

Structural and mutational characterization of a malate racemase from the LarA superfamily

Santhosh Gatreddi^{1,2}, Julian Urdiain-Arraiza³, Benoit Desguin³, Robert P. Hausinger^{1,2*}, Jian Hu^{1,4*}

¹Department of Biochemistry and Molecular Biology

²Department of Microbiology and Molecular Genetics, Michigan State University, MI 48824

³Institute of Biomolecular Science and Technology, Université Catholique de Louvain, B-1348 Louvain-La-Neuve, Belgium

⁴Department of Chemistry, Michigan State University, MI 48824

*Corresponding authors: hausinge@msu.edu; hujian1@msu.edu

Abstract

The LarA superfamily consists of nickel-dependent enzymes catalyzing racemization/epimerization reactions using a variety of α -hydroxy acids. The first-characterized LarA, a lactate racemase from *Lactobacillus plantarum*, led to the discovery of the nickel-pincer nucleotide (NPN) cofactor that is utilized by family members with alternative substrates, including malate racemase from *Thermoanaerobacterium thermosaccharolyticum* (Mar2). In this work, a higher resolution crystal structure of Mar2 was obtained with better data quality that revealed new structural and dynamic characteristics of the protein. A model of the Mar2 structure with bound cofactor and substrate was generated to uncover the common and the unique features among two distinct subgroups in the LarA superfamily. In addition, structure-guided mutational studies were used to examine the importance of residues that are modeled to interact with NPN and to explore which residues were critical for conferring specificity for malate. In particular, substitution of two residues involved in substrate binding in Mar2 to match the corresponding residues in LarA led to the acquisition of low levels of lactate racemase activity. Of additional interest, the substrate spectrum was expanded to include tartrate, an analog of malate. These new findings will help to better understand structure-function relationships of many other LarA homologs that are broadly distributed in bacterial and archaeal species.

Keywords: Malate racemase · Nickel · Nickel-pincer nucleotide · LarA · Tartrate epimerase · Mutagenesis

Introduction

The lactate racemase protein (LarA) binds the recently discovered nickel-pincer nucleotide (NPN) cofactor and catalyzes the stereo inversion of D- and L- lactic acid (Desguin et al. 2014, 2015). The NPN cofactor is sequentially synthesized by the LarB, LarE, and LarC accessory proteins from the precursor nicotinic acid adenine dinucleotide (NaAD) by consecutive carboxylation/hydrolysis, sulfur insertion, and metal transfer reactions, respectively (Desguin et al. 2014, 2015, 2018, 2017, 2016; Fellner et al. 2017; Hausinger et al. 2018). LarA homologues (LarAHs) in the LarA superfamily have been demonstrated to perform racemization or epimerization reactions using a variety of α -hydroxycarboxylic acids (Desguin et al. 2020). Among the 13 LarAH subgroups that have been identified through genomic context analysis and biochemical characterization, representatives from two groups exhibit malate racemase activity.

Malate racemase was previously reported in the cell-free extracts of *Rhodobacter capsulatus* that had been supplemented with D-malate (Martinez-Luque et al. 2001), but the enzyme responsible was not characterized. Within the LarA superfamily, Mar1 from *Desulfitobacterium hafniense* and Mar2 from *Thermoanaerobacterium thermosaccharolyticum* were shown to catalyze racemization of L-malate (Desguin et al. 2020). Mar1 exhibited a high K_M for L-malate in an *in vitro* assay (55 ± 6 mM), consistent with its natural substrate likely being a compound similar to, but distinct from, L-malate. In contrast, Mar2 exhibited a K_M of only 0.38 ± 0.04 mM, supporting the notion that L-malate is likely to be the physiological substrate of this new NPN-dependent enzyme.

The previous study of Mar2 noted that when its gene was co-expressed in *Lactococcus lactis* with those encoding the NPN synthesizing enzymes from *Lactobacillus plantarum* (an expression system that produces fully functional LpLarA (Desguin et al. 2014) when using *L. plantarum* larA), negligible levels of nickel were detected after treating the purified protein sample with the chromophoric chelator 4-(2-pyridylazo)-resorcinol (Desguin et al. 2020). The absence of nickel was further evidenced by the lack of electron density for the NPN cofactor in the Mar2 crystal structure determined at 2.38 Å (PDB ID: 6D6Z) (Desguin et al. 2020). As a result, it was necessary to add *in vitro* synthesized NPN cofactor to the apoprotein to reconstitute malate racemase activity. Compared to a rotational shift that was observed between the open and closed states of LpLarA (PDB ID: 5HUQ), a more aggressive rotation of the C-terminal domain of Mar2 relative to the N-terminal domain enabled it to adopt an extended conformation (Desguin et al. 2015, 2020). To better understand the interactions among Mar2, the NPN cofactor, and the substrate, we advanced the structural studies in this work and report a new structure of Mar2 at an improved resolution at 2.25 Å. Based on the more complete structural model, we discuss potential

implications of cofactor and substrate binding to Mar2. In addition, we extended the structural information by characterizing selected enzyme variants. We confirm the importance of NPN-binding residues and highlight residues that confer specificity for malate. Finally, we expand the substrate spectrum of Mar2 by reporting that it also exhibits epimerization activity for tartrate, a malate analog. These findings are likely to shed light on the mechanistic study of other LarAHs.

Materials and methods

Expression and purification

L. lactis NZ3900 cells bearing the pGIR211 plasmid containing *T. thermosaccharolyticum* *mar2* and *L. plantarum* *larBCDE* were grown overnight on M17 broth (BD Difco) supplemented with 0.5% w/v glucose and 10 µg/mL of chloramphenicol without shaking at 30 °C (Desguin et al. 2020). The culture was diluted to 1% with 4 L of the same medium having 5 µg/mL of chloramphenicol and incubated at 30 °C until reaching OD₆₀₀ = 0.4, at which point 1 mM NiCl₂ and 5 µg/L nisin A (to induce expression) were added, followed by incubation at 28 °C with shaking at 40 rpm. After 4 h, the culture was cooled to 4 °C and the cells were collected by centrifugation at 6,000 *g* for 10 min. The pellet was stored at -80 °C until further use.

