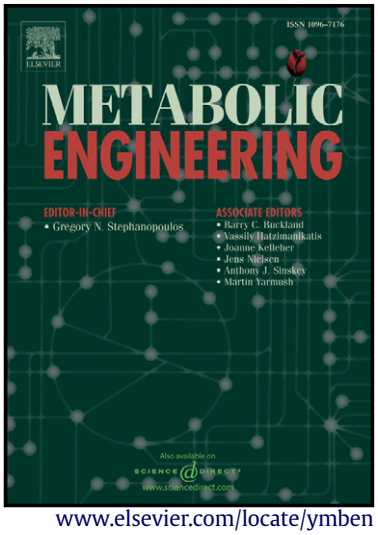
Author's Accepted Manuscript



www.elsevier.com/locate/ymben

Uncovering rare NADH-preferring ketol-acid reductoisomerases

S. Brinkmann-Chen, J.K.B. Cahn, F.H. Arnold

PII: S1096-7176(14)00106-2
DOI: http://dx.doi.org/10.1016/j.ymben.2014.08.003
Reference: YMBEN917

To appear in: *Metabolic Engineering*

Cite this article as: S. Brinkmann-Chen, J.K.B. Cahn, F.H. Arnold, Uncovering rare NADH-preferring ketol-acid reductoisomerases, *Metabolic Engineering,* http://dx.doi.org/10.1016/j.ymben.2014.08.003

Uncovering rare NADH-preferring ketol-acid reductoisomerases

Brinkmann-Chen, S.*; Cahn, J.K.B*; Arnold, F.H.†
*These authors contributed equally

California Institute of Technology, Department of Chemistry and Chemical Engineering
†Corresponding author: F. H. Arnold, California Institute of Technology, 1200 E California Blvd, MC 210-41, Pasadena, CA 91125, USA
Phone: +1 626 395 4162
Email: frances@cheme.caltech.edu
sabine@cheme.caltech.edu
jcahn@caltech.edu

Abstract

All members of the ketol-acid reductoisomerase (KARI) enzyme family characterized to date have been shown to prefer the nicotinamide adenine dinucleotide phosphate hydride (NADPH) cofactor to nicotinamide adenine dinucleotide hydride (NADH). However, KARIs with the reversed cofactor preference are desirable for industrial applications, including anaerobic fermentation to produce branched-chain amino acids. By applying insights gained from structural and engineering studies of this enzyme family to a comprehensive multiple sequence alignment of KARIs, we identified putative NADH-utilizing KARIs and characterized eight whose catalytic efficiencies using NADH were equal to or greater than NADPH. These are the first naturally NADH-preferring KARIs reported and demonstrate that this property has evolved independently multiple times, using strategies unlike those used previously in the laboratory to engineer a KARI cofactor switch.

Keywords: Ketol acid reductoisomerase; cofactor specificity; NADH; NADPH

1 Introduction

With burgeoning genomic databases and increasing ease of gene synthesis, metabolic engineers can now readily mine nature's rich collection of enzymes. However, finding a sequence with specific desired properties can be difficult, particularly when only a few members of a protein family have been characterized and a detailed understanding of the structure-function relationship is lacking. The ketol-acid reductoisomerase (KARI, EC 1.1.1.86, also known as acetohydroxyacid isomeroreductase (AHAIR)) enzymes have attracted much interest for production of amino acids and biofuels (Atsumi et al., 2008a; Bastian et al., 2011; Brinkmann-Chen et al., 2013; Hasegawa et al., 2012; Liu et al., 2010). These oxidoreductases catalyze the second step in the branched chain amino-acid (BCAA) biosynthesis pathway (Chunduru et al., 1989), conversion of (*S*)-2-acetolactate (*S*2AL) to (*R*)-2,3-dihydroxyisovalerate (*R*DHIV) via a methyl shift coupled to a reduction with concomitant oxidation of a nicotinamide adenine dinucleotide cofactor. The BCAA pathway is present in many organisms but not in mammals. Because of this, microbial production of branched-chain amino acids for animal feed or human supplements is a multimillion-dollar business (Becker and Wittmann, 2012; Vogt et al., 2014). The BCAA pathway has also been engineered to produce isobutanol, a potential source of renewable chemicals and fuels (Atsumi et al., 2008b).

All wild-type KARIs characterized and described in the literature have displayed a strong preference for nicotinamide adenine dinucleotide phosphate (NADPH) over nicotinamide adenine dinucleotide (NADH) (Bastian et al., 2011). Because intracellular levels of NAD(H) are much higher than NADP(H), particularly under fermentative conditions,

NADH-dependent oxidoreductases are strongly preferred in pathways for large-scale biocatalytic processes (Bastian et al., 2011). In the engineered isobutanol production pathway, replacement of the natural *E. coli* KARI (Ec_IlvC) and the alcohol dehydrogenase (ADH) with NADH-preferring engineered proteins increased the yield to nearly 100% of theoretical and improved titer and specific productivity (Bastian et al., 2011).

