ChemBioChem

Combining Chemistry and Biology

Chemistry Europe

European Chemical Societies Publishing

Accepted Article

Title: Engineering Glucose Dehydrogenase to Favor Totally Synthetic Biomimetic containing carboxyl group

Authors: Jieyu Zhou, Xiangyuan Gu, Yichun Zhu, Zhenyan Tao, and Ye Ni

To be cited as: *ChemBioChem* **2023**, e202300066

Link to VoR: https://doi.org/10.1002/cbic.202300066

WILEY VCH

10.1002/cbic.202300066

WILEY-VCH

RESEARCH ARTICLE

Engineering Glucose Dehydrogenase to Favor Totally Synthetic Biomimetic Cofactors containing carboxyl group

Jieyu Zhou, [a] Xiangyuan Gu, [a] Yichun Zhu, [a] Zhenyan Tao, [b] Ye Ni*[a]

[a] Dr. J. Zhou, X. Gu, Y. Zhu, Prof. Dr. Y. Ni
Key laboratory of industrial Biotechnology, Ministry of Education, School of Biotechnology,
Jiangnan University,
Wuxi 214122, Jiangsu, China.
http://biocat.jiangnan.edu.cn/
E-mail: yni@jiangnan.edu.cn

[b] Z. Tao
State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, and School of Life Sciences and Biotechnology,
Shanghai Jiao Tong University,
Shanghai 200240, China.
Supporting information for this article is given via a link at the end of the document.

Abstract: The utilization of unnatural nicotinamide cofactors for reactions catalyzed by oxidoreductases has gained increasing interest. Totally synthetic nicotinamide cofactor biomimetics (NCBs) are cost‐effective and convenient to synthesize. Thus, it has becoming increasingly important to develop enzymes that accept NCBs. Here, we have engineered *Ss*GDH to favor a newly synthesized unnatural cofactor 3-carbamoyl-1-(4-carboxybenzyl) pyridin-1-ium (BANA$^+$). Using *in situ* ligand minimization tool, sites 44 and 114 were identified as hotspots for mutagenesis. All the double mutants demonstrated 2.7–7.7-fold improvements in catalytic activity, and the best double mutant E44D/E114L exhibited 10.6-fold increased catalytic efficiency toward BANA$^+$. These results provide valuable information for the rational engineering of oxidoreductases with versatile NCBs-dependency, as well as the design of novel biomimetic cofactors.

Introduction

Nicotinamide cofactor-dependent oxidoreductases are used in a wide range of industrial and biotechnological applications[1, 2]. In enzymatic catalysis, the nicotinamide cofactor plays a fundamental role in electron and energy transfer. However, these cofactors display weak resistance to extreme conditions, such as high temperature, organic solvents, and acidic/alkaline pH, which limits the potential of oxidoreductases in industrial applications[3]. Therefore, engineering cofactor preference toward biomimetic cofactors has become particularly important for NAD(P)-dependent oxidoreductase, which is beneficial for the development of *in vitro* enzymatic synthesis system in terms of availability and costs.