The cell pellet was thawed and resuspended in 80 mL of lysis buffer containing 100 mM Tris (pH 7.5), 300 mM NaCl, 0.5 mM Na₂SO₃, 1 mM phenylmethylsulfonyl fluoride, 10 µg/ml of lysozyme, and 2 µg/ml of DNase. Cell lysis was performed by sonicating the sample for 15 min and the supernatant was separated by centrifuging the lysate at 18,500 *g* for 1 h. Clarified supernatant was loaded onto 2 ml of Strep-Tactin Sepharose (IBA) resin equilibrated with 10 ml of wash buffer (WB) composed of 100 mM Tris (pH 7.5), 300 mM NaCl, and 0.05 mM Na₂SO₃. After loading, the resin was washed with 20 ml of WB followed by elution with 15 ml of WB containing 5 mM desthiobiotin (Sigma-Aldrich).

Crystallization and structure solution

To prepare the protein for crystallization, the pooled fractions from Strep-Tactin affinity chromatography were further purified by gel filtration chromatography using a Superdex 200 Increase 10/300 GL column that was equilibrated with 50 mM Tris (pH 7.5), 125 mM NaCl, and 0.05 mM Na₂SO₃. The eluted fractions from a monomeric peak were concentrated with Amicon ultra-15 centrifugal filter units.

In attempts to obtain the substrate-bound crystals of Mar2, the protein was concentrated to 10 mg/mL, mixed with 3.8 mM L-malate, and incubated at 4° C overnight. In some cases, nicotinic acid mononucleotide (NaMN) was added. Crystallization screening was performed via the sitting drop vapor diffusion method using a Mosquito® robot (SPT Labtech Inc.). Plate-shaped crystals appeared after 24 h in several conditions. Crystals suitable for diffraction were reproduced by hanging drop vapor diffusion using an optimized condition containing 100 mM HEPES (pH 8.0), 200 mM NaCl, and 30% PEG 3350. The crystals were cryo protected in a solution containing 100 mM HEPES (pH 8.0), 200 mM NaCl, and 35% PEG 3350, then flash frozen in liquid nitrogen.

Diffraction data for the Mar2 crystals were collected at the Life Sciences Collaborative Access Team (LS-CAT) beamline 21-ID-F at the Advanced Photon Source, Argonne National Laboratory. The data were indexed with iMOSFLM (Battye et al. 2011) and scaled with SCALA (Evans 2006) from the CCP4 program suite. The initial phases were obtained by performing molecular replacement in PHENIX (Liebschner et al. 2019) using the model 6D6Z. Model building was performed using COOT (Emsley et al. 2010) and refinement was carried out in phenix.refine (Liebschner et al. 2019). The maps were generated using Phenix and structure images were generated using Pymol v 2.4.1 (Schrodinger) (DeLano 2002). The structure coordinates were deposited in the PDB with accession ID 7S91.

Generation of a structural model of Mar2 with bound NPN and substrate

To generate a structural model of Mar2 in the closed conformation, the N- and C-domains of Mar2 were separately superimposed onto the corresponding domains of *LpLarA* in the closed conformation (PDB ID: 5HUQ, chain B) using COOT (Emsley et al. 2010). The coordinates for the NPN cofactor were imported from *LpLarA* into the Mar2 structure, after which D- and L-malate were built at the position occupied by sulfate in the structure of *LpLarA*. Minor clashes generated from the modeling were removed by manually adjusting the side chain rotamers of the clashed residues.

Site-directed mutagenesis and racemization/epimerization activity measurement

pGIR319 was constructed by PCR amplification of *mar2*, digestion with restriction enzymes NcoI and NheI, and ligation into a similarly digested plasmid pGIR076 (Fellner et al. 2017). Site-directed mutations were generated by using the QuikChange kit (Agilent) and confirmed by DNA sequencing (**Table S1**). TOP10 *E. coli* cells were induced to express Mar2 when the OD₆₀₀ reached 0.5 with 0.2 % L-arabinose and the cells were harvested 4 h later. Cell lysates were obtained by mixing a cell suspension (100-fold concentrated) with glass beads and homogenizing

with a cell disruptor, followed by centrifugation at 20,000 *g* for 15 min at 4 °C. 5 µL of cell lysate was incubated with 2.5 µL of *in vitro* synthesized NPN (Desguin et al. 2016) and 30 mM D-malate in 50 µL final volume for 10 min at 45 °C. The reactions with purified enzymes were performed with 25 µL of *in vitro* synthesized NPN (Desguin et al. 2016) and 200 mM D-lactate in 50 µL final volume for 30 min at 35 °C. The reactions were then stopped by incubation at 90 °C for 5 min, and L-malate or L-lactate was assayed spectrophotometrically with the corresponding Megazyme assay kits.

To detect tartrate epimerization activity, variable amounts of L- and D-tartrate (5 mM to 400 mM), 5 µM of *in vitro* synthesized NPN and 1 µM of purified Mar2 were incubated for 15 min at 37 °C in a 100 µL final volume in a reaction buffer containing 50 mM potassium phosphate, pH 6.0. After heat-inactivation for 10 min at 90 °C, 100 µL of water was added, and the sample was loaded onto a polyvinyl alcohol-coated capillary (55 cm in length, internal diameter of 50 µm, from Agilent) and run for 20 min on a Capel 105 M from Lumex Instrument at 20 °C using -25 kV. The background electrolyte was 10 mM benzoic acid/L-histidine, pH 5.0, with 5 mM vancomycin as chiral separator. Reaction products were detected at 230 nm.