In a previous study that aimed to develop a general recipe for engineering KARIs with reversed cofactor specificity, we used available KARI structure data to identify the amino acid residues in the β2αB-loop of the Rossmann fold that distinguish between the two cofactors (Brinkmann-Chen et al., 2013). A sequence alignment of Swiss-Prot-annotated KARI sequences allowed us to divide this diverse enzyme family into three groups based on β2αB-loop length (6-, 7-, and 12-residue loops) and to develop a simple recipe for switching the cofactor specificity of each major KARI enzyme subfamily from NADPH to NADH. This engineering work provided valuable information on the determinants of cofactor binding and also led us to question whether nature might have already undertaken a similar engineering task to create an NADH-preferring KARI. Despite recent advances in bioinformatic cofactor specificity prediction (Geertz-Hansen et al., 2014), few attempts have been made to find alternate cofactor utilization profiles within large enzyme families (Di Luccio et al., 2006). Because no method existed to predict the cofactor specificity of uncharacterized KARIs based on their primary sequences, we used knowledge gained from our previous work to exhaustively search known KARI sequences for KARIs with β2αB-loops predicted to improve utilization of NADH. Here,

we report the discovery of the first bi-specific and the first naturally NADH-preferring KARIs. The catalytic efficiencies of two of the naturally NADH-preferring KARIs reported here exceed those of any previously engineered variants. We suggest that rare proteins such as these, with properties desirable for metabolic engineering, are available in nature and may be uncovered using knowledge gleaned from structural and mutational studies.

2 Materials and Methods

2.1 General

Biological media were purchased from Research Products International (Mt. Prospect, IL, USA), NAD(P)H from Codexis, Inc. (Redwood City, CA, USA), oligonucleotides and gBlocks from Integrated DNA Technologies (San Diego, CA, USA), DNA polymerases, restriction enzymes, and T4 ligase from New England Biolabs (Ipswich, MA, USA). (*S*)-2-acetolactate (*S*2AL) was provided by Gevo, Inc. (Denver, CO, USA). DNA sequencing was performed by Laragen (Los Angeles, CA, USA). Standard molecular biology methods were taken from Maniatis *et al*. (Sambrook, 1989).

2.2 Cloning, variant construction, expression, and kinetic assays

The genes encoding KARIs were obtained as gBlocks, codon-optimized for *E. coli*. For each gene, the gBlocks were assembled either via PCR using T7 promoter and terminator primers and Phusion polymerase following the manufacturer's instructions or via Gibson cloning (Gibson, 2011). All KARIs were cloned into pET22(b)+ between restriction sites *Nde*I and *Xho*I in frame with the C-terminal his-tag for expression in *E. coli* BL21 E. cloni Express cells from Lucigen (Middleton, WI, USA). Heterologous protein

expression and purification were conducted as described (Bastian et al., 2011). Protein concentration was determined via the Bradford assay (Bio-Rad, Hercules, CA, USA). KARI activities were assayed by monitoring NAD(P)H consumption in the presence of (*S*)-2-acetolactate (*S*2AL) at 340 nm in a plate reader (Tecan Infinite M200, San Jose, CA, USA). The assay buffer contained 100 mM potassium phosphate pH 7, 1 mM DTT, 200 μM NAD(P)H, 2.5 mM S2AL, and 10 mM $MgCl_2$. For kinetic assays, we monitored the consumption of the cofactors via their fluorescence emission at 440 nm with excitation at 340 nm.

2.3 Sequence Alignment

All sequences annotated as EC 1.1.1.86 and non-fragmentary were downloaded from UniProtKB (Apweiler et al., 2013) on 11/1/2013, a total of 8,043 sequences. The sequences were aligned using MAFFT version 7 (Katoh and Standley, 2013). Duplicate sequences were removed, as were those with less than 20% identity to the *Slackia exigua* KARI. This left 3,383 unique KARI sequences, with an alignment length of 933 positions. The β2αB-loop was identified based on the sequences with solved structures deposited in the RCSB PDB (Bairoch and Apweiler, 1999) (*E. coli, S. exigua, S. oleracea, O. sativa, P. aeruginosa*, and *A. acidocaldarius* (Brinkmann-Chen et al., 2013; Ahn et al., 2003; Biou et al., 1997; Leung and Guddat, 2009; Thomazeau et al., 2000; Tyagi et al., 2005; Wong et al., 2012; Cahn et al., 2014)). Further analysis was completed using a custom-written Python script (supplemental information).

3 Results and Discussion

To develop a recipe for changing the cofactor specificity of any KARI from NADPH to NADH we previously used an alignment of 643 Swiss-Prot-annotated KARI sequences (Brinkmann-Chen et al., 2013). Here, we expanded this alignment to include un-reviewed UniProt Knowledgebase (UniprotKB) (Apweiler et al., 2013) sequences and searched this expanded sequence space for β2αB-loops deviating from the classic NADPH-preferring KARI motifs. To do so, all 8,043 sequences annotated as EC 1.1.1.86 from UniProtKB were aligned using MAFFT; duplicate and low-identity sequences were removed as described in the Methods section. The β2αB-loop and its context were then identified based on KARI crystal structures (Brinkmann-Chen et al., 2013; Ahn et al., 2003; Biou et al., 1997; Leung and Guddat, 2009; Thomazeau et al., 2000; Tyagi et al., 2005; Wong et al., 2012; Cahn et al., 2014). The distribution of amino acids at each position, including positions present in only a very small number of sequences, is shown in Figure 1. The four phosphate-binding residues are highly conserved, but there is considerable variation at most of the other positions of the loop as well as in the length of the loop. Figure 2 shows these positions in the NADPH co-crystal structure of *Slackia exigua* KARI, a 7-residue loop KARI whose sequence is near the consensus sequence.

