



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 16, 2026 – 10:20 am BST

PDB ID : 30KS / pdb_000030ks
Title : Structure of the Ancestral Haloarchaeal (Anc80) Malate dehydrogenase with chloride and potassium ions.
Authors : Girard, E.; Engilberge, S.; Madern, D.
Deposited on : 2026-04-30
Resolution : 1.90 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

<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	:	4-5-2 with Phenix2.0
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

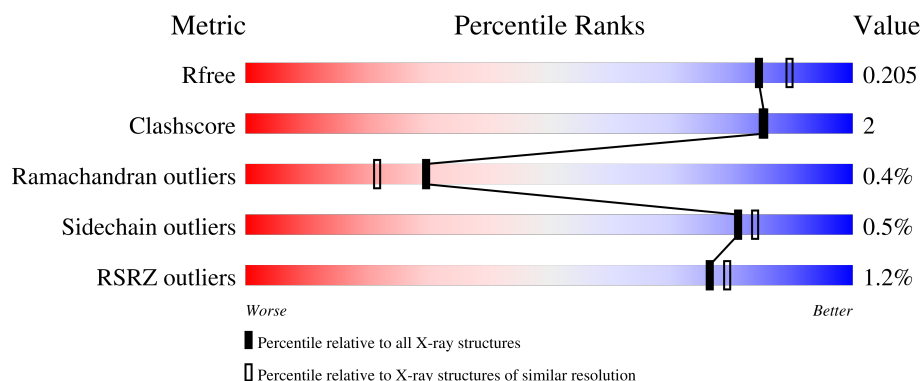
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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.



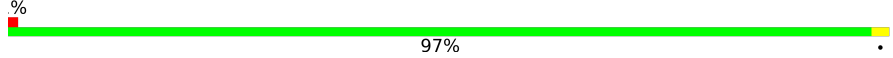
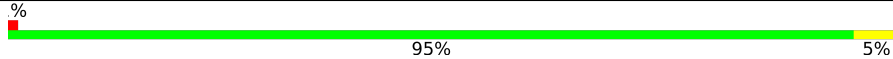
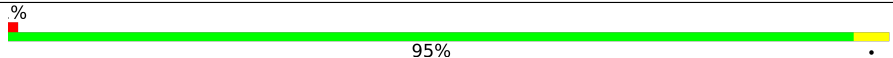
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	7789 (1.90-1.90)
Clashscore	190562	8410 (1.90-1.90)
Ramachandran outliers	187476	8333 (1.90-1.90)
Sidechain outliers	187428	8333 (1.90-1.90)
RSRZ outliers	180081	7790 (1.90-1.90)

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.

Mol	Chain	Length	Quality of chain
1	A	304	95% .
1	B	304	% 95% 5%
1	C	304	2% 94% 6%
1	D	304	2% 94% 6%
1	E	304	95% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	304	 97%
1	G	304	 95% 5%
1	H	304	 95%

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	CL	B	407	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 20327 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	303	Total	C	N	O	S	0	5	0
			2300	1423	387	486	4			
1	B	303	Total	C	N	O	S	0	6	0
			2311	1430	387	489	5			
1	D	303	Total	C	N	O	S	0	2	0
			2292	1417	387	484	4			
1	C	303	Total	C	N	O	S	0	2	0
			2288	1415	387	482	4			
1	E	303	Total	C	N	O	S	0	3	0
			2294	1419	388	483	4			
1	F	303	Total	C	N	O	S	0	1	0
			2283	1410	387	482	4			
1	G	303	Total	C	N	O	S	0	2	0
			2284	1410	387	483	4			
1	H	303	Total	C	N	O	S	0	1	0
			2283	1410	387	482	4			

