Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 12 (2004) 4067–4074
Bioorganic & Medicinal Chemistry

A bacterial selection for the directed evolution of pyruvate aldolases

Jennifer S. Griffiths,[a] Manoj Cheriyan,[b] Jayme B. Corbell,[a] Luka Pocivavsek,[a] Carol A. Fierke[b,*] and Eric J. Toone[a,*]

[a]*Department of Chemistry, Duke University, Box 90317, Durham, NC 27708, USA*
[b]*Department of Chemistry, University of Michigan, Box 1055, Ann Arbor, MI 48109, USA*

Received 23 January 2004; revised 18 May 2004; accepted 26 May 2004
Available online 24 June 2004

Abstract—A novel bacterial in vivo selection for pyruvate aldolase activity is described. Pyruvate kinase deficient cells, which lack the ability to biosynthetically generate pyruvate, require supplementation of exogenous pyruvate when grown on ribose. Supplementation with pyruvate concentrations as low as 50 μM rescues cell growth. A known substrate of the KDPG aldolases, 2-keto-4-hydroxy-4-(2′-pyridyl)butyrate (KHPB), also rescues cell growth, consistent with retroaldol cleavage by KDPG aldolase and rescue through pyruvate release. An initial round of selection against 2-keto-4-hydroxyoctonate (KHO), a nonsubstrate for wild-type aldolase, produced three mutants with intriguing alterations in protein sequence. This selection system allows rapid screening of mutant enzyme libraries and facilitates the discovery of enzymes with novel substrate specificities.

1. Introduction

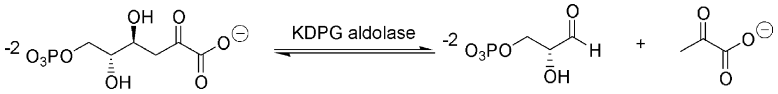
The development of enzymes with novel, predetermined activities remains a major challenge in chemical biology. Relationships between the primary amino acid sequence and the behavior of the corresponding folded protein, including properties such as substrate specificity, level of stereochemical discrimination, and rate of turnover, remain obscure. 'Directed evolution', a process in which a predetermined phenotype is selected from a library of random genotypes, obviates the need for a priori knowledge of structure–activity relationships and has emerged as a powerful tool both for the creation of useful proteins and the study of structure–function relationships in protein catalysis.[1–8] The process involves three steps: the generation of diversity through random mutagenesis, the evaluation of diversity through screening or selection for the desired function, and the recovery and detailed characterization of the mutants. When these steps are done in a recursive fashion, the protein can be rapidly optimized toward the phenotype of interest. Evolutionary approaches have been used to prepare enzymes with novel substrate specificities,[9–12] regioselectivities,[13] enantioselectivities,[14] thermostabilities,[15] and increased activities in organic solvents.[16,17]

Random incorporation of point mutations and random recombination of homologous genes, the two fundamental strategies used for the creation of genetic diversity, broadly mimic the evolutionary process in asexual and sexual evolution, respectively. Error-prone PCR produces genetic diversity through the random incorporation of point mutations.[18–22] While random mutagenesis accumulates favorable point mutations over several rounds of mutagenesis, it also introduces deleterious mutations—presumably at a rate greater than that for incorporation of favorable mutations—that are carried through subsequent rounds of mutation and selection. The familiar 'Müllers ratchet' paradox of population biology is alleviated by in vitro recombination processes that provide a mechanism for the excision of unfavorable mutations. Commonly utilized strategies include DNA shuffling,[23,24] staggered extension process (StEP),[25] and iterative truncation (ITCHY).[26,27]

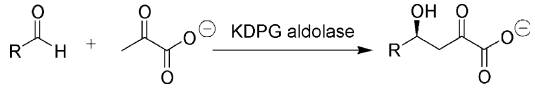
Whereas several strategies for the creation of large diversity libraries exist, methods for selection of a desired phenotype from libraries of as many as 10^{10} members remains the major challenge of evolutionary

Keywords: KDPG aldolase; Directed evolution; Enzyme catalysis; Selection.
* Corresponding authors. Tel.: +1-919-681-3484; fax: +1-919-668-5483 (E.J.T.); tel.: +1-734-936-2678; fax: +1-734-647-4865 (C.A.F.); e-mail addresses: fierke@umich.edu; toone@chem.duke.edu

0968-0896/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2004.05.034



Scheme 1. KDPG aldolase reversibly cleaves KDPG into D-glyceraldehyde 3-phosphate and pyruvate as part of the Entner–Doudoroff pathway.



Scheme 2. General synthetic activity of KDPG aldolases. Electrophilic substrates include 2-pyridine carboxaldehyde (R = pyridine), lactaldehyde ($R = CH_2OHCH_3$), and glyceraldehyde ($R = CH_2OHCH_2OH$). Nonreactive compounds, such as benzaldehyde (R = phenyl) and valeraldehyde (R = *n*-butyl), lack polar functionality at C2 and C3.[32]

approaches.[8,28,29] One of two general strategies for the identification of efficacious mutants is typically employed. In a screening strategy each library member is examined for the desired trait. The individual examination of distinct proteins limits the library size that can be evaluated to roughly 10^4 members. A facilitated (or in vivo screen couples the desired activity to an event detectable directly in cells, without the requirement for protein isolation. Such approaches typically make use of fluorogenic or chromogenic substrates, and extend the library size that can be evaluated to approximately 10^8.[29] In a selection strategy, enzyme efficacy is coupled to host survival, and desirable phenotypes are thus the only mutations observed. Selection strategies are more difficult to develop than screens, but allow for the evaluation of much larger libraries, perhaps as large as 10^{10}.

