Published in final edited form as:

Biochemistry. 2018 November 13; 57(45): 6434–6442. doi:10.1021/acs.biochem.8b00913.

Functional and Structural Resilience of the Active Site Loop in the Evolution of *Plasmodium* Lactate Dehydrogenase

Jacob D. Wirth, Jeffrey I. Boucher[†], Joseph R. Jacobowitz[‡], Scott Classen[§], and Douglas L. Theobald^{*}

Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02454, United States

Abstract

The malarial pathogen *Plasmodium falciparum* (*Pf*) is a member of the Apicomplexa, which independently evolved a highly specific lactate dehydrogenase (LDH) from an ancestral malate dehydrogenase (MDH) via a five-residue insertion in a key active site loop. *Pf*LDH is widely considered an attractive drug target because of its unique active site. The conservation of the apicomplexan loop suggests that a precise insertion sequence was required for the evolution of LDH specificity. Aside from a single critical tryptophan, W107f, the functional and structural roles of residues in the loop are currently unknown. Here we show that the loop is remarkably robust to mutation, as activity is resilient to radical perturbations of both loop identity and length. Thus, alternative insertions could have evolved LDH specificity as long as they contained a tryptophan in the proper location. *Pf*LDH likely has great potential to develop resistance to drugs designed to target its distinctive active site loop.

Graphical Abstract

^{*}**Corresponding Author:** Department of Biochemistry, Brandeis University, 415 South St., Waltham, MA 02454. Telephone: 781-736-2303. dtheobald@brandeis.edu.

[†]J.I.B.: Ginko Bioworks, 27 Drydock Ave., 8th Floor, Boston, MA 02210.

[‡]J.R.J.: Department of Biology, Massachusetts Institute of Technology, Building 68-132, 31 Ames St., Cambridge, MA 02142.

[§]S.C.: Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720.

Author Contributions

J.D.W., J.I.B., and J.R.J. contributed equally to this work.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.8b00913.

Kinetic parameters for all enzymes described (Tables S1–S5 and Figures S1–S5) and crystallographic statistics for *Pf*LDH-W107A (Table S6, PDB entry 4PLZ) (PDF)

The authors declare no competing financial interest.

Apicomplexa are obligate, intracellular eukaryotic parasites of animals responsible for many human diseases, including malaria, cryptosporidiosis, babesiosis, cystoisosporiasis, cyclosporiasis, and toxoplasmosis. The organisms responsible for malaria are of the genus *Plasmodium*, with *Plasmodium falciparum* being responsible for the most lethal form of malaria. *P. falciparum* proceeds through a complex life cycle in two different hosts. *P. falciparum* sporozoites infect humans via a mosquito bite. These sporozoites invade liver cells, where they multiply asexually into merozoites. The merozoites rupture the liver cells and enter the bloodstream, where they invade erythrocytes and replicate. The replication of *P. falciparum* in erythrocytes is known as the blood stage of infection, which is largely responsible for the clinical symptoms of malaria.¹

During the blood stage, *P. falciparum* must respire anaerobically to regenerate NAD^+ ,^{1,2} an essential electron acceptor in glycolysis. In the presence of oxygen, ATP production is maximized by metabolizing glucose to CO_2 and H_2O via glycolysis, the citric acid cycle, and the electron transport chain, which regenerates NAD^+ for use in glycolysis. In the absence of oxygen, NAD^+ must be regenerated through pyruvate fermentation to prevent a stall in glycolysis and to generate sufficient amounts of ATP for cellular function. Lactate dehydrogenase (LDH) couples the regeneration of NAD^+ from NADH to the reduction of pyruvate to lactate. Because the blood stage of malarial infection occurs under anaerobic conditions, the *P. falciparum* LDH (*PfLDH*) is essential for the pathogen's survival as the only means to regenerate NAD^+ .²

Apicomplexan LDHs evolved from an ancestral α -proteobacterial malate dehydrogenase (MDH), independently of canonical LDHs such as those found in the metazoan hosts of apicomplexan parasites.³ Despite convergent evolutionary origins, *PfLDH* and canonical LDHs both share a similar catalytic mechanism. The enzymatic reduction of pyruvate to

lactate proceeds via the following steps: (1) starting in the “loop open” conformation, NADH binds the apoenzyme, (2) R171 orients pyruvate in the active site to form the ternary complex, (3) the substrate specificity loop closes over the active site, allowing R109 to bind the substrate and stabilize the transition state, (4) NADH reduces pyruvate to lactate by hydride transfer, (5) the substrate specificity loop opens to release lactate, and (6) NAD⁺ is released to regenerate the apoenzyme. D168 activates the catalytic H195, allowing for a proton transfer during the reduction of pyruvate to lactate (Figure 1). Movement of the substrate specificity loop is the rate-limiting step in catalysis.⁴

The LDH substrate specificity loop also plays an important role in substrate recognition. This loop contains a residue at position 102, commonly termed the “specificity residue”, which distinguishes among the R groups of different 2-ketoacid substrates. In all known MDHs, the specificity residue is an arginine, which forms a salt bridge with the methylene carboxylate group of malate and oxaloacetate. In canonical LDHs, the specificity residue is a glutamine, which contacts and recognizes the lactate and pyruvate methyl group. However, the convergent apicomplexan LDHs did not evolve by a mutation at position 102. Rather, apicomplexan LDHs evolved from an ancestral MDH via a unique five-amino acid insertion in the substrate specificity loop that switches substrate specificity from malate/oxaloacetate to lactate/pyruvate.³ The insertion lengthens the loop, but otherwise, the structures of *Pf* LDH and canonical LDHs and MDHs are highly similar (Figure 2A).

