Contents lists available at SciVerse ScienceDirect

### **Bioorganic & Medicinal Chemistry**

journal homepage: www.elsevier.com/locate/bmc

# Directed evolution of a pyruvate aldolase to recognize a long chain acyl substrate

Manoj Cheriyan<sup>a</sup>, Matthew J. Walters<sup>b</sup>, Brian D. Kang<sup>a</sup>, Laura L. Anzaldi<sup>c</sup>, Eric J. Toone<sup>b,c</sup>, Carol A. Fierke<sup>a,d,\*</sup>

<sup>a</sup> Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

<sup>b</sup> Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA

<sup>c</sup> Department of Chemistry, Duke University, Durham, NC 27708, USA

<sup>d</sup> Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

#### ARTICLE INFO

Article history: Received 23 June 2011 Revised 19 August 2011 Accepted 26 August 2011 Available online 30 August 2011

Keywords: Biocatalysis Protein engineering Thermostable enzyme Substrate specificity Random mutagenesis

#### 1. Introduction

#### Protein engineering is a rapidly growing field of study where naturally occurring enzymes are altered to improve activity for industrial and laboratory applications. The advantages of biocatalysts over traditional chemical synthesis methodologies are their exquisite chemo-, stereo- and regio-selectivity, high rates of catalysis, and reactivity under mild and chemically benign conditions.<sup>1</sup> The former properties are highly desirable for pharmaceutical applications where the production of enantiomerically pure compounds is critical, while the latter property offers cost savings for large-scale industrial applications.

Much research has focused on the development of aldolases as synthetic reagents because of the centrality of carbon–carbon bond formation in synthetic chemistry.<sup>2–4</sup> We have focused our efforts on evolving *Escherichia coli* (EC) 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase and a closely related isoform from the thermophilic bacteria *Thermatoga maritima* (TM). These aldolases preferentially catalyze *si*-stereo facial aldol addition reactions between pyruvate and a range of aldehydic electrophiles<sup>5–7</sup> (Fig. 1). The types of structures afforded by the enzyme have proven useful in the synthesis of nikkomycins.<sup>8,9</sup> The substrate tolerance of KDPG

\* Corresponding author.

E-mail address: fierke@umich.edu (C.A. Fierke).

#### ABSTRACT

The use of biological catalysts for industrial scale synthetic chemistry is highly attractive, given their cost effectiveness, high specificity that obviates the need for protecting group chemistry, and the environmentally benign nature of enzymatic procedures. Here we evolve the naturally occurring 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolases from *Thermatoga maritima* and *Escherichia coli*, into enzymes that recognize a nonfunctionalized electrophilic substrate, 2-keto-4-hydroxyoctonoate (KHO). Using an in vivo selection based on pyruvate auxotrophy, mutations were identified that lower the  $K_M$  value up to 100-fold in *E. coli* KDPG aldolase, and that enhance the efficiency of retro-aldol cleavage of KHO by increasing the value of  $k_{cat}/K_M$  up to 25-fold in *T. maritima* KDPG aldolase. These data indicate that numerous mutations distal from the active site contribute to enhanced 'uniform binding' of the substrates, which is the first step in the evolution of novel catalytic activity.

© 2011 Elsevier Ltd. All rights reserved.

aldolase from a number of species includes a variety of 3 and 4 carbon sugars such as erythrose and lactaldehyde with a preference for phosphorylated substrates.<sup>6,7</sup> These enzymes also accept 2-, 3-, and 4-pyridine carboxaldehyde as electrophilic substrates with low efficiency,<sup>6,7</sup> but do not catalyze aldol addition reactions with unfunctionalized aldehydes such as butyraldehyde and valeraldehyde. Since the synthesis of many valuable natural products require carbon–carbon bond formation using unfunctionalized aldehydes (e.g., syrigomycin<sup>10</sup> and lyngbyabellin<sup>11</sup>), and since phosphorylated compounds are inconvenient for most synthetic schemes,<sup>12</sup> we sought to develop enzymes with improved catalytic activity towards long chain acyl substrates.

To broaden the substrate tolerance of the KDPG aldolases, we employed directed evolution to selectively enhance the ability of the enzyme to catalyze aldol addition with 2-keto-4-hydroxyoctonoate (KHO), an unnatural electrophile that represents a first step towards aldol synthesis with long chain acyl aldehydes. As previously reported, our selection uses a pyruvate kinase (*pykA* and *pykF*)



Figure 1. KDPG aldolase-catalyzed reversible aldol reaction.



<sup>0968-0896/\$ -</sup> see front matter  $\circledcirc$  2011 Elsevier Ltd. All rights reserved. doi:10.1016/j.bmc.2011.08.056

deficient cell line that is unable to grow on minimal media supplemented with ribose as the sole carbon source<sup>13,14</sup> (Fig. S1). Addition of exogenous pyruvate rescues the auxotrophic phenotype. Furthermore, we have previously demonstrated that cleavage of an unnatural substrate, 2-keto-4-hydroxy-4-(2'-pyridyl)butyrate (KHPB), catalyzed by a KDPG aldolase mutant can produce enough pyruvate to rescue cell growth, validating our selection approach.<sup>14</sup>

To examine the hypothesis that higher protein stability would allow for sampling of a wider range of sequence space while maintaining a properly folded structure, we carried out directed evolution experiments on both the *T. maritima* and *E. coli* KDPG aldolase enzymes. The *E. coli* and *T. maritima* KDPG aldolases have 34% identity (pairwise alignment at http://blast.ncbi.nlm.nih.gov); the sequence differences afford the TM KDPG aldolase both thermostability and minor differences in substrate specificity.<sup>6</sup> Using random mutagenesis and selection we identified KDPG aldolase mutants that improve the efficiency for cleavage ( $k_{cat}/K_{M}$ ) of KHO by up to 25-fold. The majority of mutations identified are located far from the active site, suggesting that subtle long-range remodeling of the active site is the mechanism of adaptation, consistent with Albery and Knowles' model of 'uniform binding'.<sup>15</sup>