2-Keto-3-deoxy-6-phosphogluconate (KDPG) aldolases are enzymes of demonstrated utility for the synthesis of various natural products.[30–32] In vivo, these enzymes function in the Entner–Doudoroff glycolytic pathway to catalyze the reversible retro-aldol cleavage of KDPG to pyruvate and glyceraldehyde 3-phosphate (Scheme 1).[33] As synthetic catalysts, KDPG aldolases provide substituted 2-keto-4-hydroxybutyrates with the *S*-configuration at the newly formed stereogenic carbon.[31] The enzymes show limited substrate tolerance with regard to the electrophilic component, primarily requiring a polar functionality at C2 or C3 (Scheme 2). KDPG aldolases typically have higher specificity for the nucleophilic component, accepting only oxobutyrate, and fluoropyruvate in addition to the natural substrate. To extend the synthetic utility and better understand the structural basis of substrate specificity in KDPG aldolases, we have initiated directed evolution studies of the *Escherichia coli* and *Thermotoga maritima* KDPG aldolases. Previous directed evolution efforts have used a screen based on the coupled oxidation of the β-hydroxyketone product with the reduction of NAD^+ catalyzed by horse liver alcohol dehydrogenase (HLADH).[34,35] Here we report the development of a selection strategy for the evolution of the substrate specificity of KDPG aldolase based on a pyruvate kinase-deficient strain of *E. coli*.

2. Results and discussion

Our selection approach exploits the *E. coli* PB25 cell line originally isolated and characterized by Ponce and colleagues.[36,37] The genes for both pyruvate kinase isozymes (*pykA* and *pykF*) have been interrupted by insertional mutagenesis with antibiotic resistance markers (*pykA::kan*, *pykF::cat*), making the strain both auxotrophic for pyruvate during growth on ribose as the sole carbon source and resistant to kanamycin and chloramphenicol. The cells are rescued by pyruvate augmentation, showing growth on minimal ribose media containing as little as 50 μM pyruvate. Our selection is based on rescue of this auxotroph by retro-aldol cleavage (catalyzed by KDPG aldolase mutants) of an unnatural 2-keto-4-hydroxybutyrate adduct to form pyruvate (Fig. 1). The principle of microscopic reversibility requires that enzymes selected for activity in the cleavage direction will also catalyze the reaction in the synthesis direction.

The PB25 cell line was constructed from JM101 [*supE thi*Δ(*lac-proAB*) (F′ *traD36 proAB lacI*^{q}*Z*ΔM15)].[36] In this strain the proline biosynthetic genes, *proA* and *proB*, are deleted and the biosynthesis of thiamine is prevented. The ability to grow on minimal media is achieved by transformation with the transposable element F′ *traD36 proAB lacI*^{q}*Z*ΔM15 that restores proline biosynthesis, and supplementation with exogenous thiamine in the media. The F′ plasmid contains a constitutively expressed *lac* repressor that can be used in subsequent screens to reduce the amount of aldolase expressed by repressing the upstream *lac* promoter on the pUC-derived aldolase plasmids. However, for our initial selection experiments we isolated PB25 cells lacking the F′ episome by selecting for the requirement of proline supplementation for growth on minimal media. PB25 cells lacking the F′ plasmid fail to grow on minimal media containing thiamine and either ribose or glucose as the carbon source.[38] Supplementation with proline and thiamine facilitates growth on glucose, but not on ribose plates (Table 1). Addition of pyruvate at concentrations above 50 μM to minimal media supplemented with proline and thiamine allows growth on ribose (Table 2), suggesting that in vivo production of pyruvate through KDPG aldolase catalyzed retro-aldol cleavage should also rescue cell growth (Fig. 1).

To ensure robust expression of the altered enzymes under the selection conditions, the *E. coli* and *T. maritima eda* genes encoding KDPG aldolase were subcloned and inserted into a modified pUC vector. This plasmid contains a selectable ampicillin marker and expression of the gene is controlled by a *lacZ* promoter. The pUC vector also contains the pMB1 replicon *rep* and is present at high copy number in the cell, allowing for

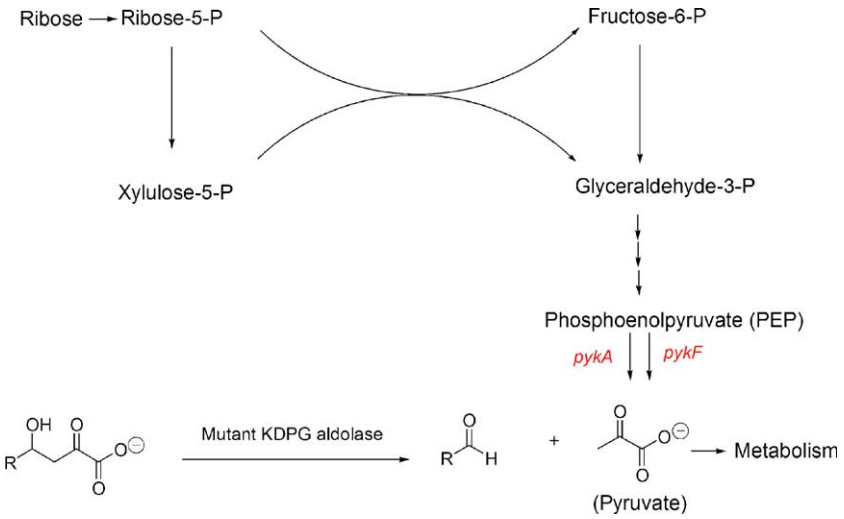


Figure 1. Abbreviated ribose metabolism pathway and selection strategy in *E. coli* PB25 cell line.[36] The genes encoding the two pyruvate kinase isozymes *pykA* and *pykF* have been interrupted by antibiotic resistance markers. Only cells possessing a mutant aldolase that can catalyze the retro-aldol cleavage to make pyruvate will grow.