There are 192 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	21	LEU	ILE	conflict	UNP A0A4D6KBI5
A	32	PHE	TYR	conflict	UNP A0A4D6KBI5
A	61	VAL	ILE	conflict	UNP A0A4D6KBI5
A	65	GLY	THR	conflict	UNP A0A4D6KBI5
A	101	GLU	ASP	conflict	UNP A0A4D6KBI5
A	158	SER	ALA	conflict	UNP A0A4D6KBI5
A	159	GLN	GLU	conflict	UNP A0A4D6KBI5
A	163	ALA	THR	conflict	UNP A0A4D6KBI5
A	164	PRO	GLU	conflict	UNP A0A4D6KBI5
A	189	VAL	ILE	conflict	UNP A0A4D6KBI5
A	192	THR	ALA	conflict	UNP A0A4D6KBI5
A	195	GLU	ASP	conflict	UNP A0A4D6KBI5
A	203	GLU	ALA	conflict	UNP A0A4D6KBI5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	207	ASP	GLU	conflict	UNP A0A4D6KBI5
A	223	GLN	GLU	conflict	UNP A0A4D6KBI5
A	239	HIS	ARG	conflict	UNP A0A4D6KBI5
A	259	ASP	GLY	conflict	UNP A0A4D6KBI5
A	271	ASN	ASP	conflict	UNP A0A4D6KBI5
A	290	ASP	ALA	conflict	UNP A0A4D6KBI5
A	291	ASP	GLU	conflict	UNP A0A4D6KBI5
A	293	ALA	SER	conflict	UNP A0A4D6KBI5
A	297	SER	ALA	conflict	UNP A0A4D6KBI5
A	298	ASP	GLU	conflict	UNP A0A4D6KBI5
A	304	ALA	-	expression tag	UNP A0A4D6KBI5
B	21	LEU	ILE	conflict	UNP A0A4D6KBI5
B	32	PHE	TYR	conflict	UNP A0A4D6KBI5
B	61	VAL	ILE	conflict	UNP A0A4D6KBI5
B	65	GLY	THR	conflict	UNP A0A4D6KBI5
B	101	GLU	ASP	conflict	UNP A0A4D6KBI5
B	158	SER	ALA	conflict	UNP A0A4D6KBI5
B	159	GLN	GLU	conflict	UNP A0A4D6KBI5
B	163	ALA	THR	conflict	UNP A0A4D6KBI5
B	164	PRO	GLU	conflict	UNP A0A4D6KBI5
B	189	VAL	ILE	conflict	UNP A0A4D6KBI5
B	192	THR	ALA	conflict	UNP A0A4D6KBI5
B	195	GLU	ASP	conflict	UNP A0A4D6KBI5
B	203	GLU	ALA	conflict	UNP A0A4D6KBI5
B	207	ASP	GLU	conflict	UNP A0A4D6KBI5
B	223	GLN	GLU	conflict	UNP A0A4D6KBI5
B	239	HIS	ARG	conflict	UNP A0A4D6KBI5
B	259	ASP	GLY	conflict	UNP A0A4D6KBI5
B	271	ASN	ASP	conflict	UNP A0A4D6KBI5
B	290	ASP	ALA	conflict	UNP A0A4D6KBI5
B	291	ASP	GLU	conflict	UNP A0A4D6KBI5
B	293	ALA	SER	conflict	UNP A0A4D6KBI5
B	297	SER	ALA	conflict	UNP A0A4D6KBI5
B	298	ASP	GLU	conflict	UNP A0A4D6KBI5
B	304	ALA	-	expression tag	UNP A0A4D6KBI5
D	21	LEU	ILE	conflict	UNP A0A4D6KBI5
D	32	PHE	TYR	conflict	UNP A0A4D6KBI5
D	61	VAL	ILE	conflict	UNP A0A4D6KBI5
D	65	GLY	THR	conflict	UNP A0A4D6KBI5
D	101	GLU	ASP	conflict	UNP A0A4D6KBI5
D	158	SER	ALA	conflict	UNP A0A4D6KBI5
D	159	GLN	GLU	conflict	UNP A0A4D6KBI5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	163	ALA	THR	conflict	UNP A0A4D6KBI5
D	164	PRO	GLU	conflict	UNP A0A4D6KBI5
D	189	VAL	ILE	conflict	UNP A0A4D6KBI5
D	192	THR	ALA	conflict	UNP A0A4D6KBI5
D	195	GLU	ASP	conflict	UNP A0A4D6KBI5
D	203	GLU	ALA	conflict	UNP A0A4D6KBI5
D	207	ASP	GLU	conflict	UNP A0A4D6KBI5
D	223	GLN	GLU	conflict	UNP A0A4D6KBI5
D	239	HIS	ARG	conflict	UNP A0A4D6KBI5
D	259	ASP	GLY	conflict	UNP A0A4D6KBI5
D	271	ASN	ASP	conflict	UNP A0A4D6KBI5
D	290	ASP	ALA	conflict	UNP A0A4D6KBI5
D	291	ASP	GLU	conflict	UNP A0A4D6KBI5
