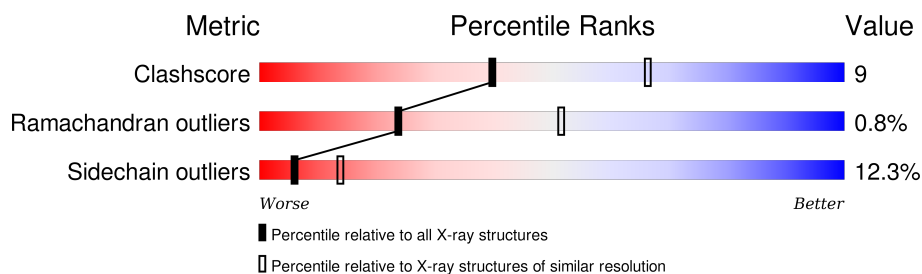
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)

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 ≥ 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	741	
1	B	741	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MLT	A	901	X	-	-	-
4	MLT	A	902	X	-	-	-
4	MLT	B	1902	X	-	-	-

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 11696 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 Probable malate synthase G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	718	Total	C	N	O	S	0	0	0
			5486	3448	964	1052	22			
1	B	715	Total	C	N	O	S	0	0	0
			5474	3441	964	1047	22			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

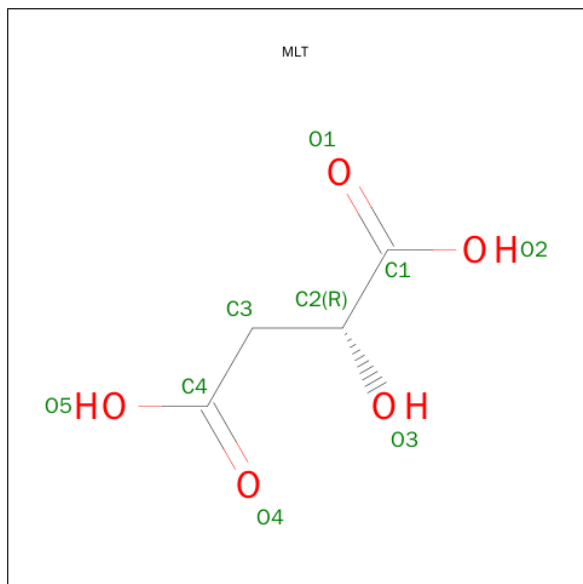
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



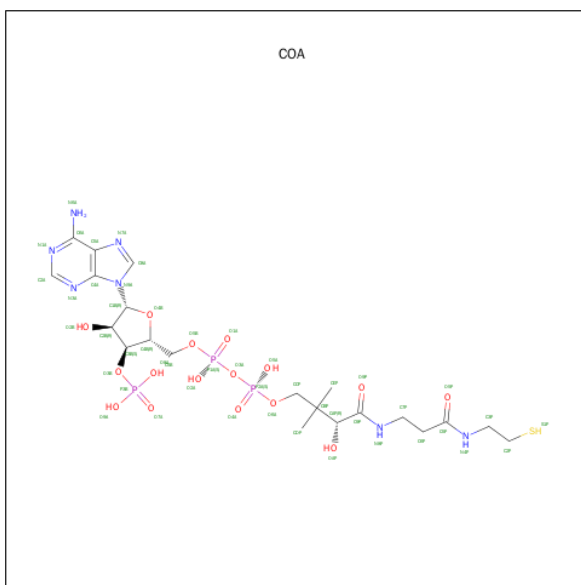
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is MALATE ION (three-letter code: MLT) (formula: $C_4H_6O_5$).



