General approach to reversing ketol-acid reductoisomerase cofactor dependence from NADPH to NADH

Sabine Brinkmann-Chen[a,1], Tilman Flock[a,1,2], Jackson K. B. Cahn[a], Christopher D. Snow[a,3], Eric M. Brustad[a,4], John A. McIntosh[a], Peter Meinhold[b], Liang Zhang[a], and Frances H. Arnold[a,5]

[a]California Institute of Technology, Pasadena, CA 91125; and [b]Gevo, Inc., Englewood, CO 80112

Edited by Judith P. Klinman, University of California, Berkeley, CA, and approved May 21, 2013 (received for review March 29, 2013)

To date, efforts to switch the cofactor specificity of oxidoreductases from nicotinamide adenine dinucleotide phosphate (NADPH) to nicotinamide adenine dinucleotide (NADH) have been made on a case-by-case basis with varying degrees of success. Here we present a straightforward recipe for altering the cofactor specificity of a class of NADPH-dependent oxidoreductases, the ketol-acid reductoisomerases (KARIs). Combining previous results for an engineered NADH-dependent variant of *Escherichia coli* KARI with available KARI crystal structures and a comprehensive KARI-sequence alignment, we identified key cofactor specificity determinants and used this information to construct five KARIs with reversed cofactor preference. Additional directed evolution generated two enzymes having NADH-dependent catalytic efficiencies that are greater than the wild-type enzymes with NADPH. High-resolution structures of a wild-type/variant pair reveal the molecular basis of the cofactor switch.

branched-chain amino acid pathway | cofactor imbalance

Ketol-acid reductoisomerases (KARI; EC 1.1.1.86) are a family of nicotinamide adenine dinucleotide phosphate (NADPH)-dependent oxidoreductases that catalyze an alkyl-migration followed by a ketol-acid reduction of (S)-2-acetolactate (S2AL) and 2-aceto-2-hydroxybutyrate to yield (R)-2,3-dihydroxy-isovalerate and (R)-2,3-dihydroxy-3-methylvalerate, respectively (1), which are essential intermediates in the biosynthesis of branched-chain amino acids (BCAAs) (2, 3). The demand for these essential amino acids, used in the preparation of animal feed, human dietary supplements, and pharmaceuticals, is currently estimated to exceed 1,500 tons per year (4). In addition, the BCAA pathway has been engineered to produce fine chemicals and biofuels, including 1-butanol and isobutanol (5, 6). Under the anaerobic conditions preferred for large-scale fermentations, biosynthesis of BCAAs and other products that use this pathway is limited by the pathway's cofactor imbalance and reduced cellular production of NADPH (7, 8). One approach to overcoming the cofactor imbalance is to engineer KARI to use NADH generated in glycolysis, thereby enabling anaerobic production of BCAA pathway products (7, 8).

Efforts to switch the cofactor specificity of oxidoreductases from NADPH to NADH have been made with varying degrees of success (8–17). The three reports of cofactor-switched KARIs (7, 8, 15) from two different organisms show few commonalities in terms of residues targeted for engineering. A general recipe for switching KARI cofactor specificity would allow metabolic engineers to take advantage of the natural sequence diversity of the KARI family, with concomitant diversity in properties such as expression level, pH tolerance, or thermal stability. By combining a systematic analysis of all reviewed and manually annotated [SwissProt (18)] KARIs, information from our previous work on switching the cofactor specificity of the *Escherichia coli* KARI (7), and available KARI structures, we have identified a subset of residues in the β2αB-loop of the Rossmann fold that distinguish NADPH and NADH. Here we provide a nuanced guide to engineering KARI cofactor specificity and apply it to three different KARIs that are representatives of the three canonical KARI β2αB-loop lengths. We also demonstrate that wild-type-like activity using this cofactor, required for industrial applications, can be achieved by directed evolution once cofactor preference has been reversed. High-resolution structures of a wild-type KARI and its cofactor-switched variant with the respective cofactors bound demonstrate how the switch was achieved.

Results and Discussion

In previous work, we described *E. coli* KARI variant Ec_IlvC^{6E6} with four mutations (Ala71Ser, Arg76Asp, Ser78Asp, and Gln110Val) that resulted in a 54,000-fold reversal in cofactor specificity for NADH over NADPH (7). This variant was also highly active when using NADH (85% of wild-type activity using NADPH). Structural analysis of wild-type Ec_IlvC with and without bound cofactor [PDB code 1YRL (19) and PDB code 3ULK (20)] showed that three of the four Ec_IlvC^{6E6} mutations (Ala71Ser, Arg76Asp, and Ser78Asp) are in a loop connecting the β2 sheet and the αB helix of the Rossmann fold (21), hereafter referred to as the β2αB loop. Arg76 and Ser78 are in direct contact with the 2′-phosphate of NADPH. The existence of β2αB loops of varying lengths obscured sequence patterns of NADPH specificity in KARI multiple sequence alignments (1, 8, 15, 22). To find the commonalities among KARI β2αB loops, we systematically analyzed the loop regions of the entire enzyme class (E.C. 1.1.1.86). We generated a multiple sequence alignment of all 643 SwissProt (18) annotated and reviewed KARI sequences (analysis in Table S1) and used structural data to identify and refine the β2αB loop region in the alignment. An N-terminal excerpt (Fig. 1*A*) of the alignment of a few representative KARIs shows the well-known, conserved GxGxxG motif (23) and the loop region 18 amino acids downstream, which is diverse in both length and amino acid sequence.

Author contributions: S.B.-C., T.F., J.K.B.C., C.D.S., P.M., and F.H.A. designed research; S.B.-C., T.F., J.K.B.C., and L.Z. performed research; J.A.M. contributed new reagents/analytic tools; S.B.-C., T.F., J.K.B.C., C.D.S., E.M.B., and F.H.A. analyzed data; S.B.-C., T.F., J.K.B.C., C.D.S., E.M.B., and F.H.A. wrote the paper.