D	293	ALA	SER	conflict	UNP A0A4D6KBI5
D	297	SER	ALA	conflict	UNP A0A4D6KBI5
D	298	ASP	GLU	conflict	UNP A0A4D6KBI5
D	304	ALA	-	expression tag	UNP A0A4D6KBI5
C	21	LEU	ILE	conflict	UNP A0A4D6KBI5
C	32	PHE	TYR	conflict	UNP A0A4D6KBI5
C	61	VAL	ILE	conflict	UNP A0A4D6KBI5
C	65	GLY	THR	conflict	UNP A0A4D6KBI5
C	101	GLU	ASP	conflict	UNP A0A4D6KBI5
C	158	SER	ALA	conflict	UNP A0A4D6KBI5
C	159	GLN	GLU	conflict	UNP A0A4D6KBI5
C	163	ALA	THR	conflict	UNP A0A4D6KBI5
C	164	PRO	GLU	conflict	UNP A0A4D6KBI5
C	189	VAL	ILE	conflict	UNP A0A4D6KBI5
C	192	THR	ALA	conflict	UNP A0A4D6KBI5
C	195	GLU	ASP	conflict	UNP A0A4D6KBI5
C	203	GLU	ALA	conflict	UNP A0A4D6KBI5
C	207	ASP	GLU	conflict	UNP A0A4D6KBI5
C	223	GLN	GLU	conflict	UNP A0A4D6KBI5
C	239	HIS	ARG	conflict	UNP A0A4D6KBI5
C	259	ASP	GLY	conflict	UNP A0A4D6KBI5
C	271	ASN	ASP	conflict	UNP A0A4D6KBI5
C	290	ASP	ALA	conflict	UNP A0A4D6KBI5
C	291	ASP	GLU	conflict	UNP A0A4D6KBI5
C	293	ALA	SER	conflict	UNP A0A4D6KBI5
C	297	SER	ALA	conflict	UNP A0A4D6KBI5
C	298	ASP	GLU	conflict	UNP A0A4D6KBI5
C	304	ALA	-	expression tag	UNP A0A4D6KBI5
E	21	LEU	ILE	conflict	UNP A0A4D6KBI5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	32	PHE	TYR	conflict	UNP A0A4D6KBI5
E	61	VAL	ILE	conflict	UNP A0A4D6KBI5
E	65	GLY	THR	conflict	UNP A0A4D6KBI5
E	101	GLU	ASP	conflict	UNP A0A4D6KBI5
E	158	SER	ALA	conflict	UNP A0A4D6KBI5
E	159	GLN	GLU	conflict	UNP A0A4D6KBI5
E	163	ALA	THR	conflict	UNP A0A4D6KBI5
E	164	PRO	GLU	conflict	UNP A0A4D6KBI5
E	189	VAL	ILE	conflict	UNP A0A4D6KBI5
E	192	THR	ALA	conflict	UNP A0A4D6KBI5
E	195	GLU	ASP	conflict	UNP A0A4D6KBI5
E	203	GLU	ALA	conflict	UNP A0A4D6KBI5
E	207	ASP	GLU	conflict	UNP A0A4D6KBI5
E	223	GLN	GLU	conflict	UNP A0A4D6KBI5
E	239	HIS	ARG	conflict	UNP A0A4D6KBI5
E	259	ASP	GLY	conflict	UNP A0A4D6KBI5
E	271	ASN	ASP	conflict	UNP A0A4D6KBI5
E	290	ASP	ALA	conflict	UNP A0A4D6KBI5
E	291	ASP	GLU	conflict	UNP A0A4D6KBI5
E	293	ALA	SER	conflict	UNP A0A4D6KBI5
E	297	SER	ALA	conflict	UNP A0A4D6KBI5
E	298	ASP	GLU	conflict	UNP A0A4D6KBI5
E	304	ALA	-	expression tag	UNP A0A4D6KBI5
F	21	LEU	ILE	conflict	UNP A0A4D6KBI5
F	32	PHE	TYR	conflict	UNP A0A4D6KBI5
F	61	VAL	ILE	conflict	UNP A0A4D6KBI5
F	65	GLY	THR	conflict	UNP A0A4D6KBI5
F	101	GLU	ASP	conflict	UNP A0A4D6KBI5
F	158	SER	ALA	conflict	UNP A0A4D6KBI5
F	159	GLN	GLU	conflict	UNP A0A4D6KBI5
F	163	ALA	THR	conflict	UNP A0A4D6KBI5
F	164	PRO	GLU	conflict	UNP A0A4D6KBI5
F	189	VAL	ILE	conflict	UNP A0A4D6KBI5
F	192	THR	ALA	conflict	UNP A0A4D6KBI5
F	195	GLU	ASP	conflict	UNP A0A4D6KBI5
F	203	GLU	ALA	conflict	UNP A0A4D6KBI5
F	207	ASP	GLU	conflict	UNP A0A4D6KBI5
F	223	GLN	GLU	conflict	UNP A0A4D6KBI5
F	239	HIS	ARG	conflict	UNP A0A4D6KBI5
F	259	ASP	GLY	conflict	UNP A0A4D6KBI5
F	271	ASN	ASP	conflict	UNP A0A4D6KBI5
F	290	ASP	ALA	conflict	UNP A0A4D6KBI5