As previously reported (Brinkmann-Chen et al., 2013), KARIs can be divided into three sub-families based on the length of the β2αB-loop. Of the 3,383 unique KARIs in our expanded alignment, 2,262 (66.8%) had 7-residue loops, 593 (17.5%) had 6-residue loops, and 512 (15.1%) had 12-residue loops, which is in good agreement with previous

alignment results. Eleven sequences were missing the β2αB-loops (these were unlabeled fragmentary sequences), leaving just six KARIs with other loop lengths.

Based on previous structural studies and mutational analysis (Brinkmann-Chen et al., 2013; Hasegawa et al., 2012; Rane and Calvo, 1997), we have proposed that cofactor specificity is controlled by up to four amino acid residues in the loop: the first two positions, the antepenultimate position, and the ultimate position, as indicated by red arrows in Figure 1 and labeled for the *S. exigua* KARI in Figure 2. We searched the expanded sequence alignment for proteins having an acidic residue at any of these four positions. In our previous work, introduction of acidic residues into the β2αB-loop was crucial for switching the cofactor specificity (Bastian et al., 2011; Brinkmann-Chen et al., 2013); this has also been shown to be a major determinant of cofactor specificity in nature (Baker et al., 1992; Carugo and Argos, 1997; Pletnev et al., 2004; Van Petegem et al., 2007) and in engineered proteins (Khoury et al., 2009). For the six KARI sequences with non-standard β2αB-loop length we could not predict which residues would interact directly with the cofactor. We therefore visually inspected those loops for the presence of multiple acidic residues in their N-terminal or C-terminal ends. Ultimately, from the alignment of 3,383 KARI sequences, 58 were identified by this method as being possibly NADH-preferring, including four of the six KARIs with nonstandard β2αB-loop lengths. To avoid over-representing KARIs from closely related hosts, the full sequences of these KARIs were clustered using the UniProt alignment tool, and for each of the 26 clusters one representative sequence was selected to create the list of putative NADH-dependent KARIs shown in Figure 3. Most of the clusters were singletons or contained only a

handful of sequences; the two largest clusters are represented by *Metallosphaera sedula* and *Sulfolobus islandicus*.

From this list, we successfully expressed, purified, and assayed the activity toward both cofactors of eight of the putative NADH-utilizing KARIs, highlighted in Figure 3. The kinetic values in Table 1 show that none of these KARIs has the typical preference for NADPH. Four of the eight – *Hydrogenobaculum* sp. KARI (Hs_KARI), *Ignisphaera aggregans* KARI (Ia_KARI), *Metallosphaera sedula* KARI (Ms_KARI), and *Syntrophomonas wolfei* KARI (Sw_KARI) – showed roughly equal catalytic efficiency using NADPH and NADH (catalytic efficiency ratio of NADH/NADPH ~ 1), while four – *Archaeoglobus fulgidus* KARI (Af_KARI), *Desulfococcus oleovorans* KARI (Do_KARI), *Thermacetogenium phaeum* KARI (Tp_KARI), and an uncultured archaeon KARI (Ua_KARI) – were NADH-preferring, with catalytic efficiency ratios (NADH/NADPH) ranging from 13 to 152 (Table 1). The NADH K_M or K_H values of these KARIs were equal to or lower than the NADPH values of the NADPH-preferring KARI from *E. coli*, Ec_IlvC. Ia_KARI in particular stood out, with very low K_M values for both cofactors. Two of the naturally NADH-preferring KARIs, Tp_KARI and Ua_KARI, have catalytic efficiencies on NADH that exceeded the best engineered NADH-preferring KARI, Ec_IlvC$^{P2D1-A1}$, by 2.8-fold and 1.2-fold, respectively. Both Tp_KARI and Ua_KARI have very low (~1 μM) K_H for NADH (Table 1). These features render these two excellent candidates for anaerobic BCAA or isobutanol fermentations. Ia_KARI may also be of value due to its ability to utilize both cofactors and its good activity at low cofactor concentrations.

In our previously engineered NADH-preferring KARIs (Bastian et al., 2011; Brinkmann-Chen et al., 2013), most, if not all, NADPH-dependent activity was abolished by the mutations. In contrast, nature achieved NADH cofactor utilization with NADPH K_M or K_H values equivalent to NADPH-preferring KARIs, leading to lower specificities than many of our engineered KARIs (Brinkmann-Chen et al., 2013). We previously demonstrated that NADPH-preferring KARIs can have their specificity reversed by mutation of the ultimate and antepenultimate residues of the β2αB-loop to aspartate in the case of the 7- or 12-residue loops, or, in the case of the 6-residue loop, mutation of the ultimate position to aspartate and of the second to proline. Interestingly, none of the naturally NADH-utilizing KARIs we found follow this recipe. Aspartate is present in only four out of 26 ultimate positions and one out of 21 antepenultimate positions, although Sw_KARI has a glutamate at both positions. Of all 26 sequences, only Sw_KARI has acidic residues at more than one of the four key positions. Only two of the sequences have proline at the second position, and neither KARI has a 6-residue β2αB-loop or an aspartate at the ultimate position. These differences may reflect different selective pressures (e.g. specificity reversal in the engineered enzymes versus utilization of NADH in the natural enzymes). They could also arise from the very different evolutionary paths taken.

