

Volume 74 (2018)

Supporting information for article:

Conformational changes with substrate binding revealed by structures of *Methylobacterium extorquens* malate dehydrogenase

Javier M. González, Ricardo Marti-Arbona, Julian C.-H. Chen, Brian Broom-Peltz and Clifford J. Unkefer

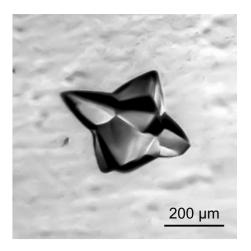


Figure S1 Crystal of *Mex*MDH.

	PDB	UniProt	Organism	n	Protein
	1HYG	Q60176	Methanocaldococcus jannaschii		L-2-HCDH
	3FI9	B2RM04	Porphyromonas gingivalis	2	MDH
	4111	Q4Q7X6	Leishmania major	2	MDH
	4TVO	P9WK13	Mycobacterium tuberculosis	2	MDH
	1CIV	P46489	Flaveria bidentis	2	MDH
	7MDH	P17606	Sorghum bicolor	2	MDH
	3D5T	Q3JKE9	Burkholderia pseudomallei	2	MDH
1 1 1 1 1 1 1 1 1 1	1B8P	Q9ZF99	Aquaspirillum arcticum	2	MDH
	4KDE	P10584	Thermus thermophilus	2	MDH
	4UUL	O96445	Trichomonas vaginalis	2	LDH
65	4000	Q27819	Trichomonas vaginalis	2	c-MDH
	4MDH	P11708	Sus scrofa	2	cm-MDH
	1EMD	P61889	Escherichia coli	2	MDH
99	4E0B	Q8DEC2	Vibrio vulnificus	2	MDH
100	1SEV	P19446	Citrullus lanatus	2	g-MDH
86 - 465	2DFD	P40926	Homo sapiens	2	m-MDH
	3TL2	Q6HSF4	Bacillus anthracis	4	MDH
64	3NEP	Q2S289	Salinibacter ruber	4	MDH
	4BGV	Q6L0C3	Picrophilus torridus	4	MDH
4	1GV0	P80039	Chlorobium tepidum	4	MDH
100	1GUZ	P0C890	Prosthecochloris vibrioformis	4	MDH
69	3P7M	Q8G942	Francisella tularensis	4	MDH
	5ULV	A9W386	Methylobacterium extorquens	4	MDH
99	3GVH	Q2YLR9	Brucella abortus	4	MDH
	4PLT	A0A075B5H0	Apicomplexa	4	MDH
	2A92	Q4PRK9	Plasmodium vivax	4	LDH
	1T2E	Q27743	Plasmodium falciparum	4	LDH
	4PLF	A0A075B5G8	Apicomplexa	4	LDH
4 53	1SOW	Q27797	Toxoplasma gondii	4	LDH
99	1PZH	P90613	Toxoplasma gondii	4	LDH
	2HJR	Q5CYZ3	Cryptosporidium parvum	2	MDH
	2D4A	Q9YEA1	Aeropyrum pernix	4	MDH
	3U95	B9KAM3	Thermotoga neapolitana	2	GH
100	1VJT	Q9WZL1	Thermotoga maritima	2	α-G
	1A5Z	P16115	Thermotoga maritima	4	LDH
	1LLD	E8ME30	Bifidobacterium longum	4	LDH
	3DL2	Q8IX04	Homo sapiens	2	Uq-c
	1V6A	Q9W7K5	Cyprinus carpio	4	LDH
	2V65	O93541	Champsocephalus gunnari	4	LDH
	5K0Z	P00337	Gallus gallus	4	LDH
	5LDH	P00336	Sus scrofa	4	LDH
54	2LDX	P00342	Mus musculus	4	LDH
	40JN	P00338	Homo sapiens	4	LDH
	1LDM	P00341	Squalus acanthias	4	LDH
	1HYH	P14295	Weissella confusa	4	L-2-HICDH
	1Y6J	A3DCA4	Clostridium thermocellum	1	LDH
	3PQE	P13714		4	LDH
83	1LDB	P00344	Geobacillus stearothermophilus	4	LDH
	3H3J	Q5HJD7	Staphylococcus aureus	4	LDH
	1EZ4	P56511	Lactobacillus pentosus	4	LDH
	1LLC	P00343	Lactobacillus casei	1	LDH
		V5XPB8	Enterococcus mundtii	4	LDH
	4Q3N	A0A0B5KUB4		4	LDH
	2E37	Q5SJA1	Thermus thermophilus	4	LDH
		P50933	Deinococcus radiodurans	4	LDH
 95	2\/6R			-r	
95	2V6B 4BGU			Δ	
95	2V6B 4BGU 1D3A	Q9P9L2 Q07841	Haloferax volcanii Haloarcula marismortui	4 4	MDH MDH