Crystal structures of canonical LDHs with the substrate specificity loop in the closed conformation show the Q102 specificity residue contacting substrate analogues (Figure 2B). In contrast, the *Pf*LDH crystal structure shows W107f contacting a substrate analogue in the closed conformation, suggesting that W107f is the apicomplexan LDH “specificity residue” rather than the lysine at position 102 (Figure 2B). Using point mutations and alanine scanning of the substrate specificity loop, we previously showed that W107f is essential for *Pf*LDH enzyme activity and substrate recognition, while mutations of K102 have a negligible effect.³

*Pf*LDH is widely considered to be an attractive therapeutic target because of its essential role during pathogen replication in the blood stage and its distinct active site architecture, which provides an opportunity to selectively inhibit the parasite enzyme versus the host LDH. Development of small molecule inhibitors against *Pf*LDH is currently an active area of research,^{2,5–13} and selective antibody inhibitors that target the unique *Pf*LDH substrate specificity loop have been developed.¹⁴ Sequence comparison of apicomplexan LDHs suggests the specificity loop region is highly conserved (Figure 3A). Hence, selective inhibitors of *Pf*LDH are also likely to be effective against various *Plasmodium* strains and other apicomplexan LDHs (for example, as with *Babesia*¹⁵).

The apparent sequence conservation of the substrate specificity loop poses an evolutionary problem: Was the conserved insertion sequence necessary for evolution of LDH specificity from the ancestral MDH? Furthermore, crystal structures show that much of the apicomplexan LDH loop bulges out into solution, lacking any clear structural role without contacting the rest of the protein. Was the entire five-residue insertion necessary, or could a shorter insertion have sufficed? To answer these questions, we probed the functional

constraints on residue identity and loop length within the *Pf*LDH substrate specificity loop by perturbation analysis. Loop conservation was quantified by estimating the evolutionary substitution rates across the loop, which identified several slow-evolving positions likely to be functionally constrained. With this information, we generated an ensemble of *Pf*LDH loop mutants and characterized their steady-state kinetics. We also made a series of mutants with varying loop lengths to characterize the functional dependence on loop length. Surprisingly, neither sequence identity nor loop length is under a strong functional constraint, as long as W107f is unperturbed.

METHODS

Site-Directed Mutagenesis.

The *P. falciparum* lactate dehydrogenase (LDH) gene (gi#: 124513226) was synthesized and subcloned into pET-11b by GenScript (Piscataway, NJ) with a C-terminal six-His tag. All mutations and truncations were generated using a QuikChange Lightning kit (Agilent, Santa Clara, CA) with primers synthesized by Integrated DNA Technologies (Coralville, IA). All sequences were confirmed by Sanger Sequencing at Genewiz (Cambridge, MA). Conventional LDH numbering is based on the canonical dogfish LDH sequence, which has no insertion.¹⁶ The residues in the apicomplexan insertion are thus numbered using the residue that precedes the insertion, K107, with letters appended to each residue sequentially: G106, K107a, S107b, D107c, K107d, E107e, W107f, and N108.

Protein Expression.

Plasmids were transformed into BL21(DE3)pLysS *Escherichia coli* cells (Invitrogen, Grand Island, NY). Cells were grown at 37 °C with 225 rpm agitation in 2×YT medium supplemented with 30 mM potassium phosphate (pH 7.8) and 0.1% (w/v) glucose with cell growth monitored at OD₆₀₀. When cultures reached an OD₆₀₀ of 0.5–0.8, cells were induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside and incubated for 4 h at 37 °C with 225 rpm agitation. Cells were pelleted by centrifugation and stored at –80 °C.

Purification.

Cell pellets were resuspended on ice in 25 mL of 50 mM sodium phosphate (pH 8.0), 300 mM sodium chloride, 10 mM imidazole buffer, and 375 units of Pierce universal nuclease (Thermo Scientific, Rockford, IL) per 1.5 L of culture pellet. Resuspended cells were sonicated on ice for 2 min at 35% amplitude (30 s on, 20 s off). The cell lysate was centrifuged for 20 min at 25000g to separate soluble and insoluble fractions. The supernatant was 0.22 μm syringe filtered. The filtered lysate was purified by affinity chromatography with a 5 mL HisTrap FF nickel column (GE Healthcare, Piscataway, NJ). The protein was eluted and fractionated with a 10 to 500 mM imidazole gradient. Fractions were inspected for protein content by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and then pooled and concentrated in Amicon Ultracel-10K centrifugal filters (Millipore, Billerica, MA). The concentrated protein was buffer exchanged into 50 mM Tris (pH 7.4), 100 mM sodium chloride, 0.1 mM EDTA, and 0.01% (w/v) azide via a PD-10 desalting column (GE Healthcare). The final enzyme concentrations were calculated by

absorbance spectroscopy using Beer's law with 280 nm extinction coefficients and molecular weights given by ExPASy's ProtParam program.