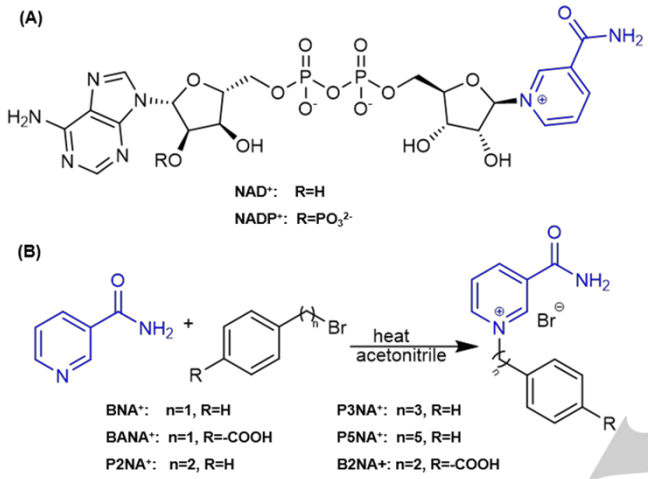
The most common biomimetic cofactors retain only the nicotinamide moiety associated with NAD(P)+/NAD(P)H, and the oxidized forms can be conveniently synthesized by nucleophilic substitution reactions using nicotinamide and haloalkanes (Scheme 1)[4]. They are regarded as totally synthetic nicotinamide cofactor biomimetics (NCBs)[5]. The nicotinamide moiety maintains the catalytic function, while the aromatic hydrocarbon helps to anchor the cofactor in the cofactor binding domain. In a natural cofactor, adenosine monophosphate (AMP) is used as the anchoring group responsible for the accurate binding to the enzyme by forming a hydrogen bond network and van der Waals interactions[6, 7]. In contrast, the described totally synthetic biomimetic cofactors are generally difficult to be accepted by most natural enzymes, because the simple anchoring groups lead to a lack of interactions that provide orientation, and in turn could not induce structural changes in enzymes[8]. Nevertheless, there are some natural enzymes that have been found to utilize biomimetic cofactors. For example, an array of biomimetic cofactors could be utilized more efficiently by a xenobiotic reductase from *Pseudomonas putida*[9]. Glucose dehydrogenase (*Ss*GDH) from *Sulfolobus solfataricus* was found to possess specific activities with 1-benzyl-3-carbamoylpyridin-1-ium (BNA$^+$), 3-carbamoyl-1-phenethylpyridin-1-ium (P2NA$^+$), and 3-carbamoyl-1-(3-phenylpropyl) pyridin-1-ium (P3NA$^+$)[10]. It is worth noting that *Ss*GDH is the first reported non-flavin enzyme that could use these totally synthetic biomimetic cofactors. The FAD-dependent enzyme *Lp*Nox from *Lactobacillus pentosus* as well as some group A flavoprotein hydroxylases displays activity with BNAH[11]. Although the biocatalytic application of biomimetic cofactors has been demonstrated, the number of successful cases is still relatively small.

Switching cofactor specificity toward biomimetic cofactor by enzyme engineering is a useful strategy[12-14]. The P450-BM3 mutant R966D/W1046S could reduce cytochrome with BNAH while the wild-type had no detected activity[15]. By screening of cofactor-binding pocket, the *Ss*GDH mutant I192T/V306I was obtained with a 10-fold higher activity when using P2NA$^+$ as a cofactor compared with the wild-type enzyme[10]. In a recent study by Zachos et al., a penta-mutant of *Ss*GDH was obtained by semi-rational design and directed evolution, which displayed the activity of 254 U/g toward a novel cofactor 3-carbamoyl-1-(5-phenylpentyl) pyridin-1-ium (P5NA$^+$)[16]. These studies demonstrate the feasibility of protein engineering. However, a lot of time and labor is needed during the mutants screening. Therefore, a combination of experimental and computational protein engineering might offer a more efficient and cost-effective process. Computational engineering could facilitate the creation of large virtual libraries of variants, which could then be evaluated and ranked automatically, thereby reducing the need for experimental screening[17, 18]. Based on the computational design, *Bs*GDH was engineered with a 10^7-fold cofactor specificity switch toward the semi-synthetic cofactor nicotinamide mononucleotide

10.1002/cbic.202300066

(NMN$^+$) over natural NADP$^+$[19]. This study inspired us to apply computational methods in enzyme engineering for the utilization of totally synthetic NCBs.

In this study, *Ss*GDH was engineered to accommodate the carboxyl group-containing totally synthetic BANA$^+$. Mutational hotspots were selected accurately by *in situ* ligand minimization, and mutants with increased catalytic activity and efficiency were obtained by screening only two saturation mutagenesis libraries. Then virtual saturation mutagenesis libraries were constructed in silico, and mutational binding affinity and protein stability were calculated to explore the synergistic effect of the mutation. This experimental and computational protein-engineering approach provides valuable information for rational enzyme engineering.