Continued on next page...

Continued from previous page...

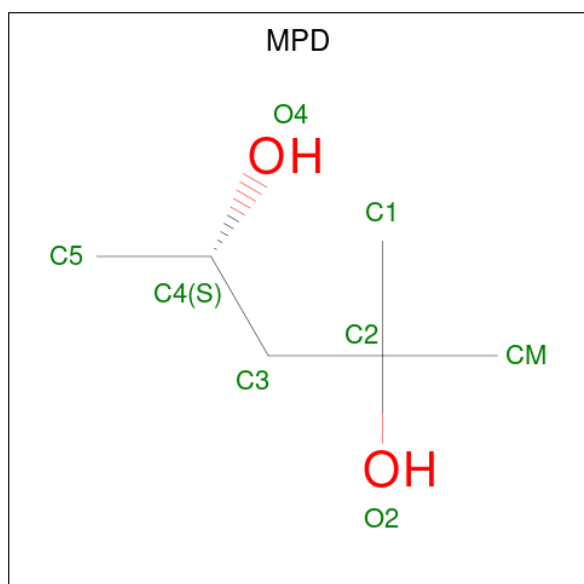
Chain	Residue	Modelled	Actual	Comment	Reference
F	291	ASP	GLU	conflict	UNP A0A4D6KBI5
F	293	ALA	SER	conflict	UNP A0A4D6KBI5
F	297	SER	ALA	conflict	UNP A0A4D6KBI5
F	298	ASP	GLU	conflict	UNP A0A4D6KBI5
F	304	ALA	-	expression tag	UNP A0A4D6KBI5
G	21	LEU	ILE	conflict	UNP A0A4D6KBI5
G	32	PHE	TYR	conflict	UNP A0A4D6KBI5
G	61	VAL	ILE	conflict	UNP A0A4D6KBI5
G	65	GLY	THR	conflict	UNP A0A4D6KBI5
G	101	GLU	ASP	conflict	UNP A0A4D6KBI5
G	158	SER	ALA	conflict	UNP A0A4D6KBI5
G	159	GLN	GLU	conflict	UNP A0A4D6KBI5
G	163	ALA	THR	conflict	UNP A0A4D6KBI5
G	164	PRO	GLU	conflict	UNP A0A4D6KBI5
G	189	VAL	ILE	conflict	UNP A0A4D6KBI5
G	192	THR	ALA	conflict	UNP A0A4D6KBI5
G	195	GLU	ASP	conflict	UNP A0A4D6KBI5
G	203	GLU	ALA	conflict	UNP A0A4D6KBI5
G	207	ASP	GLU	conflict	UNP A0A4D6KBI5
G	223	GLN	GLU	conflict	UNP A0A4D6KBI5
G	239	HIS	ARG	conflict	UNP A0A4D6KBI5
G	259	ASP	GLY	conflict	UNP A0A4D6KBI5
G	271	ASN	ASP	conflict	UNP A0A4D6KBI5
G	290	ASP	ALA	conflict	UNP A0A4D6KBI5
G	291	ASP	GLU	conflict	UNP A0A4D6KBI5
G	293	ALA	SER	conflict	UNP A0A4D6KBI5
G	297	SER	ALA	conflict	UNP A0A4D6KBI5
G	298	ASP	GLU	conflict	UNP A0A4D6KBI5
G	304	ALA	-	expression tag	UNP A0A4D6KBI5
H	21	LEU	ILE	conflict	UNP A0A4D6KBI5
H	32	PHE	TYR	conflict	UNP A0A4D6KBI5
H	61	VAL	ILE	conflict	UNP A0A4D6KBI5
H	65	GLY	THR	conflict	UNP A0A4D6KBI5
H	101	GLU	ASP	conflict	UNP A0A4D6KBI5
H	158	SER	ALA	conflict	UNP A0A4D6KBI5
H	159	GLN	GLU	conflict	UNP A0A4D6KBI5
H	163	ALA	THR	conflict	UNP A0A4D6KBI5
H	164	PRO	GLU	conflict	UNP A0A4D6KBI5
H	189	VAL	ILE	conflict	UNP A0A4D6KBI5
H	192	THR	ALA	conflict	UNP A0A4D6KBI5
H	195	GLU	ASP	conflict	UNP A0A4D6KBI5
H	203	GLU	ALA	conflict	UNP A0A4D6KBI5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	207	ASP	GLU	conflict	UNP A0A4D6KBI5
H	223	GLN	GLU	conflict	UNP A0A4D6KBI5
H	239	HIS	ARG	conflict	UNP A0A4D6KBI5
H	259	ASP	GLY	conflict	UNP A0A4D6KBI5
H	271	ASN	ASP	conflict	UNP A0A4D6KBI5
H	290	ASP	ALA	conflict	UNP A0A4D6KBI5
H	291	ASP	GLU	conflict	UNP A0A4D6KBI5
H	293	ALA	SER	conflict	UNP A0A4D6KBI5
H	297	SER	ALA	conflict	UNP A0A4D6KBI5
H	298	ASP	GLU	conflict	UNP A0A4D6KBI5
H	304	ALA	-	expression tag	UNP A0A4D6KBI5