Eight of the 26 putative NADH-utilizing KARIs possess acidic residues at the first position of the loop, six at the second, three at the antepenultimate position, and five at the ultimate position. It is interesting that the plurality of these putative NADH-utilizing

KARIs have acidic residues the first position of the β2αB-loop. Although none of our previous cofactor-specificity reversals required an acidic residue at this position, Hasegawa *et al*. introduced a mutation to glutamate at this position, suggested by comparison with the NADH-dependent dihydrolipoamide dehydrogenases, in an attempt to reverse the cofactor specificity of the *C. glutamicum* KARI (Hasegawa et al., 2012). The resulting mutant enzyme was specific for NADH, but had less activity with NADH than the wild-type enzyme. It is also worth noting that all five KARIs with 6-membered β2αB-loops have an aspartate at the antepenultimate position, but that the structure of the KARI from *Alicyclobacillus acidocaldarius* (Cahn et al., 2014), an NADPH-preferring enzyme with a 6-membered β2αB-loop, shows this position to be structurally homologous with the non-interacting pre-antepenultimate position of the 7-membered loop KARIs. Thus it is unclear whether and how this aspartate would interact with the cofactor. Outside the β2αB-loop, the 58 putative NADH-utilizing KARIs have no unique conserved residues.

All of the KARIs described here are members of the Class I (short chain) KARI subfamily, which is believed to be evolutionarily older than the Class II (long chain) KARIs such as *E. coli* KARI (Ahn, et al., 2003), and all originate from microbial hosts adapted to extreme conditions (thermophiles, acidophiles, and halophiles) (Auernik et al., 2008; Beeder et al., 1994; Hattori et al., 2000; McInerney et al., 1981; Niederberger et al., 2006; Romano et al., 2013). Additionally, with the exception of the archaeote *M. sedula*, all hosts are anaerobes. However, adaptation to anoxic environments does not automatically imply the presence of an NADH-preferring KARI, as among the thousands of sequences there are many NADPH-preferring KARIs from anaerobe hosts. For each of

the eight characterized KARIs, we examined the three KARIs with the highest overall sequence identity and predicted their cofactor specificity using the sequence of their cofactor binding loops (Table 2).

For six of the eight newly characterized KARIs, the majority of these nearest neighbors, which generally came from closely related organisms, have loops sequences that suggest a conventional NADPH binding mode. For the remaining two KARIs, Ia_KARI and Ms_KARI, which are 64% identical to each other, the nearest neighbors have very similar sequences and were filtered out during the clustering step. That these enzymes have used differing sets of substitutions for cofactor utilization and that each is phylogenetically isolated from the others but close to non-reversed sequences suggests that NADH-utilization has arisen independently in KARIs of several different organisms.

In many protein families, the few characterized representatives have come from commonly studied mesophilic organisms. The properties of these enzymes are often assumed to be representative of the enzyme family as a whole. However, a rational, knowledge-guided search of all available sequences can uncover novel properties in more distantly related organisms, many of which can be valuable tools for metabolic engineers. In this study, we have shown that the presence of acidic residues at conserved phosphate-binding positions can be used to identify candidate genes encoding NADH-preferring proteins in the industrially important KARI enzyme family. With minor modifications, we expect this approach to be useful in finding enzymes with differing cofactor requirements in other oxidoreductase families.

Acknowledgments

J.K.B.C. acknowledges the support of the Resnick Sustainability Institute (Caltech). This publication is funded by the Gordon and Betty Moore Foundation through Grant GBMF2809 to the Caltech Programmable Molecular Technology Initiative.

Author contributions

SBC, JKBC, and FHA designed research. SBC and JKBC performed research and analyzed data. SBC, JKBC, and FHA wrote the paper.

References:

Ahn H.J., Eom S.J., Yoon H.J., Lee B.I., Cho H.J., Suh S.W, 2003. Crystal structure of class I acetohydroxy acid isomeroreductase from *Pseudomonas aeruginosa*. J. Mol. Biol. 328, 505 – 515.

Apweiler R., Martin M.J., O'Donovan C., Magrane M., Alam-Faruque Y., Alpi E., Antunes R., Arganiska J., Casanova E.B., Bely B. and others, 2013. Update on activities at the Universal Protein Resource (UniProt) in 2013. Nucleic Acids Res. 41, D43 – D47.

Atsumi S., Cann A.F., Connor M.R., Shen C.R., Smith K.M., Brynildsen M.P., Chou K.J.Y., Hanai T., Liao J.C., 2008a. Metabolic engineering of *Escherichia coli* for 1-butanol production. Metab. Eng. 10, 305 – 311.

Atsumi S., Hanai T., Liao J.C., 2008b. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. Nature 451, 86 – 89.

Auernik K.S., Maezato Y., Blum P.H., Kelly R.M., 2008. The genome sequence of the metal-mobilizing, extremely thermoacidophilic archaeon *Metallosphaera sedula* provides insights into bioleaching-associated metabolism. Appl. Environ. Microb. 74, 682 – 692.

Bairoch A., Apweiler R., 1999. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1999. Nucleic Acids Res. 27, 49 – 54.

Baker P.J., Britton K.L., Rice D.W., Rob A., Stillman T.J., 1992. Structural consequences of sequence patterns in the fingerprint region of the nucleotide binding fold - Implications for nucleotide specificity. J. Mol. Biol. 228, 662 – 671.