Kinetic Assays.

All reagents were purchased from Sigma-Aldrich (St. Louis, MO). Initial rates for the conversion of pyruvate to lactate were monitored on a CARY 100 Bio spectrophotometer (Agilent) by following the decrease in absorbance at 340 nm associated with the conversion of NADH to NAD⁺. Kinetic assays were performed with variable pyruvate concentrations and a constant NADH concentration (200 μ M) in a 50 mM Tris (pH 7.5), 50 mM potassium chloride buffer at 25 °C. The enzyme concentration ranged between 1 μ M and 1 nM. Initial rates were plotted against substrate concentration and fit using KaleidaGraph to the Michaelis–Menten equation, $v/[E]_t = k_{cat}[S]/(K_M + [S])$, or a substrate inhibition equation, $v/[E]_t = k_{cat}[S]/(K_M + [S] + [S]^2/K_i)$.

Protein Crystallization, Data Collection, and Structure Determination.

All reagents were from Sigma-Aldrich. Conditions were optimized on the basis of promising conditions identified from screens with Crystal Screen and Crystal Screen 2 (Hampton Research, Aliso Viejo, CA). The 4 μ L drops (2 μ L of 600 μ M or 20 mg/mL protein stock and 2 μ L of well solution) were spotted on coverslips (Hampton Research) and equilibrated against 1 mL of the well solution by hanging-drop vapor diffusion at room temperature. Crystals were cryo-protected by being soaked in a 15% (w/v) dextrose solution for 3 min, transferred to a 30% (w/v) dextrose solution, and then immediately flash-frozen in liquid N₂.

Data sets were collected at the SIBYLS beamline (12.3.1 Lawrence Berkeley National Laboratory, Berkeley, CA) and indexed, integrated, and scaled with XDS/XSCALE. PHENIX's AutoMR program was used to determine the structures by molecular replacement with a previously determined *Pf*LDH model (PDB entry 1T2D). Models were improved by rounds of manual refinement in Coot and automated refinement with PHENIX's phenix.refine program. Model quality was determined using MolProbity in PHENIX. All structural alignments and images were generated using PyMOL.

Evolutionary Conservation Estimation.

Apicomplexan LDH sequences were collected from the NCBI database using the BLASTP searches with four query sequences, UniProt entries MDHC_PIG, Q76NM3_PLAF7, C6KT25_PLAF, and MDH_WOLPM, for full coverage of the superfamily. Non-LDH sequences were removed along with sequences shorter than 290 or longer than 340 residues. The final data set contained 277 sequences. A multiple-sequence alignment of the data set was generated using Muscle and analyzed using PhyML¹⁷ and ConSurf.¹⁸ The relative conservation of each amino acid position within the alignment was estimated by quantifying the relative substitution rate using empirical Bayes maximum likelihood methods.

RESULTS

Contributions from Individual Specificity Loop Residues.

We previously used alanine scanning mutagenesis to assess the functional contribution of individual residues within the specificity loop.³ The boundaries of the scan were determined from a superposition of the human lactate dehydrogenase (LDH) and *Pf*LDH. The main chains of the specificity loops in the two structures conformationally diverge between positions 101 and 108, a span of 12 residues. These 12 residues were individually mutated to alanine in *Pf*LDH and characterized kinetically to evaluate the effects of each substitution on catalysis.

The alanine scan surprisingly revealed that all residues except W107fA tolerate mutation with little to no effect on their catalytic efficiency [$k_{\text{cat}}/K_{\text{M}}$ (Figure 3B), as reported previously³]. W107fA has a 4 order of magnitude drop in k_{cat} , indicating that W107f is essential for wild-type levels of strong catalysis. Notably, k_{cat} remains relatively unaffected in all of the other mutants (within a factor of 2 of the wild-type value). W107fA also has an increased K_{M} , 50 times greater than that of the wild type. The remaining alanine mutants all have K_{M} values within an order of magnitude of that of wild-type *Pf*LDH (Figure 3B).

The crystal structure of W107fA [PDB entry 4PLZ (Table S6)] was determined by molecular replacement and was found to have the same space group and cell dimensions as that of wild-type *Pf*LDH (PDB entry 1T2D).³ The structure of W107fA is nearly identical to that of the wild-type protein [root-mean-square deviation of 0.95 Å² (Figure 4A)]. The only significant difference between W107fA and the wild type is in the orientation of the specificity loop. The wild-type loop is in the closed conformation, indicated by a kinked helix running along the back of the protein (Figure 4B). The W107fA mutant is in the open, noncatalytic conformation, indicated by an unkinked, linear helix conformation and disordered residues within the loop, much like the apo form in which the loop is open (Figure 4B).