Results and discussion

Overall structure of Mar2

The structure of Mar2 solved previously at 2.38 Å did not contain the cofactor even though the enzyme was co-expressed with the NPN synthesizing enzymes in *L. lactis* (Desguin et al. 2020). To better understand the structure-function relationships of Mar2, we attempted to co-crystallize the protein with the cofactor analogue NaMN and the identified substrate L-malate, added separately and simultaneously. Despite our extensive efforts, Mar2 failed to bind either NaMN or L-malate in any of the identified crystallization conditions. During these efforts to obtain structures of the complexes, we identified several new crystallization conditions. One such condition contained 0.1 M D/L-malate as the crystallization buffer, yet no malate-bound structure was obtained. This was the case even when racemic D/L-malate was replaced with L-malate at higher concentration (>0.1 M). Although no substrate-bound structure was obtained, an alternative crystallization condition allowed us to solve the structure at an improved resolution of 2.25 Å and with better data quality (**Fig. 1** and **Table 1**).

Similar to the previously solved structure of Mar2 (6D6Z) (Desguin et al. 2020), the new protein crystals are in space group I222 with one protein molecule per asymmetric unit. The two Mar2

structures are highly superimposable with a C α RMSD of 0.35 Å, but the improved resolution allowed us to model more residues (331-332, 346-351) that were missed in the previous structure. The new structure also revealed blocks of electron density at the supposed catalytic site that could not be interpreted as either NPN cofactor or malate (**Fig. S1**). Given that the protein was eluted from the Strep-Tactin resin using desthiobiotin (DTB), we attempted to model a DTB molecule near K181, corresponding to the lysine residue in *LpLarA* (K184) that covalently associates with the NPN cofactor. Our analysis of the corresponding difference and polder maps indicated the density does not match DTB; thus, we left the protein active site unoccupied.

Structural features revealed by the new Mar2 structure

Comparison between the new and old Mar2 structures reveals two significant structural differences (**Fig. 2a**). The first variance involves the loop containing residues 69-77 that adopts distinct conformations in the two structures (**Fig. 2b**) (Desguin et al. 2020). In *LpLarA*, the corresponding loop is critical for cofactor association as the ribose and phosphate moieties of the NPN cofactor are stabilized by a network of hydrogen bonds, particularly from residues Asp72, Thr74, and Arg75 (Desguin et al. 2015, 2020). Accordingly, this loop is referred to hereafter as the cofactor-binding loop. The structural difference can be primarily attributed to *cis-trans* isomerization of Pro75; it is in a *cis* configuration in 6D6Z (Desguin et al. 2020), whereas a *trans* configuration is observed in the new structure. Notably, the corresponding residue in *LpLarA* (Pro76) is in the *trans* configuration. As a result, the fold of the cofactor-binding loop in the new Mar2 structure is similar to its counterpart in *LpLarA* and likely adopts a conformation that similarly allows for cofactor binding. In contrast, the *cis* configuration of Pro75 in 6D6Z makes the loop potentially clash with the incoming cofactor. As indicated by analysis of B-factors (**Fig. 2c**), the cofactor-binding loop exhibits the highest degree of disorder in the N-terminal domain for both structures, indicative of significant local dynamics that are likely to influence cofactor association with Mar2. The observed conformational flexibility could be attributed to differences in the amino acid composition of the loop and the local environment, possibly leading to the lack of the cofactor. Related to these differences, the overall sequence identity between *LpLarA* and Mar2 is only 31%.

The second difference between the two Mar2 structures involves an α -helix, extended in the new structure, that is critically involved in substrate binding (**Fig. 2d**). The previous sequence and structural analysis had predicted the most probable substrate-binding residues (Tyr172, Lys173, Gln293 and Lys351) in Mar2 (Desguin et al. 2020); however, Lys351 was missing from the previous model due to the lack of electron density for the region 331-351. With the higher resolution and better data quality for the new structure, we were able to narrow the gap and now

the missing substrate-binding residue K351 is shown in an extended α -helix, allowing us to generate a structural model with bound malate together with the modeled NPN cofactor in the active site (see the next section). In addition, the improved electron density map helped to correct a register shift mistake for residues 352-360, which were mislabeled by +1 in residue number.

Model of the Mar2 structure with bound NPN cofactor and D- or L-malate

As observed in the structures of *LpLarA* (PDB IDs: 5HUQ and 6C1W), the active site in the closed conformation is sandwiched by the N- and C-terminal domains to provide the binding site for both the NPN cofactor and the substrate (Desguin et al. 2015; Rankin et al. 2018). Accordingly, the corresponding domains in Mar2 that are apart in the cofactor-unbound state must approach each other upon the binding of cofactor, leading to a closed conformation to enable catalysis (Desguin et al. 2015). Using the closed conformation of *LpLarA* (5HUQ, chain B) as template, Mar2 in the putative active conformation was generated by rotating the C-terminal domain relative to the N-terminal domain to reach the best structural superimposition with the C-terminal domain of *LpLarA*. The NPN cofactor and the substrate D- or L-malate were sequentially modeled into the active site to better understand substrate binding and catalysis (**Fig. 3**).