- Molecule 2 is (4S)-2-METHYL-2,4-PENTANEDIOL (CCD ID: MPD) (formula: C₆H₁₄O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 8 6 2	0	0
2	A	1	Total C O 8 6 2	0	0
2	D	1	Total C O 8 6 2	0	0
2	E	1	Total C O 8 6 2	0	0
2	F	1	Total C O 8 6 2	0	0
2	H	1	Total C O 8 6 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	H	1	Total	C	O	0	0
			8	6	2		

- Molecule 3 is POTASSIUM ION (CCD ID: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	3	Total	K	0	0
			3	3		
3	B	3	Total	K	0	0
			3	3		
3	D	3	Total	K	0	0
			3	3		
3	C	3	Total	K	0	0
			3	3		
3	E	3	Total	K	0	0
			3	3		
3	F	3	Total	K	0	0
			3	3		
3	G	3	Total	K	0	0
			3	3		
3	H	4	Total	K	0	0
			4	4		

- Molecule 4 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total	Cl	0	0
			3	3		
4	B	4	Total	Cl	0	0
			4	4		
4	D	3	Total	Cl	0	0
			3	3		
4	C	3	Total	Cl	0	0
			3	3		
4	E	1	Total	Cl	0	0
			1	1		
4	F	1	Total	Cl	0	0
			1	1		
4	G	2	Total	Cl	0	0
			2	2		
4	H	2	Total	Cl	0	0
			2	2		

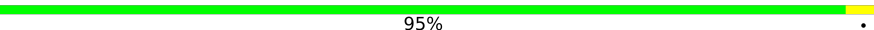
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	303	Total 303	O 303	0	0
5	B	257	Total 257	O 257	0	0
5	D	293	Total 293	O 293	0	0
5	C	194	Total 194	O 194	0	0
5	E	241	Total 241	O 241	0	0
5	F	176	Total 176	O 176	0	0
5	G	191	Total 191	O 191	0	0
5	H	237	Total 237	O 237	0	0

3 Residue-property plots [i](#)

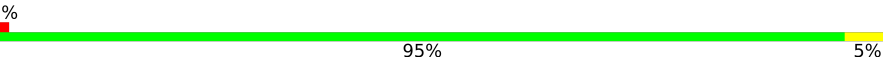
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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

Chain A:  95% .