Figure S2 Maximum-likelihood phylogenetic tree of all the MDH/LDH-like proteins available in the Protein Data Bank, filtered with a 90 % identity cutoff. The three main clades comprise dimeric malate dehydrogenases (*purple*), tetrameric malate dehydrogenases (*orange*), and tetrameric lactate dehydrogenases (*green*). Note that dimeric LDH enzymes are unusual. Bootstrap node support values are indicated as percent for nodes with \geq 50 % support. The acronyms stand for: malate and lactate dehydrogenase (MDH, LD); L-2-hydroxycarboxylate dehydrogenase (L-2-HCDH); cytosolic-, cytoplasmic-, glyoxysomal-, mitochondrial-MDH (respectively c-, cm-, g-, and m-MDH); glycoside hydrolase (GH); α -glucosidase (α -G); and L-2-hydroxyisocaproate dehydrogenase (L-2-HICDH). The divergent MDH enzymes from *Haloferax volcanii* and *Haloarcula marismortui* were selected as outgroup (*gray*).

M.extorquens_A9W386	1	αl 20000000 10 20	ε.εεε <u>β2</u> 30	<u>000000</u> 40
M.extorquens_A9W386 B.abortus_Q2YLR9 F.tularensis_Q8G942 Apicomplexa_A0A075B5H0 Cr.parvum_Q5CYZ3 P.torridus_Q6L0C3 B.anthracis_Q6HSF4 C.aurantiacus_P80040 S.ruber_Q2S289 C.tepidum_P80039 P.vibrioformis_P0C890 C.parvum_B3QPY9 A.pernix_Q9YEA1 A.fulgidus_008349 H.volcanii_Q9P9L2 H.marismortui_Q07841	MARSKIAI MARNKIAI MARKKITI LKKISYKYNTVIMRKKISI MARSKISI MARSKISI MARSKISI MKISKKVSI MKISI MKITI MKITI MKITI MKITI MKITI	LIG.SGMIGGTLAH LVG.AGNIGGTLAH LIG.AGNIGGTLAH UIG.AGQIGSTIAL VIG.AGAVGATVAQ VIG.AGFTGATTAF VIG.AGFTGATTAF VIG.AGNVGATTAF VIG.AGNVGATTAF VIG.AGNVGATTAF VIG.AGNVGATTAF VIG.AGNVGATTAF	L.LAQKELA.DVVI W.LAAKELG.DIVI C.VARQDVAKEVVM R.LAEKQLARELVI	PDI. VDGVPQGKA PDI. AEGTPQGKG PDI. AQGMPNGKA PDI. IEGVPQGKA PDI. IEGVPQGKA PDI. VEGMPQGKA PDI. VEGVPQGKA PDI. VEGVPQGKA DI. VEGIPQGKA DI. KEGVPQGKA DI. VEGIPQGKA DV. VEGIPQGKG DV. VEGIPQGKG DV. VEGIPQGKG DV. VEGIPQGKA DV. VEGIPQGKA
M.extorquens_A9W386	α2 <u>0000000000</u> 50 ★ 60	β3 η1 ▶ <u>2000</u> 70	β4 вο т:	r <u>000000</u> 90