Conflict of interest statement: F.H.A. is cofounder of Gevo, Inc.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors for wild-type *Slackia exigua* Se_KARI and mutant *S. exigua* Se_KARIDDV have been deposited in the Protein Data Bank (PDB), www.pdb.org (PDB ID codes 4KQW and 4KQX, respectively).

[1]S.B.-C. and T.F. contributed equally to this work.

[2]Present address: MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, United Kingdom.

[3]Present address: Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, CO 80523.

[4]Present address: Department of Chemistry and Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.

[5]To whom correspondence should be addressed. E-mail: frances@cheme.caltech.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1306073110/-/DCSupplemental.

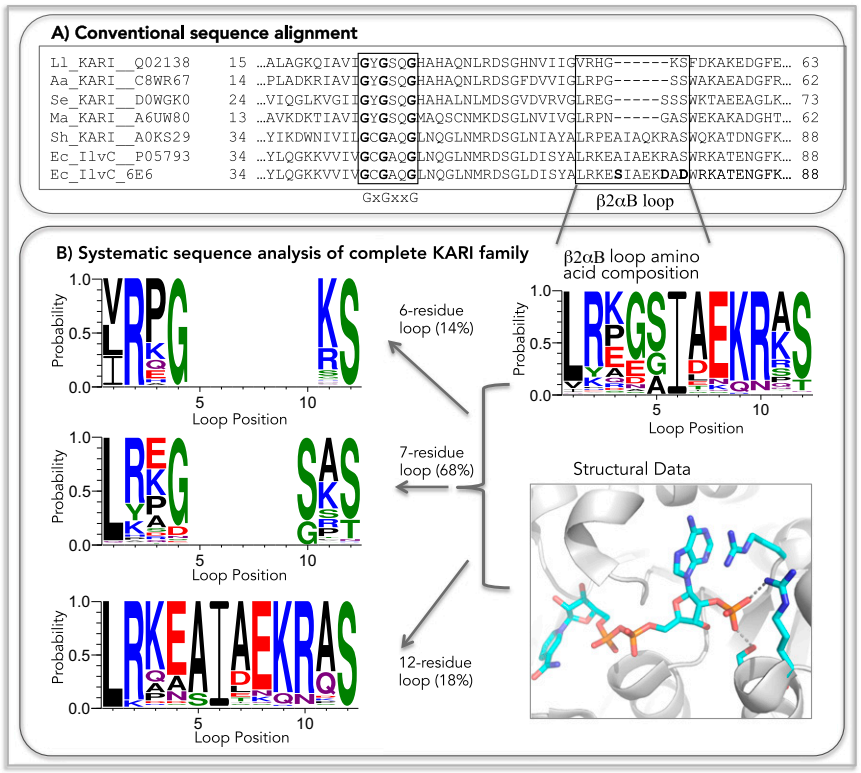

Fig. 1. (*A*) Excerpt of a multiple sequence alignment of *L. lactis* KARI (Ll_KARI, Q02138), *A. acidocaldarius KARI* (Aa_KARI, C8WR67), *S. exigua* KARI (Se_KARI, D0WGK0), *M. aeolicus* KARI (Ma_KARI, A6UW80), *Shewanella* sp. KARI (Sh_KARI, A0KS29), *E. coli* KARI (Ec_IlvC, P05793), and Ec_IlvC^{6E6}. The GxGxxG motif is highlighted in bold and boxed. The β2αB loop is boxed. Mutations of variant Ec_IlvC^{6E6} are shown in bold. (*B*) Subalignments according to β2αB-loop length uncover three different conservation patterns we suggest are important for 2′-phosphate interaction. Colors in the sequence logo refer to hydrophobic (black), positively charged (blue), negatively charged (red), and polar (green) residues.

On the basis of the sequence alignment alone, the diversity of this region might seem to argue against a conserved function of loop residues (Fig. 1*A*). However, analysis of the 643 KARI sequences shows three different loop lengths: six (14%), seven (68%), and 12 residues (18%) (Fig. 1*B*). On subalignment of KARIs according to loop length, common conservation patterns emerge: a positively charged residue (73% Arg and 9% Lys) usually appears at the N-terminal end of the β2αB loop, and small, polar residues such as serine (85%) or threonine (11%) predominate at the C-terminal end. The last two residues in the six-residue loops are conserved (Lys/Arg and Ser), and the last three residues follow the pattern SxS in seven-residue and RxS in 12-residue loops.

Comparison of the β2αB loop segment of all available KARI structures [seven structures from four different organisms (19, 20, 24–27)] provides an explanation for the trends found in the sequence alignment. Three of the seven structures have NADPH cocrystallized, and in all those cases, the small polar C-terminal residues interact with the NADPH 2′-phosphate. The conserved N-terminal Arg is homologous to Arg68 in Ec_IlvC^{6E6}, which has recently been shown to form a cation–pi interaction with the NADPH adenine ring (20) and likely plays a dual role in binding the nucleobase, as well as the 2′-phosphate. On the basis of these findings, we postulate that the β2αB loop is a major determinant of NADPH cofactor specificity in wild-type KARIs. This analysis of the β2αB loop provides the basis for transferring previously reported cofactor-switching mutations in Ec_IlvC^{6E6} to corresponding residues in KARIs that have similar 12-residue loops, as well as KARIs with six- and seven-residue loops.