Conservative W107f Mutations.

Because mutating a tryptophan to an alanine is a dramatic mutation, we wished to investigate the tolerance of position 107f to more conservative changes. W107f was mutated to various hydrophobic and aromatic residues, in addition to a glutamine that is the canonical LDH specificity residue.¹⁹

All W107f mutants had reduced enzymatic activity, with $k_{\text{cat}}/K_{\text{M}}$ values for pyruvate ranging from 50 times lower than to 5 orders of magnitude below that of the wild type. Both k_{cat} and K_{M} for pyruvate are affected. The large aromatic replacements, W107fY and W107fF, have the weakest effects, with 50- and 100-fold reductions in pyruvate activity, respectively (Figure 5). While these aromatic mutations have a substantial effect, the resulting activities are likely physiologically relevant as they are similar to those of other wild-type LDHs in the apicomplexan phylum, which have $k_{\text{cat}}/K_{\text{M}}$ values ranging from 10⁴ to 10⁷ M⁻¹ s⁻¹.^{3,20–26}

The other W107f mutants have more diminished activities. The smaller histidine, while still an aromatic residue, cripples the enzyme a further 50 times below W107fF. The values for the nonplanar hydrophobic amino acid mutants are all 4 orders of magnitude or more below the wild-type pyruvate $k_{\text{cat}}/K_{\text{M}}$, nearly on par with that of W107fA. Glutamine at position 107f is as detrimental to pyruvate activity as the nonplanar hydrophobics, despite being the specificity residue in canonical LDHs. The substitutions have substantial effects on both k_{cat} and K_{M} except for W107fY and W107fF, both of which have a k_{cat} within half of that of the wild type, and W107fH, which has a K_{M} nearly equal to that of the wild type (Figure 5).

Conservation of Loop Identity.

Our alanine-scan mutants suggest that most specificity loop residues are functionally unconstrained. This raises the question of why the loop sequence is apparently well-conserved. A high degree of sequence conservation can arise for two main reasons, either (1) a functional constraint to preserve the sequence or (2) insufficient time for the sequences to diverge. Phylogenetic rate analysis can control for evolutionary relatedness. Low rates indicate functional importance due to purifying selection for sequence identity, and high rates suggest a weak functional constraint. Our rate analysis indicates K102, K107a, E107e, and W107f within the loop are slow-evolving (relative rates of <0.4) and likely functionally important (Figure 6).

We designed two *Pf*LDH mutant backgrounds to determine the role of the conserved residues. “Poly-A” and “Poly-S” mutants were constructed with all positions 101–107e mutated to alanine and serine, respectively (Figure 7A). The alanine scan indicates that any position other than W107f can be mutated individually with a minimal effect on activity. Mutations in combination may have a more significant effect, and the Poly-A and Poly-S constructs test this by mutating each position simultaneously. The more polar Poly-S should provide a less drastic sequence change for a solvent-exposed surface loop and may indicate if changes in Poly-A function are due to the increased hydrophobicity. W107f was left unchanged in both mutant backgrounds because the alanine scan and the W107f substitution data show that the identity of position 107f is highly constrained. Remarkably, the Poly-A and Poly-S mutants retain LDH activity within approximately 2 orders of magnitude of that of wild-type *Pf* LDH (Figure 7B), a level of activity observed in other apicomplexan LDHs such as the *Toxoplasma gondii* LDHs.²⁵

We designed two additional mutants within the Poly-A background to determine the contributions from slow-evolving positions. If the loop residues are conserved because of functional constraint, there should be a partial rescue of activity when adding them back into the Poly-A mutant. The first construct “Cons” contains *Pf*LDH wild-type residues at any loop position evolving slowly (defined as a relative rate of <1¹⁷): T101, K102, P104, K107a, S107b, D107c, E107e, and W107f (Figure 7A). Adding back the conserved residues partially restored activity by a factor of 10 increase in $k_{\text{cat}}/K_{\text{M}}$ compared to that of Poly-A, but it is still 10-fold lower than wild-type *Pf* LDH activity. The second mutant “MaxCons” contains only the wild-type residues that evolve as slowly as W107f (relative rate of <0.4): K102, K107a, and E107e. MaxCons behaved like Cons despite having three fewer wild-type

residues, suggesting that K102, K107a, E107e, and W107f are the most functionally important loop residues (Figure 7B).

Tolerance of the Specificity Loop to Truncation.

The five-residue loop insertion is unique to the apicomplexan LDHs, which makes their specificity loop longer than what is seen in canonical LDHs.¹⁹ The existence of highly active canonical LDHs with shorter specificity loops raises the question of whether the large *Pf* LDH loop is functionally necessary. Can *Pf*LDH retain activity with a shortened specificity loop? We therefore designed and tested a series of loop deletion mutants, ranging from one to five residues, to assess the dependence of *Pf*LDH activity on loop length (Figure 8A). Loop deletion mutants are named in the form of M(X) where X is the single-letter amino acid code for the missing residues. Additionally, the “GSA-linker” construct incorporates a generic glycine-serine-alanine motif that connects T101 to W107f, making the shortest loop possible that retains W107f (based on the distance between the backbone C_α atoms of T101 and W107f). The GSA-linker is effectively a six-residue deletion.