#### 2. Results

According to the Haldane relationship,<sup>16</sup> the efficiency for catalysis  $(k_{cat}/K_M)$  of the forward and reverse reactions is related by the overall equilibrium constant for the reaction, K<sub>eq</sub>. Thus changes in the efficiency of retro-aldol cleavage for a given substrate also reflect the efficiency of the enzyme to perform aldol addition reactions. Similarly, since  $K_{eq}$  is unchanged by mutation of the enzyme, enhancements in the efficiency of retro-aldol cleavage reflect a comparable improvement in the synthetic direction. So the steady-state kinetic parameters for retro-aldol cleavage of KHO catalyzed by wild-type KDPG aldolase were measured to provide a baseline for evaluation of improvements in catalytic efficiency made during mutagenesis. The kinetic parameters for  $k_{cat}/K_{M}$  for the cleavage of KHO catalyzed by the T. maritima and E. coli KDPG aldolases are  $2.2 \pm 0.3 \text{ M}^{-1} \text{ s}^{-1}$  (34 °C) and  $14 \pm 4 \text{ M}^{-1} \text{ s}^{-1}$  (25 °C), respectively. Compared to the cleavage of the natural substrate, KDPG, the catalytic efficiency for cleaving KHO is decreased by 10<sup>4</sup>-10<sup>6</sup>-fold, indicating that KHO is a very poor substrate for these enzymes, presumably due in part to unfavorable interactions between the hydrophobic chain of KHO and the hydrophilic phosphate binding site of the enzyme.<sup>9</sup> Furthermore, the initial velocity has a nearly linear dependence on the concentration of KHO up to 50 mM for both enzymes, indicating that the value of  $K_{\rm M}$  is significantly larger than 50 mM. (Concentrations of KHO above 50 mM lead to noticeable precipitation of the coupling enzyme.) A fit of the Michaelis-Menten equation to these data provided estimated values for  $k_{cat}$  of  $0.4 \pm 0.05 \text{ s}^{-1}$  (TM) and 2.0 ± 0.6 s<sup>-1</sup> (EC), and  $K_{\rm M}$  of 200 ± 30 mM (TM) and 150 ± 50 mM (EC). These values for  $K_{\rm M}$  are a lower limit and could reflect the solubility of the substrate. Although  $K_{M}$  frequently includes multiple terms and does not simply reflect binding affinity, the high value of this parameter coupled with the low turnover number suggests that wild-type KDPG aldolase binds KHO in a productive conformation with low affinity. Furthermore, the low value of  $k_{cat}$  suggests that KHO bound to KDPG aldolase is not positioned to maximize catalytic efficiency. With such low catalytic activity, overexpression of either T. maritima or E. coli KDPG aldolase is not sufficient to rescue growth of the PB25 cell line on ribose/KHO media.<sup>14</sup>

To identify mutants of KDPG aldolase with enhanced catalytic activity towards retro-aldol cleavage of KHO, we prepared libraries of *T. maritima* and *E. coli* KDPG aldolases by random mutagenesis and selected for cells that showed growth on KHO. Libraries were

generated by error prone PCR, allowing for rapid incorporation of one to five amino acid changes over the 0.6 kilobase gene, resulting in an overall library complexity of  $10^{14}$ . Although it is not currently feasible to exhaustively screen such a large library using in vivo selection methodologies, literature examples suggest that meaningful improvements in catalytic activity can be obtained by screening libraries of  $10^3-10^5$  clones.<sup>17,18</sup>

#### 2.1. Directed evolution and selection of T. maritima KDPG aldolase

Libraries of *T. maritima* aldolase variants were generated by error-prone PCR, using a commercially available low fidelity polymerase (Mutazyme II, Stratagene). Compared to mutagenesis methods using *Taq* DNA polymerase with  $Mn^{2+}$ , Mutazyme II provides a more even distribution of mutations at all nucleotides.<sup>19</sup> Plasmids that confer growth of PB25 cells on ribose were identified by plating on ribose/KHO minimal media at 34 °C (optimal growth temperature for PB25 cells). Colonies that grew within 120 h of plating were isolated. Based on transformation controls, the approximate size of the total number of variants screened in this selection was  $4 \times 10^7$ .

Plasmid DNA from colonies recovered in the first round of selection was isolated and the mutant eda genes were sequenced (shown in Table 1, TM 1-3). These plasmids were retransformed into PB25 cells and subjected again to the selection conditions to confirm that the growth phenotype was dependent on the plasmid encoded eda gene. Only plasmids that engendered growth were analyzed further. These isolated mutants contain 2-9 altered nucleotides in the KDPG gene leading to 2-7 amino acid changes (Table 1). Unexpectedly, one selected plasmid contained a TM eda gene with a frameshift mutation (TM 3) in the 2nd codon of the gene. Expression of nearly full-length protein is still observed for this mutant since the third codon is a redundant start. However, a Western blot of whole cell lysates of each first round mutant indicated that the selected mutations decrease expression of the mutants up to 10-fold as compared to the wild-type enzyme (Fig. 2A. Table 1).