In the model structure of the Mar2 complex, the NPN cofactor occupies nearly the same cavity as what was observed in *LpLarA*. The residues thought to be critical for the binding of the cofactor, including Lys181 that is topologically equivalent to Lys184 forming a thioamide bond with the NPN in *LpLarA*, His197 coordinating the nickel ion, Pro185 stacking with the pyridinium ring via an aromatic-proline interaction, Asp71 and Thr73 that hydrogen bond with the ribose, and Arg74 and Ser177 associating with the phosphate, are conserved and structurally aligned with the corresponding residues in *LpLarA* (**Fig. 3a** and **3b**). Asn167 in Mar2 also appears to be involved in binding the phosphate of the cofactor, whereas the topologically equivalent residue in *LpLarA* (Phe170) does not. The two catalytic histidine residues, His108/His174 in *LpLarA* and His108/His171 in Mar2, are structurally aligned, although the distance between the histidine pair in Mar2 is longer by 2.5 Å (10.6 Å vs 8.1 Å) presumably due to the absence of the cofactor and/or the substrate. Given that the large majority of the key residues in *LpLarA* for cofactor binding and catalysis are conserved and structurally aligned with the corresponding residues in Mar2, we postulate that Mar2 utilizes the same mechanism to catalyze racemization reactions to interconvert D- and L-malate.

To model D- or L-malate on top of the modeled NPN cofactor, the C1 carboxyl group of malate was placed in a position close to the sulfate in *LpLarA* so that Arg74, Gln291, and Lys294 were able to interact with it (**Fig. 3a** and **3b**). The counterparts of these residues in *LpLarA* (Arg75,

Gln295, and Lys298) directly participate in sulfate binding. The C2 atom was positioned exactly above C4 of the pyridinium ring in the cofactor to facilitate the hydride transfer reaction according to the proposed mechanism of *LpLarA* (Rankin et al. 2018). The 2-OH of D-malate was modeled to directly interact with His171 (His174 in *LpLarA*), whereas the 2-OH of L-malate appears to best interact with Tyr290 (Tyr294 in *LpLarA*). However, it is possible that the hydroxyl group of L-malate also forms a hydrogen bond with His108 (also His108 in *LpLarA*) in the cofactor-bound state when the two catalytic histidine residues approach each other. Four unique residues conserved only in Mar2 and its close homologs (**Fig. 3c**) include Tyr172 and Lys173, from the N-terminal domain, and Gln293 and Lys351, from the C-terminal domain, that form a polar and positively-charged patch facing the C4 carboxyl group of malate. The corresponding residues in *LpLarA* are hydrophobic or small polar residues (Phe175, Phe176, V297, and Thr359) that are probably involved in recognizing the methyl group of lactic acid, the substrate of that enzyme. Another residue potentially interacting with the C4 carboxyl group is His350, which replaces Trp358 in *LpLarA*. The current model also shows that the nickel ion from the NPN cofactor appears to be close enough to form an ionic bond with the C4 carboxylic acid (~3.5 Å) to further increase substrate-binding affinity. The putative anion-aromatic interaction between the C1 carboxylic acid and the pyridinium of the cofactor and the potential C4 carboxylic acid-nickel interaction may explain why a substrate bound structure could not be obtained in the absence of the cofactor.

Structure-guided mutagenesis and activity measurement of the variants

Next, based on the structure model, we generated site-directed substitutions on selected residues to explore the reaction mechanism. Alanine substitutions were generated for Asp71 and His197, which are modeled to interact with the NPN cofactor, and for His108 which is believed to be a catalytic residue (**Fig. 4a**). The corresponding residues in *LpLarA* (Asp72, His108, and His200) have been shown to be required for activity (Desguin et al. 2015). As expected, none of the three variants showed detectable activity, supporting the notion that Mar2 catalyzes racemization using the same mechanism as *LpLarA*. Three lysine residues in the active site of Mar2 were substituted to examine their roles in reaction. Lys294 has been modeled to form a salt bridge with the C1 carboxylic acid of malate (**Fig. 3a**), and this conserved residue in *LpLarA* (Lys298) interacts with the bound sulfate (**Fig. 3b**). Consistent with the proposed role of Lys294 in substrate binding, the K294A variant showed no activity. Lys173 and Lys351 are less conserved in the LarA family (**Fig. 3c**) but were postulated to be involved in conferring substrate specificity (Desguin et al. 2020). To test this hypothesis, the lysine residues were substituted with the corresponding amino acids in

known lactate racemases (*LpLarA* or *LarA* from *Thermoanaerobacterium thermosaccharolyticum*, *TtLarA*), and the resulting variants (K173F, K351T, K351E, and their combinations) were examined for malate racemase (Mar) activity. While the Mar activity of the K173F variant was modestly reduced, the K351T and K351E variants completely lost activity. Accordingly, the positive charge of Lys351 appears to be indispensable, whereas Lys173 only plays a minor role in malate racemization (**Fig. 4a**). Of great interest, when the K173F and K351T (or K351E) substitutions were combined, a small but substantial lactate racemase (Lar) activity was detected for the purified enzymes (**Fig. 4b**), supporting the notion that these residues are involved in determining substrate specificity. Although the lactate racemization activities of the double variants are quite low, likely due to the disruption of the active site caused by the substitutions, this result demonstrates that the substrate specificity of Mar2 can be modulated as desired.

Expansion of the substrate specificity of Mar2

Given the structural similarity between malate and tartrate, we tested whether the latter can be processed by Mar2. Using D- or L-tartrate as substrate, we were able to detect meso-tartrate after incubation with the purified Mar2, indicative of epimerization of either chiral center by Mar2 (**Fig. 5a**). Kinetics study showed that the D- and L- enantiomers were processed with similar values of k_{cat} , but the enzyme slightly prefers the latter as indicated by the smaller K_M and therefore larger specificity constant (k_{cat}/K_M) (**Fig. 5b**). As the k_{cat}/K_M values for tartrate (0.74 and 0.42 s⁻¹·mM⁻¹ for L- and D-tartrate, respectively) are two orders of magnitude smaller than those for malate (185 s⁻¹·mM⁻¹) (Desguin et al. 2020), it is unlikely that tartrate is a natural substrate of Mar2. Nevertheless, the expanded substrate specificity demonstrates that the active site of Mar2 is flexible enough to accommodate structurally similar chemicals as substrates. Indeed, in the structure model of Mar2/malate complex (**Fig. 3a**), the residues surrounding C3 of malate, which would be associated with the hydroxyl group in tartrate, are either polar or charged. With a minor structural rearrangement to avoid steric clash, it would be possible to allow tartrate binding and processing in the same manner as malate in the active site.

