Biochemical, structural and dynamical characterizations of the lactate dehydrogenase from Selenomonas ruminantium provide information about an intermediate evolutionary step prior to complete allosteric regulation acquisition in the super family of lactate and malate dehydrogenases

Quentin Bertrand, Sandrine Coquille, Antonio Iorio, Fabio Sterpone, Dominique Madern

Biochemical, structural and dynamical characterizations of the lactate dehydrogenase from *Selenomonas ruminantium* provide information about an intermediate evolutionary step prior to complete allosteric regulation acquisition in the super family of lactate and malate dehydrogenases.

Quentin Bertrand[1, 2,*], Sandrine Coquille[1,*], Antonio Iorio[3], Fabio Sterpone[3] and Dominique Madern[1]

1-Univ. Grenoble Alpes, CEA, CNRS, IBS, 38000 Grenoble, France.

2-Laboratory of Biomolecular Research, Biology and Chemistry Division, Paul Scherrer Institut, Villigen, Switzerland.

3- CNRS, Université de Paris, UPR 9080, Laboratoire de Biochimie Théorique, Paris, France; Institut de Biologie Physico-Chimique-Fondation Edmond de Rothschild, PSL Research University, Paris, France.

* These authors contributed equally.

Corresponding author:

Dr Dominique Madern, Institut de Biologie Structurale, Campus EPN CS 10090, 71 avenue des Martyrs, 38044 Grenoble Cedex 9; France.

Email: dominique.madern @ ibs.fr. Phone: 33- (0) 4 57 42 85 71.

Keywords. Allosteric regulation; lactate dehydrogenase; crystal structure; molecular dynamics; quaternary structure.

Abstract

In this work, we investigated the lactate dehydrogenase (LDH) from *Selenomonas ruminantium* (*S. rum*), an enzyme that differs at key amino acid positions from canonical allosteric LDHs. The wild type (Wt) of this enzyme recognises pyuvate as all LDHs. However, introducing a single point mutation in the active site loop (I85R) allows *S. Rum* LDH to recognize the oxaloacetate substrate as a typical malate dehydrogenase (MalDH), whilst maintaining homotropic activation as an LDH. We report the tertiary structure of the Wt and I85R LDH mutant.

The Wt *S. rum* enzyme structure binds NADH and malonate, whilst also resembling the typical compact R-active state of canonical LDHs. The structure of the mutant with I85R was solved in the Apo State (without ligand), and shows no large conformational reorganization such as that observed with canonical allosteric LDHs in Apo state. This is due to a local structural feature typical of *S. rum* LDH that prevents large-scale conformational reorganization. The *S. rum* LDH was also studied using Molecular Dynamics simulations, probing specific local deformations of the active site that allow the *S. rum* LDH to sample the T-inactive state. We propose that, with respect to the LDH/MalDH superfamily, the *S. rum* enzyme possesses a specific structural and dynamical way to ensure homotropic activation.

1. Introduction

Lactate dehydrogenases (LDHs) (EC 1.1.1.27) and Malate dehydrogenases (MalDHs) (EC 1.1.1.37) belong to a wide group of 2-ketoacid:NAD(P)-dependent dehydrogenases that catalyze the reversible conversion of 2-hydroxyacids to the corresponding 2-ketoacids (Holbrook et al., 1975). Both enzymes operate in central metabolism. LDHs achieve their function at the final stage of aerobic glycolysis and MalDHs are involved in the tri-carboxylic acid (TCA) cycle. They display the same protein fold (Rossmann fold) and a similar chemistry underlying their catalytic mechanism (Birktoft and Banaszak 1983; Clarke et al., 1986; Hart et al., 1987 a, b; Clarke et al., 1988; Waldman et al., 1988). When the most chemically competent catalytic state is reached, LDH catalyzes the direct transfer of a hydride ion from the pro-R face of NADH to the C2 carbon of pyruvate (Pyr) to produce lactate; whereas MalDH converts oxaloacetate (OAA) into malate (Burgner and Ray, 1984; Fersht, 1985).

Numerous crystallographic structures of LDHs and MalDHs have been published (Iwata et al., 1994, Auerbach et al., 1998; Dalhus et al., 2002; Irimia et al., 2003; Coquelle et al., 2007; Arai et al., 2010; Coquelle et al., 2010; Ikehara et al., 2014; Kolappan et al., 2015; González et al., 2018, Roche et al., 2019; Iorio et al., 2021), allowing to describe their amino acids involved in NADH and substrate binding. For clarity, we refer to the numbering of amino acid accordingly to the one used for LDH (Eventoff et al., 1977). Pyruvate and oxaloacetate are oxoacids, which have a common negatively charged carboxylate extremity that is screened within the catalytic site by the positively charged lateral chain of R171, a universally conserved substrate-binding residue in all LDHs and MalDHs (Birkoft et al., 1982). When the most reactive conformational substate of the enzyme allowing the Michaelis complex formation is achieved, the mobile active-site loop moves down and close the catalytic site (Iwata et al., 1994, Coquelle et al., 2007). This phenomenon facilitates the catalytic site dehydration and a stronger anchoring of the substrate by hydrogen bonds with R109 (universally conserved in LDH and MalDH) and additional interactions induced by amino acid at position 102, Q for LDH and R for MalDH (Iwata et al., 1994; Coquelle et al., 2007). Canonical LDH structures with their ligands (HOLO states) revealed that the polar side chain of Q102, located on the mobile active-site loop, wrap the methyl of pyruvate allowing stabilizing this substrate in the appropriate orientation for catalysis within the catalytic site (Iwata et al., 1994). While in MalDHs, the lateral chain of R102 contributes to the screening of the second carboxylate extremities of OAA (Birkoft et al., 1982). The amino acid at position 102 is therefore considered as the most important substrate-

discriminating residue between LDHs and MalDHs. Site-directed Mutagenesis experiments have demonstrated that the mutation Q102 to R on the mobile active-site loop transforms a LDH into a highly efficient MalDH (Wilks et al., 1988). The functional conversion from MalDH to LDH due the mutation of R102 to Q has been also attempted; however, this has been much less successful with respect to the catalytic efficiency that stayed quite low (Cendrin et al., 1993; Boernke et al., 1995, Katava et al., 2020). The reason for this lack of complete functional reversibility between these enzymes is very likely due to long-range epistatic effects. Epistasis is due to interactions between amino acid networks of a given protein that silent or enhance the consequence of a mutation depending on the presence or absence of other amino acids (Harms and Thornton 2010, Olson et al., 2014; Mitton et al., 2021).

Even if the LDH/MalDH super family is divided into two main functional groups, phylogenetic and biochemical studies have shown that it is actually populated by several subgroups that display specific signature sequences and different biochemical properties (Madern, 2002; Madern et al., 2004, Boucher et al., 2014, Roche et al., 2019, Katava et al., 2020; Brochier-Armanet and Madern 2021; Iorio et al., 2022). Briefly, the super family contains the clades of MalDH 1 and 2 that are dimeric, a group of *stricto sensu* MalDH 3 mostly tetrameric, the *stricto sensu* LDH and a group whose sequence occupies an intermediate phylogenetic relationship between *stricto sensu* MalDH3 and LDH enzymes (Brochier-Armanet and Madern 2021).The intermediate group (IG) consists of several subgroups that reflect a large reservoir of sequences prone to evolve towards i) the functional conversion from MalDH to LDH and ii) allosteric regulation from non-allosteric MalDH-3 (Brochier-Armanet and Madern 2021). Note that most of the MalDHs from Archaea belong to the intermediate group.

Most of the bacterial LDHs are tetrameric enzymes, which exhibit a sigmoidal enzymatic profile with pyruvate, in the absence of fructose 1,6-bisphosphate (FBP), typical of homotropic allosteric activation of the reaction. When monitored in the presence of FBP, the enzymatic activity profile of allosteric LDHs display a hyperbolic profile demonstrating heterotropic allosteric activation (Arai et al., 2011).

With eukaryotic LDHs, the relationship to allostery is not yet resolved. Some works propose that they are non-allosteric enzymes (Pesce et al. 1967; Everse and Kaplan, 1973; LeVan and Goldberg, 1991; Holland et al. 1997), whereas others suggest they are reminiscent allosteric properties (Katava et al., 2017; Iacovino et al., 2022).

The protein folding and association pathway has been analyzed for the super family; it involves a series of sequential steps. In the first step, monomers in a molten globule state become more compact upon the formation of active dimeric species; in the second step the dimeric species condense to form active tetramers (Madern et al., 2000 and references therein). Abundant structural information has been obtained from the crystal structures of tetrameric LDHs and MalDH (Iwata et al., 1994, Auerbach et al., 1998; Dalhus et al., 2002; Coquelle et al., 2007; Coquelle et al., 2010; Ikehara et al., 2014; Kolappan et al., 2015; González et al., 2018, Roche et al., 2019; Iorio et al., 2021). Three molecular 2-fold axes named P, Q, and R (Rossmann et al., 1973) relate the tetramer subunits. LDHs have four active sites and two FBP-binding sites that are located at the AD and BC interfaces (AD-like interface) through the P-axis. The set of amino acid that participate to these sites, are absent in MalDHs type 3 and in MalDHs from Archaea, so that these enzymes are acknowledged as non-allosteric. In LDHs and MalDHs, the active site of each subunit lies near the interface along the Q-axis.

Studying present-day enzymes from the intermediate group is a relevant strategy to unravel the evolutionary steps that led to functional diversity and emergence of allosteric regulation within the LDH / MalDH super family. We consequently decided to investigate properties of an enzyme from *Selenomonas ruminantium* (*S. rum*) (Uniprot Q9EVR0), in which there is neither Q nor R at the substrate discriminating position 102, but an I. Even if this enzyme does not display a Q, it was characterized as a homotropically-activated LDH (Brochier-Armanet and Madern, 2021). In this follow-up study, the wild-type *S. rum* LDH crystal structure was solved. Its comparison with the structure from a canonical LDH that displays both homotropic and heterotropic activation reveals subtle local reorganization that explain why heterotropic activation by FBP is not possible. We enriched the characterization of the *S. rum* enzyme by a molecular dynamics simulation study. We have also analyzed the consequences of I to Q and I to R mutations at the substrate discriminating position 102, and solved the crystal structure of the former mutant. The present work offers new insights into the processes giving rise to the evolution of allostery in the super family of malate and lactate dehydrogenases.

2. Materials and methods.

2.1. Protein expression and purification.

Overexpression system and purification procedure for the Wild-type *S. rum* LDH was described previously (Brochier-Armanet and Madern, 2021).The same protocol was used for the two

mutants described in this study. The synthetic corresponding genes were constructed by Genecust. The purified proteins were stored à 40°C in 50 mM Tris-HCl pH7. 2, 50 mM NaCl at a concentration of 10 mg/mL.