Table 1. Growth conditions for the PB25pUC-ECEDA cells[a]

Additional supplements	Glucose	Ribose	Ribose + pyruvate (5 mM)
Amp[b]/kan[b]	–	–	n/a
Amp[b]/pro[c]/B$_1$[d]	+++++	–	+++++
Amp[b]/kan[b]/Cam[e]/pro[c]/B$_1$[d]	+++	–	+++

[a] Agar media contains M9 salts and the appropriate sugar at 0.4%. '–' indicates no growth after 120 h; '+++' indicates growth after 72 h; '+++++' indicates growth after 24 h.
[b] 50 μg mL^{-1}.
[c] 20 μg mL^{-1}.
[d] 1 μg mL^{-1}.
[e] 30 μg mL^{-1}.

Table 2. Effect of pyruvate supplementation on the growth of PB25pUC-ECEDA cells

Type of plate[a]	PB25pUC-ECEDA[b]
M9/ribose	–
M9/ribose + 5 mM pyruvate	++++
M9/ribose + 0.5 mM pyruvate	++++
M9/ribose + 0.05 mM pyruvate	+
M9/ribose + 0.005 mM pyruvate	–
M9/glucose	++++
LB	+++++

[a] All plates are supplemented with ampicillin (50 μg mL^{-1}), kanamycin (50 μg mL^{-1}), chloramphenicol (30 μg mL^{-1}). M9 plates were also supplemented with proline (20 μg mL^{-1}), thiamin (1 μg mL^{-1}), and the appropriate sugar (0.4%).
[b] '–' indicates no growth after 120 h; '+' indicates growth after 120 h; '++++' indicates growth after 48 h; '+++++' indicates growth after 24 h.

high levels of constitutive protein expression in the absence of high levels of *lac* repressor. Transformation of the PB25 cell line lacking the F′ episome (see discussion above) with the pUC vectors containing an inserted KDPG aldolase gene allows high levels of protein expression levels without isopropylthiogalactoside induction. The plasmids encoding *E. coli* and *T. maritima* KDPG aldolase genes are hereafter referred to as pUC-ECEDA and pUC-TMEDA, respectively. In both constructs a 6xHis tag is added to KDPG aldolase (on the N-terminus and C-terminus for the *E. coli* and *T. maritima* enzymes, respectively), which allows for rapid purification of mutant aldolases. The plasmid pUC-ECEDA2 was also constructed with a C-terminal 6XHis tag.

PB25 cells contain resistance markers for kanamycin (kan) and chloramphenicol (cam), and the pUC plasmid containing the *eda* gene confers resistance to ampicillin (amp), in theory facilitating stringent selection of growth on all three antibiotics. In practice, the presence of multiple antibiotics significantly inhibits cell growth (Table 1). Colonies were observed after 24 h of growth on M9/glucose plates in the presence of ampicillin alone (50 μg mL^{-1}), while 72 h was required for growth on the same media containing all three antibiotics (50 μg mL^{-1} amp, 50 μg mL^{-1} kan, and 30 μg mL^{-1} cam). While eliminating kanamycin and/or chloramphenicol from the media promotes cell growth, it could in principle lead to the selection of false-positives from reversion of the cell-line phenotype. However, the background frequency of cell reversion remains low in all of our experiments, as indicated by the low number of PB25 colonies transformed with the wild-type KDPG aldolase plasmid that grow on the selection plates.

To determine the minimum concentration of pyruvate required for growth, PB25 cells containing the pUC-ECEDA plasmid were cultured on M9/ribose plates containing various concentrations of pyruvate. Pyruvate concentrations as low as 0.5 mM produce visible cell growth at 48 h (Table 2). Plates supplemented with 0.05 mM (50 μM) pyruvate support growth of very small colonies after 120 h, while 5 μM augmentation produced no growth after 120 h. Similar results were observed for

growth of PB25 cells containing the pUC-TMEDA plasmid (data not shown). Although the equilibrium constant for KDPG aldolase activity lies strongly in favor of synthesis,[39] cleavage of selection substrates in vivo is coupled to pyruvate consumption so the selection should function effectively for even weakly active compounds.

The ability to screen large libraries is reduced by the poor viability of PB25 cells when grown on selective media. Serial dilutions of PB25 cells grown on M9/kan/cam/glucose plates show about 50% fewer colonies than those plated onto LB/kan/cam plates. Even fewer colonies are observed on minimal media plates compared to LB plates after transformation protocols or under stringent selection conditions (data not shown). To enhance growth on selection media, supplementation of M9 plates with vitamins (10 μg mL^{-1} each of pyridoxal, D-biotin, cyanocobalamin, niacinamide, D-pantothenate, folic acid, and riboflavin), metals ($FeCl_3$ and $ZnSO_4$), and/or nucleotides (40 μg mL^{-1} each of adenosine, guanosine, thymidine, cytosine, and uracil) was examined. Supplementation of M9/ribose plates with vitamins and nucleic acids failed to rescue the auxotrophic phenotype. However, these supplements did modestly enhance the number of colonies observed when PB25 cells are plated onto minimal media plates containing glucose as compared to the LB control (Table 3).

