



Full wwPDB X-ray Structure Validation Report i

Apr 28, 2025 – 01:02 PM EDT

PDB ID : 2PWZ / pdb_00002pwz
Title : Crystal structure of the apo form of E.Coli malate dehydrogenase
Authors : Soderberg, C.A.G.; Clarke, T.A.; Richardson, D.J.; Hemmings, A.M.
Deposited on : 2007-05-14
Resolution : 2.20 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
A user guide is available at
<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>
with specific help available everywhere you see the i 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](#) ①) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0rc1
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.006 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.43.1

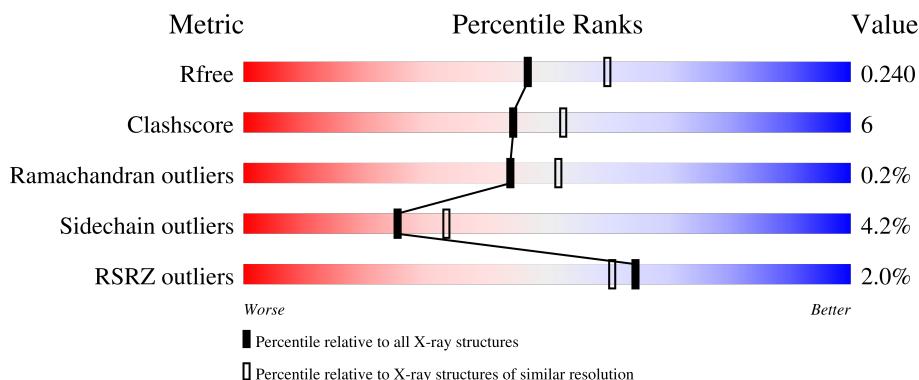
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

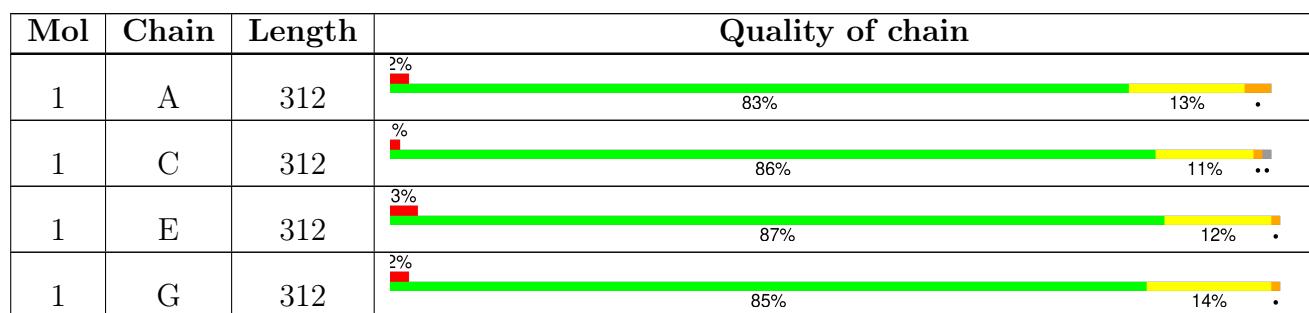
The reported resolution of this entry is 2.20 Å.

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}	164625	5791 (2.20-2.20)
Clashscore	180529	6634 (2.20-2.20)
Ramachandran outliers	177936	6560 (2.20-2.20)
Sidechain outliers	177891	6561 (2.20-2.20)
RSRZ outliers	164620	5791 (2.20-2.20)

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 ≥ 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.