M(A), M(PG), and M(PGK) exhibit between 10- and 100-fold decreases in k_{cat}/K_M relative to that of the wild type (Figure 8B), within the range of other wild-type LDH activities in the apicomplexan phylum. The M(PGKS) and M(SDKE) truncations are more deleterious but remain marginally active. M (APGKS) and the GSA-linker are crippled enzymes, with k_{cat}/K_M values that are reduced approximately 5 orders of magnitude compared to that of the wild type.

The loop deletions primarily affect K_M , with only minor effects on k_{cat} . One- to three-residue deletions decrease k_{cat} to no less than half of that of wild-type *Pf* LDH, and both four-residue deletions decrease k_{cat} by 1 order of magnitude. Only the M (APGKS) and GSA-linker mutants have large decreases in k_{cat} . However, all K_M values of deletion mutants are 30–40 times greater than that of the wild type, independent of the number of residues deleted (Figure 8B).

Deleting three residues appears to be the largest loop truncation that is still well tolerated, with a k_{cat}/K_M of $>10^4$. We next determined the kinetic effects of the position of the deleted residues in the loop by constructing and characterizing three-consecutive-residue loop truncations in all possible registers. All three-residue truncations had similar kinetics, with an approximately 100-fold decrease in k_{cat}/K_M , relative to that of the wild type, yet still within the physiologically relevant range of activity (Figure 9).

DISCUSSION

W107f Is under a Strong Functional Constraint.

Alanine scanning mutagenesis of the specificity loop revealed that only W107f contributes significantly to LDH activity (Figure 4). Additionally, most of the alternative residues at the 107f position fail to replicate wild-type levels of activity. Only the other large aromatic residues, tyrosine and phenylalanine, retain physiologically relevant activity, though it is still below that of wild-type *Pf*LDH (Figure 6). These results show that a planar aromatic

residue at position 107f was essential for the evolution of LDH activity, but a tryptophan residue in particular was apparently unnecessary.

In wild-type *Pf*LDH, W107f is buried in the active site, making several interactions potentially involved in loop closure and substrate recognition. These include packing of the W107f indole ring against the side chain of R109 in the loop, an H-bond from the indole nitrogen to the backbone carbonyl of T106, and various hydrophobic interactions with residues on the opposite side of the active site cleft, such as A236, L237, and V240. Furthermore, the edge of the W107f indole ring packs against the methyl group of the lactate substrate.

Unlike that of the wild type, the crystal structure of W107fA is in the open conformation. This observation suggests that W107f contributes favorable interactions necessary for loop closure because the W107fA mutant equilibrium is shifted to the loop open state (Figure 5). The rate-limiting step in pyruvate turnover by *Pf* LDH is loop closure,⁴ and the decrease in $k_{\text{cat}}/K_{\text{M}}$ for W107fA results primarily from a decrease in k_{cat} . Taken together, these facts suggest that the k_{cat} of the W107fA mutant decreased because loop closure is slower or loop opening is faster. Given that the pyruvate $K_{\text{M}} = (k_{\text{off}} + k_{\text{cat}})/k_{\text{on}}$ (under saturating NADH⁴) and the difference in the mutant and wild-type K_{M} values is much smaller than the difference in the k_{cat} values, there must be a compensatory change in the binding constants of the substrate: either k_{off} increased or k_{on} decreased. We propose that loss of favorable interactions from W107f results in a loop that prefers to remain open in a conformation unfavorable for substrate binding, and thus, substrate is ejected before hydride transfer takes place.

Conservation and Functional Contribution of Specificity Loop Residues.

Six of the 10 loop residues evolve more slowly than average and thus may be functionally important (Figure 7A). The contribution of the three slowest-evolving (K102, K107a, and E107e) is however modest, a 10-fold increase in $k_{\text{cat}}/K_{\text{M}}$ compared to that of a loop consisting entirely of alanines and W107f (Figure 8). Activity is not further rescued by incorporating additional slow-evolving wild-type residues.

Overall, *Pf*LDH is roughly 50% identical in sequence to LDHs from closely related Apicomplexa, such as *Eimeria maxima* and *T. gondii*,^{3,25–27} yet the specificity loop region appears to be well-conserved. However, we find no obvious reason that the sequence and length of the specificity loop should be conserved, and in fact, quantitative rate analysis indicates that much of the loop is not particularly slowly evolving. While the most conserved residues appear to contribute marginally to LDH activity, it is possible the loop is important for another, unknown function. The loop sequence identity could be pleiotropically constrained if the enzyme is “moonlighting”^{28,29} and performs multiple functions for the organism (as seen in some vertebrate LDHs^{30,31}).

The Specificity Loop Can Be Dramatically Truncated.

The *Pf* LDH specificity loop can be truncated by as many as four residues and still maintain significant levels of activity (Figure 9). K_{M} is affected by loop truncations more than k_{cat} , with all truncation mutants exhibiting a K_{M} 1 order of magnitude greater than that of the

wild type, suggesting impaired substrate binding. Assuming k_{on} is diffusion-limited and is unlikely to be affected, the loop truncation effects are probably mediated by an increase in substrate k_{off} due to a moderately unfavorable truncated loop conformation.