Transfer of a 12-Residue Cofactor Switch Solution to Seven- and Six-Residue β2αB Loops: Se_ KARI from *Slackia exigua* and Ll_KARI from *Lactococcus lactis*. We examined the transferability of mutations previously reported for Ec_IlvC^{6E6}, which has a 12-residue loop, to KARIs with shorter loops using Se_KARI from *S. exigua* (seven residues) and Ll_KARI from *L. lactis* (six residues) (Fig. 1). On inspection of the sequence alignment in Fig. 1*A*, we identified residues Ser61 and Ser63 of Se_KARI as corresponding to residues Arg76 and Ser78 in Ec_IlvC (Arg76Asp and Ser78Asp in Ec_IlvC^{6E6}). Ec_IlvC mutation Ala71Ser had no match in the Se_KARI amino acid sequence because of the latter enzyme's shorter β2αB loop. Grafting both aspartates from Ec_IlvC^{6E6} into Se_KARI resulted in variant Se_KARI^{DD} (Ser61Asp and Ser63-Asp) and a 7,800-fold reversal of cofactor specificity from NADPH to NADH. Se_KARI^{DD} (Table 1) had an eightfold decreased k_{cat} with NADPH (from 0.8 to 0.1 s^{-1}), whereas the k_{cat} with NADH increased from 0.4 to 1.0 s^{-1}. The mutations increased the K_M for NADPH 880-fold, but only 2.5-fold for NADH. Overall, the catalytic efficiency for NADH remained the same, whereas catalytic efficiency for NADPH was reduced 7,300-fold to 0.11 $mM^{-1}s^{-1}$ (Table 1). That switching Se_KARI cofactor specificity was achieved with only two mutations suggests that KARIs with seven-residue and 12-residue β2αB loops share similar cofactor specificity determinants involving interactions with small, polar residues at the end of the β2αB loop.

Ll_KARI is a representative of the 14% of the 643 KARI sequences with a six-residue β2αB loop. Whereas in 12- and seven-residue β2αB-loop KARIs the antepenultimate and ultimate residues are highly conserved, in six-loop KARIs the ultimate conserved serine is usually preceded by a positively charged residue at the penultimate position. We hypothesized that residues Lys52 and Ser53 of Ll_KARI were equivalent in function to residues Arg76 and Ser78 in Ec_IlvC. Ll_KARI^{DD} variant with mutations Lys52Asp and Ser53Asp, however, expressed at an extremely low level and exhibited no measurable activity with either cofactor, which could be a result of the destabilizing effects of the two adjacent Asp residues. Similar results were obtained with Ll_KARI^{ED}. These two Asp mutations are separated by an additional residue in KARIs with the longer β2αB loops. The failed transfer of *E. coli* cofactor switch mutations to Ll_KARI suggests that KARIs with six-residue loops require a modified approach.

To develop a recipe for switching cofactor specificity in KARIs having a six-residue β2αB loop, we generated single-site-saturation

Table 1. Biochemical properties of Ec_IlvC (12 residues), Sh_KARI (12 residues), Se_KARI (seven residues), Ma_KARI (seven residues), Ll_KARI (six residues), Aa_KARI (six residues), and their cofactor-switched variants

Enzyme	Mutations	K_M for cofactors [μM]		k_{cat} for cofactors [s^{-1}]		k_{cat}/K_M [$mM^{-1}s^{-1}$]	
		NADH	NADPH	NADH	NADPH	NADH	NADPH
Ec_IlvC (7)	—	1,075	41	0.3	3.6	0.3	88
Ec_IlvC^{6E6} (7)	**Ala71Ser, Arg76Asp, Ser78Asp**, Gln110Val	30	650	2.3	0.2	74	0.4
Ec_IlvC$^{P2D1-A1}$	**Ala71Ser, Arg76Asp, Ser78Asp**, Gln110Val, Asp146Gly, Gly185Arg, Lys433Glu	26	>1,400	4.3	0.54	165	0.4
Sh_KARI	—	415	1.0	1.1	4.5	2.6	4,500
Sh_KARI^{DD}	**Arg76Asp, Ser78Asp**	90	>1,000	1.30	0.10	14	0.17
Sh_KARI^{6E6}	**Ala71Ser, Arg76Asp, Ser78Asp**, Gln110Val	75	600	2.40	0.30	32	0.5
Se_KARI	—	45	1.0	0.41	0.8	9	800
Se_KARI^{DD}	**Ser61Asp, Ser63Asp**	113	880	0.97	0.10	9	0.11
Se_KARI^{DDV}	**Ser61Asp, Ser63Asp**, Ile95Val	47	>1,000*	1.01	0.25	22	0.25
*Ma_KARI	—	59	17.3	0.3	0.7	4.6	43
*Ma_KARI^{DD}	**Gly50Asp, Ser52Asp**	26	80	0.15	0.004	6.2	0.05
Ll_KARI	—	285	13	0.10	0.8	0.43	65
Ll_KARI^{LPLD}	**Val48Leu, Arg49Pro, Lys52Leu, Ser53Asp**	108	1,000	0.40	0.08	3.7	0.08
Ll_KARI^{LPED}	**Val48Leu, Arg49Pro, Lys52Glu, Ser53Asp**	128	1,180	0.35	0.06	2.7	0.05
Ll_KARI^{2G6}	**Val48Leu, Arg49Pro, Lys52Leu, Ser53Asp**, Glu59Lys, Thr182Ser, Glu320Lys	15	749	1.07	0.35	71	0.47
*Aa_KARI	—	28	18	0.26	0.66	9	37
*Aa_KARI^{PLD}	**Arg48Pro, Ser51Leu, Ser52Asp**	43	>1,000	0.032	0.013	0.7	<0.013
*Aa_KARI^{PLDA}	**Arg48Pro, Ser51Leu, Ser52Asp,** Arg84Ala	27	>1,000	0.03	0.01	1.1	<0.01

All enzymes were his_6-tagged and purified before characterization. The mutations are given relative to each wild-type sequence. Each value represents the average of three independent measurements. Enzyme activities were determined in 100 mM potassium phosphate at pH 7 with 1 mM DTT, 200 μM NADPH or NADH, 10 mM S2AL, and 10 mM $MgCl_2$. Concentrations of the purified enzymes were determined using the Bradford assay. The Michaelis–Menten constants for the cofactors were measured with appropriate dilutions of NADPH and NADH in the presence of saturating concentrations of substrate S2AL. Mutations located within the β2αB-loop are highlighted in bold. For errors, please refer to Table S2.