M.extorquens_A9W386 B.abortus_Q2YLR9 F.tularensis_Q8G942 Apicomplexa_A0A075B5H0 Cr.parvum_Q5CYZ3 P.torridus_Q6L0C3 B.anthracis_Q6HSF4 C.aurantiacus_P80040 S.ruber_Q2S289 C.tepidum_P80039 F.vibrioformis_P0C890 C.parvum_B3QPY9 A.pernix_Q9YEA1 A.fulgidus_008349 H.volcanii_Q9P9L2 H.marismortui_Q07841	LD ILEGAPHWGYDLD IKG LDMLEASPVQGFDANIIG LDLYEASPIEGFDVRVTG LDMRESSPIHGFDTRVTG LDMYESGPVGLFDTKVTG LDMYETGPVGLFDTKITG LDMYETGPVGLFDTKITG LDMYETGPVGLFDTKITG	AND YAAIEG TNDYKDLEN TNNYEDIKG ENNYEYLQN FCTADESKYAEMKG TSDYADTAD TNNYADTAD SNDYADTAD SNDYADTAD SNDYADTAD SNDYADTAD SNSYEDMRG GADYSLLKG GADYSLLKG	A D V V I V T A G V P RK P (S D V V I V T A G V P RK P (S D V V I I T A G I P RK P (S D V V I I T A G I A RK P (S D V V V I T A G I A RK P (S D V V V I T A G I A RK P (S D V V I T A G I P RK P (S D I V I I T A G L P RK P (S D I V I I T A G L P RK P (S D I V I I T A G L P RK P (S D I V V I T A G L P RK P (S D I V V I T A G I P RK P (S D I V V V I T A G I P RK P (S D I V V V I T A G I P R (S D V V V I T A G I P R (S D V V V I T A G I P R (S D V V V I T A G I P R (S D V V V I T A G I P R (P (GMSRDDLLG G.S.MSRDDLLG G.S.MSRDDLLG G.S.SDKEWSRDDLLT GMTRSDLLT GMSRDDLL GMSRDDLLA GMSRDDLLA GMTREDLLS GMTREDLLM GMTREDLLM GMTREDLLM GMTREDLAH GMTRLDLAH GMTRLDLA
M.extorquens_A9W386	α3 2020202020202000 100 110	β5 120 2000	α4 η2 <u>00000000 000</u> 130 140	β6 α5 <u>00000000</u> 15 [±]
M.extorquens_A9W386 B.abortus_Q2YLR9 F.tularensis_Q8G942 Apicomplexa_A0A075B5H0 Cr.parvum_Q5CYZ3 P.torridus_Q6L0C3 B.anthracis_Q6HSF4 C.aurantiacus_P80040 S.ruber_Q2S289 C.tepidum_P80039 P.vibrioformis_P0C890 C.parvum_B3QPY9 A.pernix_Q9YEA1 A.fulgidus_O08349 H.volcanii_Q9P9L2 H.marismortui_Q07841	INIKVME AVGAGIKEHAPI INIKVME QVGAGIKKYAPI INIKVMQ TVGEGIKHNCPI VNAKIMKDVAENIKKYCPI VNAKIMKSVGEAIKKYSPI TNSKIMKSITRDIAKHSPI VNADITRACISQAAPLSPI KNTEIVGGVTEQFVEGSPI KNAGIVKEVTONIMKHSKI KNAGIVKEVTDNIMKHSKI ANANTMADLAEKIKAYAKI KNAGIVKEVTDNIMKHSKI ANANTMADLAEKIKAYAKI	DAFVICITNPLDAM AFVICITNPLDAM NAFVICITNPLDVM NAFVICITNPLDVM NAFVICITNPLDAM ESKIVVVTNPADIM NAIIVVLTNPVDAM NAVIIWVNNPLDAM DSTIIVVANPLDVM NPIIIVVSNPLDIM NPIIIVVSNPLDIM NPIIIVVSNPLDIM DAIVVITNPVDAM CSKILVVTNPMDVM	VWALQKFSGLPTNK VWALQKFSGLPAHK VNMLQKFSGVPDNK VYVLHKYSGLPHNK VYYFKEKSGIPANK AYAIYKASGISPER TYSVFKEAGFPKER TYLAAEVSGFPKER TYVAYEASGFPKER THVAWQKSGLPKER THVAWVRSGLPKER THVAWVRSGLPKER TYVMYKKGFPRER TYIMYKESGKPRNE NRHLYETGDRDRHK	VVGMAGVLDSARFR VVGMAGVLDSARFR VCGMAGVLDSARFR ICGMSGVLDSARFR IGLGGSLDSTRFR IGLGGSLDSTRFR IGLGQAGVLDAARFR VIGQAGVLDAARYR VIGMAGVLDAARFR VIGMAGVLDAARFR VIGMAGVLDAARFR VIGMAGVLDAARFR VIGMAGVLDAARFR VIGMAGVLDAARFR VIGMAGVLDAARFR VIGMAGVLDAARFR