Loop deletion mutants may expose the active site to the bulk solvent and increase the frequency of nonproductive binding events, which may explain why each of the three-residue deletion mutants has an approximately equal $k_{\text{cat}}/K_{\text{M}}$ regardless of the location in the specificity loop or the identity of the residues removed. The size of the substrate loop is more important than its identity during catalysis. However, there is a lower limit to how short the loop can be to remain functional.

Evolutionary Requirements and Implications for Drug Resistance.

Our results show that *Pf*LDH activity is extremely robust to mutations in the specificity loop that alter either its sequence identity or its length. In fact, only W107f appears to be particularly important for enzyme function, yet other aromatic residues can function nearly as well as tryptophan. Ten of 11 possible positions are mutated in the Poly-A and Poly-S constructs, and both are functional enzymes with physiologically relevant activity. Deletions of one, two, and three amino acids throughout the loop result in functional enzymes. These results suggest that the exact insertion sequence and length were unnecessary for the evolution of a novel LDH from the ancestral MDH. Many possible loop insertions could have led to evolution of a functional LDH, regardless of sequence identity or length, with the only necessity being an aromatic amino acid properly positioned in the active site.

Drug resistance can develop if the protein target is not functionally constrained. There are many possible specificity loops that can produce physiologically relevant LDH activities in *P. falciparum*. *Pf*LDH thus has great potential to yield mutations that could abolish any inhibitory effect of a drug that targets the active site loop because of the lack of a significant functional constraint. The specificity loop region is conserved in the phylum, and thus, targeting the specificity loop in other Apicomplexa would likely be met with the same difficulties.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank all of the other members of our lab. Though they did not directly contribute data, their constant support throughout the project was essential.

Funding

This work was supported by National Institutes of Health Grants R01GM096053 (D.L.T.) and R01GM105404 (S.C.).

ABBREVIATIONS

LDH	lactate dehydrogenase
-----	-----------------------

MDH	malate dehydrogenase
NADH/NAD⁺	nicotinamide adenine dinucleotide
PDB	Protein Data Bank
<i>Pf</i>	<i>P. falciparum</i>

REFERENCES

- (1). Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, and Haynes JD (2002) A proteomic view of the Plasmodium falciparum life cycle. *Nature* 419 (6906), 520–6. [PubMed: 12368866]
- (2). Royer RE, Deck LM, Campos NM, Hunsaker LA, and Vanderjagt DL (1986) Biologically-Active Derivatives of Gossypol -Synthesis and Antimalarial Activities of Peri-Acylated Gossylic Nitriles. *J. Med. Chem* 29 (9), 1799–801. [PubMed: 3528492]
- (3). Boucher JJ, Jacobowitz JR, Beckett BC, Classen S, and Theobald DL (2014) An atomic-resolution view of neofunctionalization in the evolution of apicomplexan lactate dehydrogenases. *eLife* 3, e02304 DOI: 10.7554/eLife.02304.
- (4). Shoemark DK, Cliff MJ, Sessions RB, and Clarke AR (2007) Enzymatic properties of the lactate dehydrogenase enzyme from Plasmodium falciparum. *FEBS J* 274 (11), 2738–48. [PubMed: 17459101]
- (5). Gomez MS, Piper RC, Hunsaker LA, Royer RE, Deck LM, and Makler MT (1997) Substrate and cofactor specificity and selective inhibition of lactate dehydrogenase from the malarial parasite P-falciparum. *Mol. Biochem. Parasitol* 90 (1), 235–46. [PubMed: 9497046]
- (6). Read JA, Wilkinson KW, Tranter R, Sessions RB, and Brady RL (1999) Chloroquine binds in the cofactor binding site of Plasmodium falciparum lactate dehydrogenase. *J. Biol. Chem* 274(15), 10213–8. [PubMed: 10187806]
- (7). Brady RL, and Cameron A (2004) Structure-based approaches to the development of novel anti-malarials. *Curr. Drug Targets* 5 (2), 137–49. [PubMed: 15011947]
- (8). Cameron A, Read J, Tranter R, Winter VJ, Sessions RB, and Brady RL (2004) Identification and activity of a series of azole-based compounds with lactate dehydrogenase-directed anti-malarial activity. *J. Biol. Chem* 279 (30), 31429–39. [PubMed: 15117937]
- (9). Connors R, Schambach F, Read J, Cameron A, Sessions RB, and Vivas L (2005) Mapping the binding site for gossypol-like inhibitors of Plasmodium falciparum lactate dehydrogenase. *Mol. Biochem. Parasitol* 142 (2), 137–48. [PubMed: 15978953]
- (10). Vivas L, Easton A, Kendrick H, Cameron A, Lavandera JL, and Barros D (2005) Plasmodium falciparum: stage specific effects of a selective inhibitor of lactate dehydrogenase. *Exp. Parasitol* 111 (2), 105–14. [PubMed: 16098967]
- (11). Choi S-r., Pradhan A, Hammond NL, Chittiboyina AG, Tekwani BL, and Avery MA (2007) Design, Synthesis, and Biological Evaluation of Plasmodium falciparum Lactate Dehydrogenase Inhibitors. *J. Med. Chem* 50 (16), 3841–50. [PubMed: 17636950]
- (12). Granchi C, Bertini S, Macchia M, and Minutolo F (2010) Inhibitors of lactate dehydrogenase isoforms and their therapeutic potentials. *Curr. Med. Chem* 17 (7), 672–97. [PubMed: 20088761]
- (13). Saxena S, Durgam L, and Guruprasad L (2018) Multiple epharmacophore modelling pooled with high-throughput virtual screening, docking and molecular dynamics simulations to discover potential inhibitors of Plasmodium falciparum lactate dehydrogenase (PfLDH). *J. Biomol. Struct. Dyn*, 1–17.
- (14). Kaushal NA, and Kaushal DC (2014) Production and characterization of monoclonal antibodies against substrate specific loop region of Plasmodium falciparum lactate dehydrogenase. *Immunol. Invest* 43 (6), 556–71. [PubMed: 24702659]
- (15). Bork S, Yokoyama N, Ikehara Y, Kumar S, Sugimoto C, and Igarashi I (2004) Growth-Inhibitory Effect of Heparin on Babesia Parasites. *Antimicrob. Agents Chemother* 48 (1), 236–41. [PubMed: 14693545]