Conclusions

In the present study, we conducted structural characterization of Mar2 from *T. thermosaccharolyticum* by solving its structure with a higher resolution and better data quality. The conformational flexibility of the cofactor-binding loop is deemed to be a crucial factor affecting the NPN association with the enzyme even when the binding pocket is widely open and fully

accessible. The significant local dynamics caused by the *cis-trans* isomerization of Pro75 may account for the lack of the NPN in the protein during coexpression of the gene for Mar2 with those encoding the cofactor-producing enzymes. The dynamics may also result in the inability to load the NPN analog (NaMN) into the protein during crystallographic experiments. The structural model of the Mar2-NPN-malate complex highlights both the common and the unique features in substrate recognition for distinct subgroups in the LarA superfamily, which, when combined with mutational studies, indicates that Mar2 and *Lp*LarA share the same reaction mechanism while the active site residues unique to each subgroup are likely involved in determining substrate specificity. The demonstrated expandability and tunability of the substrate spectrum may inspire protein engineering of LarA homologs for potential applications, e.g. producing high value α -hydroxyl acids.

Acknowledgements

This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817). This work is supported by National Science Foundation Grant CHE 1807073 (to R.P.H. and J.H.), National Institutes of Health Grant GM128959 (to R.P.H. and J.H.), and Fonds de la Recherche Scientifique (FNRS) (to B.D. and J.U.). We thank Alessia Citti for her participation in the construction of the Mar2 variants.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr.* 67:271-281. <https://doi.org/10.1107/S0907444910048675>
- DeLano WL (2002) The PyMOL molecular graphics system. <http://www.pymol.org>.
- Desguin B, Goffin P, Viaene E, Kleerebezem M, Martin-Diaconescu V, Maroney MJ, Declercq JP, Soumilion P, Hols P (2014) Lactate racemase is a nickel-dependent enzyme activated by a widespread maturation system. *Nat Commun.* 5:3615. <https://doi.org/10.1038/ncomms4615>
- Desguin B, Zhang T, Soumilion P, Hols P, Hu J, Hausinger RP (2015) A tethered niacin-derived pincer complex with a nickel-carbon bond in lactate racemase. *Science.* 349:66-69. <https://doi.org/10.1126/science.aab2272>
- Desguin B, Fellner M, Riant O, Hu J, Hausinger RP, Hols P, Soumilion P (2018) Biosynthesis of the nickel-pincer nucleotide cofactor of lactate racemase requires a CTP-dependent cyclometallase. *J Biol Chem.* 293:12303-12317. <https://doi.org/10.1074/jbc.RA118.003741>
- Desguin B, Soumilion P, Hausinger RP, Hols P (2017) Unexpected complexity in the lactate racemization system of lactic acid bacteria. *FEMS Microbiol Rev.* 41:S71-S83. <https://doi.org/10.1093/femsre/fux021>
- Desguin B, Soumilion P, Hols P, Hausinger RP (2016) Nickel-pincer cofactor biosynthesis involves LarB-catalyzed pyridinium carboxylation and LarE-dependent sacrificial sulfur insertion. *Proc Natl Acad Sci U S A.* 113:5598-5603. <https://doi.org/10.1073/pnas.1600486113>
- Desguin B, Urdiaín-Arraiza J, Da Costa M, Fellner M, Hu J, Hausinger RP, Desmet T, Hols P, Soumilion P (2020) Uncovering a superfamily of nickel-dependent hydroxyacid racemases and epimerases. *Sci Rep.* 10:18123. <https://doi.org/10.1038/s41598-020-74802-6>
- Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* 66:486-501. <https://doi.org/10.1107/S0907444910007493>
- Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr.* 62:72-82. <https://doi.org/10.1107/S0907444905036693>
- Fellner M, Desguin B, Hausinger RP, Hu J (2017) Structural insights into the catalytic mechanism of a sacrificial sulfur insertase of the N-type ATP pyrophosphatase family, LarE. *Proc Natl Acad Sci U S A.* 114:9074-9079. <https://doi.org/10.1073/pnas.1704967114>
- Hausinger RP, Desguin B, Fellner M, Rankin JA, Hu J (2018) Nickel-pincer nucleotide cofactor. *Curr Opin Chem Biol.* 47:18-23. <https://doi.org/10.1016/j.cbpa.2018.06.019>
- Liebschner D *et al* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol.* 75:861-877. <https://doi.org/10.1107/S2059798319011471>
- Madeira F *et al* (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47:W636-W641. <https://doi.org/10.1093/nar/gkz268>
- Martínez-Luque M, Castillo F, Blasco R (2001) Assimilation of D-malate by *Rhodobacter capsulatus* E1F1. *Curr Microbiol.* 43:154-157. <https://doi.org/10.1007/s002840010279>
- Rankin JA, Mauban RC, Fellner M, Desguin B, McCracken J, Hu J, Varganov SA, Hausinger RP (2018) Lactate racemase nickel-pincer cofactor operates by a proton-coupled hydride transfer mechanism. *Biochemistry.* 57:3244-3251. <https://doi.org/10.1021/acs.biochem.8b00100>
- Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42:W320-324. <https://doi.org/10.1093/nar/gku316>

Table 1. Crystallographic statistics

Data collection	Mar2
Beamline	LS-CAT 21-ID-F
Wavelength (Å)	0.97872
Space group	I 2 2 2
Unit cell a, b, c (Å); α , β , γ (°)	72 100 132; 90, 90, 90
^a Resolution (Å)	29.38 – 2.25 (2.37 – 2.25)
Unique reflections	23,233 (3,344)
^a Redundancy	4.0 (4.1)
^a Completeness (%)	99.9 (100)
^a $I/\sigma I$	7.3 (1.7)
^{a,b} R_{merge}	0.114 (0.844)
^{a,c} R_{pim}	0.063 (0.475)
^d $CC_{1/2}$	0.994 (0.549)
Refinement	
Protein atoms	3,050
Cl	2
H ₂ O molecules	88
^e R_{work}/R_{free}	0.206/0.239
B -factors (Å ²)	41.6
Protein atoms	41.6
Cl	53.30
H ₂ O molecules	38.5
R.m.s. deviation in bond lengths (Å)	0.008
R.m.s. deviation in bond angles (°)	1.05
Ramachandran plot (%) favored	96.7
Ramachandran plot (%) outliers	0
Rotamer outliers	0
PDB ID	7S91

^aHighest resolution shell is shown in parentheses.