M.extorquens_A9W386 M.extorquens_A9W386 B.abortus_Q2YLR9 F.tularensis_Q8G942 Apicomplexa_A0A075B5H0 Cr.parvum_Q5CYZ3 P.torridus_Q6L0C3 B.anthracis_Q6HSF4 C.aurantiacus_P80040 S.ruber_Q2S289 C.tepidum_P80039 P.vibrioformis_P0C890 C.parvum_B3QPY9 A.pernix_Q9YEA1 A.fulgidus_008349 H.volcanii_Q9P9L2 H.marismortui_Q07841	Y F LSEEFNV TF LADELNV YF LAEKLNV CNLSRALGV TF LAQELNV TF IAQELNV TF IAQELNV SF IAEELDV SF IAEELDV SF IAMELGV SF IAMELGV Y I SQRFDV	SVEDVTVF SVQQVQAY SPNDVQAM KPSDVSAI SFEDVNAF SVEDVQAM SVRDVQAL SNQDVTAC SMQDINAC SSKQDINAC SSFKSVNAI NIRRAW	VLGGHGDI VLGGHGD VIGGHGD VIGGHGDI VIGGHGDI LMGGHGDI LMGGHGDI VLGGHGDI VLGGHGDI VLGGHGDI VLGGHGDI VLGGHGDI IIGEHGDI IIGEHGDI	MVPLTR MVPLTR MVPLTR MVPLTR MVPLTR MVPLTR MVPLVR MVPLVR MVPLPR MVPVVK MVPVVK MVPVVK MVPVVK MFVAKSJ QVPVFS	2 190 X STVAGVP X STVAGVS X CTVGGIP X SNVSGIP X STISGIP X STISGIP X TVAGIP X TTVAGIP X TTVAGIP X TTVAGIP X TTVAGIP X TTVAGIP X TTVAGIP X TTVAGIP X TVAGIP X TVAGIP X TVAGIP X TVAGIP X TVAGIP X TVAGIP X TVAGIP X TVAGIP X TVAGIP	200000 200 LTDLVKLGV LPDLVKMGV LEQLVKEGF LTEFIKQGV LSDFVEQGF IEDLL VSEFI VPQLI ISDLL ISDLL LEHLM PA.	α7 210 210 VITQEKLDAMVER VISQDKLDKIIQR KLKQERLDAIVSR VITQEEIDEIVER CITHSQINEIIKK PREKIDEIVKR PREKIDEIVKR APDRLAQIVER APDRLAQIVER APDRLAQIVER APDRLAVIVER SAERIAELVER PAETIDKLVER PAETIDKLVER SKEEIEEVVSE EVDWEAVEND FSADEKEEILGD FSGDEKEQLLGD
M.extorquens_A9W386 B.abortus_Q2YLR9 F.tularensis_Q8G942 Apicomplexa_A0A075B5H0 Cr.parvum_Q5CYZ3 P.torridus_Q6L0C3 B.anthracis_Q6HSF4 C.aurantiacus_P80040 S.ruber_Q2S289 C.tepidum_P80039 P.vibrioformis_P0C890 C.parvum_B3QPY9 A.pernix_Q9YEA1 A.fulgidus_008349 H.volcanii_Q9P9L2 H.marismortui_Q07841	TRNAGGEIV TAFGGGEIV TRFGGGEIV TRKGGGEIV TRKGGGEIV TRKGAGGEIV TRTGAEIV TRTGGAEIV	230 NLLKTGSA GLLKTGSA ALLKTGSA NLKTGSA CLLKTGSA NLYKTGSA GLLGNGSA NLLKTGSA DLMGT.SA NHLKQGSA EHLKQGSA EHLKQGSA E.LRGYSS K.RKGAT E.RKGAT	F YAP AAS F YAP AAS Y YAP AAS Y YAP AAS F YAP GIS Y YAP AAS Y YAP AAS Y YAP AAA F YAP AAA F YAP AAS F YAP ATS F YSP GSS F YAP AS I F GP AVA QW GP ATG	I TAMAES I QMAES I QMAES I QMAES I LEMAES VVAMAQA AVMVES VVEMVES VEMVES VEMVES VVEMVES VVEMVES VVEMVES VVEMVES VVEMVES VVEMVES VVEMVES	250 (1RDKKRV (1KDKKRV (1KDKKRV (1KDKKRV (1KDCKRV (1KDCKRV (1KDCKRV (1KDCKRV (1KDCKRV (1LCCKRV) (1LCCKRV (1LCCKRV (1LCCKRV (1LCCKRV) (1LCCKRV (1LCCKRV) (1LCCKRV (1LCCKRV)	LPVAAQLS LPCAAKVKA LPCSAYLE LVCSTYLT	GQ YGVK.D AGM YGLDED GQ YGVK.D GQ YNVN.N GEHSKTYLVN.N GE YGLN.D GQ YGLN.D GQ YGLN.D GQ YGLN.D GE YGID.G GQ YGID.G GQ YGID.K GE YGIE.N GE YGIE.N