- (16). Eventoff W, Rossmann MG, Taylor SS, Torff HJ, Meyer H, and Keil W (1977) Structural Adaptations of Lactate-Dehydrogenase Isozymes. *Proc. Natl. Acad. Sci. U. S. A* 74 (7), 2677–81. [PubMed: 197516]
- (17). Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, and Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol* 59 (3), 307–21. [PubMed: 20525638]
- (18). Celniker G, Nimrod G, Ashkenazy H, Glaser F, Martz E, and Mayrose I (2013) ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. *Isr. J. Chem* 53 (3–4), 199–206.
- (19). Nicholls DJ, Miller J, Scawen MD, Clarke AR, Holbrook JJ, Atkinson T, and Goward CR (1992) The importance of arginine 102 for the substrate specificity of *Escherichia coli* malate dehydrogenase. *Biochem. Biophys. Res. Commun* 189 (2), 1057–1062. [PubMed: 1472016]
- (20). Ohshima T, and Sakuraba H (1986) Purification and characterization of malate dehydrogenase from the phototrophic bacterium. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol* 869(2), 171–7.
- (21). Hartl T, Grossebuter W, Gorisch H, and Siezowski JJ (1987) Crystalline NAD/NADP-dependent malate dehydrogenase; the enzyme from the thermoacidophilic archaeobacterium *Sulfolobus acidocaldarius*. *Biol. Chem. Hoppe-Seyler* 368 (1), 259–268. [PubMed: 3109450]
- (22). Wynne SA, Nicholls DJ, Scawen MD, and Sundaram TK (1996) Tetrameric malate dehydrogenase from a thermophilic *Bacillus*: cloning, sequence and overexpression of the gene encoding the enzyme and isolation and characterization of the recombinant enzyme. *Biochem. J* 317 (Part 1), 235–245. [PubMed: 8694770]
- (23). Langelandsvik AS, Steen IH, Birkeland NK, and Lien T (1997) Properties and primary structure of a thermostable L-malate dehydrogenase from *Archaeoglobus fulgidus*. *Arch. Microbiol* 168 (1), 59–67. [PubMed: 9211715]
- (24). Madern D, Ebel C, Mevarech M, Richard SB, Pfister C, and Zaccai G (2000) Insights into the molecular relationships between malate and lactate dehydrogenases: structural and biochemical properties of monomeric and dimeric intermediates of a mutant of tetrameric L-[LDH-like] malate dehydrogenase from the halophilic archaeon *Haloarcula marismortui*. *Biochemistry* 39 (5), 1001–10. [PubMed: 10653644]
- (25). Dando C, Schroeder ER, Hunsaker LA, Deck LM, Royer RE, Zhou X, et al. (2001) The kinetic properties and sensitivities to inhibitors of lactate dehydrogenases (LDH1 and LDH2) from *Toxoplasma gondii*: comparisons with pLDH from *Plasmodium falciparum*. *Mol. Biochem. Parasitol* 118, 23–32. [PubMed: 11704270]
- (26). Madern D, Cai X, Abrahamsen MS, and Zhu G (2003) Evolution of *Cryptosporidium parvum* Lactate Dehydrogenase from Malate Dehydrogenase by a Very Recent Event of Gene Duplication. *Mol. Biol. Evol* 21 (3), 489–497. [PubMed: 14694073]
- (27). Madern D (2002) Molecular Evolution Within the L-Malate and L-Lactate Dehydrogenase Super-Family. *J. Mol. Evol* 54 (6), 825–40. [PubMed: 12029364]
- (28). Jeffery CJ (1999) Moonlighting proteins. *Trends Biochem. Sci* 24 (1), 8–11. [PubMed: 10087914]
- (29). Huberts DHEW, and van der Klei IJ (2010) Moonlighting proteins: An intriguing mode of multitasking. *Biochim. Biophys. Acta, Mol. Cell Res* 1803 (4), 520–525.
- (30). Hendriks W, Mulders JW, Bibby MA, Slingsby C, Bloemendal H, and de Jong WW (1988) Duck lens epsilon-Crystallin and lactate dehydrogenase B4 are identical: a single-copy gene product with two distinct functions. *Proc. Natl. Acad. Sci. U. S. A* 85 (19), 7114. [PubMed: 3174623]
- (31). Piatigorsky J (1998) Multifunctional Lens Crystallins and Corneal Enzymes: More than Meets the Eye. *Ann. N. Y. Acad. Sci* 842(1), 7–15. [PubMed: 9599288]