To test the utility of the PB25 cell line in directed evolution experiments, a model selection was carried out using 2-keto-4-hydroxy-4-(2′-pyridyl)butyrate (KHPB) (Fig. 1, R = pyridine) to rescue cell growth. This substrate is prepared enzymatically from pyruvate and 2-pyridinecarboxaldehyde using wild-type *E. coli* KDPG aldolase (Scheme 2, R = pyridine). Microscopic reversibility requires that this adduct will be cleaved by wild-type KDPG aldolase to produce pyruvate and should therefore rescue the auxotrophic PB25 cell line, provided that the molecule enters the cell (Fig. 1).[32]

PB25 cells were electroporated with either the pUC-ECEDA or pUC-TMEDA plasmids and plated on M9/ribose plates supplemented with 5 mM KHPB (Table 4). Visible colonies were observed on KHPB plates after incubation for 72 and 48 h for the PB25 cells containing pUC-ECEDA and pUC-TMEDA, respectively, demonstrating that supplementation with KHPB can rescue the cells. No colonies were observed when cells were plated on minimal media ribose plates (no KHPB) after incubation up to 120 h indicating that cell-line reversion does not occur with great frequency. As a positive control, growth was observed for cells plated on minimal glucose plates after 48 h of incubation.

Table 3. Number of colonies observed after plating PB25 cells on M9 plates supplemented with vitamins, metals, and nucleotides compared to LB plates

Type of supplement	M9/glucose[a]	LB[a]
None	++	++++
Vitamin + metal solutions	+++	N/A
Nucleotide solution	+++	N/A
Vitamin + metal + nucleotides	+++	N/A

[a] All plates are supplemented with kan (25 μg mL^{-1}), cam (17 μg mL^{-1}), and M9 plates were also supplemented with proline (20 μg mL^{-1}), thiamin (1 μg mL^{-1}), and the appropriate sugar (0.2%). The composition of the added solutions is described in the materials and methods section. '++' indicates that the number of colonies was 25–50% the number observed on LB plates; '+++' indicates 50–75%; '++++' indicates 75–100%.

Table 4. Growth of PB25 cells transformed with a plasmid encoding KDPG aldolase on media supplemented with 4-hydroxy-2-keto-4-(2′-pyridyl)butyrate (KHPB, 5 mM)

Type of plate[a]	pUC-ECEDA	pUC-TMEDA	pUC19
M9/glucose	++++	++++	++++
M9/ribose/KHPB	+++	++++	+++
M9/ribose	–	–	–

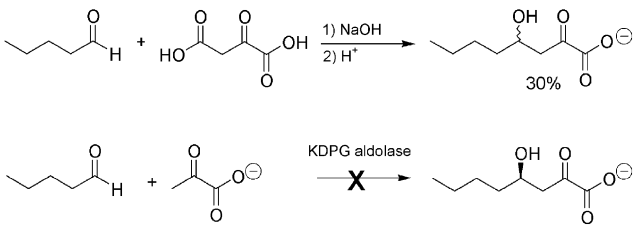
'–' indicates no growth after 120 h; '+++' indicates growth after 72 h; and '++++' indicates growth after 48 h.

[a] M9 agar plates were supplemented with proline (20 μg mL^{-1}), thiamin (1 μg mL^{-1}), antibiotics: 25 μg mL^{-1} each kan and amp, 17 μg mL^{-1} cam, and the appropriate sugar (0.4%).

As a final control, pUC19 plasmid lacking the *eda* gene was electroporated into the pyruvate kinase deficient PB25 cells. These PB25pUC19 cells also grow on M9/Ribose plates augmented with KHPB, likely due to the endogenous KDPG aldolase in PB25 cells. However, the possibility of an alternative metabolic pathway for the production of pyruvate from KHPB cannot be definitively ruled out.

To further test this selection strategy, we examined the growth of PB25 cells on media supplemented with a compound derived from valeraldehyde and pyruvate, 2-keto-4-hydroxy-octanoate (KHO). The synthesis of this substrate is shown in Scheme 3. Because valeraldehyde is not a substrate for wild-type KDPG aldolase,[32] endogenous enzyme cannot convert 2-keto-4-hydroxyoctanoate to pyruvate to rescue cell growth. Consistent with this observation, minimal ribose plates containing 5 mM KHO do not support growth after 120 h of incubation for PB25 cells alone or when transformed with either pUC-ECEDA or pUC-TMEDA suggesting that this compound will be an appropriate substrate for selection experiments.

Using this new selection method, pyruvate aldolase activity with a diverse range of aldehydes can be evolved. KDPG aldolase activity in the synthetic direction using benzaldehyde and its derivatives as substrates



Scheme 3. Chemical synthesis of 2-keto-4-hydroxy-octanoate (KHO). Valeraldehyde is not a substrate for the KDPG aldolases.

is of particular interest.[31] Previous directed evolution experiments on KDPG aldolase used a microtiter-plate HLADH/NADH assay to evolve enzymes with altered substrate specificity or reversal of enantioselectivity at the C4 hydroxyl group.[34,35] This screening method couples KDPG aldolase activity in the synthetic direction with oxidation of the resulting C4 hydroxyl group and to subsequent turnover of NAD^+ to NADH catalyzed by HLADH. NADH is detected spectrophotometrically using a fluorescence plate reader. While some success was achieved using this screen, maximum library sizes assayed using this labor-intensive method were on the order of 10^4.[35] An initial round of selection of *E. coli* KDPG aldolase mutants for growth of PB25pUC-ECEDA2 cells on minimal ribose plates supplemented with KHO allowed us to screen greater than 10^7 mutants in a single afternoon. In this initial round of selection, three colonies were observed on the selection plates after 120 h. The plasmid from each colony was isolated and the KDPG aldolase genes were sequenced demonstrating intriguing alterations in the protein sequence (Table 5, Fig. 2). Characterization of these mutants as well as further rounds of mutagenesis and selection will be reported in due course.