Bastian S., Liu X., Meyerowitz J.T., Snow C.D., Chen M.M.Y., Arnold F.H., 2011. Engineered ketol-acid reductoisomerase and alcohol dehydrogenase enable anaerobic 2-methylpropan-1-ol production at theoretical yield in *Escherichia coli*. Metab. Eng. 13, 345 – 352.

Becker J., Wittmann C., 2012. Systems and synthetic metabolic engineering for amino acid production - the heartbeat of industrial strain development. Curr. Opin. Biotech. 23, 718 – 726.

Beeder J., Nilsen R.K., Rosnes J.T., Torsvik T., Lien T., 1994. *Archaeoglobus fulgidus* isolated from hot north sea oil field waters. Appl. Environ. Microb. 60, 1227 – 1231.

Biou V., Dumas R., Cohen-Addad C., Douce R., Job D., Pebay-Peyroula E., 1997. The crystal structure of plant acetohydroxy acid isomeroreductase complexed with NADPH, two magnesium ions and a herbicidal transition state analog determined at 1.65 angstrom resolution. Embo J. 16, 3405 – 3415.

Brinkmann-Chen S., Flock T., Cahn J.K.B., Snow C.D., Brustad E.M., McIntosh J.A., Meinhold P., Zhang L., Arnold F.H., 2013. General approach to reversing ketol-acid reductoisomerase cofactor dependence from NADPH to NADH. P. Natl. Acad. Sci. USA 110, 10946 – 10951.

Cahn J.K.B., Brinkmann-Chen S., Arnold F.H., 2014. Ketol-acid reductoisomerase from *Alicyclobacillus acidocaldarius*. PDB ID 4TSK.

Carugo O., Argos P., 1997. NADP-dependent enzymes. II: Evolution of the mono- and dinucleotide binding domains. Proteins 28, 29 – 40.

Chunduru S.K., Mrachko G.T., Calvo K.C., 1989. Mechanism of ketol-acid reductoisomerase - steady-state analysis and metal-ion requirement. Biochemistry 28, 486 – 493.

Crooks G.E., Hon G., Chandonia J.M., Brenner S.E., 2004. WebLogo: A sequence logo generator. Genome Res. 14, 1188 – 1190

Di Luccio E., Elling R.A., Wilson D.K., 2006. Identification of a novel NADH-specific aldo-keto reductase using sequence and structural homologies. Biochem. J. 400, 105 – 114.

Geertz-Hansen H.M., Blom N., Feist A.M., Brunak S., Petersen T.N. Cofactory: Sequence-based prediction of cofactor specificity of Rossmann folds. Proteins doi: 10.1002/prot.24536

Gibson D.G., 2011. Chapter fifteen - Enzymatic assembly of overlapping DNA fragments, in: Christopher V. (editor), Methods in Enzymology. Academic Press, volume 498, pp. 349 – 361.

Hasegawa S., Uematsu K., Natsuma Y., Suda M., Hiraga K., Jojima T., Inui M., Yukawa H., 2012. Improvement of the redox balance increases L-valine production by *Corynebacterium glutamicum* under oxygen deprivation conditions. Appl. Environ. Microb. 78, 865 – 875.

Hattori S., Kamagata Y., Hanada S., Shoun H., 2000. *Thermacetogenium phaeum* gen. nov., sp nov., a strictly anaerobic, thermophilic, syntrophic acetate-oxidizing bacterium. Int. J. Syst. Evol. Microbiol. 50, 1601 – 1609.

Katoh K., Standley D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772 – 780.

Khoury G.A., Fazelinia H., Chin J.W., Pantazes R.J., Cirino P.C., Maranas C.D., 2009. Computational design of *Candida boidinii* xylose reductase for altered cofactor specificity. Protein Sci. 18, 2125 – 2138.

Leung E.W.W., Guddat L.W., 2009. Conformational changes in a plant ketol-acid reductoisomerase upon Mg^{2+} and NADPH binding as revealed by two crystal structures. J. Mol. Biol. 389, 167 – 182.

Liu X.H., Chen P.Q., Wang B.L., Dong W.L., Li Y.H., Xie X.Q., Li Z.M., 2010. High throughput receptor-based virtual screening under ZINC database, synthesis, and biological evaluation of ketol-acid reductoisomerase inhibitors. Chem. Biol. Drug Des. 75, 228 – 232.

McInerney M.J., Bryant M.P., Hespell R.B., Costerton J.W., 1981. *Syntrophomonas wolfei* gen. nov. sp. nov., an anaerobic, syntrophic, fatty acid-oxidizing bacterium. Appl. Environ. Microb. 41, 1029 – 1039.

Niederberger T.D., Gotz D.K., McDonald I.R., Ronimus R.S., Morgan H.W., 2006. *Ignisphaera aggregans* gen. nov., sp nov., a novel hyperthermorphilic crenarchaeote isolated from hot springs in Rotorua and Tokaanu, New Zealand. Int. J. Syst. Evol. Microbiol. 56, 965 – 971.

Pletnev V.Z., Weeks C.M., Duax W.L., 2004. Rational proteomics II: Electrostatic nature of cofactor preference in the short-chain oxidoreductase (SCOR) enzyme family. Proteins 57, 294 – 301.

Rane M.J., Calvo K.C., 1997. Reversal of the nucleotide specificity of ketol-acid reductoisomerase by site-directed mutagenesis identifies the NADPH binding site. Arch. Biochem. Biophys. 338, 83 – 89.