2.2. Enzymatic assay and protein determination

The activity of *S. rum* LDH for reduction of pyruvate to lactate was carried out at 30°C in 500 µL of 2-(N-morpholino) ethane sulfonic acid (MES) pH 6.0 and supplemented with 50 mM NaCl. MalDH activity was assessed in a mixture containing 50 mM Tris-HCl pH 7.2, 50 mM NaCl. The reaction was monitored at 340 nm by following the oxidation of NADH (0.5 mM) on a Jasco 540 spectrophotometer. To record the enzymatic profile of *S. rum* LDH, various substrate concentrations were tested. The data were analyzed using Michaelis-Menten or allosteric sigmoidal equations in GraphPad Prism version 7.03. The protein concentration was estimated from the absorbance at 280 nm using a nanodrop Thermofischer. One unit of MalDH or LDH activity corresponds to the amount of enzyme that catalyzes the oxidation of 1 micromole of NADH per min.

2.3. Crystallization and structure determination

Wild type *S. rum* LDH was first crystallised by vapour diffusion using the sitting drop method at 293 K (HTX Lab, https://htxlab.embl.fr). Manual reproduction were performed in hanging drop method with a mix composed of: 1 uL of *S. rum* LDH at 8 mg/ml supplemented with NADH at 1 mM and with 1 uL of 10 mM TbXo4 crystallophore (Engilberge et al., 2017, Engilberge et al., 2018) supplemented with 10 mM sodium bicarbonate and 1 µL of PEG 1500 21 % w/v, MIB buffer 150 mM pH 5.5 (PACT, Qiagen) reservoir solution. First crystals appeared in 5 days. Crystals were fished and placed in a cryoprotectant solution composed of PEG 1500 21 % w/v, MIB buffer 150 mM pH 5.5 supplemented with 25 % ethylene glycol right before being flash frozen under liquid nitrogen stream. Wild type *S. rum* LDH diffraction data were collected at the IBS in-house source under a nitrogen stream at 100 K and at a wavelength of 1.541790 A. The in-house setup is composed of a Microfocus sealed tube X ray-source from Xenocs (Cu Alpha radiation) and a MAR345 Imaging plate from MarResearch. Datasets were indexed, integrated, and scaled using XDS (Kabsch, 2010). Molecular replacement and automatic model building were done with the CPP4 suite using PHASER (McCoy et al., 2007), Parrot (Cowtan et al., 2010) and Buccaneer (Cowtan et al., 2006, Cowtan et al., 2008). The model used for molecular replacement was a poly-Alanine structure of

Staphylococcus aureus LacDH (RCSB PDB accession code: 3d0o). The structure was refined with multiple cycles of manual building using COOT (Emsley et al., 2010) and refinement using REFMAC (Murshudov et al., 1997). Model quality was validated with MolProbity (http://molprobity.biochem.duke.edu), (Davis et al., 2007, Chen et al., 2010) and wwPDB validation service (https://validate-rcsb-1.wwpdb.org), (Berman et al., 2003).

I85R *S. rum* LDH protein crystals were obtained using the hanging-drop vapour diffusion technique at 20 °C in 24-well crystallization plates. The drops were prepared by mixing 1.5 µl of the protein solution (at a concentration of 2–10 mg ml^{-1}) with an equal volume of crystallization solution and were equilibrated against 1 ml of crystallization solution. I85R *S. rum* LDH proteins gave diffracting crystals under the following conditions: 0.1 M BisTris propane pH 7.5, 0.2 M $NaNO_3$, 16–26% PEG 3,350 (optimized from Molecular Dimensions PACT screen, condition 77). After optimization, crystals were transferred into cryoprotective buffer (0.1 M BisTris propane pH 7.5, 0.2 M $NaNO_3$, 16–26% PEG 3,350, 20% Glycerol) before flash freezing in liquid nitrogen. Diffraction data for I85R *S. rum* LDH protein crystals were collected at ESRF beamline BM07 at 100 K (Grenoble). I85R *S. rum* LDH protein crystallizes in the space group $P2_12_12_1$ with four monomers in the asymmetric unit. A resolution of 2.8 Å was obtained. The structure of *S. rum* LDH I85R protein was determined by molecular replacement using *S. rum* LDH crystal structure (PDB ID 7NAY) as model. The data set was indexed, integrated and scaled with XDS package (Kabsch, 2010). The molecular replacement program Phaser from CCP4 package (McCoy et al., 2007) was then used for phasing the data set with 7NAY model. Model building was done using the graphic program COOT (Emsley et al., 2010). The atomic model was refined using the program Phenix (Adams et al., 2010). The refinement process included successive rounds of simulated annealing, energy minimization, B-factor and TLS refinements as well as calculation of difference Fourier electron density maps. Water molecules and ions were added in the late stage of the refinement. The final I85R *S. rum* LDH protein model show good stereochemistry as indicated by the program PROCHECK with no residue in the disallowed regions of the Ramachandran plot.

2.4. Molecular dynamics.

We performed all-atom Molecular Dynamics (MD) simulations of tetrameric assembly for *S. rum* and *Thermotoga maritima* (*T. mar*) LDHs. Simulations were performed in the NPT ensemble with the MD software Gromacs 2018.7 (Abraham et al., 2015) starting from crystallographic structures (7NAY for *S. rum* and 1A5Z for *T. mar*). The pressure was kept

constant at a value of 1.01 bar using the Parrinello-Rahman algorithm (Parrinello et al., 1981). The temperature was controlled by the Nose-Hoover algorithm (Nosé, 1984, Hoover, 1985). The simulations time step was set to 2 fs. The Particle Mesh Ewald method (Darden et al. 1993) was used to calculate the long-range electrostatic interactions. Bonds involving hydrogen atoms were constrained with the LINCS algorithm (Hess at al., 1997). The Charmm36 force field (Huang et al., 2016) was used to model protein's interactions along with the three points TIP3P water model. To neutralize the total charge of the systems, ions were added to the simulation boxes. After an initial 100 ns long equilibration phase at 300 K, an additional equilibration phase was performed at 315 K for the system with *S. rum* and at 340 K for the system with *T. mar*. Finally, production runs of 1 μs were performed for both systems and the resulting trajectories used for the analysis.

Structural superposition and representations

PyMOL was used for representation (http://www.pymol.org/pymol). To perform homogeneous comparisons between PDB files, chains LDH standards, as monomers A, B, C and D, respectively located at upper left, upper right, down right and down left positions.

3. Results and discussion

For the sake of clarity, analysis of amino acid positions (Fig. S1) are frequently presented using the "normalized" numbering system proposed for lactate dehydrogenases (nnL) (Eventoff et al., 1977). A) According to this numbering, important active site residues are labeled Q102, R109, D168, R171, H195 and T246. B). Important amino acid positions that contribute to homotropic activation capacity are labeled H68 and I250 (Iorio et al., 2022). C) Several amino acids participate in FBP-binding, including R173, H188 and Y190 (Iwata et al., 1994). In the linear numbering of *S. rum* LDH, equivalent positions are i) 85, 93, 153, 156, 180 and 234. ii) 51 and 238. iii) 158, 173 and 175. For structural comparison, the linear numbering is used.

The new picture that describes the evolutionary relationship within the LDH/MalDH super family is schematized in Fig. 1.

3.1. The overall structure of wild-type S. rum LDH.

Even if the protein stock solution did not contained any additional compound, we found NADH and malonate molecules bound to the Wt enzyme. This is analyzed below. Consequently, the *S. rum* LDH crystal structure corresponds to a HOLO rather than APO form. Data collection and refinement statistics are shown in **Table S1**. The atomic coordinates are deposited at the PDB under accession number 7NAY. *S. rum* LDH crystallizes in the space group I222 with a monomer in the asymmetric unit. Crystal growth was obtained using the hanging drop method and vapor diffusion using mother liquor containing 21 % PEG 1500, 150 mM Malonate-Imidazole-Boric Acid pH 5 (MIB-PACT), 10 mM Sodium Bicarbonate pH 8 and 10mM of a lanthanide-based adjuvant (Crystallophore, Tb-XO4). This compound has unique nucleating and phasing properties giving it useful properties for crystallography (Engilberge et al., 2017, Jiang et al., 2020). The monomeric structure is shown in Fig. 2a and the final tetrameric structure of *S. rum* LDH is shown in Fig.2b. As expected, the overall architecture of *S. rum* LDH monomer is identical to all MalDHs and LDHs previously reported.

This assembly corresponds to the oligomeric state previously determined in solution (Brochier-Armanet and Madern, 2021). Subunit contacts that maintain the tetramer state occur through interfaces along the P, Q and R axes. *In silico* structural analysis done with the Super family server (Wilson et al., 2009) using the *S. rum* LDH primary structure predicts two domains: a NAD(P)-binding Rossmann-fold domain (residues 2 to 149) and a LDH C-terminal-like domain (residues 148 to 313).

During refinement, unusual electronic densities were appearing in 2Fo-Fc and Fo-Fc maps around sulfur atoms of cysteines 55 and 281. Those cysteines are solvent accessible and incompatibles with any disulfide bond, indicating a possible oxidation. The insertion of oxidized cysteines, CSD in position 55 and OCS in position 281 was validated by the refinement procedure. To confirm the presence of such unusual residues we performed mass spectrometry experiments on fresh protein as well as on 2 weeks old crystals and 1 month-old crystals. Mass spectrometry confirmed that no additive oxygen was present on the fresh protein. However, the presence of 5 additive oxygens was confirmed on the protein in the 2 weeks old and month-old crystals. The oxidization of cysteine 55 and 281 is therefore a phenomenon that happens during crystallization and is probably due to the crystallization condition (Fig S3).

3.2. S. rum LDH scaffold differs from canonical LDH by local structural changes.

LDHs from *Thermus* species are the most extensively studied allosteric enzymes that belong to the clade of *stricto sensu* LDH (Taguchi, 2017; Iorio et al., 2021). Several representative crystal structures of the T-inactive (APO) and R-active (HOLO) states are available at the Protein Data Bank. We used the enzyme from *Thermus caldophilus* (*T. cal*), or *Thermus thermophilus* (*T. the*); their PDB accession codes are 3VPG/3VPH and 2V6M/2V7P, respectively.

Because the *S. rum* LDH structure was obtained with ligands, we used *T. cal* LDH HOLO state (3VPH) for a relevant comparison. A PyMOL superposition of the monomers reveals that, in spite of a similar fold, there exist three noticeable local topological differences on the *S. rum* LDH (colored in blue in Fig. 3 A) that may explain the functional and regulatory differences with respect to enzymes from the *stricto sensu* LDH clade.

First, compared to *T. cal* LDH, in addition to the difference in the nature of the substrate discriminating residue at position 102 (nnL), the mobile active site loop in *S. rum* LDH has a two amino acid residues insertion (P91 and D92) (Fig.3, 1B, 1C). Consequently, the mobile active site loop of *S. rum* LDH is distorted compared to the one observed in LDHs from *THERMUS* species. This is analyzed below.