Table 5. Mutations in the KDPG aldolase gene selected by growth of PB25pUC-ECEDA cells on minimal media supplemented with KHO

Mutant designation	Mutations
EC-1	P77S, G183R
EC-2	V22L, N168S, R170S, K202M
EC-3	G70S, Q89R, K202N

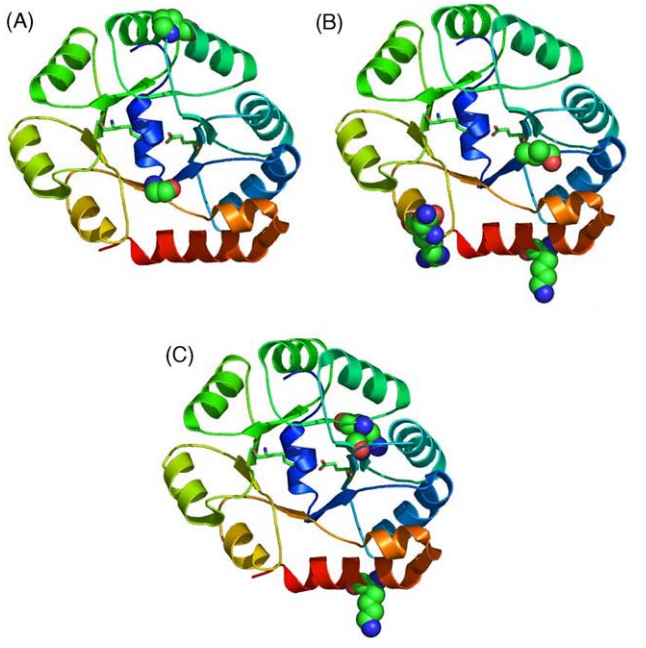


Figure 2. Structures of wild-type *E. coli* KDPG aldolase showing positions of mutated residues. Active site residues Lys 133 and Glu 45 are shown as sticks. Mutated residues are shown as spheres. (A) EC-1 (Pro 77, Gly 183), (B) EC-2 (Val 22, Asn 168, Arg 170, Lys 202), (C) EC-3 (Gly 70, Gln 89, Lys 202).

3. Experimental

3.1. Materials

^1H and ^{13}C NMR spectra were recorded on a Varian 300 MHz NMR spectrometer. The PB25 cell line was a generous gift from Dr. Francisco Bolivar at Universidad Nacional Autónoma de México.[36]

3.2. pUC-ECEDA and pUC-ECEDA2 construction

A pUC19 plasmid was modified by site-directed polymerase chain reaction (PCR) mutagenesis (Stratagene) to insert an *Nco* I site upstream of the poly-cloning site, using the following PCR primers: 5′-GCTTGGCGTAACCATGGTCATA-3′ and 5′-TATGACCATGGTTACGCCAAGC-3′ (Invitrogen). The resulting plasmid was digested with *Nco* I and *Bam* H I and the linearized plasmid was purified on a 1% agarose gel.[38] The *E. coli* KDPG aldolase gene was isolated by digestion of pET-ECEDA with *Nco* I and *Bam* H I and purified using a 1% agarose gel.[40] The DNA fragments were extracted from the gel using the Qiagen Gel Purification Kit and were then ligated together using T4 DNA ligase (NEB). The ligation product was transformed into calcium competent XL-10 Gold cells (Stratagene) and cells containing plasmids were selected by growth on LB/amp plates. The sequence of the complete *eda* gene was confirmed by automated dideoxy sequencing (Duke DNA Analysis Facility).

The pUC-ECEDA2 plasmid was constructed to incorporate *Sac* I and *Xho* I restriction sites flanking the start and stop of the gene. PCR amplification of the EDA gene from the pET-ECEDA plasmid used the following primers: 5′-TCGAGCTCTACCATGAAAAACTGGAAAACTTCCG-3′ and 5′-ATGCAGTGACTCGAGTTCCAGCTTAGCGCCTTCTACAGCTTC-3′. The PCR product was digested with *Sac* I and *Xho* I and purified on a 1% agarose gel. The pUC-TMEDA vector was digested with *Sac* I and *Xho* I and the pUC fragment purified on a 1% agarose gel. The DNA fragments were ligated together using T4 DNA ligase (NEB) and transformed into competent Smart Cells (Gene Therapy Systems). After isolation of the plasmid DNA, the sequence of the *eda* gene was verified by automated dideoxy nucleotide sequencing (University of Michigan Sequencing Core).

3.3. pUC-TMEDA construct

The *T. maritima eda* gene was amplified from pTM-eda2[30] using the following primers that contain *Sac* I and *Sph* I sites, respectively; 5′-CTTTAAGAAGGAGCTCTACCATGGC-3′ and 5′-GGCTTTGTTAGCATGCCGGATCTC-3′ (Invitrogen). The PCR product was digested with *Sac* I and *Sph* I (NEB), purified using a 1% agarose gel,[38] and extracted from the gel using Millipore Ultrafree DA filters. This fragment was ligated to *Sac* I/*Sph* I-digested pUC18 plasmid using T4 DNA ligase (NEB). The resulting plasmid was

transformed into calcium competent XL1-Blue cells (Stratagene) and selected by growth on LB/amp plates. Automated dideoxy nucleotide sequencing (University of Michigan Sequencing Core) was used to verify the sequence of the cloned *eda* gene.