*Ma_KARI, Aa_KARI, and their variants show cooperative behavior, and their kinetics follow the Hill equation instead of the Michaelis–Menten equation: affinity is described as K_H and catalytic efficiency as k_{cat}/K_H to compare them with the other KARIs. Hill coefficients are provided in Table S2.

mutagenesis libraries at each of the six loop residues in Ll_KARI (Val48, Arg49, His50, Gly51, Lys52, and Ser53), expecting that the β2αB loop also is key to specificity in this KARI family. On screening for activity with both cofactors, we found no single mutation that resulted in an NADH-preferring variant (rate of NADH consumption > rate of NADPH consumption at saturating substrate conditions). Mutations were identified at Arg49 (Pro) and Val48 (Leu) that increased activity with both cofactors. Saturation mutagenesis at His50, Gly51, Lys52, and Ser53 did not yield variants with improved NADH activity. Convinced that NADPH specificity in KARIs with a six-residue β2αB loop is conveyed in a similar manner as in KARIs with seven- and 12-residue loops, we built a dual-site library by saturation mutagenesis at Lys52 and Ser53 while also incorporating the Val48Leu and Arg49Pro mutations. Screening this library to ~80% coverage, we identified two variants, Ll_KARI^{LPLD} (mutations Val48Leu, Arg49Pro, Lys52Leu, and Ser53Asp) and Ll_KARI^{LPED} (mutations Val48Leu, Arg49Pro, Lys52Glu, and Ser53Asp), with 46-fold and 54-fold specificity for NADH over NADPH (Table 1), representing specificity shifts of 6,600 (Ll_KARI^{LPLD}) and 7,700 (Ll_KARI^{LPED}).

These cofactor-switched KARIs with six-residue loops contained mutations to four of the six loop residues, suggesting that KARIs with six-residue loops are slightly more difficult templates for engineering the cofactor switch. Although mutation of the conserved arginine at the beginning of the β2αB loop was not required to switch cofactor specificity in seven- and 12-loop KARIs, variant Ll_KARI^{LRLD} with Pro reverted to Arg showed that the Arg contributes to NADPH cofactor preference in six-residue loop KARIs. Reversion of Pro49 to Arg not only reduced the cofactor K_M values (twofold for NADH and threefold for NADPH) (Table S2) but also decreased the k_{cat} on NADH (fourfold), thereby reducing the 46-fold preference for NADH over NADPH to only sixfold. Reversion of Leu48 to Val (Ll_KARI^{VPLD}) lowered the catalytic efficiency on NADH threefold (because of a reduction in k_{cat}), whereas the NADPH K_M was reduced twofold. Although polar residues at the C-terminal end of the loop and charged residues at the N-terminal end of the loop continue to support NADPH specificity, removing the polar residues at the C-terminal end of the loop might not be sufficient to switch the specificity of six-residue loop KARIs in favor of NADH. In the shorter six-residue loop KARIs, the N-terminal, positively charged Arg residue appears to form stronger interactions with the 2′-phosphate than in the seven- and 12-residue loop KARIs, possibly because of its closer packing and proximity to the cofactor. A six-amino acid loop KARI structure with cofactor would help elucidate the detailed mechanism.

Recovering Catalytic Activity of Cofactor-Switched Enzymes. Shifting cofactor preference often decreased the overall activity (i.e., catalytic efficiency using NADH) relative to the wild-type enzyme using NADPH, which is also true for other cofactor-switched enzymes (9–11, 13, 15). To demonstrate that the activity of the cofactor-switched variants can be improved to match or even surpass the wild-type enzyme, we randomly mutated Se_KARI^{DD}, Ll_KARI^{LPLD}, and Ec_IlvC^{6E6} and screened for higher total activity in cell lysate. For Se_KARI, we identified variant Se_KARI^{DDV} with mutation Ile95Val. Interestingly, this mutation corresponds to Gln110Val, which was previously found in Ec_IlvC^{6E6}, and is speculated to confer general activation by optimizing cofactor orientation for catalysis. With Ll_KARI, we isolated variant Ll_KARI^{2G6}, which contained three new mutations: Glu59Lys, Thr182Ser, and Glu320Lys. These mutations effectively restored the enzyme to wild-type levels of activity; the catalytic efficiency of Ll_KARI^{2G6} with NADH was

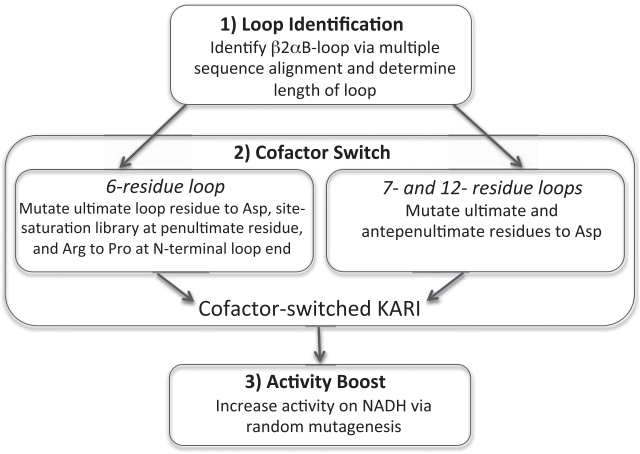

Fig. 2. Cofactor switch guide for the KARI enzyme family. Steps include (*1*) identification of the loop, (*2*) determination of β2αB-loop length and mutation based on loop length, and (*3*) improvement of overall activity on NADH via random mutagenesis.