Secondly, due to an insertion of two amino acids, the area bridging αC helix and βC strand is significantly different in *S. rum* LDH and in 2VPH (Fig. 3,2B,2C). In *S. rum* LDH; the extremum of αC is locally unfolded, a structural feature that allows the connecting loop to be more extended than in 2VPH. In canonical LDH, the αF helix moves its position by 15° between the T-and R states allowing R171 to sample the in and out configurations (Taguchi 2017). In parallel, H68 located on αC from the adjacent monomer (Q axis) displays coordinated side chain motions preventing (or not) the side chain of R171 to access the catalytic site. A calculation of the hydrogen bonding pattern using VADAR reports that the local structural reorganization in *S. rum* LDH favors establishing five inter subunits H-bonds, which are absent in *T. cal* LDH (Fig S2). This structural feature strongly suggests a decrease of local entropy in *S. rum* LDH, unfavorable for αF and αC helices reorganization as in canonical LDH.

Thirdly, in *S. rum* LDH the area comprising part of α1G/α2G kinked helix and the connecting region to αT has a noticeably different topology compared to 2VPH (indicated in blue, Fig.3, 3B, 3C). In allosteric LDHs, these elements participate in the allosteric core (Alco),

hydrophobic-rich area, which mechanically links the FBP-binding site to the substrate-binding residue R171(nnL) (Ikehara et al., 2014, Iorio et al., 2021). Allosteric activation of enzymes relies on the capacity to convey a signal *via* a network of contacts from a binding-site far from the substrate binding-site. In the specific case of allosteric LDHs, studies have revealed that Alco has a strong capacity to control their dynamical properties (Iorio et al., 2021). R171(nnL) side chain occupancy in the substrate-binding site is thus directly influenced by communication within Alco (Iorio et al., 2021). In *T. cal* LDH, R169(nnL-171) is strongly coupled to Alco owing to the side chain of the neighboring amino acid F170(nnL-172). In *S. rum* LDH because of the structural differences with *T. cal* LDH, the F157(nnL172) side chain has a different orientation within Alco (Fig. 3, 3B, 3C). The comparison suggests therefore that a LDH from the intermediate group and a strict sensu allosteric LDH, differ by their capacity to influence direct dynamics of the universally R171(nnL) substrate-binding residue. Other differences observable in this area are presented below.

3.3. Catalytic site differences between the S. rum enzyme and a canonical LDH.

According to its sigmoidal pyruvate saturation profile (Brochier-Armanet and Madern, 2021) *S. rum* LDH samples both T and R states in solution. Numerous investigations using site-directed mutagenesis and crystallographic structures of LDHs have demonstrated that the nature of amino acid residues located at some key positions, together with movements of some secondary structure elements are involved in substrate recognition and catalysis (Clarke et al., 1986; Wilks et al., 1988; Cendrin et al., 1993; Iwata et al., 1994; González et al., 2018, Brochier-Armanet and Madern 2021). Unfortunately, we were not successful in growing APO state crystals of the wild type enzyme that would have been useful to describe a putative T-inactive state associated to large conformational change. In canonical LDHs, a specific R171(nnL) side chain conformation is associated with the T and R states, being either outside or inside the active site, respectively (Coquelle et al., 2007; Colletier et al., 2012; Taguchi 2017). In the activated R-state of allosteric LDHs, the amine moiety of R171(nnL) side chain interacts with the carboxylate moiety of pyruvate. We failed to get crystals representative of the putative T-inactive state of Wt *S. rum* LDH in which the R156 (nnL-171) would have been located outside the catalytic site. The Wt *S. rum* LDH close view of the catalytic site shows that the NH_2 extremity of R156 (nnL-171) protrudes within the catalytic site as it is the case with canonical LDHs (Fig. 4).

However, the comparison between the HOLO-state of *S. rum* LDH with *T. the* LDH, suggests another mechanism allowing the *S. rum* LDH to sample a T-inactive like state without large conformational state.

We present a comparison between the catalytic site, including the mobile active site loop of *S. rum* LDH and *T. the* LDH taken as representative of a canonical allosteric LDH (Fig. 4A-B). Substrate binding in LDHs is mainly due to four residues found at positions 102, 109, 171, and 246 (Fig. 4B).

In *S. rum* LDH, the mobile active site loop (MASL) differs from the one encountered in canonical LDH by three amino acid mutations, i) an amino acid insertion of two residues (P91 and D92) and ii) the presence of an I at equivalent position 102 amino acid (Fig. 4A). Therefore, the malonate molecule that occupies the same location as oxamate in canonical LDH is not totally covered by the mobile loop. We propose that because of the distorted MASL, the side chain of I85(nnL-102) in *S. rum* LDH cannot establish close contact to substrate. At nnL positions 109, 171, and 246, there are no differences.

The reaction catalyzed by LDH comprises the direct transfer of a hydride ion from NADH to pyruvate, abetted by the protonation of substrates' keto oxygen (Holbrook et al., 1975). The proton is given by the universally conserved H195. Within the active site, H195 is polarized owing to its interaction with D168 (Birktoft and Banaszak, 1983). The *S. rum* LDH shows a glutamic acid (E153) at equivalent position 168 (Fig.4a) suggesting that its catalytic H180(nnL-195) has a slightly different polarization state.

In agreement with the repartition of the group within the super family of LDH and MalDH, a negatively charged residue (E) is always present, in canonical LDH, at universal position 199 whereas in MalDHs type 3, there is always a neutral residue at position 199 (M) Brochier-Armanet and Madern, 2021). With *S. rum* LDH, at this equivalent position there is also a neutral residue (A184) (Fig.4a). We propose that it corresponds to a reminiscent "MalDH-like" amino acid as expected for an enzyme from the Ig group. Analysis of contacts and inter atomic distances using VADAR and ARPEGIO (Willard et al., 2003; Jubb et al., 2015), shows the E199 OE1 atom making a hydrogen bond with D168 OD1 within the catalytic site of *T. the* LDH. Such an H-bond most likely participates in keeping the D168 side chain in a configuration

favorable for efficient polarization of H195. Calculation indicates that in *S. rum* LDH the ND1 atom of the catalytic histidine establishes two hydrogens bonds with each oxygen atoms of E153(nnL168) extremities, whereas in *T. the* LDH the ND1 atom interacts solely with D168 OD2. It is reasonable to propose that in *S. rum* LDH, the absence of negatively charged residues, replaced by A184, decreases the local electro negativity of the catalytic site and therefore, alters the polarization state of the catalytic histidine compared to canonical LDH.

3.4. AD-like interface topology in S. rum LDH prevents FBP binding sites formation.

Wt *S. rum* LDH lacks the heterotropic activation by FBP (Brochier-Armanet and Madern, 2021). In *T. cal* LDH structure, two FBP molecules are trapped at the AD-like interface between active dimers where R173, H188 and Y190 from two contiguous subunits interact with the two phosphate groups of FBP (Taguchi 2017). These amino acids at equivalent positions are indicated in orange on Fig. 3B-C. *S. rum* LDH. Sequence and structure analysis indicate that both R158(nnL-173) and Y175(nnL-190) are conserved, whereas there is no conservation at position H188(nnL) due to N173. This incomplete signature sequence in *S. rum* LDH explains why the enzyme does not recognize FBP. In addition to this amino acid change, it is possible to see that the local topology of the putative *S. rum* FBP binding-site is different from the one observed in canonical allosteric LDH. Different views help to understand the phenomenon (Fig. 5). Inspection of the AD dimer in different orientations around the Q axis shows that the small extension in *S. rum* LDH (analyzed previously, Fig. 3) creates additional contacts and extend the buried surface of the AD like interface (P axis). For example, the side chain of Y57 from adjacent *S. rum* LDH monomers protrudes within the central solvent accessible cavity, so that the volume of this cavity decreases compared to *T. cal* LDH one (exemplified by comparing the central panel of Fig. 5A- B). The *S. rum* LDH structure reveals that four malonate molecules occupy the putative FBP-binding sites located at the AD-like interface (Fig. 5A- B third panels from top). This unexpected observation due to the recruitment of precipitating agents reveals an incomplete binding site that can be considered in a locked state compared to a complete FBP-BS in which the activator can diffuse to its binding site. The comparison clearly shows that in the tetrameric state, αC helices from monomer A and D comes in close contact to the putative FBP-BS made by monomer B and C and *vice versa*. Consecutively, in *S. rum* LDH, side chains of I57 and Y57 create strong steric hindrance that prevent the putative FBP-BS to be connected to the solvent-accessible central cavity.

So that, neither the Alco topology, nor the amino sequence of putative FBP-binding site and favorable accessibility for FBP are achieved in *S. rum* LDH; despite its close resemblance with allosteric LDHs.

3.5. Effect of I85Q and I85R mutations on substrate recognition.

As previously shown by Brochier-Armanet and Madern (2021) the wild type *S. rum* LDH recognizes pyruvate as substrate, even if the typical Q at position102 (nnL) found on its mobile active site loop (MASL) as in LDHs is lacking (Fig. 4A). We decided to investigate the effect of introducing the typical LDH (Q) or MalDH (R) amino acid on *S. rum* MASL. The resulting mutants I85Q (102nL) and I85R(102 nnL) were purified as described for the Wt enzyme and their enzymatic properties monitored. We expected enhanced LDH enzymatic properties with the I85Q *S. rum* enzyme due to the mutation. The data showed that the I85Q *S. rum* mutant was always able to recognize pyruvate as substrate with a sigmoid profile (Fig. 6B). However, the replacement has not enhanced its capacity to recognize pyruvate. The affinity for pyruvate is slightly lowered with a K_{half} value ($S_{0.5}$) of 4 mM compared to 2 mM with the Wt enzyme (Table1). The maximal velocity values using pyruvate is also lowered (Table1). We tested that I85Q *S. rum* enzyme does not recognize OAA. We then tested, the I85R mutation and observed that the resulting enzyme can be considered as a functional MalDH that recognizes OAA with an affinity values of 1.45 mM (Fig. 6C and Table 1). Here again, compared to the Wt enzyme, the resulting I85R *S. rum* mutant has a reduced maximal turn over when using OAA.

The crystal structure of *S. rum* LDH has revealed a distorted MBAL in *S. rum* LDH compared to canonical LDHs and MalDHs. The measurements have demonstrated that introducing a Q (LDH specific) or R (MalDH specific) instead of I85 (102nnL) in the MBAL of *S. rum* LDH was detrimental to catalytic efficiency mainly because the turnover is lowered. We propose that the longer side-chain of Q and R compared to I, are less suitable amino acids for open-closure motions of a distorted MBLA

3.6. Comparison between the Apo and Holo crystal structures of S. rum LDH shows no large structural reorganization.

We decided to continue our structural investigation using the I85R *S. rum* LDH mutant that can use OAA as substrate.