^b $R_{merge} = \sum_{hkl} \sum_j |I_j(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_j I_j(hkl)$, where I is the intensity of reflection.

^c $R_{pim} = \sum_{hkl} [1/(N-1)]^{1/2} \sum_j |I_j(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_j I_j(hkl)$, where N is the redundancy of the dataset.

^d $CC_{1/2}$ is the correlation coefficient of the half datasets.

^e $R_{work} = \sum_{hkl} | |F_{obs}| - |F_{calc}| | / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} is the observed and the calculated structure factor, respectively. R_{free} is the cross-validation R factor for the test set of reflections (5% of the total) omitted in model refinement.

Figure legends

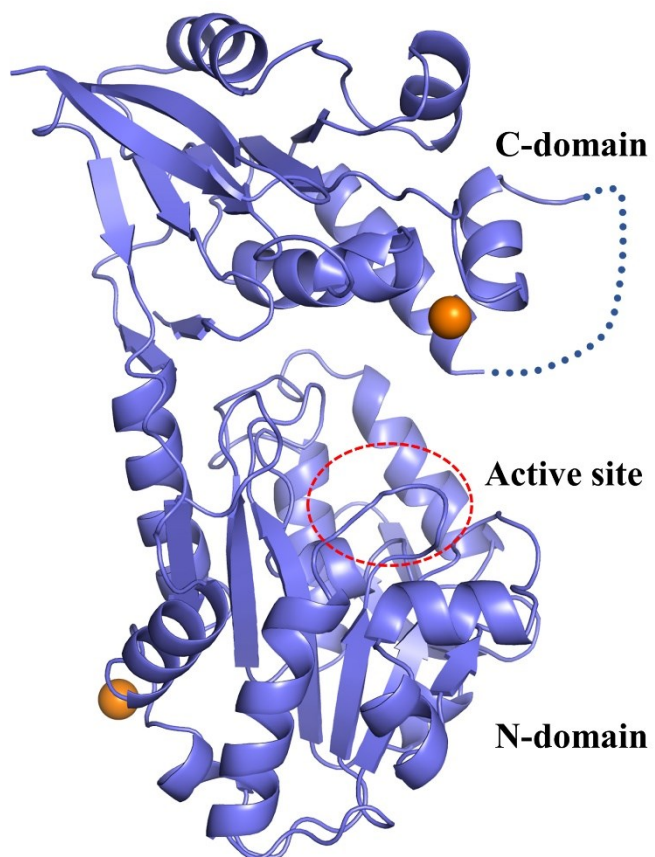


Fig. 1 Structure of the malate racemase Mar2. The protein is shown in cartoon mode with bound chloride shown as orange spheres. The dashed line indicates a region of missing residues (333-345). The putative active site is highlighted in the red circle.