~10% higher than wild-type Ll_KARI with NADPH (71 vs. 65 $mM^{-1}s^{-1}$). The Ec_IlvC^{6E6} random mutant library yielded variant Ec_IlvC$^{P2D1-A1}$, also with three additional mutations (Asp146Gly, Gly185Arg, and Lys433Glu) and an approximately twofold greater catalytic efficiency on NADH than the wild-type on NADPH (Table 1). The random mutations in Ec_ILvC$^{P2D1-A1}$ and Ll_KARI^{2G6} are surface mutations, and their effects are difficult to rationalize. These three examples demonstrate that the activity of a cofactor-switched enzyme can be improved, even to levels that exceed wild-type activity, and that activating mutations can be found outside of the specificity-determining β2αB loop.

Cofactor Switch Guide for the KARI Enzyme Family. We propose the following guide for switching KARI cofactor specificity, which does not require a priori knowledge of the KARI structure (Fig. 2). The first step is the identification of the β2αB loop and its length via sequence alignment against the KARIs reported in this work or a multiple sequence alignment of all KARIs. If the target KARI has a 12- or seven-residue loop, replacement of the last and third-to-last residue of the loop with aspartates is likely to achieve a switch in cofactor specificity. In the case of the six-residue loop, a modified approach is required. The last polar loop residue should be mutated to aspartate, and the conserved charged residue near the N-terminus of the loop should be mutated to Pro. Simultaneously, the penultimate loop residue should be targeted for site-saturation mutagenesis. This approach led to reversed cofactor preference in both test cases. Last, to achieve wild-type-like activity for NADH, additional mutations that fine-tune cofactor orientation, as exemplified by Gln110Val or Ile95Val, may be introduced, and further enhancement of activity can be achieved by random mutagenesis and screening.

Application of the Cofactor Switch Guide. We tested the proposed protocol on three additional KARIs representing the β2αB-loop lengths (12, seven, and six residues) and composition. With the addition of these three KARIs, we covered the different phylogenetic subbranches of the KARI enzyme family. Representative of 12-residue β2αB-loop KARIs was Sh_KARI from *Shewanella*

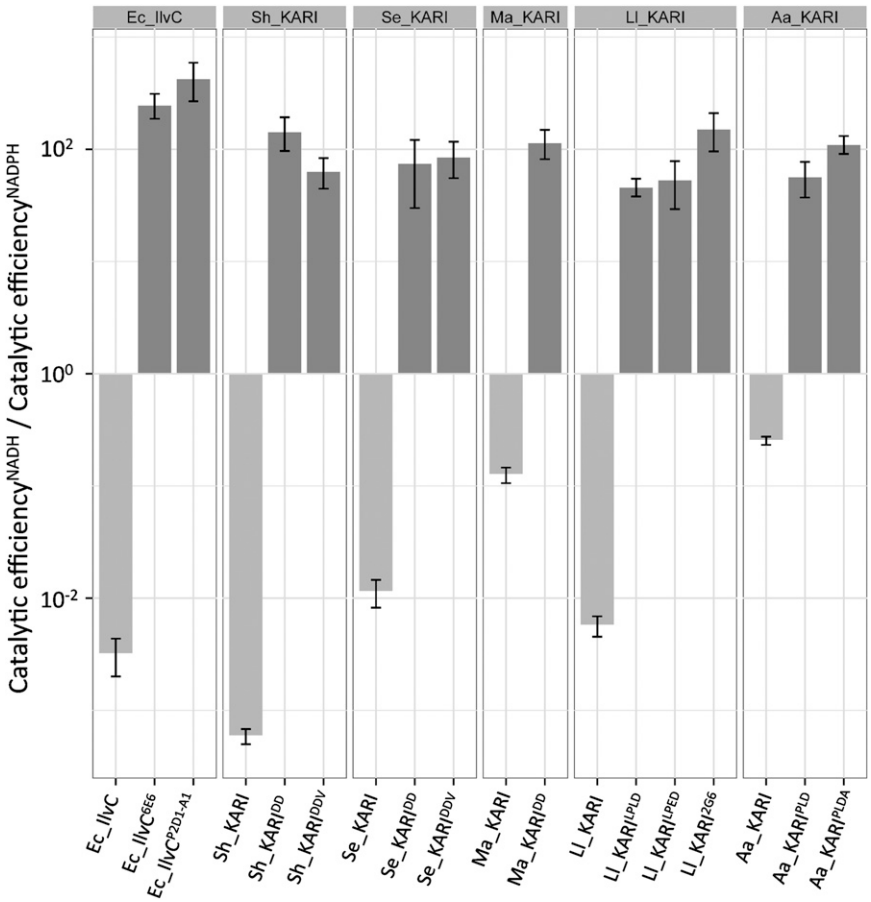

Fig. 3. Catalytic efficiency NADH/catalytic efficiency NADPH, on log scale, for six wild-type KARIs (light gray) and their cofactor-switched variants (dark gray).