The I85R(102nnL) *S. rum* LDH structure showed it did not contain i) any malonate molecule within the catalytic site or in the FBP-BS like pockets or ii) any NADH as observed with the Wt enzyme. The I85R (102nnL) *S. rum* LDH structure can therefore be considered as representative of the Apo state. In *stricto sensu* allosteric LDHs, Apo and Holo crystal structures have revealed large conformational reorganization representative of the T- or R- states including some helix sliding between monomers and a compaction of the R-state compared to the T-state (Reviewed by Taguchi 2017). With canonical allosteric LDH the RMSD between T-and R- states is in the range 1.7 to 2.2 Å. In contrast, the structural superimposition of the I85R (102nnL) mutant and the Wt *S. rum* LDH with a RMSD of 0.3Å for a monomer indicates that no major structural reorganization take place (Fig. 7A). As it has been frequently reported in crystal structures of LDH and MalDH, the MASL is not defined in the mutant structure. This is also the case with the loop connecting αT and α1G-α2G helices that is not defined in monomers B and C. Inspection of the mutant structure revealed that with monomers A and D, E209 is engaged in crystallographic contact with K95 from adjacent tetramer avoiding fluctuations of the connecting loop. We then compared the FBP-BS like cavity between the I85R(102nnL) mutant and the Wt enzyme (Fig. 7B). The close-up views show that some local lateral chain reorganizations take place when the site is not occupied by malonate and EDO. It is exemplified by I56, M151, Y175 and L257 in monomer C in which a water molecule occupies the site (Fig. 7C-D). In the Apo I85R(102nnL) *S. rum* LDH, each FBP-BS-like cavity behaves differently with respect to binding capacity and side chain position. We found that a phosphate and a chloride ion are present in monomers A and D, respectively (Fig. S4). We compared the Alco region in the I85R (102nnL) *S. rum* LDH and observed a slight side-chain reorganization of some hydrophobic residues in each monomer (sup mat). This could be the consequence of two things. First, the absence of ligands may increase the local dynamics of the enzyme compared its bound state with malonate. Second, due to packing within the crystal lattice, contacts between monomers are not equivalent and therefore, can slightly influence their respective dynamics.

The comparison between the *S. rum* LDH in Apo and Holo states has demonstrated this enzyme has a very limited capacity to explore large conformational reorganizations such as those encountered in canonical allosteric LDHs. This is mainly due to specific small local structural features that i) prevent the true allosteric activator binding site of LDH to exist and ii) hamper the local dynamics to propagate up to the substrate binding residue R156(nnL-171), which consequently cannot explore the T-inactive state as in canonical LDH. The absence of large

structural reorganization does not mean that subtle changes of conformational dynamics cannot exist between each state of the *S. rum* LDH, as it was observed with the side chain of some amino acids within the Alco and pseudo FBP BS. Despite the absence of a mechanism similar to the one existing in canonic LDH, the S. *rum* LDH displays a homotropic activation profile, demonstrating that its capacity to sample the T-inactive state should be achieved by another process.

3.7. Catalytic site dynamics of S. rum LDH and canonical allosteric LDHs are different.

In order to grasp the conformational fluctuations of the *S. rum* LDH we performed a Molecular Dynamics (MD) comparison with a canonical LDH displaying both homotropic and heterotropic activation. MD simulations have been proven to be very effective for inspecting the relationship among protein conformational flexibility, allostery and functionality (Sterpone et al., 2009; Sterpone and Melchionna 2012; Kalimeri et al., 2013; Katava et al., 2016; Katava et al., 2017; Katava et al., 2020; Maffucci et al., 2020, Iorio et al., 2021; 2022). The structure of *Thermotoga maritima* LDH (*T. mar* LDH) (1A5Z) was used as reference. The MD simulations were done at temperatures close to the functional temperatures, with 314 and 345K for *S. rum* and *T. mar* LDHs, respectively. Previous MD simulations using canonical allosteric LDHs have shown that allosteric large R171(nnL) side chain fluctuations are strong markers to characterize the propensity to explore the low substrate affinity (T-) and high substrate affinity (R-) states (Iorio et al., 2021, Iorio et al ., 2022).

In order to characterize the conformational space explored by R171(nnL), side chain distances, d1 and d2, were recorded with two amino acid positions located within the catalytic site, I250(nnL) and T246(nnL), respectively (Fig. 8). In a given monomer, when d1 and d2 values stay close to 7.5 and 4.5 Å, this indicate that the R171(nnL) side chain protrudes within the catalytic site (Cs-in) a favorable situation for substrate binding. In contrast, when the distance increases, the R171(nnL) side chain will be located outside the catalytic site (Cs-out). The distance variation (d3) between T246(nnL) with I250(nnL) was also considered. With *S. rum* LDH, the distance fluctuations (d1 and d2) around 7 and 5Å are very similar among the various monomers, indicating that the R171(nnL) side chain mostly occupies the catalytic site. *T. mar* LDH, d1 and d2 measurements show that i) the various monomers behave differently with ii) variations of wider amplitude up to 15 Å, demonstrating the R171 (nnL) side chain may explore Cs-in and Cs-out configurations, confirming previous MD studies on allosteric LDHs (Iorio et al., 2021). In contrast to d1 and d2 distances, when d3 is taken into account, the observation

with respect to the catalytic site dynamics in each enzyme is different. With *S. rum* LDH, in most of the monomers, d3 fluctuates between 5 and 10 Å, whereas it stays rather stable around 5Å in *T. mar* LDH.

The data clearly showed that the *S. rum* LDH displays a catalytic site dynamical behavior different from the one observed with canonical LDHs (This work and Iorio et al., 2021).

5. Conclusion

The concept of protein evolvability, linking the capacity of proteins to evolve and conformational dynamics, (Tokuriki and Tawfik, 2009) is the fundamental approach we applied to solve the question about the origin of allostery in LDHs.

Allosteric bacterial LDHs fit well the Monod-Wyman-Changeux transition model, which stipulates that the T-inactive and R-active states of an enzyme coexist independently of allosteric effectors (reviewed in Taguchi, 2017). Based on an extensive phylogenetic approach, biochemical, structural and dynamical characterizations, it has been possible to demonstrate that the clade of *stricto sensu* LDHs evolved from non-allosteric MalDHs type 3 via a reservoir group of sequences with various intermediate properties (Katava et al., 2017; Roche et al., 2019, Katava et al., 2020, Brochier-Armanet and Madern, 2021; Iorio et al 2021, 2022). Because non-allosteric MalDHs are always in a R-active state, understanding the evolution of allostery in LDHs requires understanding of i) how the capacity to sample the T-inactive state is achieved and ii) how the resulting T-R equilibrium is fine tuned.

The reservoir group identified by Brochier-Armanet and Madern (2021) is divided into several subgroups of sequences displaying different capability to evolve new functional and regulatory properties. With respect to functionality, MalDH and canonical LDH are able to discriminate oxaloacetate and pyruvate thanks to the presence of two amino acids. Those residues are R102(nnL) and A/S 246(nnL) in MalDHs, and Q102(nnL) and T246(nnL) in LDHs (Bur et al., 1989; Binay et al., 2013; Borchier-Armanet and Madern, 2021). In the intermediate group of enzymes, it was determined that MalDH functionality was abolished and replaced by a capacity to use pyruvate as substrate (Brochier-Armanet and Madern, 2021).

This was exemplified by the characterization of the gene product (UniProt ID Q9EVR0) from *Selenomonas ruminantium* that recognizes pyruvate whereas it lacks the strictly conserved Q102 (nnL) in the *stricto sensu* LDH clade but has T246(nnL) (Brochier-Armanet and Madern,

2021). In canonical LDH, this mobile active site loop covers the catalytic site upon substrate binding. Here, owing to the structure solved in a ternary complex state with NADH and malonate it was possible to see that even if the *S. rum* LDH mobile loop organization differs from those in canonical LDH it still allows substrate analog binding. This is because the substrate binding site topology within the catalytic site is the same as that in canonical LDH with a strict amino acid conservation at nnL positions R109, H195, T246 and I250. The role of I250(nnL) has been recently investigated. During the transition from MalDHs to LDHs, the P250I(nnL) replacement has created favorable dynamical conditions allowing homotropic activation to appear (Iorio et al., 2022). This phenomenon was analyzed to be the consequence of αF helix increased mobility that allowed the R171(nnL) side-chain to sample configurations outside the catalytic site. This amino acid replacement considered as an allosterizing mutation occurred in the intermediate group prior to its fixation in canonical LDH (Iorio et al., 2022). The comparison between the I85R(nnL-102) and Wt *S. rum* LDH, representative of the Apo and Holo states, showed that there is no noticeable structural reorganization as in canonical LDH, strongly suggesting that specific structural features can counterbalance the effects of allosterizing mutations. Our structural comparison with canonical LDH demonstrated that it is the case. In the allosteric transition between the T- and R- states of canonical LDHs, the R171(nnL) side-chain switch inside the active site occurs through a change of contact between the α2F helix and the αC helix at the Q-axis related monomer. Consequently, movements of H68(nnL) located on the αC helix controls the conformational sub state sampling of R171(nnL) through steric hindrance (Iwata et al., 1994, Colletier et al 2012). In *T. cal* LDH taken as a canonical allosteric enzyme, the C-terminus extremum of αC helix is not involved in any interaction with adjacent monomers. In contrast, in *S. rum* LDH, a small amino acid insertion modifies the αC helix topology and locally changes the hydrogen bonds pattern at the interface forming dimers AB and CD (Q-axis). We propose that these local modifications are sufficient to prevent tertiary and quaternary conformational reorganization of *S. rum* LDH compared to those observed in canonical LDHs.

In addition, we observed that the αC helix local extension has had consequences on the P-related interface that joins the AB and CD dimers into a tetrameric assembly. Because the extension protrudes into the large central solvent accessible cavity, it restricts the *S. rum* LDH capacity to make a true allosteric effector-binding site competent for FBP accommodation. Furthermore, the *S. rum* LDH sequence indicates that the sequence signature (nnL) R173, H188 and Y190 that participate to FBP binding at the P-related interface in canonical allosteric LDHs is

incomplete. These observations explain why the *S. rum* LDH does not display heterotropic activation as with canonical LDH.

With *S. rum* LDH, the activation profile for substrate recognition is sigmoid, demonstrating a homotropic activation capacity. Since *S. rum* LDH does not sample large conformational change and R171(nnL) side chain fluctuations, its capacity to exist as an inactive state is due to another mechanism. MD simulation shows that the catalytic site can be distorted in a different way compared to canonical LDH. The capacity to experience the T-inactive state without large conformational change has been previously observed in a MalDH from the intermediate group that can recognize oxaloacetate and pyruvate (Roche et al., 2019). Compared to canonical allosteric LDH, both enzymes display differences in surface contact between protomers that reshape the solvent accessible cavity and prevent large motion to exist.

The *S. rum* LDH characterization confirms that i) the acquisition of homotropic activation during the evolution of allostery in the MalDH/LDH super family, is rooted within the intermediate group and ii) this phenomenon relies on different mechanisms (Katava et al., 2017, Roche et al., 2019; Brochier-Armanet and Madern, 2021; Iorio et al., 2022). Due to several capacity such as i) the use of pyruvate even without Q102, ii) a homotropic activation behavior and iii) the formation of a pseudo effector-binding pocket suitable to accommodate a smaller, negatively charged molecule other than FBP, we consider that within the intermediate group the *S. rum* enzyme corresponds, to date, to closest relative of canonical allosteric LDHs. It does not exclude that some others even more closely LDH-related "intermediate" enzymes being to be discovered, also exists.