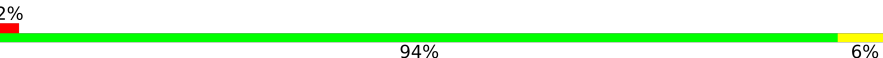


- Molecule 1: Malate dehydrogenase

Chain B:  95% 5%

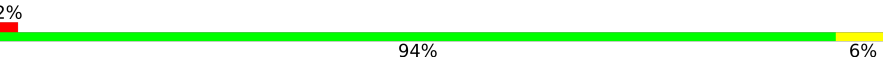


- Molecule 1: Malate dehydrogenase

Chain D:  94% 6%

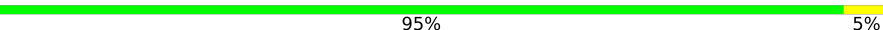


- Molecule 1: Malate dehydrogenase

Chain C:  94% 6%



- Molecule 1: Malate dehydrogenase

Chain E:  95% 5%



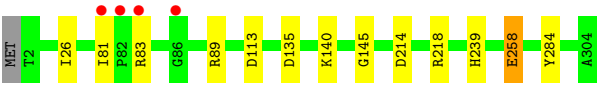
- Molecule 1: Malate dehydrogenase



• Molecule 1: Malate dehydrogenase



• Molecule 1: Malate dehydrogenase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	80.87Å 168.29Å 93.24Å 90.00° 95.20° 90.00°	Depositor
Resolution (Å)	47.91 – 1.90 47.91 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.1 (47.91-1.90) 98.4 (47.91-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.33 (at 1.90Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.169 , 0.204 0.170 , 0.205	Depositor DCC
R_{free} test set	9589 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	25.6	Xtriage
Anisotropy	0.660	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 38.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	20327	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.12% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: CL, K, MPD

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	0.18	0/2352	0.39	0/3203
1	B	0.17	0/2366	0.37	0/3221
1	C	0.14	0/2331	0.35	0/3174
1	D	0.16	0/2335	0.39	0/3179
1	E	0.15	0/2340	0.35	0/3187
1	F	0.14	0/2323	0.33	0/3163
1	G	0.14	0/2326	0.34	0/3167
1	H	0.15	0/2323	0.34	0/3163
All	All	0.15	0/18696	0.36	0/25457

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2300	0	2184	8	0
1	B	2311	0	2196	12	0
1	C	2288	0	2168	13	0
1	D	2292	0	2172	14	0
1	E	2294	0	2176	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	2283	0	2157	2	0
1	G	2284	0	2163	8	0
1	H	2283	0	2158	7	0
2	A	16	0	28	0	0
2	D	8	0	14	0	0
2	E	8	0	14	2	0
2	F	8	0	14	1	0
2	H	16	0	28	2	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0
3	C	3	0	0	0	0
3	D	3	0	0	0	0
3	E	3	0	0	0	0
3	F	3	0	0	0	0
3	G	3	0	0	0	0
3	H	4	0	0	0	0
4	A	3	0	0	0	0
4	B	4	0	0	1	0
4	C	3	0	0	0	0
4	D	3	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	G	2	0	0	0	0
4	H	2	0	0	0	0
5	A	303	0	0	3	0
5	B	257	0	0	8	0
5	C	194	0	0	4	0
5	D	293	0	0	7	0
5	E	241	0	0	7	0
5	F	176	0	0	1	0
5	G	191	0	0	1	0
5	H	237	0	0	4	0
All	All	20327	0	17472	79	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 2.

The worst 5 of 79 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:D:289:MET:SD	5:D:778:HOH:O	2.14	1.05

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:297:SER:O	5:E:501:HOH:O	1.77	1.00
1:B:289:MET:SD	5:B:749:HOH:O	2.20	0.99
1:C:83:ARG:NH2	5:C:501:HOH:O	2.04	0.91
4:B:407:CL:CL	5:B:757:HOH:O	2.31	0.86