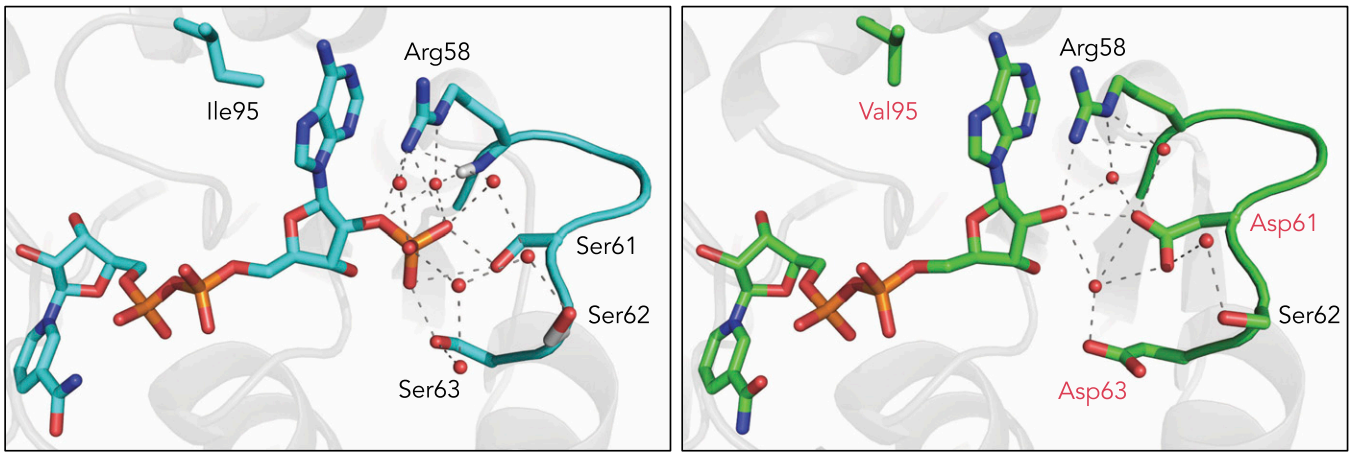

Fig. 4. Crystal structures of Se_KARI wild-type enzyme with cocrystallized NADPH (*Left*, cyan) and variant Se_KARI^{DDV} with cocrystallized NADH (*Right*, green). The β2αB loop is highlighted, and side chains involved in defining cofactor-specificity are shown as sticks. Introduced mutations (Ser61Asp, Ser63Asp, and Ile95Val) are shown with red labels. In Se_KARI^{DDV}, Ser61Asp and Ser63Asp compensate for the missing 2′-phosphate and electrostatically and sterically exclude NADPH. Mutation Ile95Val allows the adenine moiety to shift 1 Å inward. The backbone of Arg58 follows this movement, so that in both structures the cation–pi interaction with the adenine moiety is preserved.

sp. *Methanococcus aeolicus* Ma_KARI (28) exemplifies the seven-residue type, and Aa_KARI from *Alicyclobacillus acidocaldarius* has a six-residue loop (Table 1). We used this last enzyme to test the transferability of the Q110/I95 position by making and screening a site-saturation library at position Arg84.

By introducing a customized set of two to four mutations based on the guide in Fig. 2, we obtained variants with the desired cofactor specificity for KARI family members sharing as little as 20% sequence identity (Table S3). Low catalytic efficiency in cofactor-switched variants can be remedied by directed evolution, as demonstrated for Ec_IlvC$^{P2D1\text{-}A1}$ and Ll_KARI^{2G6}. Mutations (Ala or Val) at positions corresponding to Ec_IlvC's Q110 improved activity in three different KARIs. Overall, catalytic efficiency ratios (NADH/NADPH) of more than 400-fold (Fig. 3) with catalytic efficiencies up to 188% of those of the corresponding wild-type enzymes using NADPH were achieved, corresponding to shifts in cofactor specificity of 200–200,000-fold.

Molecular Determinants of Cofactor Specificity in KARIs. We solved the crystal structures of Se_KARI wild-type enzyme (1.39 Å, PDB 4KQW) and variant Se_KARI^{DDV} (1.8 Å, PDB 4KQX) with their respective cofactors. The crystallographic parameters are summarized in Table S4. These structures confirm that only the β2αB loop is involved in interactions with the respective 2′-moiety, and thus is responsible for specificity. In the wild-type structure, three residues form direct interactions with the 2′-phosphate of NADPH: Arg58, Ser61, and Ser63. The structures support the suggested dual role of Arg58: the positively charged guanidinium moiety is 3.5 Å from the adenine moiety of the cofactor in both structures, forming cation–pi stacking interactions (29), as reported for *E. coli* KARI (20). At the same time, this side chain could form a salt bridge to the negatively charged 2′-phosphate of NADPH (possibly also involving the oxygen of the phosphoester bond) and a hydrogen bond to the 2′-OH in NADH.

The residues that were mutated to alter cofactor preference, Ser61 and Ser63, are in a position to hydrogen bond directly to at least a single oxygen atom of the phosphate. The high-resolution structures revealed water molecules surrounding the 2′-phosphate group, enabling additional, indirect interactions with the side chains of Arg58, Ser61, and Ser63 (Fig. 4). The side chain of Ser62 stabilizes this network of water molecules, as does the Arg58 backbone. In the mutant structure, the β2αB loop is moved slightly closer toward the cofactor. Mutations Ser61Asp and Ser63Asp would interrupt the serine hydrogen bonding interactions and also result in electrostatic repulsion to the 2′-phosphate of NADPH. The carboxyl groups of the two aspartates compensate for the missing 2′-phosphate by filling the pocket and substituting its negative charge. In addition, the β carboxyl group of Ser61Asp is at an ideal distance for hydrogen bonding to the 2′-hydroxyl group of NADH. As in the wild-type structure, water molecules link the 2′-hydroxyl moiety of the ribose sugar with Arg58 and Asp63 by hydrogen bonds.

Whereas these mutations shifted the KARI cofactor preference, improved catalytic activity was achieved by substituting an additional residue that is not part of the loop, Ile95, with Val. This mutation retained the hydrophobic nature while allowing the adenine moiety to shift about 1 Å toward the side chain of residue 95. This compensates exactly the distance the β2αB loop is flipped inward in the mutant structure at position Arg58 and preserves the cation–pi stacking of the adenine moiety and the side chain of Arg58, which is in the same rotamer conformation in the wild-type and in the cofactor-switched mutant. We propose that this movement compensates for the slightly different conformations of NADPH and NADH and readjusts the catalytically active nicotinamide moiety of NADH to take on a more favorable position for electron transfer. A similar activating mechanism is speculated for Ec_IlvC^{6E6}'s Gln110Val mutation (7). All other interactions of Se_KARI^{DDV} with the cofactor, for instance, involving the GxGxxG motif, remain the same. The remaining loop residues Leu57, Glu59, and Gly60 are not involved in binding the cofactor.