From comparative studies between large families of oligomeric proteins, it has been suggested that changes that reduced contact surfaces between protomers in oligomeric proteins facilitates new pathways of evolution (Marsh and Teichmann, 2014; Meng et al., 2018; Abrusán and Marsh, 2019). If so, it suggests that it would be possible to identify new intermediate enzymes with a larger solvent accessible cavity than in *S. rum* LDH, allowing a binding-pocket efficient for FBP recognition to be formed and a quasi-achieved allosteric regulation as in canonical allosteric LDHs. Such an enzyme would correspond to the closest relic of the last evolutionary step prior to the birth of canonical allosteric LDH.

Acknowledgements

The authors also acknowledge funding by Agence Nationale de la Recherche (ANR) grant to DM (program AlloSpace ANR-21-CE44-0034-01). This work used the platforms of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

FS acknowledges the financial support by the Initiative d'Excellence program from the French State (Grant DYNAMO, ANR-11-LABX-0011-01, and CACSICE, ANR-11-EQPX-0008). The work was performed using HPC resources from GENCI [CINES, TGCC, IDRIS] (grant X20206818) and from LBT.

We acknowledge the French Biology/Health Panel Review Committee for provision of synchrotron radiation beamtime at beamline BM07-FIP2 of the ESRF, Grenoble, France.
We thank Dr Sylvain Engilberge for support during the experiment.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Author contributions

DM conceived the project. DM performed sample preparation and enzymatic measurements, QB and SC conducted crystallization trials, diffraction data collection and structure refinement. AI and FS conducted molecular dynamics simulation. All the authors analyzed data, discussed the data and wrote the paper.

The crystallographic data are deposited at the Protein Data Bank. Accession numbers 7NAY and 8Q3C

Legends

Fig. 1. Schematic drawing of the evolutionary relationship between MalDH-type 3 and LDH sequences. The left part shows the clades of MalDH type 3 (Purple) and *stricto sensu* LDH (Blue) as defined by Brochier-Armanet and Madern (2021). The R and Q correspond to amino acids found at position 102 (nnL) in each clade (universal LDH numbering). The small grey circle illustrates the intermediate group of sequences as defined by Brochier-Armanet and Madern (2021). The right part proposes an extended phylogenetic representation of the intermediate group. The archaeal and bacterial sequences are in red and green, respectively. No R/Q means that these amino acids are never found at the 102 (nnL) position. The red arrow indicates the location of the *Selenomonas ruminantium* sequence that carries an I at position 85 (102 -nnL).

Fig. 2. Structure of the Apo form of tetrameric *S. rum* LDH. (A) *S. rum* LDH monomer representation. α-helices are colored in red whereas β-strands and loops are in yellow and green, respectively. Secondary structures are labeled accordingly to the LDH nomenclature (Coquelle et al., 2007). (B) Cartoon representation of the four monomers labeled A to D. Subunits of the tetrameric assembly are related by three molecular dyads, *P, Q* and *R*.

Figure 3. Pairwise ribbon drawing superposition between *S. rum* and *T. cal* LDHs. (A). Most of the secondary elements are in ribbon representation to the exception of kinked helix α1G–α2G and αT helix that are shown in cartoon representation. The Coenzyme is colored in orange. The substrate analogs are in red. The catalytic histidine is shown in stick representation. The three noticeable local structural deviations of the *S. rum* monomer with respect to 2VPH are indicated in blue. 1, the active site mobile loop; 2, extremity of helix αC and 3, allosteric core area. The corresponding close-up views are shown in the right columns. 1. (B) *S. rum* LDH, (C) *T. cal* LDH.

Fig. 4. Close-up views showing a comparison of the *S. rum* LDH catalytic site with a canonical LDH. (A) *S. rum* LDH. Amino acids are shown in stick representation. Compared to canonical LDH, the two additional amino acids are indicated in blue. Secondary structural elements of monomer A and B are in green and cyan, respectively. Amino acid differences at equivalent position in canonical LDH are in red. The malonate molecule is abbreviated mal. Some amino acids and the malonate are represented with dots. This representation allows to observe differences of compressibility in the active site. (B) *T. the* LDH (2V7P). The NADH is shown in orange. The arrow shows that in *S. rum* LDH, the side chain of I85(nnL-102) is distant from the malonate.

Fig. 5. Different views of AD like interfaces in *S. rum* LDH and a canonical LDH showing the differences in FBP-binding capacity. Top panel: view of AD dimers for (A) *S. rum* LDH. (B) *T. cal* LDH. The red circles show the αC helix insertion with respect to the AD interface. Area corresponding to the FBP-BS is in purple. Intermediate top panel: close-up views of the αC helix through the P axis. It corresponds to a rotation of 90° around the Q axis. Intermediate bottom panel: slice of the tetramer. The solvent-accessible surface is shown in grey. The slices allow evaluating the respective size of the solvent -accessible central cavity in each enzyme. Ligands are shown in ball and stick repesentation. Bottom panel: close-up view of area involved in FBP-binding. The ligands are in red. The three amino acids involved in FBP-recognition in heterotropicaly activated LDH are in orange. Amino acid differences are indicated in red.

Fig. 6. Enzymatic activity profiles of *S. rum* LDH and mutants. Measurements were done at 35°C in the presence of the indicated concentrations of substrate. (A) *S. rum* LDH wild-type. (B) *S. rum* LDH I85Q. (C) *S. rum* LDH I85R.

Fig. 7. Comparison between I85R (102nnL) **and Wt *S. rum* LDHs**. (A) Ribbon drawing superposition of monomer A. Smudge and light green for Wt and mutant, respectively. (B) Slice through the Q-axis allowing seeing amino acids that delineate the cavity containing the malonate and EDO molecules (shown in balls) in Wt *S. rum* LDH. (C) and (D) close-up views with the Wt and I85R (102nnL) *S. rum* LDHs, respectively. Amino acids are in stick representation. Those that display side chain reorganization are indicated in *italics*. A water molecule is indicated by W.

Fig. 8. Comparative molecular dynamics properties between *S. rum* LDH and a canonical allosteric LDH. On the top, the distance in Å between three amino acids located within the *S. rum* LDH catalytic site are shown in dashed lines. The universal numbering for LDH is indicated. The expected effect with respect to R171(nnL) conformational sub state occupancy is indicated. The bottom panels show distances (d1, d2 and d3) variations in each monomer for *S. rum* and *T. mar* LDHs. The structural numbering is used.

References

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J., C., Hess, B., Lindahl, E., (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1, 19-25.

Abrusán G, Marsh JA. (2019) Ligand-Binding-Site Structure Shapes Allosteric Signal Transduction and the Evolution of Allostery in Protein Complexes. Mol Biol Evol. 36(8):1711-1727.

Adams, P. D. et al.(2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta. Crystallogr. D. Biol. Crystallogr.* 66, 213–221

Arai K, Ishimitsu T, Fushinobu S, Uchikoba H, Matsuzawa H, Taguchi H. (2010) Active and inactive state structures of unliganded Lactobacillus casei allosteric L-lactate dehydrogenase. Proteins. 78(3):681-94.

Arai, K., Ichikawa, J., Nonaka, S., Miyanaga, A., Uchikoba, H., Fushinobu, S., Taguchi, H. (2011). A molecular design that stabilizes active state in bacterial allosteric l-lactate dehydrogenases. J Biochem.150, 579-591.

Auerbach G, Ostendorp R, Prade L, Korndörfer I, Dams T, Huber R, Jaenicke R. (1988) Lactate dehydrogenase from the hyperthermophilic bacterium thermotoga maritima: the crystal structure at 2.1 A resolution reveals strategies for intrinsic protein stabilization. Structure. 6(6):769-81.

Berman, H., Henrick, K., Nakamura, H. (2003). Announcing the worldwide Protein Data Bank. Nature Struct. Mol. Biol. 10, 980-980.

Binay, R.B. Sessions, N.G. Karaguler, A double mutant of highly purified Geobacillus stearothermophilus lactate dehydrogenase recognises l-mandelic acid as a substrate, Enzyme Microb Technol, 52 (2013) 393-399.

Birkoft, J. J. Fernley, R. T., Bradshaw, R. A., Banaszak L. J. (1982). Amino acid sequence homology among the 2-hydroxyacid dehydrogenase, mitochondrial and cytoplasmic malate dehydrogenase form a homologous system with lactate dehydrogenase. Proc. Natl. Acad. Sci. USA 79, 6166-6170.

Birktoft, J. J., Banaszak, L. J. (1983). The presence of a histidine-aspartic acid pair in the active site of 2-hydroxyacid dehydrogenases. X-ray refinement of cytoplasmic malate dehydrogenase. J. Biol. Chem. 258, 472-482.

Boernke, W. E., Sanville-Millard, C., Wilkins-Stevens, P., Kakar, S. N., Stevens, F. J., Donnelly, M.I. (1995). Stringency of substrate specificity of *Escherichia coli* malate dehydrogenase. Arch. Biochem. Biophys. 332, 43-52.

Boucher, J. I., Jacobowitz, J. R., Beckett, B. C., Classen, S., Theobald, D. L. (2014). An atomic-resolution view of neofunctionalization in the evolution of apicomplexan lactate dehydrogenases. ELife 3.

Brochier-Armanet C, Madern D. (2021). Phylogenetics and biochemistry elucidate the evolutionary link between l-malate and l-lactate dehydrogenases and disclose an intermediate group of sequences with mix functional properties. Biochimie. 191:140-153. doi: 10.1016/j.biochi.2021.08.004. Epub 2021 Aug 18. PMID: 34418486.

Bur, T. Clarke, J.D. Friesen, M. Gold, K.W. Hart, J.J. Holbrook, J.B. Jones, M.A. Luyten, H.M. Wilks, On the effect on specificity of Thr246----Gly mutation in L-lactate dehydrogenase of Bacillus sterothermophilus, Biochemical and biophysical research communications, 161 (1989) 59-63.

Burgner, J. W., Ray, W. J. (1984). On the origin of lactate dehydrogenase induced rate effect. Biochemistry 23, 3636-3648.

Cendrin, F., Chroboczek, J., Zaccai, G., Eisenberg, H.,Mevarech, M. (1993). Cloning, sequencing and expression in *Escherischia coli* of the gene coding for malate dehydrogenase of the extremely halophilic archaebacterium *Haloarcula marismortui*. Biochemistry 32, 4308-4313.

Chen, V.B., Arendall, W. B., Headd, J.J., Keedy, D. A., Immormino, R. M., Kapral, G. J. Murray, L. W., Richardson, J. S., Richardson, DC. (2010). *MolProbity*, all-atom structure validation for macromolecular crystallography. Acta Cryst. D. 66, 12-21.

Clarke, A. R., Wigley, D. B., Chia, W. N., Barstow, D., Atkinson, T., Holbrook J. J. (1986). Site-directed mutagenesis reveals role of mobile arginine residue in lactate dehydrogenase catalysis. Nature. 324, 699-702.