Scheme 1. Chemical structures of natural nicotinamide cofactors (A). Synthetic route for the synthesis of totally synthetic nicotinamide cofactor biomimetics (B)

Results and Discussion

Identification of key residues of *Ss*GDH

Using BANAH as a cofactor instead of natural NADH in reactions catalyzed by XenA, Falcone et al. observed a 6.5 to10-fold higher catalytic efficiency (k_{cat}/K_m)[20]. This prompted us to construct a recycling system of this biomimetic cofactor for efficient biocatalytic processes. Glucose dehydrogenase from *Sulfolobus solfataricus* (*Ss*GDH) was selected as an ideal candidate for protein engineering due to its ability in utilizing several totally synthetic compounds as cofactors, such as BNA$^+$, P2NA$^+$, and P3NA$^+$[11]. Additionally, the double mutant I192T/V306I with a 10-fold higher activity toward P2NA$^+$ embodied its great potential for evolution. Compared with BNA$^+$, BANA$^+$ has an additional carboxyl moiety at the C14 position, which could be mainly responsible for the extremely low activities with wild-type (WT) *Ss*GDH (0.6 mU/mg) and mutant I192T/V306I (0.4 mU/mg). A docking simulation of BANA$^+$ into the cofactor binding site of WT *Ss*GDH (PDB: 2CDB) was carried out to identify potential hotspots for mutagenesis. However, there was no plausible BANA$^+$-bound conformation generated by either CDOCKER or Libdock or Flexible docking. Based on the reaction mechanism[21], the distance between C1 of the glucose and C4 of the nicotinamide ring needs to be around 3.7 Å for an ideal hydride transfer. Additionally, the ideal distance from C1 of the glucose to catalytic zinc and zinc-coordinated water molecule needs to be around 3.7 Å and 3.4 Å, respectively (Figure 1A). In the study of Nowak et al., the highest catalytic efficiency of WT *Ss*GDH was found with P2NA$^+$, which was 2.5-fold higher than with BNA$^+$. Here, docking results of P2NA$^+$ complexed with *Ss*GDH revealed that 11 of 50 generated conformations were considered to in line with the generally accepted mechanism (Figure S1A). For all ligand poses, CDOCKER scores were reported as negative value (-CDOCKER energy), where a higher value indicates a more favorable binding (Table S1). Upon alignment of the optimal docking conformation and the *Ss*GDH-glucose-NADP$^+$ complex crystal (PDB: 2CDB), the nicotinamide ring of P2NA$^+$ and NADP$^+$ are nearly overlapping (Figure S1B). This might partially explain the moderate activity (5.8 mU/mg) detected with P2NA$^+$ in wet experiments. Thus, we manually modified the ligand to BANA$^+$ based on the optimal docking conformation of P2NA$^+$ (the pose with the highest CDOCKER score), followed by *in situ* ligand minimization, in which the original NADP$^+$-binding site residues were allowed to move freely. Figure 1B&1C show that both E44 and E114 exhibit repulsive interactions with the negatively charged carboxylate moieties in the optimized conformation. It is rationally presumed that catalytic activity could be improved by eliminating these unfavorable interactions. Therefore, E44 and E114 were selected for site mutagenesis.

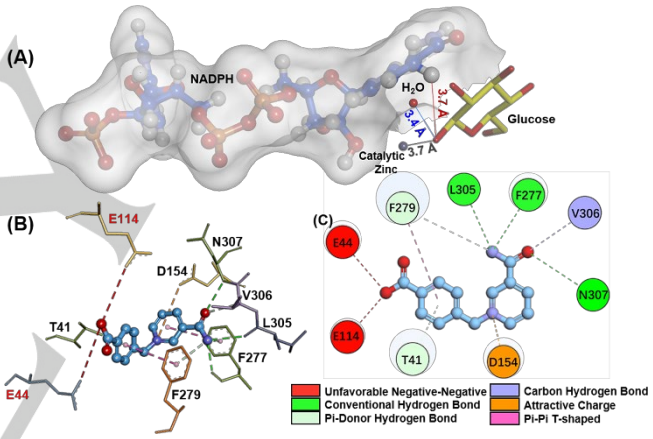


Figure 1. (A) The locations of cofactor NADP$^+$ and glucose at the catalytic state (PBD: 2CDB) in the crystal structure of *Ss*GDH. The red line represents the distance between C1 of the glucose and C4 of the nicotinamide ring. The blue line represents the distance between C1 of the glucose and the O atom of zinc-coordinated water. The gray line represents the distance between C1 of the glucose and the catalytic zinc. (B) Pose of BANA$^+$ in WT *Ss*GDH after *in situ* ligand minimization. (C) Interactions between BANA$^+$ and WT *Ss*GDH.