To examine whether the selection procedure identified KDPG aldolases with increased catalytic efficiency, we analyzed the kinetic parameters for KHO cleavage catalyzed by these proteins. Each KDPG aldolase mutant was expressed in PB25 cells grown on rich media and purified using a C-terminal  $6 \times$  His-tag. Proteins of approximately 24 kDa, as determined by SDS–PAGE analysis, were obtained for all TM mutants. The degree of fitness towards cleavage of the KHO substrate was analyzed by measuring the stea-

Tuble I				
T. maritima KDPG aldolase	mutants selected	for enhanced	catalysis of I	KHO cleavage

Table 1

Name	Mutations	Relative expression level
TM 1	G34R/V172A/R190I	++++
TM 2	K9T/G83E	+++
TM 3	<sup>c</sup> Deletion of first two AA,	++
	L16P/V31A/S82C/G83V/F99L/W167R	
TM 4 <sup>a</sup>	V136M	++
TM 5 <sup>a</sup>	A30V	++
TM 6 <sup>b</sup>	I52F/V135M/T185A	+
TM 7 <sup>b</sup>	P109S	+
TM 8 <sup>b</sup>	N150I	+

++++ 100%; +++ 50%; ++ 10%; + 5% expression relative to wild-type TMA in TMEDA-pUC plasmid.

<sup>a</sup> These proteins were selected in a pUC plasmid where the expression of wildtype TMA is 10% of the expression in the TMEDA-pUC plasmid. <sup>b</sup> These proteins were selected in a pUC plasmid where the expression of wild-

These proteins were selected in a pUC plasmid where the expression of wild-type TMA is  $\sim 2\%$  of the expression in the TMEDA-pUC plasmid.

<sup>c</sup> Signifies a frameshift mutation where expression of the gene is initiated at the third amino acid.



**Figure 2.** (A) Western blot of His-tagged KDPG aldolase mutants after the first round of selection probing with an anti-His antibody and; (B) dependence of the initial velocity for pyruvate formation on the concentration of KHO catalyzed by wild-type *T. maritima* KDPG aldolase (open circle) and mutants TM 5 (open square) and TM 8 (filled diamond), measured using a lactate dehydrogenase coupled assay at 34 °C.

 Table 2

 Kinetic parameters for KHO cleavage catalyzed by selected TM KDPG aldolase mutants<sup>a</sup>

-					
	Name	$k_{\rm cat}({ m s}^{-1})$	$K_{\rm M}({ m mM})$	$k_{\rm cat}/K_{\rm M}~({ m M}^{-1}~{ m s}^{-1})$	Relative $k_{cat}/K_{M}$
	wt-TMA	$0.40 \pm 0.05$	200 ± 30	2.2 ± 0.3	1
	TM 1	$0.07 \pm 0.01$	9 ± 2	9.3 ± 1.8	4
	TM 2	$0.5 \pm 0.1$	90 ± 20	5.5 ± 1.0	2.5
	TM 3	$0.14 \pm 0.10$	$20 \pm 14$	7 ± 4	3
	TM 4	$0.36 \pm 0.02$	23 ± 4	16 ± 2	7
	TM 5	$0.75 \pm 0.10$	14 ± 5	54 ± 10	25
	TM 6	$0.13 \pm 0.02$	8 ± 3	17 ± 6	8
	TM 8	$1.2 \pm 0.3$	49 ± 19	25 ± 5	11

<sup>a</sup> Initial velocity for pyruvate formation was measured using a lactate dehydrogenase coupled assay at 34 °C. Steady state kinetic parameters were determined by fitting of the Michaelis–Menten equation to the initial velocity as a function of the substrate concentration.

dy-state kinetic parameters at the selection growth temperature, 34 °C. Values of  $k_{cat}/K_{M}$  for KHO cleavage increased for all three mutants by 2–4-fold (Table 2).

The most improved mutants, TM 3 and TM 1, increase the value of  $k_{cat}/K_M$  by three- and fourfold and decrease the value of  $K_M$  by at least 10 and 22-fold, respectively, compared to the corresponding parameters for KHO cleavage catalyzed by wild-type enzyme. Interestingly,  $k_{cat}$  decreases up to sixfold for these mutants, implying that the selective pressure in vivo is for increased catalytic efficiency ( $k_{cat}/K_M$ ) rather than turnover ( $k_{cat}$ ). This result suggests that the in vivo concentration of KHO under selective conditions is low, a conclusion consistent with the nature of KHO, a charged molecule that cannot easily pass through a lipid bilayer by simple diffusion. Based on these data we presume that the cellular concentration of KHO is in the  $\mu$ M to nM range under the selection conditions.

## 2.2. Selection of *T. maritima* KDPG aldolase mutants under increased stringency

To both increase the selection pressure and to mitigate any cellular strain created by overexpression of KDPG aldolase, we constructed plasmids that confer decreased KDPG aldolase expression by making mutations in the RNA polymerase binding region of the *lac* promoter (between -10 and -35) that have previously been shown to significantly decrease gene expression.<sup>20</sup> A TMEDA-pUC plasmid containing an A to C mutation at the -35 RNA polymerase binding region decreases TM KDPG aldolase expression about 10-fold, while a single guanosine nucleotide insertion into the 18 bp

region between the -10 and -35 region of the promoter decreases protein expression ~50-fold. Using these plasmids, libraries of TM KDPG aldolase genes were generated by error-prone PCR amplification. A library of  $1 \times 10^7$  plasmids was selected for conferring growth on ribose/KHO minimal media in PB25 cells and a total of 5 plasmids that allow for growth within 120 h were identified (Table 1, TM 4–8). In this case, expression of the selected mutant proteins is comparable to the reduced level observed for the wild-type enzyme in these plasmids.