Clarke, A. R., Wilks, H. M., Barstow, D. A., Atkinson, T., Chia, W. N., Holbrook, J. J. (1988). An investigation of the contribution made by the carboxylate group of an active site histidine-aspartate couple to binding and catalysis in lactate dehydrogenase. Biochemistry 27, 1617-1622.

Colletier, J. P., Aleksandrov, A., Coquelle, N., Mraihi, S., Mendoza-Barberá, E., Field, M., Madern, D. (2012). Sampling the conformational energy landscape of a hyperthermophilic protein by engineering key substitution. Mol. Biol. Evol. 29, 1683-1694.

Coquelle N, Talon R, Juers DH, Girard E, Kahn R, Madern D. Gradual adaptive changes of a protein facing high salt concentrations. J Mol Biol. 2010 Dec 3;404(3):493-505.

Coquelle, N., Fioravanti, E., Weik, M., Vellieux, F., Madern, D. (2007). Activity, Stability and Structural Studies of Lactate Dehydrogenases Adapted to Extreme Thermal Environments. J. Mol Biol. 374, 547–562.

Cowtan, K. (1999). Error estimation and bias correction in phase-improvement calculations. Acta Cryst. D. 55, 1555-1567.

Cowtan K. (2006).The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr. 62(Pt 9):1002-11. doi: 10.1107/S0907444906022116. Epub 2006 Aug 19. PMID: 16929101.

Cowtan K. (2008) Fitting molecular fragments into electron density. Acta Crystallogr D Biol Crystallogr. 64(Pt 1):83-9. doi: 10.1107/S0907444907033938. Epub 2007 Dec 5. PMID: 18094471; PMCID: PMC2394793.

Cowtan K. (2010). Recent developments in classical density modification. Acta Crystallogr D Biol Crystallogr. 66(Pt 4):470-8. doi: 10.1107/S090744490903947X. Epub 2010 Mar 24. PMID: 20383000; PMCID: PMC2852311.

Dalhus, B., Saarinen, M., Sauer, U. H., Eklund, P., Johansson, K., Karlsson, A., Ramaswamy, S., Bjørk, A., Synstad, B., Naterstad, K., Sirevåg, R., Eklund, H. (2002). Structural Basis for Thermophilic Protein Stability, Structures of Thermophilic and Mesophilic Malate Dehydrogenases. J. Mol. Biol. 318, 707-721.

Darden, T., York, D., Pedersen, L., (1993). Particle mesh Ewald - an N.log(N) method for Ewald sums in large systems. J Chem Phys. 98, 10089-10092.

Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W. Arendall, W. B., Snoeyink, J., Richardson, J. S., Richardson, D. C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 35, W375-W383.

Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. (2010). Features and development of Coot. Acta Cryst. D. 66, 486-501.

Engilberge, S., Riobé, F., Wagner, T., Pietro, S.D., Breyton, C., Franzetti, B., Shima, S., Girard, E., Dumont, E., Maury, O. (2018). Unveiling the Binding Modes of the Crystallophore, a Terbium-based Nucleating and Phasing Molecular Agent for Protein Crystallography. Chemistry. 24, 9739-9746.

Engilberge, S., S., Riobé, F., Di Pietro, S., Lassalle, L., Coquelle, N., Arnaud, C. A., Pitrat, D. , Mulatier, J. C., Madern, D., Breyton, C., Maury, O., Girard, E. (2017). Crystallophore, a versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. Chem Sci 8, 5909–5917.

Evans, P. (2006). Scaling and assessment of data quality. Acta Cryst. D. 62, 72-82.

Eventoff, W., Rossmann, M. G.,Taylor, S. S., Torff, H. J., Meyer, H., Keil, W., Kiltz, H. H. (1977). Structural adaptations of lactate dehydrogenase isozymes. Proc. Natl. Acad. Sci. USA 74, 2677-2268.

Everse, J. and Kaplan, NO. 1973. Lactate dehydrogenases: structure and function. *Adv. Enzym. Mol. Biol.* 37, 61–133.

Fersht, A. (1985). Enzyme structure and mechanism, 2nd edition. Freeman and Co., New York.

González, J. M., Marti-Arbona, R., Chen, J. C. H., Broom-Peltz, B., Unkefer, C. J. (2018). Conformational changes on substrate binding revealed by structures of Methylobacterium extorquens malate dehydrogenase. Acta Crys. F 74, 610–616.

Harms, M. J., Thornton, J. W. (2010). Analyzing protein structure and function using ancestral gene reconstruction. Curr. Opin. Struct. Biol. 20, 360-366.

Hart, K. W., Clarke, A. R., Wigley, D. B., Chia, W. N., Barstow, D. A., Atkinson, T., Holbrook, J. J. (1987). The importance of arginine 171 in substrate binding by *Bacillus stearothermophilus* lactate dehydrogenase. Bio. Bioph. Res. Com. 146, 346–353.

Hart, K. W., Clarke, A. R., Wigley, D. B., Waldman, A. D. B., Chia, W. N., Barstow, D. A., Atkinson, T., Jones, J. B., Holbrook, J. J. (1987). A strong carboxylate-arginine interaction is important in substrate orientation and recognition in lactate dehydrogenase. B. B. A. 914, 294-298.

Hess, B., Bekker, H., Berendsen, H. J. C., Fraaije, J. (1997). LINCS: a linear constraint solver for molecular simulations. J Comput Chem. 18, 1463-1472.

Holbrook, J. J., Liljas, A., Steindel, S. J., Rossmann, M. G. (1975) Lactate Dehydrogenase. P.D. Boyer (Ed.), The Enzymes (3rd Edn.), Academic Press, New York , pp. 191-292

Holland LZ, McFall-Ngai M, Somero GN. 1997. Evolution of lactate dehydrogenase-A homologs of barracuda fishes (genus Sphyraena) from different thermal environments: differences in kinetic properties and thermal stability are due to amino acid substitutions outside the active site. Biochemistry. 36(11):3207-15.

Holm, L., Laakso, L. M. (2016). Dali server update. Nucl. Ac. Res. 44, 351-355.

Hoover, W. G. (1985) Canonical dynamics: Equilibrium phase-space distributions. Phys. Rev. A 31, 695-1697.

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M, de Groot, B. L., Grubmuller, H., and MacKerell, A.D. Jr. (2016). CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. Nature Methods 14, 71-73.

Iacovino LG, Rossi M, Di Stefano G, Rossi V, Binda C, Brigotti M, Tomaselli F, Pasti AP, Dal Piaz F, Cerini S, Hochkoeppler A. (2022) Allosteric transitions of rabbit skeletal muscle lactate dehydrogenase induced by pH-dependent dissociation of the tetrameric enzyme. Biochimie. 2022199:23-35.

Ikehara, Y., Arai, K., Furukawa, N., Ohno, T., Miyake, T., Fushinobu, S., Nakajima, M., Miyanaga, A., Taguchi, H. (2014). The Core of Allosteric Motion in *Thermus caldophilus* l-Lactate Dehydrogenase. J. Biol. Chem. 289, 31550-31564.

Iorio A, Roche J, Engilberge S, Coquelle N, Girard E, Sterpone F, Madern D. (2021) Bio-chemical, structural and dynamical studies reveal strong differences in the thermal-dependent allosteric behavior of two extremophilic lactate dehydrogenases. J. Struct. Biol. 213(3), 107769.

Iorio A, Brochier-Armanet C, Mas C, Sterpone F, Madern D. (2022). Protein Conformational Space at the Edge of Allostery: Turning a Nonallosteric Malate Dehydrogenase into an "Allosterized" Enzyme Using Evolution-Guided Punctual Mutations. Mol. Biol. Evol. 1;39(9):msac186. doi: 10.1093/molbev/msac186. PMID: 36056899; PMCID: PMC9486893.

Irimia, A., Ebel, C., Madern, D;, Richard, S. B., Cosenza, L. W., Zaccai, G., Vellieux, F. M. (2003). The Oligomeric states of Haloarcula marismortui malate dehydrogenaseare modulated by solvent components as shown by crystallographic and biochemical studies. J. Mol. Biol. 326, 859-873.

Iwata, S., Kamata, K., Yoshida, S., Minowa, T., Ohta, T. (1994). T and R states in the crystals of bacterial L-lactate dehydrogenase reveal the mechanism for allosteric control. Nat. Struct. Biol. 1, 176-185.

Jiang T., Roux A., Engilberge S., Alsalman Z., Di Pietro S., Franzetti B., Riobé F., Maury O. & Girard E. (2020). Tracking crystallophore nucleating properties: setting up a database for statistical analysis. Crystal Growth & Design, 20(8), 5322-532.

Jubb HC, Higueruelo AP, Ochoa-Montaño B, Pitt WR, Ascher DB, Blundell TL.(2016) Arpeggio: A Web Server for Calculating and Visualising Interatomic Interactions in Protein Structures. J Mol Biol. 429(3), 365-371.

Kabsch W. (2010). XDS. Acta Crystallogr. D. 66,125–132.

Kalimeri M., Rahaman O., Melchionna S., Sterpone F. (2013). How Conformational Flexibility Stabilizes the Hyperthermophilic Elongation Factor G-Domain. J. Phys. Chem. B 117, 13775-13785.

Kalimeri M., Girard E., Madern D., Sterpone F. (2014). Interface Matters, the Stiffness Route to Stability of a Thermophilic Tetrameric Malate Dehydrogenase. PLoS One 9(12), e113895.

Katava M, Kalimeri M, Stirnemann G, Sterpone F. (2016). Stability and Function at High Temperature. What Makes a Thermophilic GTPase Different from Its Mesophilic Homologue. J Phys Chem B. 120(10):2721-30. doi: 10.1021/acs.jpcb.6b00306. Epub 2016 Mar 8. PMID: 26907829.

Katava, M., Maccarini, M., Villain, G., Paciaroni, A., Sztucki, M., Ivanova, O., Madern, D., Sterpone, F. (2017). Thermal activation of 'allosteric-like' large-scale motions in a eukaryotic Lactate Dehydrogenase. Sci. Rep. 7, 41092.

Katava M, Marchi M, Madern D, Sztucki M, Maccarini M, Sterpone F. (2020). Temperature Unmasks Allosteric Propensity in a Thermophilic Malate Dehydrogenase via Dewetting and Collapse. J Phys Chem B. 2020 124(6):1001-1008. doi: 10.1021/acs.jpcb.9b10776. Epub 2020 Jan 30. PMID: 31961162.

Kolappan S, Shen DL, Mosi R, Sun J, McEachern EJ, Vocadlo DJ, Craig L. (2015). Structures of lactate dehydrogenase A (LDHA) in apo, ternary and inhibitor-bound forms. Acta Crystallogr D Biol Crystallogr. 71(Pt 2):185-95.

LeVan KM, Goldberg E. 1991. Properties of human testis-specific lactate dehydrogenase expressed from Escherichia coli. *Biochem J.* 273 (Pt 3)(Pt 3):587-592.