Screening of saturation mutagenesis libraries of E44 and E114

Saturation mutagenesis libraries of E44 and E114 were constructed and screened. Compared with WT, E44N, E44D, E44Q, E44W, E114K and E114D exhibited increased activity when BANA$^+$ was used as the cofactor (Figure 2A). It is noteworthy that the highest specific activity was observed when a negatively charged glutamic acid was exchanged for glutamine on site 44, which is 3-fold over the WT. In contrast, E114Q did not exhibit enhanced activity, whereas a basic lysine substitution resulted in 1.6-fold activity than that of WT (Table S2). Notably, other basic amino acids substitution on 44 and 114 did not show advantageous effect on activity. The activity of E44W was improved by 1.57 times over WT, this might be attributed to the steric hindrance change. The substitution of aspartic acid possessed positive effects on both sites, as evidenced by a 2.5-fold increase in E44D and a 1.2-fold increase in E114D. This result is somewhat counterintuitive, as both glutamic acid and

aspartic acid are acidic amino acids that have a negatively charged side chain.

Interaction analysis based on *in situ* ligand minimization

To get better insight into the structure-activity relationship, computational analysis of interactions between BANA$^+$ and above single mutants were conducted. Among simulation algorithms, molecular docking is one of the most common procedures for predicting the structural interactions of the protein-ligand complex. Since both the carbonyl group and carboxyl group of BANA$^+$ can function as hydrogen bond acceptors (Figure 2B), a large number of irrational docking conformations that do not follow the catalytic reaction mechanism were generated. Thus, instead of molecular docking, computational models of mutants were constructed based on the WT-BANA$^+$-glucose complex structure followed by *in situ* ligand minimization program. To improve the accuracy of the models, BANA$^+$ was selected to define a sphere inside which atoms are free to move in the receptor, and residues within 6 Å around BANA$^+$ were allowed to move. Compared with semi-flexible docking and flexible simulations, this procedure could increase the accuracy of the predicted protein-ligand complex while reducing calculation time and promoting analytical efficiency.

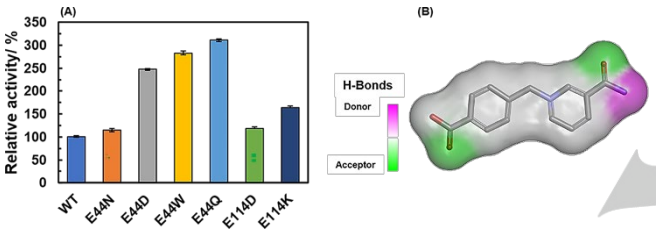


Figure 2. (A) Relative activity of *Ss*GDH single mutants. The specific activity of WT *Ss*GDH is 0.6 mU/mg. Activity measurements were performed in 100 mM HEPES pH 7.5 containing purified enzyme, 5 mM D-glucose, and 25 mM BANA$^+$, and the reaction was monitored at 358 nm. (B) Solvent hydrogen bond donor/acceptor surface of BANA$^+$.

Simulations were performed on the complex structures of E44Q and E114Q. As shown in figure 3A, the two glutamine-substituted mutants have almost the same nicotinamide moiety location, while the benzoic acid group is positioned differently. The nicotinamide moiety, which is responsible for the catalytic function, is mainly anchored by hydrophobic interactions and hydrogen bonds with I92, F277, L305, and N307 in both mutants. In E114Q, no interaction is formed between the benzoic acid group of BANA$^+$and its surrounding residues. Consequently, a stable catalytic conformation could not be formed due to the flexibility of BANA$^+$ in the active center. In contrast, hydrogen bonds can be formed between the benzoic acid group of BANA$^+$ and Q44 as well as G40 in mutant E44Q. Also, the *pi-pi* interaction formed with F279 is crucial for the stabilization of BANA$^+$. This may explain why the same glutamine substitution at positions 44 and 114 resulted in different catalytic activities. According to the interaction analysis between E44W and BANA$^+$ (Figure 3B), tryptophan forms a new *pi*-anion interaction with benzoic acid in addition to the previously predicted steric hindrance effect, suggesting this substitution favors the binding of BANA$^+$ at the cofactor pocket.