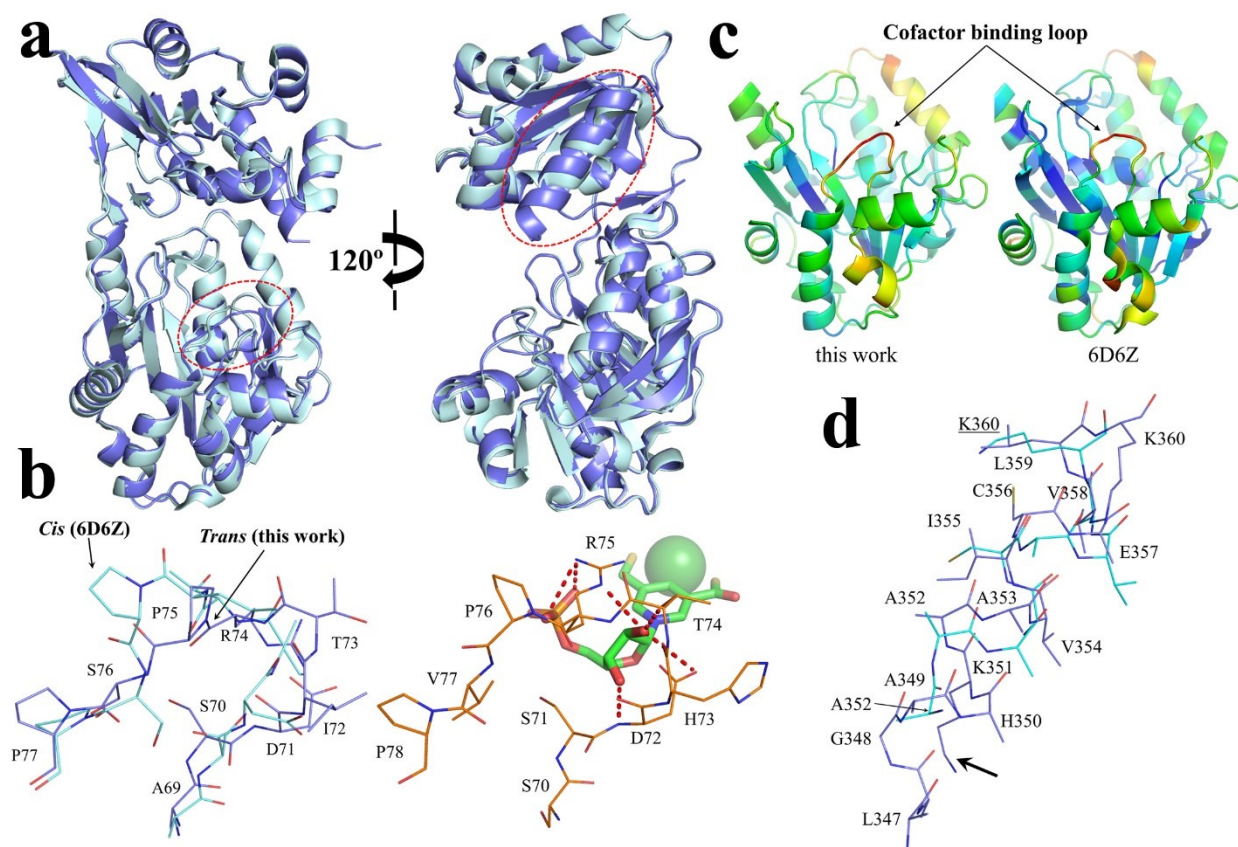


Fig. 2 Structural comparison of the Mar2 structures. **(a)** The Mar2 structure solved in this work (blue) was superimposed with the previously solved structure (PDB ID: 6D6Z, cyan). Structural differences observed in the cofactor-binding loop (left) and the α -helix involved in substrate binding (right) are indicated in red circles. **(b)** *Cis-trans* isomerization of Pro75 in the cofactor-binding loop (residues 69-77). *Left*: The loop from the current Mar2 structure and 6D6Z are represented with blue and cyan sticks, respectively. *Trans* (new) and *cis* (6D6Z) conformations of Pro75 are indicated with arrows. *Right*: the cofactor-binding loop of *LpLarA* (orange) that interacts with the NPN cofactor (stick mode in green) are depicted by red dotted lines. **(c)** Mar2 structures colored by the *B*-factor of $C\alpha$ atoms. For clarity, only the residues in the N-terminal domain (residues 41-265) are shown in cartoon mode. Blue and red indicate low and high *B*-factor, respectively. **(d)** Comparison of the α -helix involved in substrate binding through Lys351. The α -helix (residue 347-360) in the new Mar2 structure (blue) contains the predicted substrate binding Lys351 (indicated by the arrow), which was not modeled in 6D6Z (cyan) due to poor electron density map. A register shift mistake, as revealed by the mis-assigned Lys360 (underlined in 6D6Z), is corrected in the new structure model.