Figure 1. Mechanism of *Pf*LDH loop closure and catalysis. NADH and then pyruvate (red) bind to the enzyme to form the ternary complex. Specificity loop closure brings Arg109 and Trp107f into the proximity of the substrate, and NADH donates a hydride to reduce pyruvate to lactate.

Figure 2.

Superposition of human (teal, PDB entry 1I10) and *P. falciparum* (red, PDB 1T2D) LDHs. (A) An ordinary least-squares superposition of human LDH and Pf LDH (PyMol) shows a high degree of structural similarity. (B) A close-up of the loop shows the structural effect of the five-amino acid insertion (colored gray). *Pf*LDH's W107f occupies the same three-dimensional space as human LDH's Q102.

Figure 3.

(A) Multiple-sequence alignment of the apicomplexan specificity loop. The loop is defined by positions 101–108 (as described in the text) and is boxed in yellow. The insertion found only in apicomplexan LDHs is boxed in white. (B) Kinetic parameters for each alanine mutant are plotted as $\log(k_{\text{cat}})$ vs $\log(K_{\text{M}})$. The red dot denotes wild-type *Pf*LDH, and each black dot is labeled with the position that was mutated. The red line designates a $k_{\text{cat}}/K_{\text{M}}$ of $10^4 \text{ M}^{-1} \text{ s}^{-1}$, roughly the minimum for physiologically relevant activity. Points on a gray line all have the same $k_{\text{cat}}/K_{\text{M}}$. W107fA has both a lower k_{cat} and a higher K_{M} compared to those of the wild type.

Figure 4. Superposition of apo *Pf*LDH (yellow, PDB entry 2×8L), closed *Pf*LDH (red, PDB entry 1T2D), and the W107fA mutant (blue, PDB entry 4PLZ). (A) A structural comparison shows that the global LDH structure is unperturbed by the W107fA mutation. (B) The prominent difference between the structures is the position of the loop and adjacent α -helix. The closed state has a large bend in the helix correlated with a loop positioned over the active site. The mutant has an unkinked helix, similar to the apo structure, and has a loop that is too dynamic or disordered to be modeled.

Figure 5.

Kinetic parameters for the alternative W107f mutants. Upon separate analysis of the kinetic parameters, both tyrosine and phenylalanine have a k_{cat} nearly equal to that of wild-type *Pf* LDH, with the loss of activity being primarily a K_{M} effect. Other mutants have both k_{cat} and K_{M} effects, except W107fH, which has only a reduction in k_{cat} .

Figure 6. Evolutionary rates in the *Pf*LDH loop and conservation mutants. Evolutionary rates are mapped onto the sequence of the *Pf*LDH specificity loop. Positions that are colored magenta evolve slowly, while those that are colored cyan evolve more rapidly. Slow-evolving sites (rate of <0.4) are shown as sticks.

Figure 7. Loop conservation mutants. (A) Multiple-sequence alignment of the Poly-A, Poly-S, Cons, and MaxCons mutants and the wild-type *Pf*LDH sequence. (B) Both Poly mutants have large changes in K_M and a small but significant drop in k_{cat} . Cons and MaxCons show partially rescued activity. The MaxCons construct shows that only K102, K107a, and E107e are necessary for the same level of activity. The reintroduction of the highly conserved loop residues rescued activity by both increasing k_{cat} and decreasing K_M .

Figure 8.

Loop truncation mutants. (A) Protein sequence alignment of the truncation mutants. Dashes represent deleted amino acids. (B) Truncation mutants all have reduced $\log(k_{\text{cat}}/K_M)$ values compared to that of the wild type. Deleting three or fewer loop residues reduces activity 2 orders of magnitude. Removing four residues has a more pronounced effect. The five-residue truncation and GSA-linker are severely crippled. The decrease in k_{cat}/K_M for all the constructs is primarily a K_M effect, though both M(APGKS) and the GSA-linker have a significantly decreased k_{cat} .

Figure 9. Positional effects of three-residue deletions. (A) Sequence alignment of the three-residue deletion mutants. Dashes indicate the location of the gap. (B) A majority of the activity loss is due to a higher K_M , which is approximately the same for all three-residue deletion mutants.