Conclusions: General Cofactor Binding Principles for KARIs. We identified a common motif for cofactor specificity in KARIs by using structural knowledge to identify the critical role of the β2αB loop and deconvolute its three key variations in multiple sequence alignments, thereby defining a limited set of mutations that generate NADH specificity. Applying this to KARIs with different loop lengths has enabled us to develop a robust guide to switching the cofactor preference from NADPH to NADH of any enzyme in this family. This approach opens the door to exploration of a wealth of different KARI properties in the context of valuable BCAA pathways under anaerobic conditions.

Materials and Methods

Cloning and Library Construction. Strains, plasmids, and primers are listed in Tables S5 and S6. The genes encoding *S. exigua* Se_KARI, *L. lactis* Ll_KARI, *Shewanella* sp. Sh_KARI, and Sh_KARI^{6E6} were obtained from DNA2.0. The genes encoding *M. aeolicus* Ma_KARI and *A. acidocaldarius* Aa_KARI were obtained as gBlocks from Integrated DNA Technologies. For each gene, the

gBlocks were assembled via PCR, using T7 promoter and terminator primers and Phusion polymerase following the manufacturer's instructions (Thermo Scientific). Site-saturation mutagenesis libraries were made by splicing by overlap extension PCR (30), as described (7). Error-prone PCR was performed according to a published protocol (31), using commercial T7 promoter and terminator primers. All KARIs and libraries were cloned into pET22(b)+, using NdeI and XhoI in frame with the C-terminal his-tag for expression in *E. coli*. Heterologous protein expression, high-throughput expression, and purification were conducted as described (7).

Kinetic Assays and High-Throughput Screening. For the high-throughput assays, *E. coli* cells were lysed with 100 mM potassium phosphate at pH 7, 750 mg/L lysozyme, and 10 mg/L DNaseI. KARI activities were then assayed by monitoring NAD(P)H consumption in the presence of S2AL at 340 nm in a plate reader. The assay buffer contained 100 mM potassium phosphate at pH 7, 1 mM DTT, 200 μM NAD(P)H, 12.5 mM S2AL for Ll_KARI and 2.5 mM for the other KARIs, and 10 mM $MgCl_2$. The Ll_KARI error-prone PCR library was screened at 5 mM S2AL. The Ec_IlvC^{6E6} library was screened at 1 mM S2AL. The Se_KARI^{DD} library was screened at 100 μM NADH and 2.5 mM S2AL.

KARI Sequence Alignment. Manually annotated and reviewed sequence data for ketol-acid reductoisomerases (E.C. 1.1.1.86) were retrieved from the Uniprot Database (32). Clustal Omega (33, 34) was used to perform a multiple sequence alignment. The sequence logo plot (35) was created with the WebLogo 3.3 interface (36).

Crystallization and Data Collection. *N*-hydroxy-*N*-isopropyloxamate was prepared as described (37). High-throughput screening of crystallization conditions for Se_KARI and Se_KARI^{DDV} was conducted at the Beckman Molecular Observatory at the California Institute of Technology. For Se_KARI with NADPH, the best condition was an unbuffered 0.2 M di-ammonium tartrate solution containing 20% (wt/vol) PEG3350 as precipitant. For Se_KARI^{DDV} with NADH and *N*-hydroxy-*N*-isopropyloxamate as inhibitor, the best condition was an unbuffered 0.1-M potassium thiocyanate solution with 30% (wt/vol) polyethylene glycol monomethyl ether 2000 as precipitant. The crystals were soaked in Fomblin oil for cryoprotection before flash-freezing in liquid nitrogen. Diffraction data were collected using a Dectris Pilatus 6M detector on beamline 12–2 at the Stanford Synchrotron Radiation Laboratory at 100 K. Diffraction datasets were integrated with *XDS* (38) and scaled using SCALA (39).

Structure Determination and Refinement. For Se_KARI, the structure of *Pseudomonas aeruginosa* KARI [PDB code 1NP3 (24)] was used as for molecular replacement. A multiblock refinement was applied dividing the model in six subparts according to secondary structure elements (residues 1–202, 203–228, 229–252, 253–278, 279–308, and 309–337) to allow automated standard refinement with Phenix (CCP4 suite). Refinement was conducted by iterating automatic refinement with Refmac5 (CCP4 suite) and manual refinement using Coot (40). We used the refined wild-type structure as a model for molecular replacement to obtain the structure for Se_KARI^{DDV}. After placement of the inhibitor, several iterations of automated refinement with Refmac5 and manual refinement in Coot were performed. The structures were submitted to the protein database as PDB 4KQW (Se_KARI) and PDB 4KQX (Se_KARI^{DDV}).

ACKNOWLEDGMENTS. We thank Dr. Jens Kaiser and Dr. Pavle Nikolovski for their continued support. This publication was supported by the Gordon and Betty Moore Foundation through Grant GBMF2809 to the Caltech Programmable Molecular Technology Initiative, the German National Academic Foundation (to T.F.), a Ruth L. Kirschstein National Research Service Award (F32GM101792) (to J.A.M.), and a Ruth L. Kirschstein National Institutes of Health postdoctoral fellowship (F32GM087102) (to E.M.B.). The Molecular Observatory is supported by the Gordon and Betty Moore Foundation, the Beckman Institute, and the Sanofi-Aventis Bioengineering Research Program at Caltech.