Romano C., D'Imperio S., Woyke T., Mavromatis K., Lasken R., Shock E.L., McDermott T.R., 2013. Comparative genomic analysis of phylogenetically closely related *Hydrogenobaculum* sp isolates from Yellowstone National Park. Appl. Environ. Microb. 79, 2932 – 2943.

Sambrook J., Frisch, E.F., Maniatis, T., 1989. Molecular Cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, New York.

Thomazeau K., Dumas R., Halgand F., Forest E., Douce R., Biou V., 2000. Structure of spinach acetohydroxyacid isomeroreductase complexed with its reaction product dihydroxymethylvalerate, manganese and (phospho)-ADP-ribose. Acta Crystallogr. D 56, 389 – 397.

Tyagi R., Duquerroy S., Navaza J., Guddat L.W., Duggleby R.G., 2005. The crystal structure of a bacterial Class II ketol-acid reductolsomerase: Domain conservation and evolution. Protein Sci. 14, 3089 – 3100.

Van Petegem F., De Vos D., Savvides S., Vergauwen B., Van Beelumen J., 2007. Understanding nicotinamide dinucleotide cofactor and substrate specificity in class I flavoprotein disulfide oxidoreductases: Crystallographic analysis of a glutathione amide reductase. J. Mol. Biol. 374, 883 – 889.

Vogt M., Haas S., Klaffl S., Polen T., Eggeling L., van Ooyen J., Bott M., 2014. Pushing product formation to its limit: Metabolic engineering of *Corynebacterium glutamicum* for L-leucine overproduction. Metab. Eng. 22, 40 – 52.

Wong S.H., Lonhienne T.G.A., Winzor D.J., Schenk G., Guddat L.W., 2012. Bacterial and plant ketol-acid reductoisomerases have different mechanisms of induced fit during the catalytic cycle. J. Mol. Biol. 424, 168 – 179.

Table 1. Biochemical characterization of bi-specific and naturally NADH-specific KARIs and comparison to wild-type *E. coli* KARI (Ec_IlvC) and engineered Ec_IlvC^{6E6} and Ec_IlvC$^{P2D1\text{-}A1}$ (Bastian et al., 2011).

Enzyme	K_M or K_H [μM]		k_{cat} [s^{-1}]		$k_{cat}/K_{M\ or\ H}$ [$mM^{-1}s^{-1}$]		NADH/NADPH ratio of $k_{cat}/K_{M\ or\ H}$
	NADH	NADPH	NADH	NADPH	NADH	NADPH	
Ec_IlvC	1,075	41	0.3	3.6	0.3	88	0.003
Ec_IlvC^{6E6}	30	650	2.3	0.20	74	0.40	185
Ec_IlvC$^{P2D1\text{-}A1}$	26	> 1,400	4.3	0.54	165	< 0.4	> 412
Hs_KARI	39	46	0.12	0.12	3.2	2.7	1.2
Ia_KARI	< 1	<1	0.02	0.03	> 20	>25	~ 0.8
Ms_KARI	24	31	0.06	0.07	2.5	2.1	1.2
Sw_KARI	57	44	0.28	0.22	5.0	5.0	1.0
Af_KARI	5.0	26	0.1	0.04	20	1.5	13
Do_KARI	32	n. a.	0.25	n. a.	8.0	n. a.	-
Tp_KARI	<1	40	0.46	0.25	460	6.0	74
Ua_KARI	1.1	38	0.22	0.05	200	1.3	152

All enzymes were his_6-tagged and purified prior to characterization. Enzyme activities were determined in 100 mM potassium phosphate pH 7 with 1 mM DTT, 200 μM NADPH or NADH, 2.5 mM *S*2AL, and 10 mM $MgCl_2$. Concentrations of the purified enzymes were determined using the Bradford assay. The K_M or K_H values for the cofactors, corresponding to the concentration of the half-maximum activity, were measured with appropriate dilutions of NADPH and NADH in the presence of saturating concentrations of substrate *S*2AL. Mutations located within the cofactor-binding pocket of Ec_IlvC^{6E6}: A71S, R76D, S78D, and Q110V. Additional mutations in Ec_IlvC$^{P2D1\text{-}A1}$: D146G, G185R, and K433E. For standard errors and Hill coefficients, please refer to Table S1 of the SI.
n.a. = not active

Table 2. Nearest neighbors of the eight KARIs characterized in this work. Results are limited to the top three hits, and the cofactor binding loop and predicted cofactor specificity are shown.