Figure S3 Sequence alignment of selected MDH enzymes. The indicated alphanumeric codes for each microorganism name indicate the corresponding sequence ID for the UniProtKB database. Graphical representation was prepared with ESPRIPT (available at http://espript.ibcp.fr). See the main text for additional details.

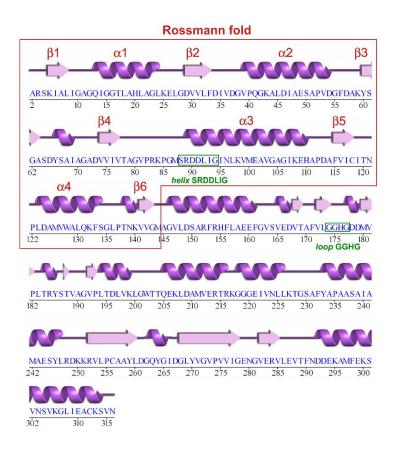


Figure S4 Topological arrangement of *Mex*MDH secondary structure elements. The image was prepared with the structure PDB 5luv (reported in this work), based on the output of PDBSUM (available at http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/).

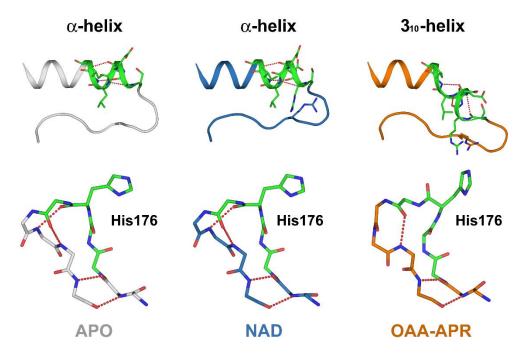


Figure S5 Cartoon representation of the conformational changes associated with NAD binding (*blue*) and OAA-APR binding (*orange*) as compared with the apo-enzyme resting state (*gray*). OAA induces the most important conformational changes (highlighted in *green*) in helix SRDDLIG (experiencing a transition between α -helix and 3_{10} -helix), and loop GGHG.

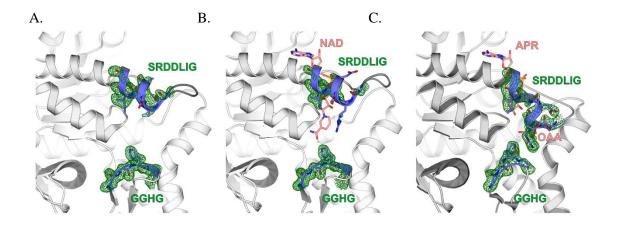


Figure S6 Polder OMIT maps (Liebschner *et al.*, 2017) for loops SRDDLIG and GGHG. (*a*) Apo, (*b*) NAD, and (*c*) OAA-APR *Mex*MDH. Relevant areas are colored in *blue*, with polder maps in *green*, and ligands in *pink*. Polder maps were calculated with Phenix 1.13-2998, contoured at 2.5 r.m.s.d. Figures were prepared with PyMOL 1.8.3.2 (Schrodinger) and Corel Draw X7 (Corel).