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	306/304 (101%)	300 (98%)	5 (2%)	1 (0%)	36	29
1	B	307/304 (101%)	298 (97%)	7 (2%)	2 (1%)	18	10
1	C	303/304 (100%)	292 (96%)	10 (3%)	1 (0%)	36	29
1	D	303/304 (100%)	297 (98%)	5 (2%)	1 (0%)	36	29
1	E	304/304 (100%)	297 (98%)	6 (2%)	1 (0%)	36	29
1	F	302/304 (99%)	293 (97%)	8 (3%)	1 (0%)	36	29
1	G	303/304 (100%)	293 (97%)	9 (3%)	1 (0%)	36	29
1	H	302/304 (99%)	296 (98%)	5 (2%)	1 (0%)	36	29
All	All	2430/2432 (100%)	2366 (97%)	55 (2%)	9 (0%)	30	22

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	113	ASP
1	G	145	GLY
1	B	145	GLY
1	C	145	GLY
1	E	145	GLY

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	246/243 (101%)	245 (100%)	1 (0%)	84	87
1	B	248/243 (102%)	246 (99%)	2 (1%)	73	75
1	C	243/243 (100%)	243 (100%)	0	100	100
1	D	244/243 (100%)	244 (100%)	0	100	100
1	E	244/243 (100%)	244 (100%)	0	100	100
1	F	242/243 (100%)	239 (99%)	3 (1%)	63	63
1	G	243/243 (100%)	242 (100%)	1 (0%)	84	87
1	H	242/243 (100%)	239 (99%)	3 (1%)	63	63
All	All	1952/1944 (100%)	1942 (100%)	10 (0%)	81	84

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	135	ASP
1	H	258	GLU
1	H	284	TYR
1	F	113	ASP
1	F	239	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	257	HIS
1	F	180	GLN
1	G	286	GLN
1	G	257	HIS
1	C	299	GLN

5.3.3 RNA ⓘ

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](#)

Of 51 ligands modelled in this entry, 44 are monoatomic - leaving 7 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	$\# Z > 2$
2	MPD	A	401	-	7,7,7	0.29	0	9,10,10	0.51	0
2	MPD	D	401	-	7,7,7	0.25	0	9,10,10	0.28	0
2	MPD	E	401	-	7,7,7	0.26	0	9,10,10	0.73	0
2	MPD	F	401	-	7,7,7	0.30	0	9,10,10	0.45	0
2	MPD	A	402	-	7,7,7	0.29	0	9,10,10	0.38	0
2	MPD	H	401	-	7,7,7	0.26	0	9,10,10	0.55	0
2	MPD	H	402	-	7,7,7	0.27	0	9,10,10	0.35	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	MPD	A	401	-	-	1/5/5/5	-
2	MPD	D	401	-	-	0/5/5/5	-
2	MPD	E	401	-	-	0/5/5/5	-
2	MPD	F	401	-	-	0/5/5/5	-
2	MPD	A	402	-	-	3/5/5/5	-
2	MPD	H	401	-	-	0/5/5/5	-
2	MPD	H	402	-	-	1/5/5/5	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	402	MPD	O2-C2-C3-C4
2	A	402	MPD	C1-C2-C3-C4
2	A	402	MPD	CM-C2-C3-C4
2	A	401	MPD	O2-C2-C3-C4
2	H	402	MPD	O2-C2-C3-C4

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	401	MPD	2	0
2	F	401	MPD	1	0
2	H	401	MPD	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, 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	303/304 (99%)	-0.45	1 (0%) 90 91	16, 25, 38, 52	5 (1%)
1	B	303/304 (99%)	-0.35	3 (0%) 79 82	16, 27, 48, 81	6 (1%)
1	C	303/304 (99%)	-0.10	7 (2%) 61 65	20, 34, 62, 101	2 (0%)
1	D	303/304 (99%)	-0.37	5 (1%) 69 72	18, 26, 46, 67	2 (0%)
1	E	303/304 (99%)	-0.27	1 (0%) 90 91	19, 30, 48, 72	3 (0%)
1	F	303/304 (99%)	0.16	4 (1%) 75 78	20, 40, 61, 84	1 (0%)
1	G	303/304 (99%)	0.18	4 (1%) 75 78	20, 38, 57, 68	2 (0%)
1	H	303/304 (99%)	-0.14	4 (1%) 75 78	19, 31, 52, 82	1 (0%)
All	All	2424/2432 (99%)	-0.17	29 (1%) 76 79	16, 31, 55, 101	22 (0%)

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	304	ALA	5.4
1	H	82	PRO	5.1
1	E	304	ALA	4.4
1	G	112[A]	ASP	4.1
1	C	304	ALA	3.5

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 oligosaccharides in this entry.