1. Dumas R, Biou V, Halgand F, Douce R, Duggleby RG (2001) Enzymology, structure, and dynamics of acetohydroxy acid isomeroreductase. *Acc Chem Res* 34(5):399–408.
2. Arfin SM, Umbarger HE (1969) Purification and properties of the acetohydroxy acid isomeroreductase of Salmonella typhimurium. *J Biol Chem* 244(5):1118–1127.
3. Umbarger HE, Davis B (1960) *The Bacteria* 3:167.
4. Park JH, Lee SY (2010) Fermentative production of branched chain amino acids: a focus on metabolic engineering. *Appl Microbiol Biotechnol* 85(3):491–506.
5. Atsumi S, et al. (2008) Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab Eng* 10(6):305–311.
6. Atsumi S, Hanai T, Liao JC (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* 451(7174):86–89.
7. Bastian S, et al. (2011) Engineered ketol-acid reductoisomerase and alcohol dehydrogenase enable anaerobic 2-methylpropan-1-ol production at theoretical yield in *Escherichia coli*. *Metab Eng* 13(3):345–352.
8. Hasegawa S, et al. (2012) Improvement of the redox balance increases L-valine production by *Corynebacterium glutamicum* under oxygen deprivation conditions. *Appl Environ Microbiol* 78(3):865–875.
9. Elmore CL, Porter TD (2002) Modification of the nucleotide cofactor-binding site of cytochrome P-450 reductase to enhance turnover with NADH *in Vivo*. *J Biol Chem* 277(50):48960–48964.
10. Kristan K, Stojan J, Adamski J, Lanisnik Rizner T (2007) Rational design of novel mutants of fungal 17beta-hydroxysteroid dehydrogenase. *J Biotechnol* 129(1):123–130.
11. Petschacher B, Leitgeb S, Kavanagh KL, Wilson DK, Nidetzky B (2005) The coenzyme specificity of *Candida tenuis* xylose reductase (AKR2B5) explored by site-directed mutagenesis and X-ray crystallography. *Biochem J* 385(Pt 1):75–83.
12. Petschacher B, Nidetzky B (2008) Altering the coenzyme preference of xylose reductase to favor utilization of NADH enhances ethanol yield from xylose in a metabolically engineered strain of *Saccharomyces cerevisiae*. *Microb Cell Fact* 7:9.
13. Shiraishi N, Croy C, Kaur J, Campbell WH (1998) Engineering of pyridine nucleotide specificity of nitrate reductase: mutagenesis of recombinant cytochrome b reductase fragment of *Neurospora crassa* NADPH:Nitrate reductase. *Arch Biochem Biophys* 358(1):104–115.
14. Khoury GA, et al. (2009) Computational design of *Candida boidinii* xylose reductase for altered cofactor specificity. *Protein Sci* 18(10):2125–2138.
15. Rane MJ, Calvo KC (1997) Reversal of the nucleotide specificity of ketol acid reductoisomerase by site-directed mutagenesis identifies the NADPH binding site. *Arch Biochem Biophys* 338(1):83–89.
16. Scrutton NS, Berry A, Perham RN (1990) Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* 343(6253):38–43.
17. Rosell A, et al. (2003) Complete reversal of coenzyme specificity by concerted mutation of three consecutive residues in alcohol dehydrogenase. *J Biol Chem* 278(42):40573–40580.
18. Boeckmann B, et al. (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* 31(1):365–370.
19. Tyagi R, Duquerroy S, Navaza J, Guddat LW, Duggleby RG (2005) The crystal structure of a bacterial class II ketol-acid reductoisomerase: domain conservation and evolution. *Protein Sci* 14(12):3089–3100.
20. Wong S-H, Lonhienne TGA, Winzor DJ, Schenk G, Guddat LW (2012) Bacterial and plant ketol-acid reductoisomerases have different mechanisms of induced fit during the catalytic cycle. *J Mol Biol* 424(3-4):168–179.
21. Rossmann MG, Moras D, Olsen KW (1974) Chemical and biological evolution of nucleotide-binding protein. *Nature* 250(463):194–199.
22. Dumas R, Curien G, DeRose RT, Douce R (1993) Branched-chain-amino-acid biosynthesis in plants: molecular cloning and characterization of the gene encoding acetohydroxy acid isomeroreductase (ketol-acid reductoisomerase) from *Arabidopsis thaliana* (thale cress). *Biochem J* 294(Pt 3):821–828.
23. Wierenga RK, De Maeyer MCH, Hol WGJ (1985) Interaction of pyrophosphate moieties with alpha-helixes in dinucleotide-binding proteins. *Biochemistry* 24(6):1346–1357.
24. Ahn HJ, et al. (2003) Crystal structure of class I acetohydroxy acid isomeroreductase from *Pseudomonas aeruginosa*. *J Mol Biol* 328(2):505–515.
25. Biou V, et al. (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 A resolution. *EMBO J* 16(12):3405–3415.
26. Leung EWW, Guddat LW (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(1):167–182.
27. Thomazeau K, et al. (2000) Structure of spinach acetohydroxyacid isomeroreductase complexed with its reaction product dihydroxymethylvalerate, manganese and (phospho)-ADP-ribose. *Acta Crystallogr D Biol Crystallogr* 56(Pt 4):389–397.
28. Xing RY, Whitman WB (1991) Characterization of enzymes of the branched-chain amino acid biosynthetic pathway in *Methanococcus spp*. *J Bacteriol* 173(6):2086–2092.
29. Gallivan JP, Dougherty DA (1999) Cation-pi interactions in structural biology. *Proc Natl Acad Sci USA* 96(17):9459–9464.
30. Kunkel TA, Roberts JD, Zakour RA (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol* 154:367–382.
31. Bloom JD, Labthavikul ST, Otey CR, Arnold FH (2006) Protein stability promotes evolvability. *Proc Natl Acad Sci USA* 103(15):5869–5874.
32. UniProt Consortium (2012) Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res* 40(Database issue):D71–D75.
33. Goujon M, et al. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38(Web Server issue):W695-9.
34. Sievers F, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.
35. Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* 18(20):6097–6100.
36. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14(6):1188–1190.
37. Aulabaugh A, Schloss JV (1990) Oxalyl hydroxamates as reaction-intermediate analogues for ketol-acid reductoisomerase. *Biochemistry* 29(11):2824–2830.
38. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125–132.
39. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1):72–82.
40. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.