3.4. Construction of error-prone PCR library

Error-prone PCR amplification of the *eda* gene was carried out using an error-prone DNA polymerase (Mutazyme, Stratagene) and primers containing the restriction sites *Sac* I and *Xho* I: (Forward) 5′-GGAAACAGCTATGACCATGATTACGAATTC-GAGCTCTACCATG-3′ and (Reverse) 5′-CTCAGT-GGTGGTGGTGGTGGTGCTCGAGTTC -3′. Plasmid pECA-pUC (10 ng), PCR primers (125 ng each), dNTPs (200 μM each), Mutazyme buffer (5 μL 10×, Strategene), and Mutazyme DNA polymerase (2.5 units, Stratagene) were combined in a total volume of 50 μL. Each cycle of PCR amplification (40 total) consisted of 94 °C for 30 s, 58 °C for 30 s, and 72 °C for 50 s, followed by a final extension of 72 °C for 2 min. The PCR-amplified DNA was purified by extracting twice with an equal volume of phenol:chloroform (E.M. Science, pH 8.0) and once with an equal volume of chloroform. The salt content of the aqueous phase was adjusted to 0.2 M in NaCl and two volumes of cold absolute ethanol were added. The DNA was incubated for 15 min at −80 °C and pelleted by centrifugation (14,000 rpm, 10 min, 4 °C). The pelleted DNA was washed once with 70% ethanol, centrifuged again, and air-dried before resuspension in 200 μL water. The PCR fragments were digested at 37 °C with *Sac* I and *Xho* I restriction enzymes (NEB, 40 U each, 15 h) and purified on a 1% agarose gel followed by electroelution into 10,000 MWCO dialysis tubing (Pierce, Snakeskin). The DNA was concentrated to ~1 mL using a Microcon spin-filter (Amicon, 30,000 MWCO), purified by phenol:chloroform extraction and precipitated by the addition of NaCl and ethanol. The PCR fragments (60 μg total) were redissolved in 600 μL water. This DNA (12 μg) was mixed with *Sac* I/*Xho* I-digested pUC18 (60 μg) in NEB T4 Buffer containing ATP (2.5 mM) and NEB T4 DNA ligase (16,000 cohesive-end units) in a total volume of 400 μL. The reaction was incubated on ice and allowed to warm until the ice melted (16 h). The DNA was extracted with phenol:chloroform and precipitated with NaCl/ethanol as described above to concentrate the DNA and remove salts. The ligated DNA was resuspended in 40 μL of water. A concentration of 0.4 μg μL^{-1} ligated DNA was estimated from analysis of the linearized plasmid DNA (*Sca* I digest) on a 1% agarose gel by comparison to a 10 ng μL^{-1} 1.1 kb DNA standard (Stratagene).

3.5. Electroporation of PB25 cells

Electrocompetent PB25 cells were prepared using the method of Hanahan et al.[41] Frozen glycerol stocks (80 μL) of electrocompetent cells were thawed on ice. Cold DNA (10 pg to 125 μg) was added and cells were electroporated in 0.1 cm electroporation cuvettes (Bio-Rad) at 1.8 kV. SOC media (1 mL) was immediately added and the cells transferred to a Falcon tube. The cuvette was washed with SOC (2 × 1 mL) and the washings combined. The cells were incubated at 37 °C, 200 rpm for 1 h. The cells were pelleted by centrifugation (3000*g*, 10 min) and resuspended in M9 media (1 mL). Cells were pelleted again and resuspended in M9 media (1 mL) before plating.

3.6. Preparation of M9 plates

Agar LB and M9 media plates were made using Sambrook's published procedure.[38] In some cases, the media was supplemented with filter-sterilized (0.22 μm) solutions of one or more of the following (final concentrations are listed): proline (pro, 20 μg mL^{-1}), thiamin (B_1, 1 μg mL^{-1}), glucose or ribose (0.4% each) ampicillin (amp; 25 or 50 μg mL^{-1}), kanamycin (kan; 25 or 50 μg mL^{-1}), chloramphenicol (cam; 17 or 30 μg mL^{-1}), pyruvate (0.005–5 mM), KHPB (5 mM), or KHO (5 mM).

3.7. Supplementation with vitamins, nucleotides, metals, and amino acids

A 100× vitamin solution was made with 1 mg mL^{-1} of each of the following vitamins: pyridoxal hydrochloride, D-biotin, cyanocobalamin, niacinamide, D-pantothenate, folic acid, and riboflavin. The solution was neutralized by the addition of NaOH and then sterilized by filtration. A 1000× metal solution containing 1 mM $FeCl_3$ and 50 mM $ZnSO_4$ was made and sterilized by filtration. The 100× nucleotide solution contained 4 mg mL^{-1} each of adenosine, guanosine, thymidine, cytosine, and uracil (pH 7). The solution was heated to 65 °C to dissolve all solids, and was then sterilized by filtration.

M9/pro/B_1 agar plates supplemented with kanamycin (25 μg mL^{-1}) and chloramphenicol (17 μg mL^{-1}) were prepared as described. The plates containing 0.2% ribose or 0.2% glucose were also supplemented with either {vitamins + metals}, {nucleotides}, or {vitamins + metals + nucleotides}. All plates were inoculated with 25 μL of a single colony of PB25 cells resuspended in 200 μL. After 48 h of incubation at 34 °C, the number of colonies was determined and compared to the number of colonies observed on a LB/kan/cam plate.