The unexpected increase in activity exhibited by E44D and E114D is of interest because they are both acidic amino acids. In E44D, a slightly shorter side chain after the aspartate substitution may relieve its unfavorable interaction with the carboxyl group of BANA$^+$. Although glutamate at 114 is present in both WT and E44D, it shows a difference in the ~~force of~~ interaction, probably due to the flexibility of the atoms and amino acid residues in the "*in situ* ligand minimization" simulation (Figure S2). Additionally, I117 and F279 also contributed to BANA+ stabilization through *pi*-alkyl and *pi-pi* interactions in E44D (Figure 3C). In E114D, the unreasonable interactions were relieved by a mutation in a similar manner. Additionally, only F279 exhibited a *pi-pi* interaction with benzoic acid, whereas no interaction between carboxyl group of the benzoic acid group and nearby residues was formed. It may explain why the activity of E114D is not improved significantly (Figure 3D).

The interaction analysis for Lys, His, and Arg substitutions at site 114 is shown in figures 3E, 3F, and 3G. The activity of E114K was the only one that improved, even though they are all basic amino acids. The substituted lysine in E114K formed a salt bridge, and a similar interaction was observed in E114R as well. In E114K, however, the benzoic acid ring of BANA$^+$ could be stabilized through *pi-pi* stacking in E114K, but not in E114R and E114H. As described in the method of "*in situ* ligand minimization", this algorithm allows flexible movement of atoms and residues in the coenzyme binding pocket, which leads to changes in the overall interaction network. In the E114K, the phenyl ring of F279 and the benzoic acid moiety of BANA$^+$ form a pi-pi interaction. However, in the E114H and E114R, a subtle change in the position of the surrounding amino acid residues leads to the formation of a hydrogen bond between amino group of F279 and the carboxyl group of F277. This hydrogen bond prevents the formation of the pi-pi interaction between F279 and BANA$^+$, which is observed in the E114K mutant. Therefore, BANA+ cannot be anchored as firmly in E114H and E114R as it is in E114K due to the absence of this important interaction. In light of these interaction analysis, all mutants with increased activity exhibited attenuated negative-negative interactions in WT. By using an *in situ* ligand minimization algorithm, the nicotinamide ring is fitted into the appropriate position of the receptor for the catalysis. Importantly, the *pi-pi*/*pi*-alkyl interactions between Phe279 and the benzoic acid ring, as well as the hydrogen bond and salt bridge formed with the carboxyl group of BANA$^+$, are crucial for the localization of BANA$^+$.