Madern, D. (2000). The putative L-lactate dehydrogenase from Methanococcus jannaschii is a NADPH-dependent L-malate dehydrogenase. Mol Microbiol. 37, 1515-1520

Madern, D. (2002). Molecular Evolution within the L-Malate and L-Lactate Dehydrogenase Super-Family. J. Mol. Evol. 54, 825-840

Madern, D., Cai, X. M., Abrahamsen, M. S., Zhu, G. (2004). Evolution of Cryptosporidium parvum lactate dehydrogenase from malate dehydrogenase by a very recent event of gene duplication. Mol. Biol. Evol. 21, 489-497.

Maffucci I , Laage D , Stirnemann G , Sterpone F . (2020). Differences in thermal structural changes and melting between mesophilic and thermophilic dihydrofolate reductase enzymes. Phys Chem Chem Phys. 22(33):18361-18373. doi: 10.1039/d0cp02738c. Epub 2020 Aug 13. PMID: 32789320.

Marsh, J. A., Teichmann, S. A. (2014). Protein flexibility facilitates quaternary structure assembly and evolution. Plos ONE 12(5), e1001870.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., Read, R. J. (2007). *Phaser* crystallographic software. J. Appl Cryst. 40, 658-674.

Meng W, Clerico EM, McArthur N, Gierasch LM.(2018). Allosteric landscapes of eukaryotic cytoplasmic Hsp70s are shaped by evolutionary tuning of key interfaces. Proc Natl Acad Sci U S A. 115(47):11970-11975.

Miton CM, Buda K, Tokuriki N.(2021). Epistasis and intramolecular networks in protein evolution. Curr Opin Struct Biol. 69:160-168.

Murshudov GN, Vagin AA, Dodson EJ. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. 1;53(Pt 3):240-55. doi: 10.1107/S0907444996012255. PMID: 15299926.

Nosé, S. (1984). A unified formulation of the constant temperature molecular dynamics methods. J. Chem. Phys. 81, 511-519.

Olson CA, Wu NC, Sun R. (2014) A comprehensive biophysical description of pairwise epistasis throughout an entire protein domain. Curr Biol. 24(22):2643-51.

Parrinello, M., Rahman, A., (1981). Polymorphic transitions in single crystals: a new molecular dynamics method. J Appl Phys. 52, 7182-7190.

Pesce A, Fondy TP, Stolzenbach F, Castillo F, Kaplan NO. (1967). The comparative enzymology of lactic dehydrogenases. 3. Properties of the H4 and M4 enzymes from a number of vertebrates. *J Biol Chem*. 242(9):2151-2167.

Roche J, Girard E, Mas C, Madern D. (2019). The archaeal LDH-like malate dehydrogenase from Ignicoccus islandicus displays dual substrate recognition, hidden allostery and a non-canonical tetrameric oligomeric organization. J Struct Biol. 208(1):7-17. doi: 10.1016/j.jsb.2019.07.006. Epub 2019 Jul 10. PMID: 31301348.

Sterpone F, Bertonati C, Briganti G, Melchionna S. (2009). Key role of proximal water in regulating thermostable proteins. J Phys Chem B. 113(1):131-7. doi: 10.1021/jp805199c. PMID: 19072709.

Sterpone F, Melchionna S. (2012). Thermophilic proteins: insight and perspective from in silico experiments. Chem Soc Rev. 41(5):1665-76. doi: 10.1039/c1cs15199a. Epub 2011 Oct 5. PMID: 21975514; PMCID: PMC3775309.

Taguschi, H. (2017). The Simple and Unique Allosteric Machinery of Thermus caldophilus lactate dehydrogenase, structure-function relathionship in bacterial allosteric LDHs. Adv. Exp. Med. Biol. 925,117-145.

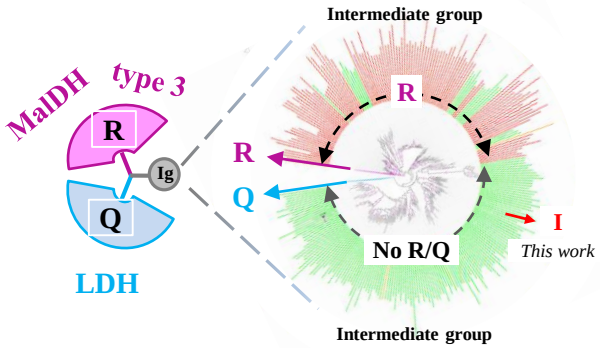
Tokuriki N, Tawfik DS. (2009). Protein dynamism and evolvability. Science. 324(5924):203-7. doi: 10.1126/science.1169375. PMID: 19359577.

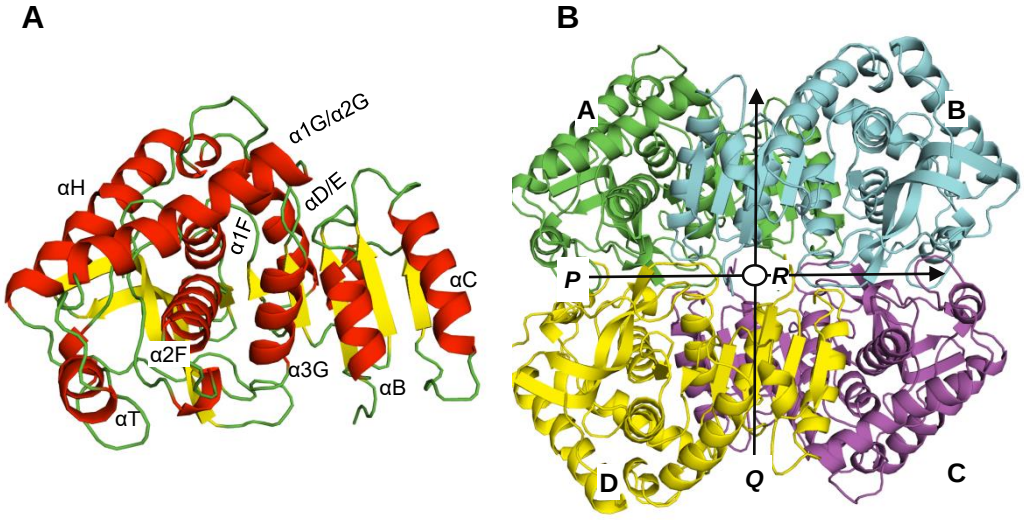
Waldman A. D. B., Hart K. W., Clarke A. R., Wigley D. B., Barstow D. A., Atkinson T., Chia W.N., Holbrook J. J. (1988). The use of a genetically engineered tryptophan to identify the movement of a domain of *Bacillus stearothermophilus* lactate dehydrogenase with the process which limits the steady-state turnover of the enzyme. Bio. Bioph. Res. Com. 150, 752-759.

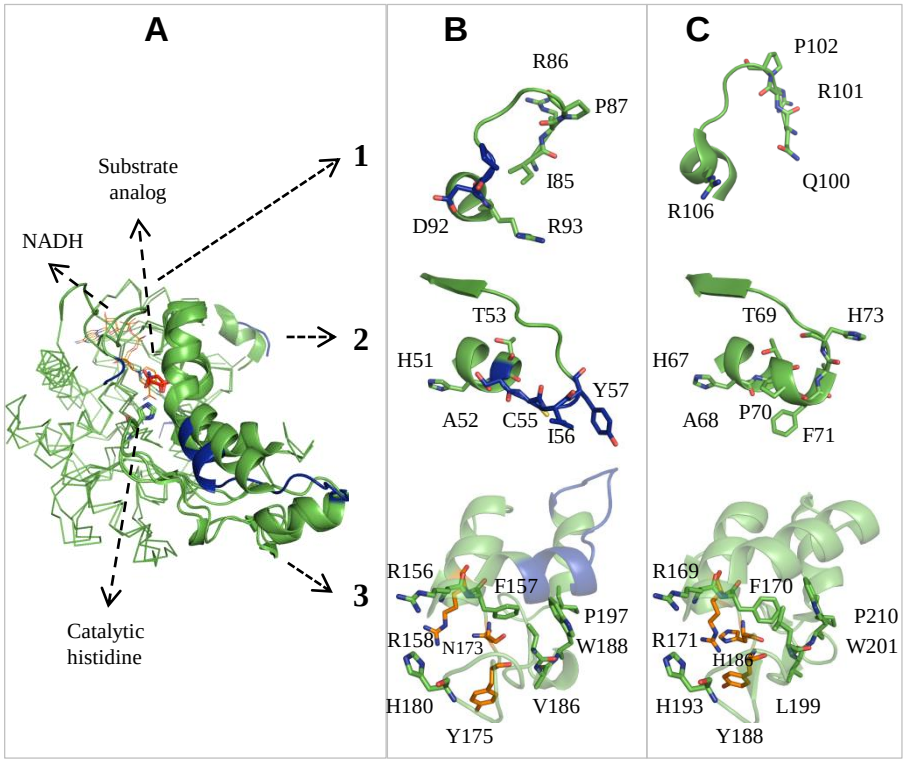
Wilks H. M., Hart K. W., Feeney R., Dunn C. R., Muirhead H., Chia W. N., Barstow D. A., Atkinson T., Clarke A. R., Holbrook J. J. (1988). A specific, highly active malate dehydrogenase by redesign of a lactate dehydrogenase framework. Science 242, 1541–1544.

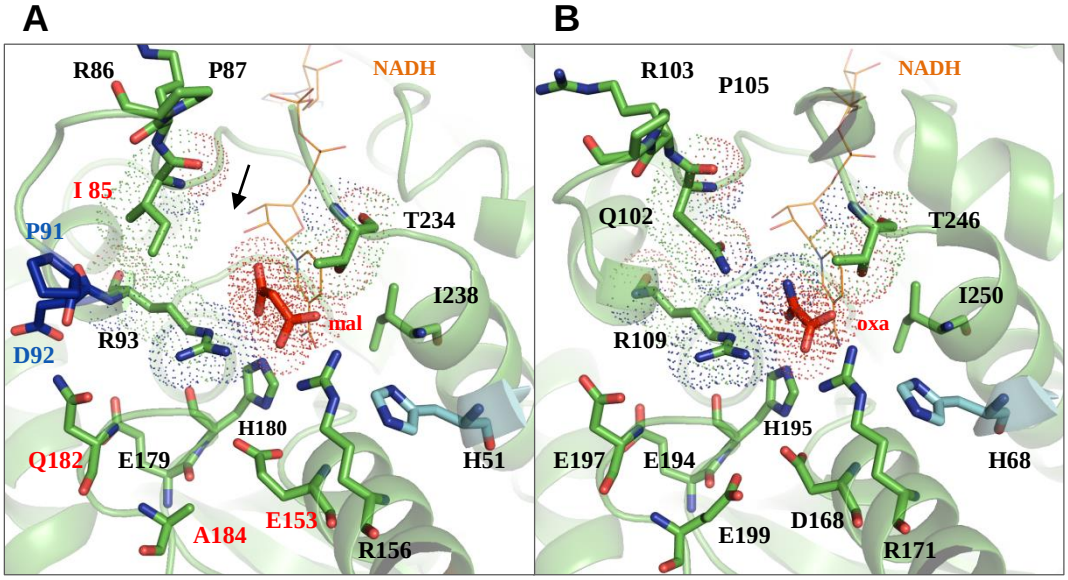
Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD, Wishart DS. (2003). VADAR: a web server for quantitative evaluation of protein structure quality. Nucleic Acids Res. 2003 31(13):3316-9.