3.8. Serial dilutions of pyruvate

M9/pro/B_1 plates containing kanamycin (50 μg mL^{-1}), ampicillin (50 μg mL^{-1}), and chloramphenicol (30 μg mL^{-1}) were prepared as described and supplemented with 5, 0.5, 0.05, or 0.005 mM sodium pyruvate. Electroporation of PB25 cells with pUC-ECEDA or pUC-TMEDA plasmid (50 ng) was carried out as described above. Cells (50 μL) were plated and incubated at 34 °C.

3.9. Model selection with 2-keto-4-hydroxy-4-(2′-pyridyl)butyrate (KHPB)

M9/pro/B_1 plates were prepared as described and supplemented with kanamycin (25 μg mL^{-1}), ampicillin (25 μg mL^{-1}), chloramphenicol (17 μg mL^{-1}), and KHPB (5 mM). Electroporation of PB25 cells with pUC-ECEDA or pUC-TMEDA (5 ng) was carried out as described above. Cells were diluted 1:10 into M9/ribose media before plating (50 μL) on the M9/kan/amp/cam/KHPB agar plates.

3.10. Selection with 2-keto-4-hydroxyoctonate (KHO)

Minimal A media plates[38] were prepared and supplemented with vitamins, metals, nucleotides, proline, and thiamin as described above, as well as carbenicillin (50 μg mL^{-1}), ribose or glucose (0.4%), and KHO (2.5 mM). Selections were carried out at 34 °C.

3.11. Synthesis of 2-keto-4-hydroxy-4-(2′-pyridyl)-butyrate (KHPB)

Sodium pyruvate (3.8 g, 34 mmol) was dissolved in KH_2PO_4 buffer (20 mM, pH 6.5, 30 mL) and sterile filtered (0.22 μm). In a separate flask, freshly distilled 2-pyridine carboxaldehyde (3.8 mL, 34 mmol) was dissolved in KH_2PO_4 buffer (20 mM, pH 6.5, 400 mL). *E. coli* KDPG aldolase (6 mg, 2400 U) was added to the 2-pyridine carboxaldehyde and the resulting solution passed through a sterile filter (0.22 μm). Pyruvate (1.13 M, 10 mL) was added and the reaction was incubated for 17 h at 21 °C. Additional pyruvate was added (1.13 M, 10 mL), and the reaction incubated another 8 h. Additional pyruvate was added (1.13 M, 10 mL), and the reaction incubated for another 17 h. The reaction was lyophilized, and the residual solid stirred for 30 min with absolute ethanol (620 mL) and filtered. The ethanol was removed in vacuo and the solid was dried under vacuum overnight. One mole equivalent of LiCl (1.5 g, 34 mmol) was dissolved in deionized water (60 mL) and added to the crude product, stirred until no solid remained, and then lyophilized. The resulting solid was dissolved in absolute ethanol (50 mL), stirred at 21 °C, and then chilled to −20 °C until a white solid formed. The solid was filtered, washed with cold ethanol (75 mL), and then dried. Purification using a 130 g C18 derivatized silica gel column (Redisep, Isco) in 1 g batches (100% H_2O, 70 mL min^{-1} flow rate, 10 mL fractions) yielded pure product (1.52 g, 7.5 mmol) that is spectroscopically identical to previously reported material.[31]

3.12. Synthesis of 2-keto-4-hydroxy-octanoate (KHO)[42]

NaOH (3 M, 9.6 mL) was cooled to 0 °C. Oxalacetic acid (1.9 g, 14.5 mM) was added slowly with stirring. The pH was adjusted to 10.0 with 3 M NaOH and valeraldehyde (2.4 g, 28 mM) dissolved in THF (3 mL) was added. The pH was followed and readjusted to 10.0 for approximately 1 h until stabilization, and then the reaction was allowed to stir at room temperature for 23 more hours. The reaction was washed with ether (3 × 10 mL) to remove excess valeraldehyde, and the aqueous layer treated with Amberlyst 131 H^+ resin in vacuo until no more bubbling was observed. The solution was filtered, and the resin washed with dH_2O (200 mL). The pH was adjusted to 7.0 and the solution lyophilized overnight. Purification using a 130 g C18 derivatized silica gel column in 1 g batches on Isco Combiflash Retrieve automated system (100% H_2O, 70 mL min^{-1} flow rate, 15 mL fractions) yielded pure product (851 mg, 4.3 mmol). ^{1}H NMR (D_2O) δ 4.21–4.11 (m, 1H), 3.02–2.82 (m, 2H), 1.58–1.48 (m, 2H), 1.42–1.26 (m, 4H), 0.94–0.86 (m, 3H) ppm. ^{13}C NMR (D_2O) δ 207.83, 172.82, 70.08, 49.43, 38.76, 29.72, 24.69, 16.11 ppm.

Acknowledgements

E.J.T. and C.A.F. acknowledge the support of the NIH (GM 61596). J.S.G. acknowledges the support of the Biological Chemistry training program (T32 GM008558). M.C. acknowledges the support of the Chemical Biology Interface training program (GM08597).