6.4 Ligands ⓘ

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, 95th 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(Å ²)	Q<0.9
4	CL	B	407	1/1	0.72	0.47	56,56,56,56	1
3	K	C	403	1/1	0.78	0.23	66,66,66,66	1
4	CL	A	408	1/1	0.81	0.15	58,58,58,58	1
2	MPD	D	401	8/8	0.81	0.14	36,47,53,59	0
3	K	F	404	1/1	0.84	0.14	66,66,66,66	1
4	CL	C	406	1/1	0.85	0.53	52,52,52,52	1
2	MPD	A	402	8/8	0.86	0.13	25,30,35,37	0
2	MPD	H	401	8/8	0.87	0.14	22,38,44,56	0
2	MPD	E	401	8/8	0.87	0.15	33,42,48,48	0
2	MPD	F	401	8/8	0.90	0.12	29,35,41,42	0
2	MPD	A	401	8/8	0.92	0.11	23,32,37,37	0
4	CL	D	407	1/1	0.92	0.50	48,48,48,48	1
3	K	F	402	1/1	0.92	0.12	48,48,48,48	0
3	K	B	403	1/1	0.93	0.12	59,59,59,59	0
3	K	D	403	1/1	0.94	0.12	49,49,49,49	0
4	CL	F	405	1/1	0.94	0.09	36,36,36,36	0
3	K	E	404	1/1	0.95	0.06	54,54,54,54	1
4	CL	B	405	1/1	0.95	0.21	46,46,46,46	0
2	MPD	H	402	8/8	0.95	0.10	22,36,40,41	0
4	CL	D	406	1/1	0.95	0.09	40,40,40,40	1
3	K	E	402	1/1	0.95	0.09	39,39,39,39	0
3	K	G	401	1/1	0.95	0.10	46,46,46,46	0
3	K	H	405	1/1	0.95	0.12	53,53,53,53	0
4	CL	A	407	1/1	0.96	0.07	41,41,41,41	1
4	CL	E	405	1/1	0.96	0.09	30,30,30,30	0
3	K	G	403	1/1	0.96	0.10	62,62,62,62	0
4	CL	G	405	1/1	0.96	0.06	49,49,49,49	1
3	K	H	406	1/1	0.97	0.09	35,35,35,35	0
4	CL	B	404	1/1	0.97	0.06	27,27,27,27	0
3	K	H	404	1/1	0.97	0.07	40,40,40,40	0
4	CL	G	404	1/1	0.97	0.06	29,29,29,29	0
4	CL	C	405	1/1	0.97	0.05	45,45,45,45	1
4	CL	H	407	1/1	0.97	0.06	31,31,31,31	0
3	K	A	405	1/1	0.98	0.06	30,30,30,30	0
3	K	A	403	1/1	0.98	0.03	29,29,29,29	0
3	K	A	404	1/1	0.98	0.05	38,38,38,38	0
3	K	D	404	1/1	0.98	0.05	32,32,32,32	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
3	K	F	403	1/1	0.98	0.04	34,34,34,34	0
3	K	C	401	1/1	0.98	0.06	37,37,37,37	0
4	CL	C	404	1/1	0.98	0.04	28,28,28,28	0
4	CL	H	408	1/1	0.98	0.08	33,33,33,33	0
3	K	E	403	1/1	0.99	0.08	33,33,33,33	0
3	K	G	402	1/1	0.99	0.03	31,31,31,31	0
3	K	D	402	1/1	0.99	0.04	29,29,29,29	0
3	K	H	403	1/1	0.99	0.07	35,35,35,35	0
4	CL	B	406	1/1	0.99	0.03	26,26,26,26	1
3	K	C	402	1/1	0.99	0.05	29,29,29,29	0
4	CL	D	405	1/1	0.99	0.03	20,20,20,20	0
3	K	B	402	1/1	0.99	0.04	36,36,36,36	0
3	K	B	401	1/1	0.99	0.06	26,26,26,26	0
4	CL	A	406	1/1	1.00	0.04	24,24,24,24	0

6.5 Other polymers [i](#)

There are no such residues in this entry.