Mutant KDPG aldolase genes were subcloned into a pET expression plasmid and the protein was expressed in XL1-Blue cells grown in rich media with IPTG induction. Mutant TM 7 precipitated during purification so kinetic data could not be obtained for this mutant. The kinetic parameters for catalysis of KHO cleavage for the other four mutants (Table 2) demonstrate that the value of  $k_{cat}/K_{\rm M}$  increases up to 25-fold, from a combination of improvements in  $K_{\rm M}$  (4 to 25-fold), as well as changes in the turnover number (0.3–3-fold). The activity of TM 5 showed the greatest improvement (Fig. 2B), due primarily to a >14-fold decrease in the value of  $K_{\rm M}$  for KHO, suggesting that new, beneficial protein-substrate contacts occur in this mutant. TM 8 shows a threefold increase in  $k_{\rm cat}$  in addition to a fourfold decrease in  $K_{\rm M}$  (Fig. 2B) making it one of the most promising mutants for biocatalysis, where maximal turnover under saturating conditions is critical.

#### 2.3. Directed evolution and selection of E. coli KDPG aldolase

For comparison, we constructed a library of *E. coli* KDPG aldolase variants and selected for enhanced cleavage of KHO. Mutagenesis was limited to two to three mutations per KDPG aldolase gene and the library was incorporated into a pUC plasmid that contains a deletion at the -35 RNA polymerase binding region (TTTACA to TTACA) to decrease protein expression and enhance the selection stringency. Plasmids were selected for conferring growth of PB25 cells within 120 h on minimal ribose/KHO plates at 34 °C. Based on transformation controls, the approximate size of the library screened in each case was  $\sim 10^7$  members. Plasmids from recovered colonies were isolated and sequenced, and used to template subsequent rounds of error-prone PCR without extensive kinetic analysis. After five sequential rounds of mutagenesis and selection, the mutant aldolase proteins from the final round of selection were purified and assayed for catalysis of cleavage of KHO.

Unfortunately, 2 of the 3 selected mutant proteins lost substantial activity during purification and did not yield reproducible kinetic results. The remaining mutant, EC 11, also showed limited stability. This mutant contained substitutions both proximal and distal to the active site that improved the value of  $K_{\rm M}$  for KHO

#### Table 3

Kinetic parameters for KHO cleavage catalyzed by selected EC KDPG aldolase mutants  $^{\rm a}$ 

Name	Mutations	$k_{\text{cat}}$ (s <sup>-1</sup> )	$K_M$ (mM)	$k_{cat}/K_{M}$ (M <sup>-1</sup> s <sup>-1</sup> )
Name EC 9	None V17L/E51K/P94L/ G124M/L173M	2±0.6	150 ± 50 —	14 ± 4 >0.1
EC 10	V17L/V23I/E51K/ P94L/F134I/L173M/D190E	_	_	>0.2
EC 11	V17L/A53T/P94L/ V118L/A137T/L173M	0.01 ± 0.005	1.5 ± 0.4	10 ± 3

<sup>a</sup> Initial velocity for pyruvate formation was measured using a lactate dehydrogenase coupled assay at 25 °C. Steady state kinetic parameters were determined by fitting of the Michaelis–Menten equation to the initial velocity as a function of the substrate concentration.

cleavage by 100-fold to 1.5 mM. However, the turnover number also decreased by a similar amount resulting in a mutant with a  $k_{\rm cat}/K_{\rm M}$  value comparable to that of wild-type enzyme (Table 3). These measurements may underestimate the value of  $k_{\rm cat}$  due to the instability of the protein.

#### 3. Discussion

Directed evolution is a powerful tool for altering the catalytic properties of enzymes where diversity libraries are created by error prone PCR or DNA shuffling methods and then subjected to iterative rounds of screening for the desired property.<sup>21,22</sup> More recently, semi-rational methods of library generation have helped minimize library complexity by focusing mutations on active site positions or mutational hotspots, and have yielded significant success.<sup>23</sup> Here we used random whole gene mutagenesis of *T. maritima* and *E. coli* KDPG aldolase to identify mutations both proximal and distal to the active site which significantly alter the efficiency for catalysis of KHO cleavage. Selections in both isoforms clearly demonstrate that improvements in  $K_{\rm M}$  and  $k_{\rm cat}/K_{\rm M}$  can be achieved with a minimal number of mutations at sites that would not have been predicted a priori based on protein structural data.

The KDPG aldolases are TIM barrel proteins that catalyze the retro-aldol cleavage of deoxy-sugars to pyruvate and the corresponding aldehyde via a Schiff base-dependent mechanism. The active site is located at the top of the pore created by the barrel structure and contains the important catalytic residues K129, E40, and R17 (Fig. 3) (K133, E45 and R49 in the *E. coli* protein).<sup>24,25</sup> Lysine 129 forms a Schiff base with the substrate, while E40 serves as an acid/base catalyst, and R49 forms critical charge-charge interactions with the carboxylate of the pyruvyl- substrate.<sup>25,26</sup> Part of the electrophile binding pocket is created by serine 179, (equivalent to S184 in the E. coli protein), a position that is important in EC KDPG aldolase for recognition of phosphorylated substrates.<sup>9</sup> The rest of the phosphate binding pocket consists of the amide nitrogen of G157, G158, and V159 (Fig. 3). Since KHO is not phosphorylated, it is likely to make poor contacts with the residues that line this pocket. Surprisingly, we did not select any mutants with alterations in these residues that enhance catalytic activity for cleavage of KHO.