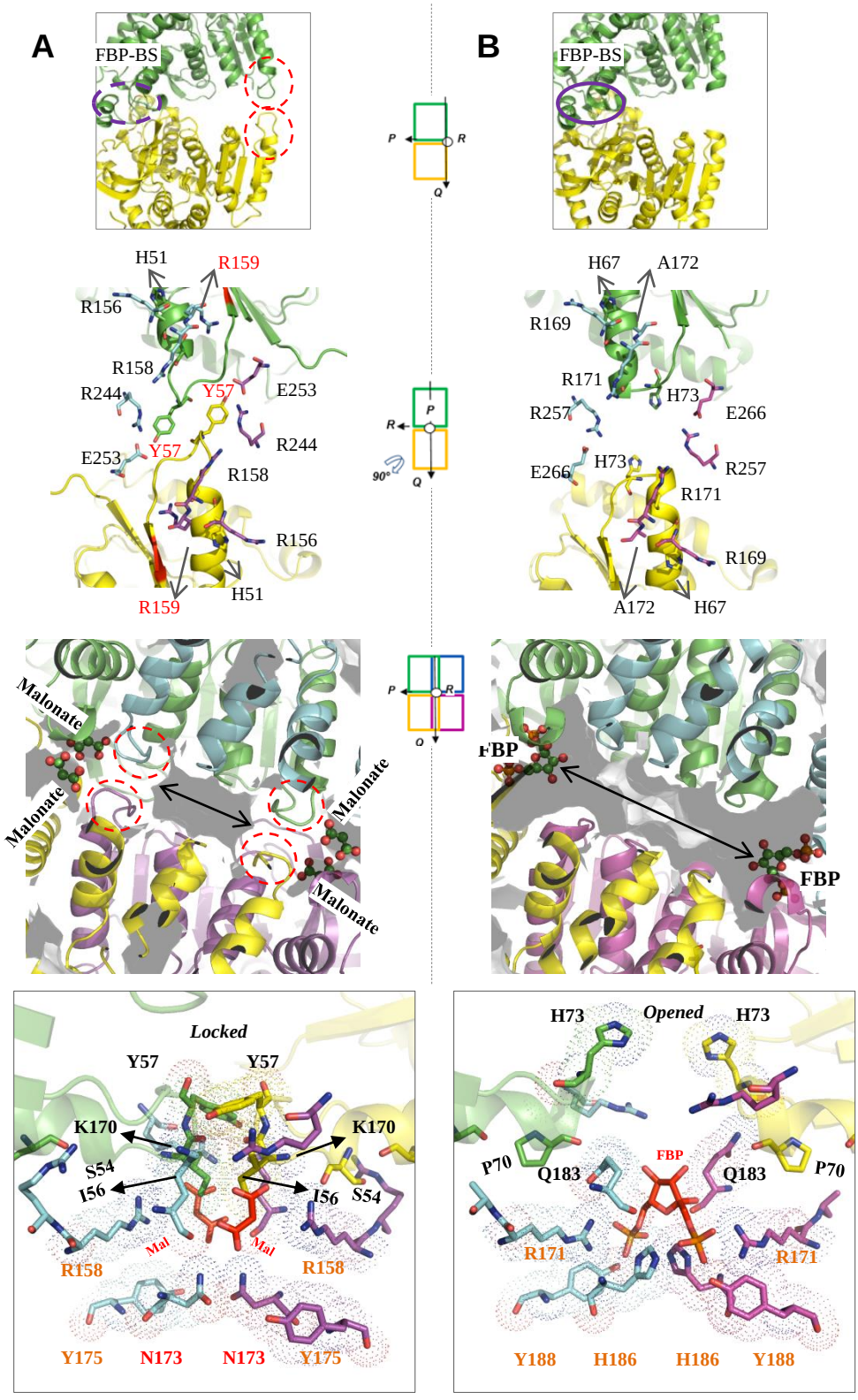
Wilson D., Pethica R., Zhou Y., Talbot C., Vogel C., Madera M., Chothia C., Gough J. (2009). SUPERFAMILY-sophisticated comparative genomics, data mining, visualization and phylogeny. Nucleic. Acids. Res. 37, 380-386.

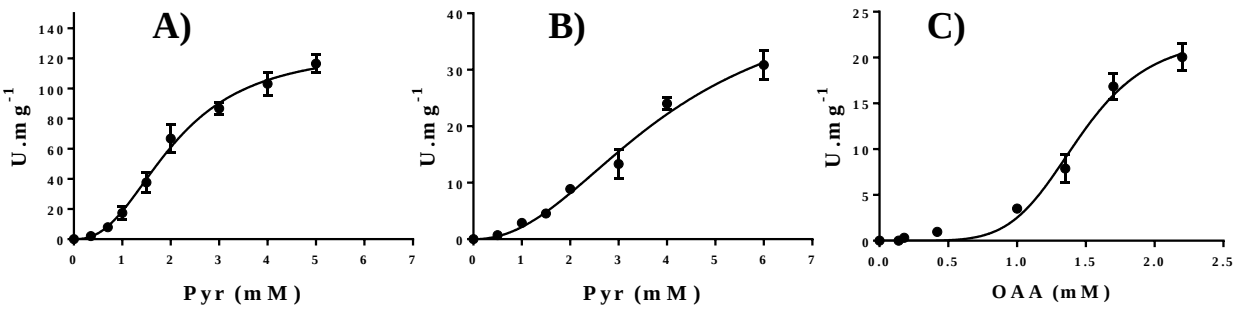

MalDH
type 3
R
Ig
Q
LDH
Intermediate group
R
R
Q
No R/Q
I
This work
Intermediate group

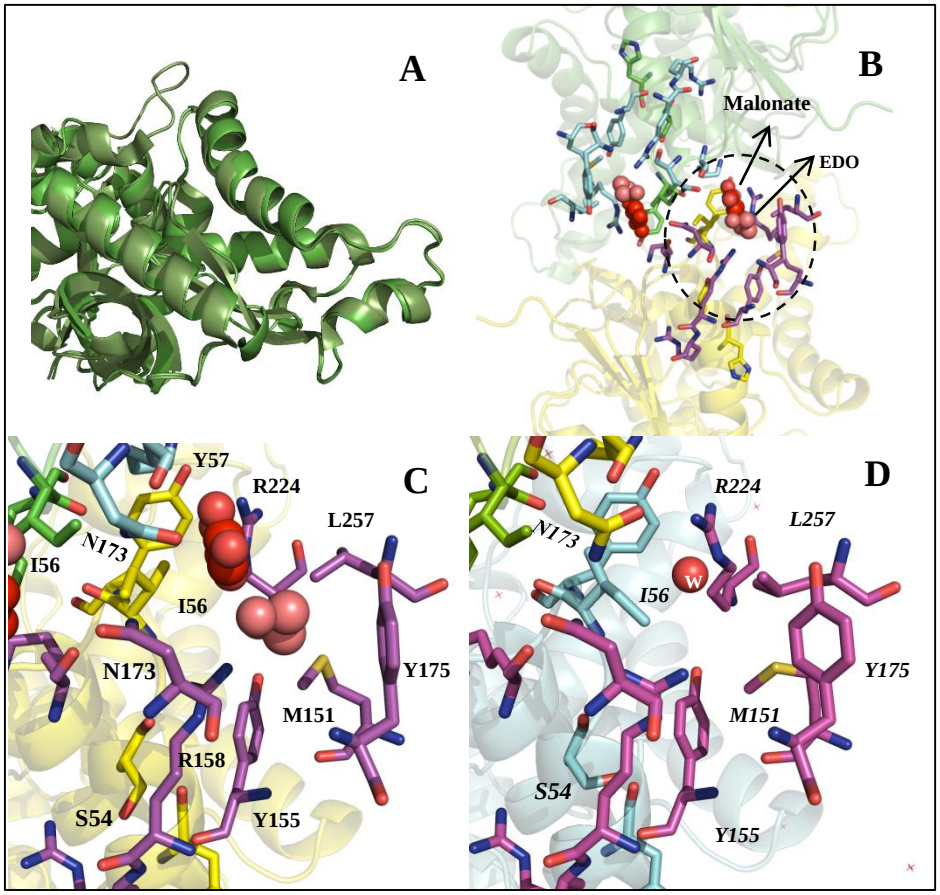

A
α1G/α2G
αH
αD/E
α1F
αC
α2F
α3G
αB
αT
B
A
B
P
R
Q
D
C

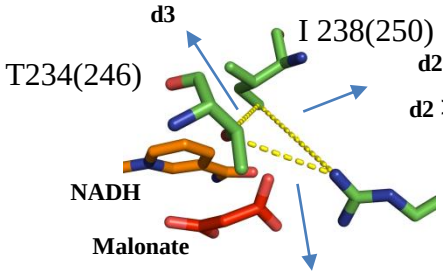

A
Substrate analog
NADH
1
2
3
Catalytic histidine
B
R86
P87
I85
D92
R93
T53
H51
Y57
A52
C55
I56
R156
F157
P197
R158
N173
W188
H180
V186
Y175
C
P102
R101
Q100
R106
T69
H73
H67
A68
P70
F71
R169
F170
P210
R171
W201
H186
H193
L199
Y188


A
R86
P87
NADH
I 85
P91
T234
I238
mal
R93
D92
H180
Q182
E179
H51
E153
A184
R156
B
R103
P105
NADH
Q102
T246
I250
oxa
R109
H195
E197
E194
H68
D168
E199
R171


A
FBP-BS
B
FBP-BS
P
R
Q
H51
R159
R156
R158
R244
Y57
E253
R244
E253
Y57
R158
R156
R159
H51
H67
A172
R169
R171
H73
R257
E266
E266
H73
R257
R171
R169
A172
H67
90°
Malonate
Malonate
Malonate
Malonate
FBP
FBP
Locked
Y57
Y57
K170
K170
S54
I56
I56
S54
Mal
Mal
R158
R158
Y175
N173
N173
Y175
Opened
H73
H73
P70
Q183
FBP
Q183
P70
R171
R171
Y188
H186
H186
Y188


A)
B)
C)
U.mg-1
Pyr (mM)
OAA (mM)


A
B
Malonate
EDO
C
Y57
R224
L257
N173
I56
I56
N173
Y175
M151
R158
S54
Y155
D
R224
L257
N173
I56
W
Y175
M151
S54
Y155



d2 < 7.5 A in favor of R171 Cs-In

d2 > 7.55 A in favor of R171 Cs-Out

d1 < 4.5 A in favor of R171 Cs-In

d1 > 4.5 A in favor of R171 Cs-Out