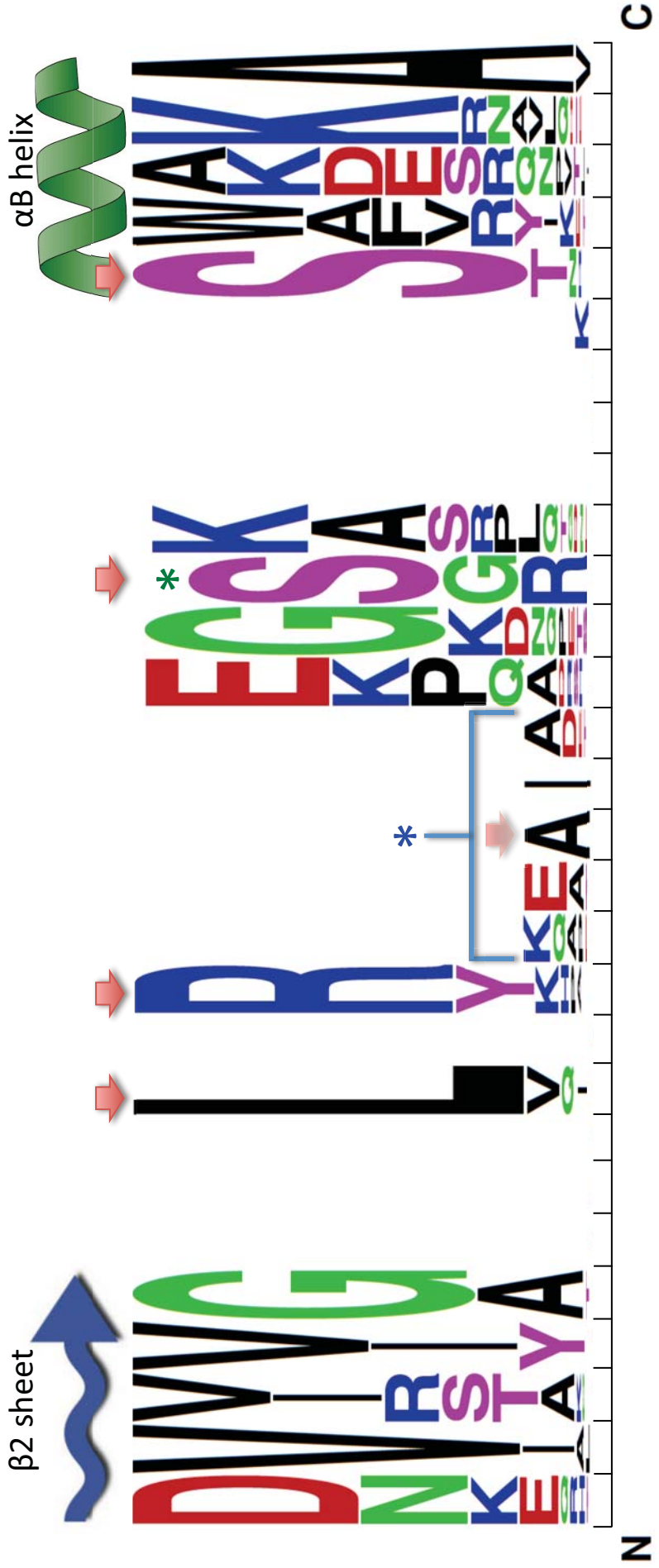
KARI (β2αb-loop sequence)	% ID	Organism	Oxygen requirements*	β2αb-loop sequence	Predicted cofactor specificity
Do_KARI (QLEGDAY)	86	Uncultured *Desulfobacterium*	Anaerobic	QMEGDAY	NADH
	80	*Desulfococcus multivorans*	Anaerobic	QREGGAS	NADPH
	78	*Desulfatibacillum alkenivorans*	Anaerobic	QRDGGKS	NADPH
Sw_KARI (LRKPFDEASEKE)	73	*Syntrophothermus lipocalidus*	Anaerobic	LRKPEDDFTTAE	Bi-specific
	72	*Desulfitobacterium dehalogenans*	Anaerobic	LRADSRR	NADPH
	71	*Syntrophobotulus glycolicus*	Anaerobic	LRKDSSR	NADPH
Af_KARI (LPEWDKAT)	77	*Archaeoglobus veneficus*	Anaerobic	LYKGSRS	NADPH
	77	*Archaeoglobus profundus*	Anaerobic	LYKGSKS	NADPH
	76	*Archaeoglobus sulfaticallidus*	Anaerobic	DKKGTRN	unclear
Tp_KARI (DIPSSEN)	72	*Caldicellulosiruptor saccharolyticus*	Anaerobic	LYHGSKS	NADPH
	72	*Caldicellulosiruptor hydrothermalis*	Anaerobic	LYHGSKS	NADPH
	71	*Caldicellulosiruptor obsidiansis*	Anaerobic	LYQGSKS	NADPH
Ua_KARI (ETEILGGNKNPS)	98	Uncultured archaeon	Anaerobic	ETEILGGNKNPS	NADH
	64	*Akkermansia muciniphila*	Anaerobic	VRPGKS	NADPH
	52	*Paenibacillus lactis*	Facul. anaerobic	LREGKS	NADPH
Ia_KARI (LERQGDS)	73	*Sulfolobus solfataricus*	Facul. anaerobic	LEREGKS	Bi-specific
	73	*Sulfolobus islandicus*	Facul. anaerobic	LEREGKS	Bi-specific
	69	*Acidianus hospitalis*	Aerobic	LEREGNS	Bi-specific
Ms_KARI (LEREGKS)	90	*Metallosphaera cuprina*	Aerobic	LEREGKS	Bi-specific
	88	*Metallosphaera yellowstonensis*	Aerobic	LEREGKS	Bi-specific
	80	*Acidianus hospitalis*	Aerobic	LEREGNS	Bi-specific
Hs_KARI (LDDKSPH)	79	*Thermocrinis albus*	Microaerophilic	LPAGSKS	NADPH
	77	*Hydrogenobacter thermophilus*	Aerobic	LPEGSKS	NADPH
	77	*Hydrogenivirga* sp.	Microaerophilic	LHEKSRS	NADPH