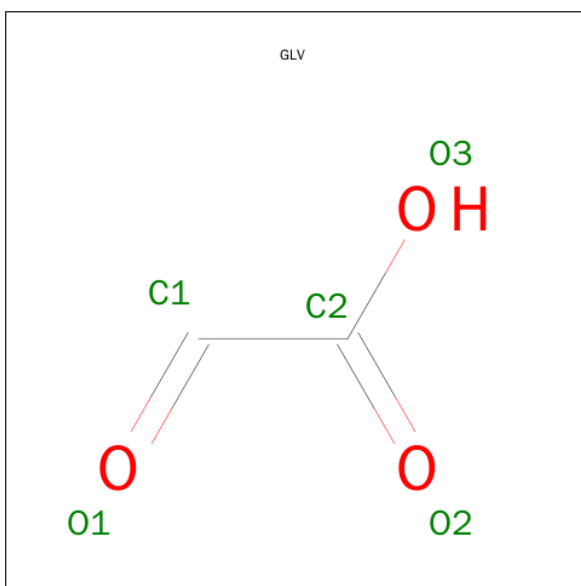
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			9	4	5		
4	A	1	Total	C	O	0	0
			9	4	5		
4	B	1	Total	C	O	0	0
			9	4	5		

- Molecule 5 is COENZYME A (three-letter code: COA) (formula: $C_{21}H_{36}N_7O_{16}P_3S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	S	
			48	21	7	16	3	1	

- Molecule 6 is GLYOXYLIC ACID (three-letter code: GLV) (formula: $C_2H_2O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O		
			5	2	3	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	348	Total 348	O 348	0	0
7	B	291	Total 291	O 291	0	0

L314	L315	R316	D317	R318	Q327	F328	T329	R333	S334	L335	R336	R339	N340	V341	L344	R345	T346	N347	D348	V351	D352	T353	D354	E357	D364	A372	L373	R374	G375	L376	R377	A378	N382	T386	N387	S388	R389	L393	R397	A404	L412	V416	E417	D418
Q424	N425				E434	E435	R436	T439	V440	K443		D451	F455	T456	N457	T458		D462	R463	H469	M472	V478	K484	R508	A509	Q510	K513	G514	M515	W516	E520	D524	P533	A537	S538	V542	P543	A547	L550	H556	Q557			
V558	D559	V560	Q565	R571	T574	L575	E576	Q577	P582	L583	A584	K585	E586	L587	K588	E589	D592	R595	V598	D599	S604	G607	V610	R611	A612	V613	D614	V622	P623	D627	V628	D633	L637	L643	R646	L653	T654	D657	V658	R659				
R664	P667	L668	R671	Q672	V677	A683	P684	D687	D688	A691	I699	Q705	F706	N707	G708	Y709	T710	E720	R724	A725	A726	E727	LYS	PRO	ALA	PRO	SER	ASP	ARG	ALA	GLY	ASP	ASP	ALA	ALA	ARG								

4 Data and refinement statistics

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

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	120.98Å 120.98Å 232.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.36 – 2.70	Depositor
% Data completeness (in resolution range)	94.3 (29.36-2.70)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.190 , 0.287	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	11696	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

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: COA, MG, GLV, SO4, MLT

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	0.42	0/5591	0.79	27/7606 (0.4%)
1	B	0.42	1/5578 (0.0%)	0.81	29/7586 (0.4%)
All	All	0.42	1/11169 (0.0%)	0.80	56/15192 (0.4%)

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	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	588	ALA	C-N	-6.08	1.20	1.34

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	589	TRP	O-C-N	-8.54	109.03	122.70
1	A	280	ASP	CB-CG-OD2	7.03	124.63	118.30
1	B	348	ASP	CB-CG-OD2	6.85	124.47	118.30
1	A	298	ASP	CB-CG-OD2	6.71	124.34	118.30
1	A	206	ASP	CB-CG-OD2	6.70	124.33	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	266	ILE	Peptide

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	5486	0	5433	80	0
1	B	5474	0	5425	112	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	10	0	0	0	0
3	B	5	0	0	0	0
4	A	18	0	7	3	0
4	B	9	0	4	0	0
5	A	48	0	32	4	0
6	B	5	0	1	0	0
7	A	348	0	0	11	0
7	B	291	0	0	11	0
All	All	11696	0	10902	193	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 9.