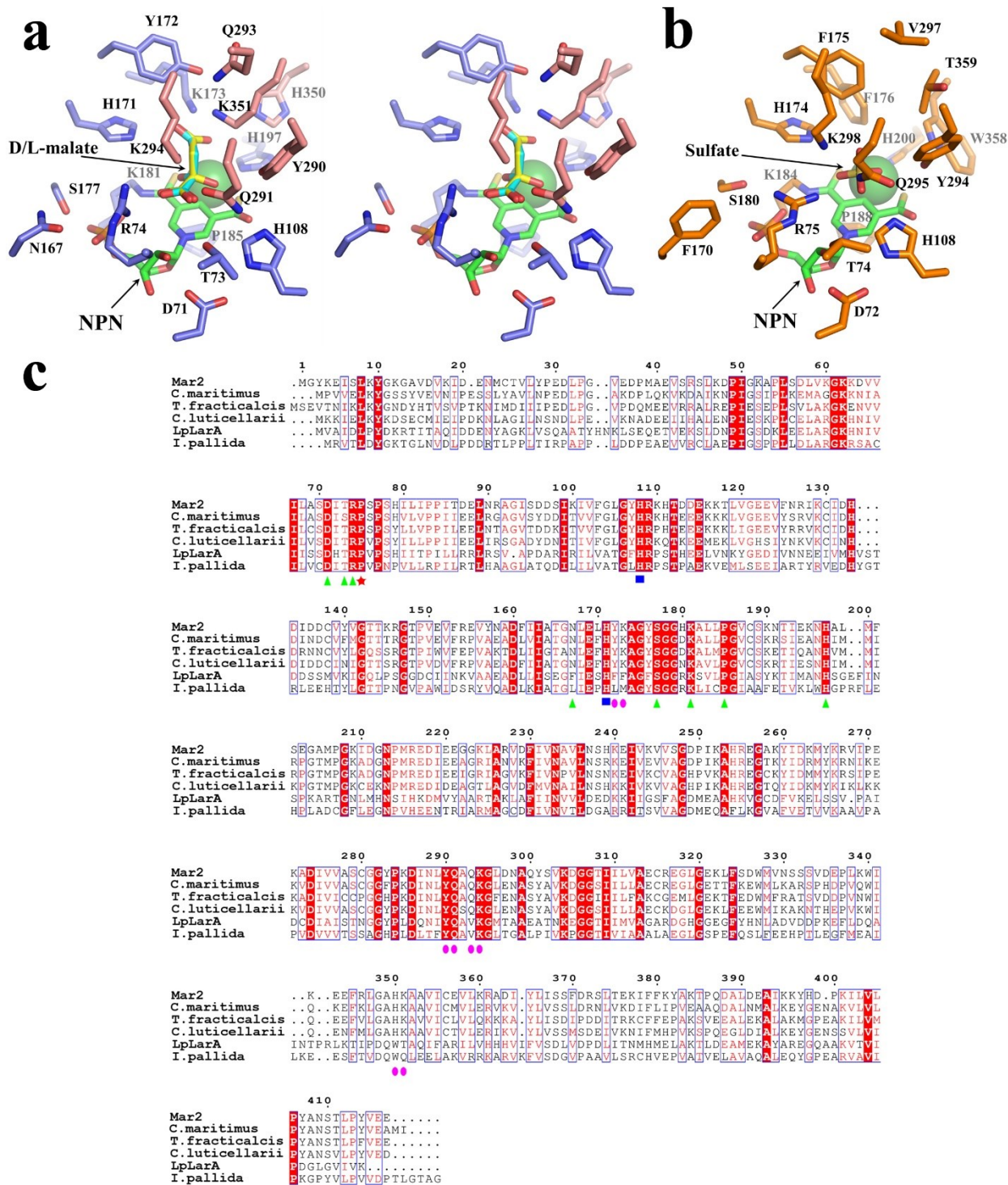


Fig. 3 Structural model of the Mar2 active site with bound NPN cofactor and D/L-malate. (a) Stereo view of the active site of Mar2. The NPN cofactor (green), D- or L-malate (yellow and cyan, respectively), and the residues involved in cofactor or substrate binding are depicted in stick mode. Only the sidechains are shown for clarity, some of which were manually adjusted to avoid clashes. The residues in the N-terminal region are shown in blue, and those in the C-terminal domain are in pink. (b) The active site of *LpLarA* with bound sulfate (5HUQ, chain B). The residues

topologically equivalent to those shown in (A) are labeled and depicted in stick mode. (c) Multiple sequence alignment of Mar2 with *LpLarA* and their close homologs. Sequences aligned are as follows: Mar2, homologs with sequence identity between 60-70% from *Calderihabitans maritimus*, *Thermoanaerobacterium fracticalcis*, and *Clostridium laticellarii*, *LpLarA*, and its close homolog *Isosphaera pallida*. The residues predicted to be involved in the binding of cofactor and D/L-malate are indicated with triangles (green) and circles (pink), respectively. The two catalytic histidine residues are denoted with squares (blue), whereas the proline residue undergoing *cis-trans* isomerization is indicated by a star (red). Sequence alignment was performed with Clustal omega (Madeira et al. 2019) and the figure was generated with ESPript 3.0 (Robert et al. 2014).

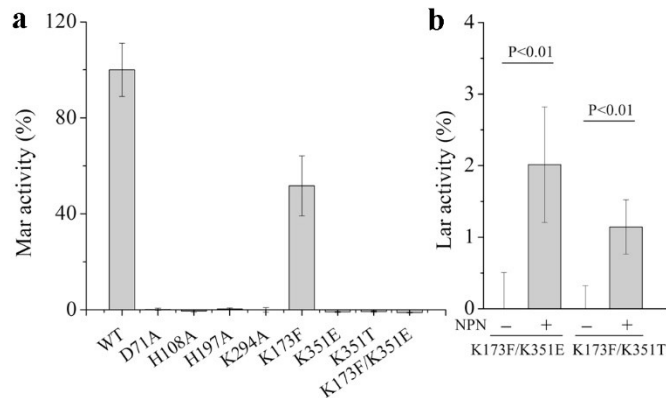


Fig. 4 Substitution of selected residues at the active site of Mar2. (a) Mar activity of the Mar2 variants in cell-free extracts (n=6-7). (b) Lar activity of the purified Mar2 variants (n=4). Activity is presented as percentage of the wild type Mar2 for Mar activity or *TtLarA* for Lar activity. The error bars represent S.D. Student's t-tests were conducted to examine whether the difference between the samples with and without added NPN were statistically significant.

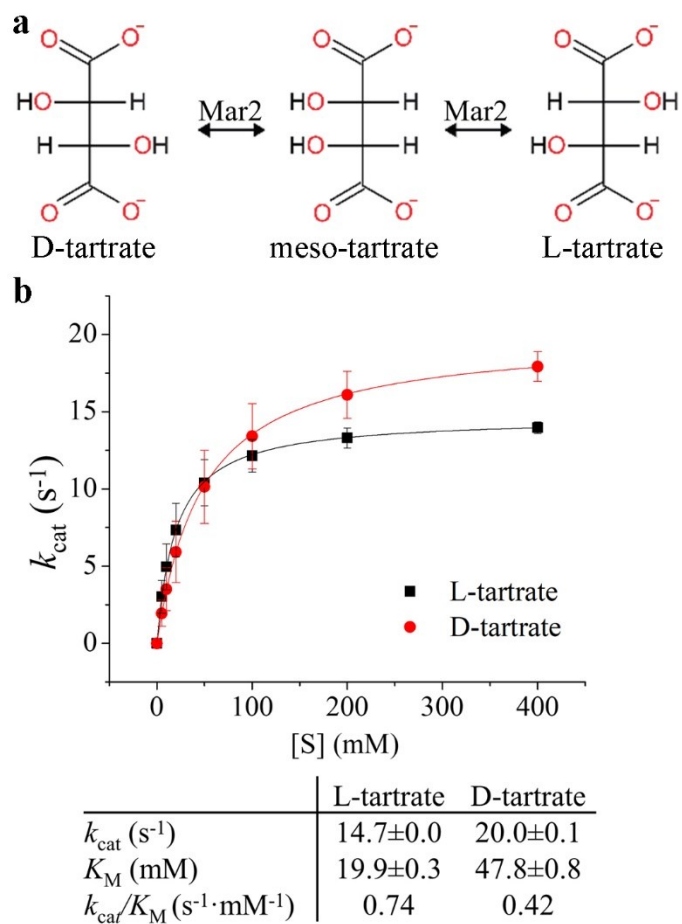


Fig. 5 Tartrate epimerization by Mar2. **(a)** Mar2 catalyzed interconversion of D-/meso-/L-tartrate. **(b)** Kinetics of tartrate epimerization by Mar2. The kinetic parameters were obtained by curve fitting using the Michaelis-Menten equation in Origin. The error bars indicate S.D. (n=3).