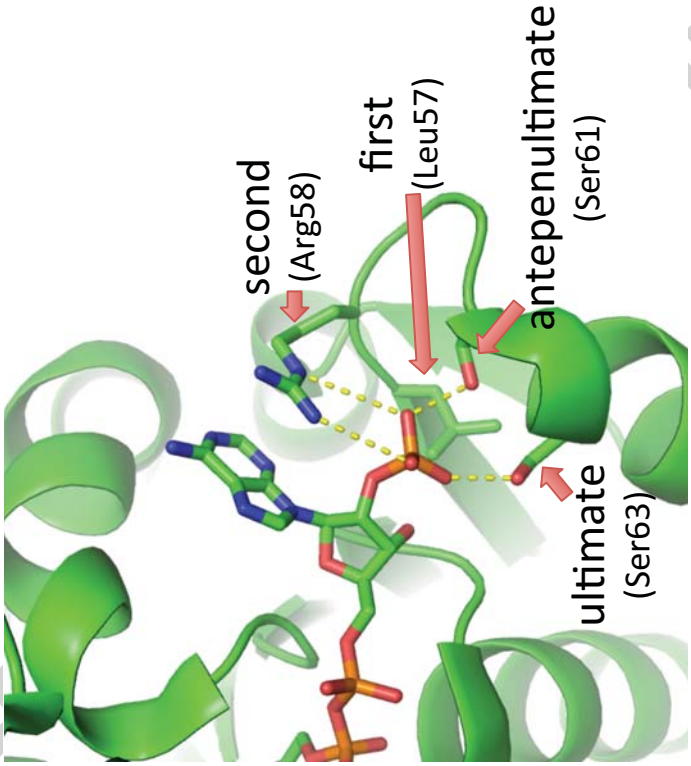
% ID: percent identity of the nearest neighbor to its respective KARI

*) The oxygen requirements of the host organisms were obtained from various microbiological databases.

Highlights

- Available KARI sequences searched for those predicted to effectively utilize NADH
- Eight novel KARIs characterized; four are bi-specific (use NADH and NADPH) and four prefer NADH
- These are the first natural NADH-utilizing KARIs reported
- Nature found different cofactor re-engineering solutions


αB helix
β2 sheet
*
*
C
N



DVVVG---L-R-----EGSK----SWAKA **Consensus of KARIs**

PVRVG---N-I-----DD-A----YRQRA	H6RL58_BLASD	*Blastococcus saxobsidens*
AVRVG---N-I-----DD-R----YFELA	A3NGR4_BURP6	*Burkholderia pseudomallei*
NVVVG---D-L-----LGNE----YSKRA	G4RLE5_THETK	*Thermoproteus tenax*
NVIIG---N-I-----ED-E----YANRA	E1QUK3_VULDI	*Vulcanisaeta distributa*
EVIVG---N-V-----ND-E----YRRRA	D9PYY3_ACIS3	*Acidilobus saccharovorans*
NVIIG---N-V-----KD-R----YHDLA	C4KG48_SULIK	*Sulfolobus islandicus*
NVIVG---QSK-----QFQR----DWDRA	B1ZV88_OPITP	*Opitutus terrae*
***EVIIG---Q-L-----EGDA----YWEKA**	**A8ZTR0_DESOH**	***Desulfococcus oleovorans***
NVIIG-AGN-K-----DRYP----DWENA	D8GTE6_CLOLD	*Clostridium ljungdahlii*
NVTVG---L-E-----RQGK----SWEKA	Q6L044_PICTO	*Picrophilus torridus*
NVKLG---L-E-----REGN----SWKKA	S0AM07_FERAC	*Ferroplasma acidarmanu*
***NVVVG---L-E-----RQGD----SWRRA**	**E0SRA9_IGNAA**	***Ignisphaera aggregans***
NVIVG---L-E-----RKGN----SWILA	L0A7Y1_CALLD	*Caldisphaera lagunensis*
***QVTVG---L-E-----REGK----SWEQA**	**ILVC_METS5**	***Metallosphaera sedula***
***NVVIG---E-TEILGGNKNP----SWEKA**	**D1JA89_9ARCH**	**uncultured archaeon**
DVVVG---L-P-----EGNE----DRPVA	C7NTA6_HALUD	*Halorhabdus utahensis*
NVIIS---E-K-----EGTA----NYKLA	K1X9E2_9BACT	uncultured bacterium
DVIAA---D-L-----KDSP----AWKRA	ILVC_METTP	*Methanosaeta thermophila*
DVIVS---E-L-----EGTE----NYELA	D9QTJ4_ACEAZ	*Acetohalobium arabaticum*
***NVVVG---L-P-----EWDKA---TWERA**	**ILVC_ARCFU**	***Archaeoglobus fulgidus***
***RVIVA---L-D-----DKSP----HRKTA**	**ILVC_HYDSO**	***Hydrogenobaculum sp.***
NVIIG---L-R-----RDLP----DWDKA	D5XAE9_THEPJ	*Thermincola potens*
***KVVVA---D-I-----PSSE----NWKKA**	**K4LVZ1_THEPS**	***Thermacetogenium phaeum***
***NVIVG---L-RKPFDEASEK----EWNAV**	**ILVC_SYNWW**	***Syntrophomonas wolfei subsp. wolfei***
DVLVS---E-I-----PNTP----NYELA	F2NIE5_DESAR	*Desulfobacca acetoxidans*
EVVVS---D-R-----AGSE----NFEKA	A0LMK6_SYNFM	*Syntrophobacter fumaroxidans*