2 Entry composition [\(i\)](#)

There are 2 unique types of molecules in this entry. The entry contains 9470 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 Malate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	312	Total 2278	C 1442	N 387	O 442	S 7	0	1	0
1	C	308	Total 2250	C 1424	N 382	O 438	S 6	0	1	0
1	E	312	Total 2272	C 1438	N 386	O 441	S 7	0	0	0
1	G	312	Total 2272	C 1438	N 386	O 441	S 7	0	0	0

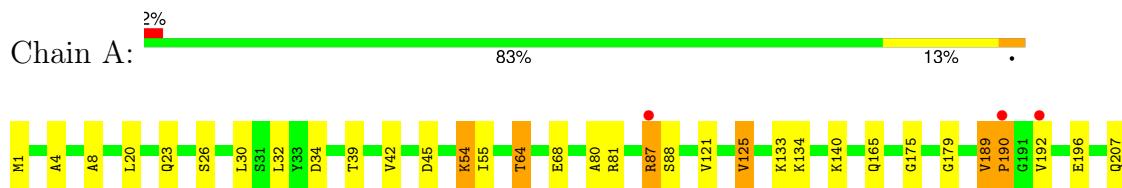
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	114	Total 114 O 114 114	0	0
2	C	104	Total 104 O 104 104	0	0
2	E	96	Total 96 O 96 96	0	0
2	G	84	Total 84 O 84 84	0	0

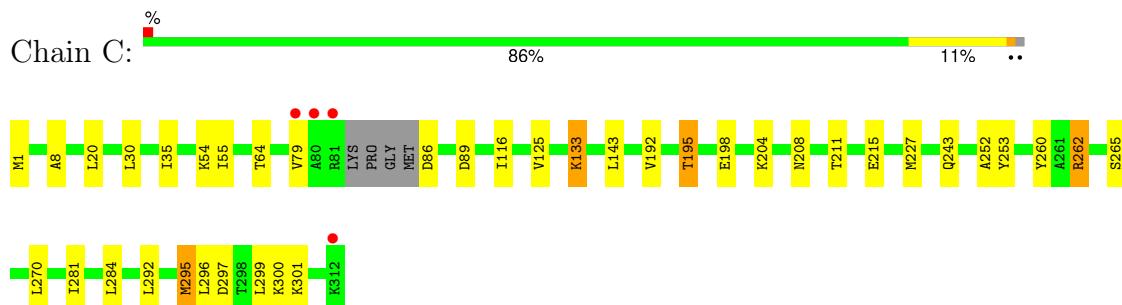
3 Residue-property plots [\(i\)](#)

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.