References and notes

1. Arnold, F. H.; Wintrode, P. L.; Miyazaki, K.; Gershenson, A. *Trends Biochem. Sci.* **2001**, *26*, 100–106.
2. Wintrode, P. L.; Miyazaki, K.; Arnold, F. H. *Biochim. Biophys. Acta—Protein Struct. Molec. Enzym.* **2001**, *1549*, 1–8.
3. Farinas, E. T.; Bulter, T.; Arnold, F. H. *Curr. Opin. Biotechnol.* **2001**, *12*, 545–551.
4. Petrounia, I. P.; Arnold, F. H. *Curr. Opin. Biotechnol.* **2000**, *11*, 325–330.
5. Kuchner, O.; Arnold, F. H. *Biofutur* **1999**, B1–B11.
6. Arnold, F. H. *Acc. Chem. Res.* **1998**, *31*, 125–131.
7. Arnold, F. H.; Volkov, A. A. *Curr. Opin. Chem. Biol.* **1999**, *3*, 54–59.
8. Taylor, S. V.; Kast, P.; Hilvert, D. *Angew. Chem., Int. Ed.* **2001**, *40*, 3311–3335.
9. Meyer, A.; Schmid, A.; Held, M.; Westphal, A. H.; Rothlisberger, M.; Kohler, H. P. E.; van Berkel, W. J. H.; Witholt, B. *J. Biol. Chem.* **2002**, *277*, 5575–5582.
10. Broo, K.; Larsson, A. K.; Jemth, P.; Mannervik, B. *J. Mol. Biol.* **2002**, *318*, 59–70.
11. Xia, G.; Chen, L. J.; Sera, T.; Fa, M.; Schultz, P. G.; Romesberg, F. E. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6597–6602.
12. Joo, H.; Lin, Z. L.; Arnold, F. H. *Nature* **1999**, *399*, 670–673.
13. Dion, M.; Nisole, A.; Spangenberg, P.; Andre, C.; Glottin-Fleury, A.; Mattes, R.; Tellier, C.; Rabiller, C. *Glycoconjugate J.* **2001**, *18*, 215–223.
14. Reetz, M. T.; Wilensek, S.; Zha, D. X.; Jaeger, K. E. *Angew. Chem., Int. Ed.* **2001**, *40*, 3589–3591.
15. Flores, H.; Ellington, A. D. *J. Mol. Biol.* **2002**, *315*, 325–337.
16. Moore, J. C.; Arnold, F. H. *Nat. Biotechnol.* **1996**, *14*, 458–467.
17. You, L.; Arnold, F. H. *Protein Eng.* **1996**, *9*, 77–83.
18. Shao, Z. X.; Arnold, F. H. *Curr. Opin. Struct. Biol.* **1996**, *6*, 513–518.

19. Cline, J.; Hogrefe, H. *Strategies Newsletter 2000* **2000**, *13*, 157–161.
20. Leung, D. W.; Chen, E.; Goeddel, D. V. *Technique* **1989**, *1*, 11–15.
21. Kuipers, O. P.; Boot, H. J.; Devos, W. M. *Nucl. Acids Res.* **1991**, *19*, 4558–4558.
22. Spee, J. H.; de Vos, W. M.; Kuipers, O. P. *Nucl. Acids Res.* **1992**, *21*, 777–778.
23. Stemmer, W. P. C. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10747–10751.
24. Stemmer, W. P. C. *Nature* **1994**, *370*, 389–391.
25. Zhao, H. M.; Giver, L.; Shao, Z. X.; Affholter, J. A.; Arnold, F. H. *Nat. Biotechnol.* **1998**, *16*, 258–261.
26. Ostermeier, M.; Shim, J. H.; Benkovic, S. J. *Nat. Biotechnol.* **1999**, *17*, 1205–1209.
27. Ostermeier, M.; Benkovic, S. J. *Biotechnol. Lett.* **2001**, *23*, 303–310.
28. Olsen, M.; Iverson, B.; Georgiou, G. *Curr. Opin. Biotechnol.* **2000**, *11*, 331–337.
29. Wahler, D.; Reymond, J. L. *Curr. Opin. Biotechnol.* **2001**, *12*, 535–544.
30. Griffiths, J. S.; Wymer, N. J.; Njolito, E.; Niranjanakumari, S.; Fierke, C. A.; Toone, E. J. *Bioorg. Med. Chem.* **2002**, *10*, 545–550.
31. Henderson, D. P.; Shelton, M. C.; Cotterill, I. C.; Toone, E. J. *J. Org. Chem.* **1997**, *62*, 7910–7911.
32. Shelton, M. C.; Cotterill, I. C.; Novak, S. T. A.; Poonawala, R. M.; Sudarshan, S.; Toone, E. J. *J. Am. Chem. Soc.* **1996**, *118*, 2117–2125.
33. Conway, T. *FEMS Microbiol. Rev.* **1992**, *103*, 1–28.
34. Fong, S.; Machajewski, T. D.; Mak, C. C.; Wong, C. H. *Chem. Biol.* **2000**, *7*, 873–883.
35. Wymer, N.; Buchanan, L. V.; Henderson, D.; Mehta, N.; Pocivavsek, L.; Toone, E. J.; Naismith, J. H. *Structure* **2001**, *9*, 1–9.
36. Ponce, E.; Flores, N.; Martinez, A.; Valle, F.; Bolivar, F. *J. Bacteriol.* **1995**, *177*, 5719–5722.
37. Ponce, E.; Martinez, A.; Bolivar, F.; Valle, F. *Biotechnol. Bioeng.* **1998**, *58*, 292–295.
38. Sambrook, J.; Russell, D. W. In *Molecular Cloning: a Laboratory Manual*, 3rd ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 2001; Vol. 3.
39. Meloche, H. P.; Wood, W. A. *J. Biol. Chem.* **1964**, *239*, 3511–3514.
40. Buchanan, L. V.; Mehta, N.; Pocivavsek, L.; Niranjanakumari, S.; Toone, E. J.; Naismith, J. H. *Acta Crystallogr. Sect. D—Biol. Crystallogr.* **1999**, *55*, 1946–1948.
41. Hanahan, D.; Jessee, J.; Bloom, F. R. *Method Enzymol.* **1991**, *204*, 63–113.
42. Cornforth, J. W.; Firth, M. E.; Gottshack, A. *Biochem. J.* **1958**, *68*, 57–60.