In the *T. maritima* selections, many (32%) of the observed mutations are located 12–16 Å from the active site lysine, on solvent exposed  $\alpha$ -helices and loops. Mutations on solvent exposed surfaces include K9T, S82C, G83V, F99L, V172A, and T185A. The proteins containing these mutations also contain additional mutations buried within the protein, and these surface mutations may constitute neutral drift. Although, we did not test these mutations individually, it is possible that some may improve protein stability.<sup>27</sup> Several mutations (A30V, V31A, I52F, and W167R) are located in buried portions of the helices and may therefore communicate



Figure 3. Schematic diagram of KHO binding to TM KDPG aldolase.

structural changes to the neighboring  $\beta$ -sheets that line the active site. The most remarkable of these mutants is TM 5, in which a single A30V mutation results in a 25-fold improvement in the value of  $k_{cat}/K_{M}$ . A30 is located on  $\alpha$ -helix 1 with the side chain of this residue directed towards the interface with helix 8 (Fig. 4A). Helix 1 attaches to the loop containing the pyruvate-binding residue R17, while helix 8 is attached to the loop containing the phosphate binding site residue S179 (equivalent position to EC S184). Substituting alanine with valine at this site may increase the distance between these two helices and alter the position of the residues on the adjoining loops to enhance catalysis of KHO cleavage (Fig. 4A).

Several recovered mutations are near the active site (L16P, P109S, and N150I) and may directly affect catalytic residues and/ or directly contact the substrate. Of note is mutation L16P (TM 3), located at the C-terminal end of  $\beta$ -sheet 1 (Fig. 4A) adjacent to R17, the arginine that interacts with the carboxylate moiety of KHO. Another interesting mutation is P109S (TM 7), which was identified as a single mutation in the selection. Proline 109 is located on the inward face of  $\beta$ -sheet 5, and causes a kink in the sheet adjacent to the active site lysine 129. P109 is located 4 Å away from the catalytic lysine; mutation at this site may reposition or altering the mobility of the K129 side chain. Unfortunately this mutation also significantly diminishes enzyme stability, complicating analysis of catalytic properties. Additional rounds of mutagenesis might identify secondary mutations that enhance protein stability.

In contrast to the generally robust mutant proteins identified from the TM KDPG aldolase selections, recovered E. coli KDPG aldolase mutants showed marked instability. As a result the observed  $k_{\text{cat}}$  and  $k_{\text{cat}}/K_{\text{M}}$  values reported for these mutants represent a lower limit for these parameters; however, the  $K_{\rm M}$  value is independent of the active protein concentration and should reflect a genuine characteristic of the isolated proteins. The selected E. coli KDPG aldolase mutations primarily lower the value of  $K_{\rm M}$  for KHO, likely reflecting enhancements in the binding affinity for this substrate. The most active mutant after five rounds of mutagenesis (EC 11) improves the value of  $K_{\rm M}$  by 100-fold; however, the value of  $k_{\rm cat}$  decreases a maximum of 200-fold, resulting in a mutant with a  $k_{cat}/K_{M}$  comparable to that of the wild-type enzyme. (Mutants EC 9 and 10 also showed lower values for  $K_{\rm M}$  in the order of 10- to 20-fold but are not reported due to rapid inactivation of the proteins in vitro.) Three mutations are common to each of the selected EC KDPG aldolase mutants: V17L, P94V, and L173M. (Fig. 4B) Residue P94 lines the pyruvate binding cavity and is located ~4.6 Å from the catalytic lysine (K133). The conversion of proline 94 to valine likely alters the structure of the substrate binding pocket, enhancing interactions with KHO. V17L and L173M are located adjacent to each other, with their side chains pointing towards each other in the buried



**Figure 4.** Buried mutations in *T. maritima* and *E. coli* KDPG aldolases. (A) Structure of *T. maritima* KDPG aldolase (PDB:1WA3<sup>25</sup>. Mutant TM 5 contains a single mutation A30V. The side chain of A30 (in red) points between the helices 1 and 8. TM 3 contained numerous mutations but only L16P is buried. The side chain of L16 is shown in red. Helices are labeled as H1-8. The side chains of K129, R17 and a bound sulfate are shown for reference. (B) Structure of *E. coli* KDPG aldolase (PDB:1EUA<sup>24</sup>). All final round EC mutants possess V17L, L173M, and P94V (in red). The side chains of V17 and L173 point towards each other at the interface between  $\beta$ -sheets 1 and 8, and  $\alpha$ -helix 8. P94 lines the pyruvate binding pocket. Residues K133, R49 and a bound sulfate are shown for reference.

interface between  $\beta$ -sheets 1 and 8, and  $\alpha$ -helix 8. Similar to the A30V mutation observed in *T. maritima* aldolase, the V17L and L173M pair likely reposition helices 1 and 8, thereby affecting substrate recognition. The surprising correlation in the location of mutations observed in the two related aldolases may imply that the interface between helices 1 and 8 is important for defining the structure of the substrate binding pocket of KDPG aldolases. The function of the additional mutations in EC 11 (A53T, V118L, and A137T) is unclear, although they may alter protein stability.