The worst 5 of 193 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:571:ARG:HG2	1:B:571:ARG:HH11	1.22	1.00
1:A:206:ASP:HB3	7:A:1138:HOH:O	1.62	0.97
1:B:425:ASN:O	1:B:428:LYS:HE2	1.65	0.96
1:B:533:PRO:HB2	1:B:558:VAL:HG11	1.47	0.96
1:A:340:ASN:HD21	1:A:368:THR:HG21	1.31	0.93

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	714/741 (96%)	667 (93%)	43 (6%)	4 (1%)	30	59
1	B	709/741 (96%)	656 (92%)	45 (6%)	8 (1%)	17	42
All	All	1423/1482 (96%)	1323 (93%)	88 (6%)	12 (1%)	24	51

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	73	ILE
1	B	585	LYS
1	B	589	TRP
1	B	267	THR
1	B	726	ALA

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	577/594 (97%)	515 (89%)	62 (11%)	8	19
1	B	577/594 (97%)	497 (86%)	80 (14%)	4	10
All	All	1154/1188 (97%)	1012 (88%)	142 (12%)	6	14

5 of 142 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	50	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	162	LYS
1	B	627	ASP
1	B	69	ARG
1	B	101	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	707	ASN
1	B	63	GLN
1	B	510	GLN
1	B	21	ASN
1	B	116	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 2 are monoatomic - leaving 8 for Mogul analysis.

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
3	SO4	A	1800	-	4,4,4	0.20	0	6,6,6	0.27	0
3	SO4	A	1802	-	4,4,4	0.18	0	6,6,6	0.09	0
4	MLT	A	901	2	1,8,8	0.50	0	2,10,10	1.34	0
4	MLT	A	902	-	1,8,8	1.00	0	2,10,10	1.99	1 (50%)
5	COA	A	903	-	40,50,50	1.75	3 (7%)	50,75,75	2.00	5 (10%)
3	SO4	B	1803	-	4,4,4	0.18	0	6,6,6	0.19	0
6	GLV	B	1901	2	1,4,4	6.82	1 (100%)	0,4,4	0.00	-
4	MLT	B	1902	-	1,8,8	0.28	0	2,10,10	0.33	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
3	SO4	A	1800	-	-	0/0/0/0	0/0/0/0
3	SO4	A	1802	-	-	0/0/0/0	0/0/0/0
4	MLT	A	901	2	1/1/3/3	0/2/8/8	0/0/0/0
4	MLT	A	902	-	1/1/3/3	0/2/8/8	0/0/0/0
5	COA	A	903	-	-	1/44/64/64	0/3/3/3
3	SO4	B	1803	-	-	0/0/0/0	0/0/0/0
6	GLV	B	1901	2	-	0/0/2/2	0/0/0/0
4	MLT	B	1902	-	1/1/3/3	0/2/8/8	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	903	COA	C2A-N1A	2.75	1.39	1.33
5	A	903	COA	C2A-N3A	3.73	1.38	1.32
6	B	1901	GLV	O1-C1	6.82	1.43	1.21
5	A	903	COA	O9P-C9P	9.46	1.41	1.23

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	903	COA	N3A-C2A-N1A	-11.39	120.17	128.89
5	A	903	COA	P2A-O3A-P1A	-4.12	121.16	132.73
5	A	903	COA	C4A-C5A-N7A	-2.45	107.22	109.48
4	A	902	MLT	C3-C2-C1	-2.32	107.86	111.19
5	A	903	COA	C1B-N9A-C4A	-2.05	123.85	126.94

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	902	MLT	C2
4	B	1902	MLT	C2
4	A	901	MLT	C2

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	903	COA	CAP-C9P-N8P-C7P

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	901	MLT	3	0
5	A	903	COA	4	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](#)

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.