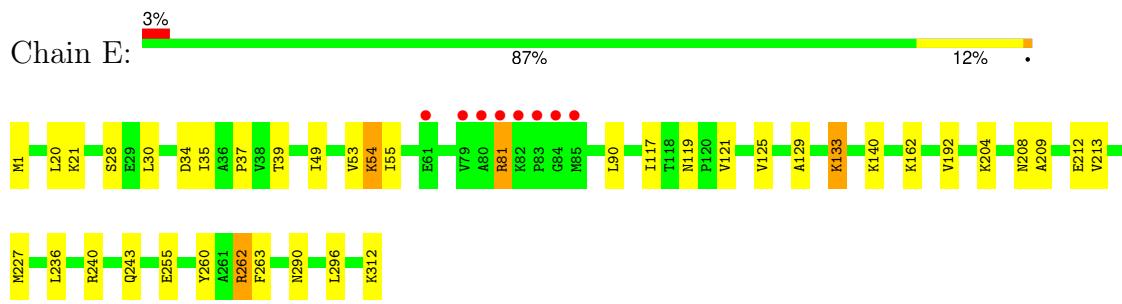
- Molecule 1: Malate dehydrogenase

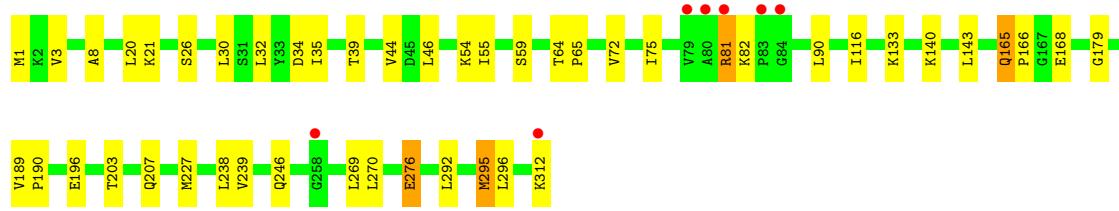


- Molecule 1: Malate dehydrogenase



- Molecule 1: Malate dehydrogenase





4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	146.47 Å 52.48 Å 170.20 Å 90.00° 102.28° 90.00°	Depositor
Resolution (Å)	50.00 – 2.20 50.00 – 2.20	Depositor EDS
% Data completeness (in resolution range)	97.6 (50.00-2.20) 97.6 (50.00-2.20)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$ ¹	2.32 (at 2.20 Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R , R_{free}	0.199 , 0.246 0.195 , 0.240	Depositor DCC
R_{free} test set	3211 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	21.6	Xtriage
Anisotropy	0.599	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 33.1	EDS
L-test for twinning ²	$< L > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	9470	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.85% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	242/241 (100%)	233 (96%)	9 (4%)	29 39
1	C	239/241 (99%)	230 (96%)	9 (4%)	28 37
1	E	241/241 (100%)	232 (96%)	9 (4%)	29 39
1	G	241/241 (100%)	228 (95%)	13 (5%)	18 23
All	All	963/964 (100%)	923 (96%)	40 (4%)	25 33

All (40) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	20	LEU
1	A	54	LYS
1	A	87	ARG
1	A	125	VAL
1	A	133	LYS
1	A	196	GLU
1	A	270	LEU
1	A	295	MET
1	C	20	LEU
1	C	35	ILE
1	C	79	VAL
1	C	133	LYS
1	C	192	VAL
1	C	195	THR
1	C	262	ARG
1	C	270	LEU
1	C	295	MET
1	E	20	LEU
1	E	35	ILE
1	E	54	LYS
1	E	81	ARG
1	E	90	LEU
1	E	125	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	133	LYS
1	E	262	ARG
1	E	290	ASN
1	G	20	LEU
1	G	21	LYS
1	G	35	ILE
1	G	54	LYS
1	G	81	ARG
1	G	82	LYS
1	G	90	LEU
1	G	133	LYS
1	G	238	LEU
1	G	270	LEU
1	G	276	GLU
1	G	295	MET
1	G	312	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	100	ASN
1	A	243	GLN
1	A	311	ASN
1	C	92	ASN
1	C	103	GLN
1	C	208	ASN
1	C	246	GLN
1	C	311	ASN
1	E	100	ASN
1	E	119	ASN
1	E	208	ASN
1	E	311	ASN
1	G	23	GLN
1	G	100	ASN
1	G	103	GLN
1	G	243	GLN
1	G	311	ASN

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 oligosaccharides 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 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, 95th 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 ²)	Q<0.9
1	A	312/312 (100%)	-0.41	6 (1%) 66 62	8, 16, 31, 47	1 (0%)
1	C	308/312 (98%)	-0.40	4 (1%) 74 71	6, 16, 31, 62	1 (0%)
1	E	312/312 (100%)	-0.30	8 (2%) 57 54	9, 19, 35, 62	0
1	G	312/312 (100%)	-0.15	7 (2%) 62 58	10, 21, 39, 62	0
All	All	1244/1248 (99%)	-0.31	25 (2%) 64 61	6, 18, 36, 62	2 (0%)

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	80	ALA	7.0
1	E	83	PRO	5.5
1	C	79	VAL	4.0
1	G	83	PRO	3.5
1	G	80	ALA	3.4
1	E	84	GLY	3.3
1	E	81	ARG	3.2
1	A	192	VAL	3.1
1	A	190	PRO	2.9
1	C	312	LYS	2.9
1	A	87	ARG	2.8
1	A	312	LYS	2.8
1	C	81	ARG	2.7
1	E	80	ALA	2.7
1	G	79	VAL	2.7
1	G	81	ARG	2.6
1	G	258	GLY	2.6
1	G	312	LYS	2.4
1	E	85	MET	2.2
1	E	61	GLU	2.1
1	G	84	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	311	ASN	2.1
1	E	82	LYS	2.1
1	A	219	GLY	2.1
1	E	79	VAL	2.0

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 [\(i\)](#)

There are no ligands in this entry.

6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.