The mutants isolated from TM and EC KDPG aldolase libraries all show considerable improvements in the  $K_M$  parameter with modest, if any, increases in  $k_{cat}$ . This observation suggests that the pressure applied by this selection is for the value  $k_{cat}/K_{M}$ , conferring a growth advantage to the cellular host. A relatively small concentration of KHO (2.5 mM) is used in the media; the intracellular concentration is presumably much lower. Therefore, it is not surprising that selections led to improvements in the  $K_{\rm M}$  parameter given that the cellular concentration of KHO is likely far lower than the wild-type  $K_{\rm M}$  value. The subtle remodeling of the binding pocket leads to enhanced binding of the substrate and enhances reactivity with the low concentration of the available substrate. However these mutations have a similar stabilizing effect on both the ground and transition states of the rate-limiting step leaving the value of  $k_{cat}$  unchanged. This type of enhancement, characterized as 'uniform binding',<sup>15</sup> is proposed as a first step in the evolution of novel enzymatic activity. Alterations in numerous interactions (hydrogen bonding, hydrophobic interactions, electrostatic effects, etc.) that are often distant from the active site can lead to uniform binding enhancements, increasing the likelihood that this form of enhancement is identified when screening libraries of modest size. Once the value of  $K_{\rm M}$  approaches the cellular concentration of the substrate, additional enhancements in catalytic efficiency can only be achieved by 'differential binding' where the transition states are stabilized relative to the substrate-bound ground states and specific catalysis of elemental steps.15

Our findings reveal significant limitations with in vivo selection methodologies to obtain enzymes suitable for biocatalytic applications: (i) when the concentration of the substrate is limiting, most improvements in catalysis are attributed to enhancements in the apparent substrate affinity rather than improving turnover at saturating substrate, and (ii) enhancements observed under in vivo conditions may not translate to the in vitro conditions used for biocatalytic transformations; and (iii) selection may lead to proteins with decreased stability. To achieve selective pressure required to increase turnover numbers, the in vivo substrate concentration must be comparable to or higher than the value of  $K_{\rm M}$ . For biocatalysis applications it is desirable that the binding affinity be weak and values of  $k_{\rm cat}$  be large. In this case, the selections must be carried out under conditions where the cellular substrate concentration is high. Unfortunately the KHO substrate is toxic to bacterial cells at concentrations above 10 mM, preventing such experiments for this application.

Our results also suggest that thermophilic, rather than the mesophilic, scaffolds may be superior for directed evolution, due to an enhanced initial stability that facilitates recovery of catalytically beneficial yet destabilizing mutations. In our hands, it was far easier to achieve an improvement in the *T. maritima* aldolase than the E. coli aldolase. Two rounds of mutagenesis yielded up to a 25-fold improvement in catalysis of KHO using the thermophilic aldolase; however after multiple rounds of mutagenesis, mutations in the E. coli enzyme improved the value of K<sub>M</sub> but had negligible impact on the value of  $k_{cat}/K_{M}$ . Moreover, the *E. coli* mutants proved to be difficult to handle in vitro, frustrating efforts to accurately analyze their kinetic parameters. The additional stability of the T. maritima scaffold presumably allows it to sample a greater number of mutations while still remaining properly folded, thereby making the screening effort more effective. A similar result was reported in the case of subtilisin.28

#### 4. Conclusions

Our data demonstrate that selection using a pyruvate auxotrophic strain can generate enzymes with enhanced kinetic efficiency towards cleavage of substrates with unnatural electrophiles. In theory, this method could select for reaction of aldolases with any unnatural aldehyde within the limits of solubility, chemical stability, competing reactivity, and acute toxicity to the cell. It is clear that significant improvements in  $K_{\rm M}$  and  $k_{\rm cat}/K_{\rm M}$  can be achieved with a minimal number of mutations at sites that would not have been identified a priori from protein structural data. However, in vivo selections may not be optimal for generating improvements in the turnover number ( $k_{\rm cat}$ ) of an enzyme, as high substrate concentrations may not be readily achieved inside the cell. These findings also point to one of the challenges of in vivo selections; the overall fitness of the organism is a function of careful tuning of expression levels and activity of multiple proteins.

Most of the selected mutants improve KHO catalysis by lowering the  $K_{\rm M}$  parameter, suggesting that the cellular steady-state concentration of KHO is low. As a result, these in vivo selections mainly identify improvements in  $K_{\rm M}$  until the value of  $K_{\rm M}$  is comparable to the cellular concentration of the substrate. Beyond this point further improvements in  $K_{\rm M}$  will not increase the catalytic efficiency of the enzyme or viability of the cells; additional improvements must derive from an increase in the value of the turnover number. For biocatalyst development, the value of  $k_{cat}$ is arguably the most important kinetic parameter, so developing methods to engineer enzymes with improved  $k_{cat}$  is desirable. The intracellular concentration of KHO is difficult to control; however the periplasmic influx of small molecules is generally unregulated<sup>29</sup> and KHO concentrations in the periplasm should parallel those in the medium. Therefore, localization of the selected enzyme to the periplasm may circumvent issues with substrate concentration. A key advantage of this approach would be that the concentration of KHO can be directly controlled by the growth conditions, presumably allowing for the rapid engineering of aldolases with improved values of  $k_{cat}$ .

#### 5. Materials and methods

#### 5.1. Library creation

Error prone PCR amplification of the T. maritima eda genes was conducted using the Genemorph mutagenesis kit (Stratagene). Primers containing Sacl and Xhol restriction sites were used (forward 5'-GGAAACAGCTATGACCATGATTACGAATTCGAGCTCTAC CATG-3' and reverse 5'-CTCAGTGGTGGTGGTGGTGGTGGTGGTGCTCGACT TC -3'). The eda genes were encoded on pUC derived plasmid TME-DA-pUC as previously reported.<sup>14</sup> Following PCR amplification, the DNA was fractionated on a 1% agarose gel and extracted using Microcon Ultrafree DA spin filters. The plasmid was simultaneously digested with SacI and XhoI in NEB Buffer 4 and BSA (New England BioLabs) and the DNA encoding the *eda* gene was purified on a 1% agarose gel and extracted with phenol/chloroform followed by precipitation with ethanol using standard protocols.<sup>30</sup> The vector for ligation reactions was made by digestion of the original TMEDApUC plasmid, (or reduced expression vectors described below) with SacI and XhoI and purification on a gel followed by phenol:chloroform extraction. Ligations were run using a 1:1.2 molar excess of insert to vector at 125 ng/µL total DNA concentration. T4 DNA Ligase (New England BioLabs) was used with the supplied buffer and 10 mM ATP. Ligation reactions were incubated on ice overnight, allowing the reaction to slowly reach room temperature. The product DNA was extracted with phenol:chloroform and precipitated with ethanol before resuspending in water to a final concentration of 625 ng/µL. The E.coli KDPG aldolase library was constructed similarly to the TM library using the following primers: forward 5'-GAAAACAGCTATGACCATGATTACGAATTC<u>GAGCTC</u>GCTCATG-3'; reverse 5'-ATCTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTG' and the pUC-ECEDA derived plasmid.14

The PB25 strain is an *E. coli* K12 JM101-derived strain where the pyruvate kinase genes, *pykA* and *pykF*, have been disrupted [*supE thi* $\Delta$ (*lac-proAB*) *pykA::kan<sup>r</sup> pykF::cat<sup>r</sup>*]. Electrocompetent PB25 cells were made according to the protocol described by Hanahan et al.<sup>31</sup> except that 0.2% glucose was added to the growth media to enhance growth of the auxtrophic cell line and stored in 80 µL aliquots at  $-80 \,^{\circ}$ C for up to 3 months. Electroporation of up to 1 µg of DNA was done using a BIORAD Gene Pulser in 0.1 mm cuvettes under standard *E. coli* transformation conditions.<sup>31</sup> The cells were immediately resuspended in a total of 4 mL of SOC media and allowed to recover for 1 h at 34 °C and 250 rpm. The cells were pelleted and resuspended in 4 mL of M9 media twice to remove trace amounts of glucose before plating on selective minimal media.

#### 5.2. Low expression KDPG aldolase vectors

Mutations were made in the *lac* promoter of the pUC plasmid using site-directed mutagenesis. The –35 region was changed from TTTACA to TTTCCA by whole plasmid PCR amplification using the following primer and its reverse complement: 5'-GCACCCCAGGCTT TCCACTTTATGC-3'. A single guanosine nucleotide was introduced in the spacer region between –10 and –35 using the following primer and its reverse complement: 5'-CACTTTATGCTCTCCGGCTCGTAT GTTG-3'. A point deletion was made in the –35 region of the *lac* promoter (TTTACA to TTACA) using the following primer and its reverse complement; 5'-GCACCCCAGGC<u>TT</u>ACACTTTATGC-3'. The PCR products were transformed into SmartCells (Genlantis) using the standard heat shock methods. Following overnight growth on LB agar plates containing 100 µg/mL ampicillin (LB/AMP), a single colony was grown up and the plasmid DNA was purified and sequenced.

The level of protein expression achieved by various plasmids was determined by isolating PB25 cells that were transformed with each plasmid. Ten large colonies were picked from LB/AMP and resuspended in 100 µL H<sub>2</sub>O to a final concentration of  $OD_{600}$  = 1.25. The cells were lysed by heating for 10 min at 95 °C after addition of SDS loading dye. The cell debris was pelleted and the supernatant was fractionated on a 12% polyacylamide gel. The proteins were transferred electrophoretically to nitrocellulose paper and blotted according to the manufacturer's procedures using Mouse anti-His antibody (Novagen) followed by Goat anti-Mouse AP-conjugate secondary antibody (Novagen). The presence of bound antibody was visualized by staining with 5-bromo-4-chloro-3'-indolyphosphate *p*-toluidine and nitro-blue tetrazolium chloride to detect phosphatase activity. The blot was quantified using the Gel-Pro Analyzer software (Media Cybernetics) to determine the level of expression.

#### 5.3. Selection conditions

Transformed cells were plated on selective M9 minimal media that contained the following: 0.2% ribose, 1  $\mu$ g/mL thiamine, 40  $\mu$ g/mL proline, 50  $\mu$ g/mL carbenecillin, and 2.5 mM KHO. The plates were further supplemented with 1  $\mu$ M FeCl<sub>3</sub>, 50  $\mu$ M ZnSO<sub>4</sub>, 0.1  $\mu$ M CaCl<sub>2</sub>, 40  $\mu$ g/mL each adenosine, guanosine, thymidine, cytosine and uracil, and 1  $\mu$ g/mL each of D-biotin, cyanocobalamin, folic acid, niacinamide, D-pantothenate, pyridoxal hydrochloride and riboflavin. These supplements improve the growth of the PB25 strain on minimal plates without negating the pyruvate auxotroph phenotype. The library size was estimated by counting the number of recovered colonies on LB plates containing carbenicillin. Selection plates were grown at 34 °C.

As controls, plasmids containing the wild-type aldolase genes were transformed into PB25 cells and plated on selective media alongside the libraries. No significant growth was seen on this plate within 120 h. All colonies that grew in the first 120 h on selection plates were picked and restreaked on selective media and then grown in 5 mL LB liquid cultures containing carbenicillin to recover the plasmid DNA. The plasmid DNA was isolated using the MoBio UltraClean DNA spin kit and was sequenced at the University of Michigan DNA sequencing core. The plasmids were retransformed and only plasmids that reproduced the growth phenotype on selection plates are reported.

#### 5.4. Kinetics

2-Keto-4-hydroxyoctonoate, synthesized and purified as described by Griffiths et al.<sup>14</sup> was used as the substrate. Aldolase activity was measured using a lactate dehydrogenase (LDH) coupled assay where the production of pyruvate was measured by the decrease in NADH absorbance at 340 nm.<sup>7</sup> All assays were done in guartz microcuvettes (70 µL) using a CARY 100 Bio UV/vis spectrometer fitted with a Peltier temperature controller. The reactions contained the following final concentrations: HEPES (100 mM, pH 7.5), NADH (250 µM), LDH (0.023 Units/µL), 100 µM to 50 mM KHO. Each reaction was initiated with addition of 1 µM enzyme. Kinetic parameters for the TM aldolase variants were measured at the selection temperature (34 °C), while the EC aldolase variants were measured at 25 °C to minimize thermal denaturation of unstable mutants. Each measurement was made three times and averaged, and the Michaelis-Menten equation was fit to the data using the curve-fitting program KaleidaGraph (Synergy Software).

#### Acknowledgments

This work was supported by the National Institutes of Health [Grant number R01 GM61596]. M.C. received support from the Chemical Biology Interface training program [GM08597]. M.J.W. was supported by the Biological Chemistry training program [T32 GM008558].

#### Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.bmc.2011.08.056.

#### **References and notes**

- Faber, K. Biotransformations in Organic Chemistry; Springer: Berlin, 2004. 1.
- Fessner, W. D.; Helaine, V. Curr. Opin. Biotechnol. 2001, 12, 574. 2.
- Bolt, A.; Berry, A.; Nelson, A. Arch. Biochem. Biophys. 2008, 474, 318. 3.
- 4. Clapes, P.; Fessner, W. D.; Sprenger, G. A.; Samland, A. K. Curr. Opin. Chem. Biol. 2010. 14. 154
- 5. Henderson, D. P.; Cotterill, I. C.; Shelton, M. C.; Toone, E. J. J. Org Chem 1998, 63, 906
- 6. Griffiths, J. S.; Wymer, N. J.; Njolito, E.; Niranjanakumari, S.; Fierke, C. A.; Toone, E. J. Bioorg. Med. Chem. 2002, 10, 545. 7
- Shelton, M. C.; Cotterill, I. C.; Novak, S. T.; Poonawala, R. M.; Sudarshan, S.; Toone, E. J. J. Am. Chem. Soc. 1996, 118, 2117.
- 8 Henderson, D. P.; Shelton, M. C.; Cotterill, I. C.; Toone, E. J. J. Org. Chem. 1997, 62, 7910
- 9 Cherivan, M.: Toone, E. I.: Fierke, C. A. Protein Sci. 2007, 16, 2368.
- Stock, S. D.; Hama, H.; Radding, J. A.; Young, D. A.; Takemoto, J. Y. Antimicro. 10 Agents Chemo, 2000, 44, 1174.
- Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. J. Nat. Prod. 2000, 63, 1437. 11 Royer, S. F.; Haslett, L.; Crennell, S. J.; Hough, D. W.; Danson, M. J.; Bull, S. D. J. 12.
- Am. Chem. Soc. 2010, 132, 11753. 13 Ponce, E.; Flores, N.; Martinez, A.; Valle, F.; Bolivar, F. J. Bacteriol. 1995, 177.
- 5719. 14. Griffiths, J. S.; Cheriyan, M.; Corbell, J. B.; Pocivavsek, L.; Fierke, C. A.; Toone, E. J.
- Bioorg. Med. Chem. 2004, 12, 4067. 15.
- Albery, W. J.; Knowles, J. R. Biochemistry 1976, 15, 5631. 16
- Haldane, J. B. S. Enzymes; Longmans: London, 1930.
- Fong, S.; Machajewski, T. D.; Mak, C. C.; Wong, C. H. Chem. Biol. 2000, 7, 873. 17
- 18 Ran, N.; Frost, J. W. J. Am. Chem. Soc. 2007, 129, 6130.
- Zhang, V. Q.; Hogrefe, H. H. Methods Mol. Biol. 2010, 634, 399. 19
- Piette, J.; Decuyper-Debergh, D.; Gamper, H. Proc. Natl. Acad. Sci. U.S.A. 1985, 82, 20. 7355.
- 21 Reetz, M. T. Angew. Chem., Int. Ed. 2011, 50, 138.
- Jackel, C.; Hilvert, D. Curr. Opin. Biotechnol. 2010, 21, 753. 22
- 23. Lutz, S. Curr. Opin. Biotechnol. 2010, 21, 734.
- Allard, J.; Grochulski, P.; Sygusch, J. Proc. Natl. Acad. Sci. U.S.A. 2001, 98, 3679. 24
- 25. Fullerton, S. W. B.; Griffiths, J. S.; Merkel, A. B.; Cheriyan, M.; Wymer, N. J.; Hutchins, M. J.; Fierke, C. A.; Toone, E. J.; Naismith, J. H. Bioorg. Med. Chem. 2006, 14, 3002.
- 26. Meloche, H. P.; Wood, W. A. J. Bio. Chem. 1964, 239, 3511.
- Chiti, F.; Stefani, M.; Taddei, N.; Ramponi, G.; Dobson, C. M. Nature 2003, 424, 27. 805.
- 28 Bloom, J. D.; Labthavikul, S. T.; Otey, C. R.; Arnold, F. H. Proc. Natl. Acad. Sci. U.S.A. 2006, 103, 5869.
- Sroga, G. E.; Dordick, J. S. Prot. Eng. 2001, 14, 929. 29
- 30. Sambrook, J.; Russell, D. W. Molecular Cloning: a Laboratory Manual; Cold Spring Harbor: Cold Spring Harbor Laboratory, 2001.
- 31. Hanahan, D.; Jessee, J.; Bloom, F. R. Method Enzymol. 1991, 204, 63.