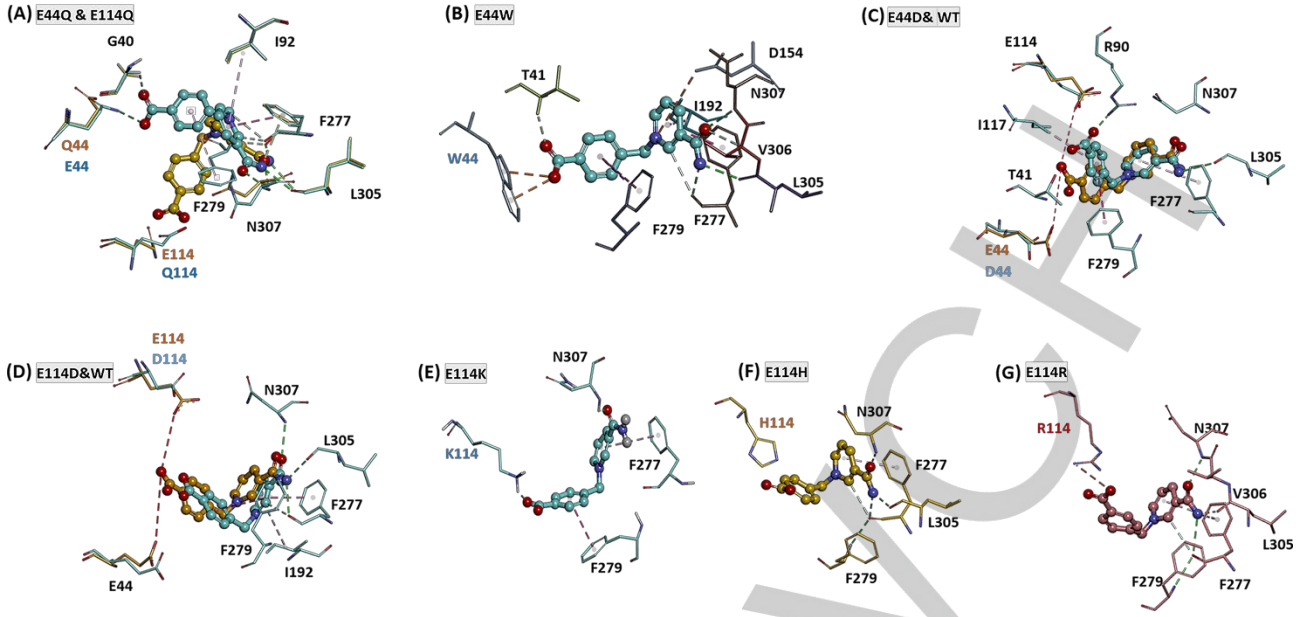


Figure 3. Interaction analysis of BANA$^+$ and *Ss*GDH mutants (A) Structure alignment of E44Q & E114Q. (B) Poses of BANA$^+$ in E44W. (C) Structure alignment of E44D & WT. (D) Structure alignment of E114D & WT. (E) Poses of BANA$^+$ in E114K. (F) Poses of BANA$^+$ in E114H. (G) Poses of BANA$^+$ in E114R. BANA$^+$ is shown as ball-and-stick model and residues in the interaction network are shown as sticks. Green dashed lines represent hydrogen bonds, and the distance between the donor and acceptor atoms is 3.4 Å by default. Pink dashed lines represent pi-pi/pi-alkyl interaction, and the distance between a cation and the centroid of a Pi ring is less than 6.0 Å by default. Purple dashed line in (B) represents pi-anion interaction, and the distance between a cation and the centroid of a Pi ring is less than 5.0 Å by default. Red dashed line in(C) and (D) represents the unfavorable negative-negative interactions in WT *Ss*GDH.

Analysis of the synergistic effect of site E44 and E114

To further improve the BANA$^+$ availability for *Ss*GDH, six combinatorial mutants were constructed, however, none of them exhibited improved catalytic activity than the single mutants (Table S2). It appears that simple combination of beneficial single mutations is not an effective approach for evolving utilization of BANA$^+$. Therefore, to probe the synergistic effect of sites E44 and E114, both virtual and experimental saturation mutagenesis libraries of site 114, by using E44D, E44W, and E44Q as templates, were constructed and screened. For each virtual library, the effect of mutations on the binding affinity and protein stability were calculated. Mutations show values below 0.5 kcal/mol for stabilizing, between -0.5 and 0.5 kcal/mol for neutral, and over 0.5 kcal/mol for destabilizing[22-25][21-24]. The mutation energy of binding (MEB) was used to assess the binding affinity of BANA$^+$ in the protein complex, and the mutation energy of stability (MES) was used to evaluate the difference of the free energy of folding between the structure of mutants and their corresponding templates.

Based on calculations, no double mutant containing E44W exhibited MEB value less than 0.5 kcal/mol, whereas nearly half of these mutants exhibited decreased binding affinity between BANA$^+$ and protein (Figure 4A). In the library with E44Q as template, MEB values of -0.5–0.9 kcal/mol indicate that the mutation has little effect on affinity for most mutants. Only E44Q/E114D showed an increased affinity (-0.52 kcal/mol), however, the increased MES value of E44Q/E114D (0.72 kcal/mol) implied a decrease in the stability of protein structure (Figure 4B). In contrast, over half of mutants showed decreased MES values that were considered to be "stabilizing" for library based on E44D. According to the predicted MEB values, it is envisaged that a synergistic effect could achieved when E44D was used as template (Figure 4C), and it was echoed by wet experiments. In fact, no double mutants with increased activity were obtained when either E44W or E44Q using as template, whereas all five double mutants with increased activity were found in the library based on E44D. Among them, E44D/E114L shows the highest catalytic activity, which is 7.67-fold over WT, and a 3–6-fold increase in activity was also observed in the other three mutants of E44D/E114R, E44D/E114S, E44D/E114T, and E44D/E114V (Figure 4D).

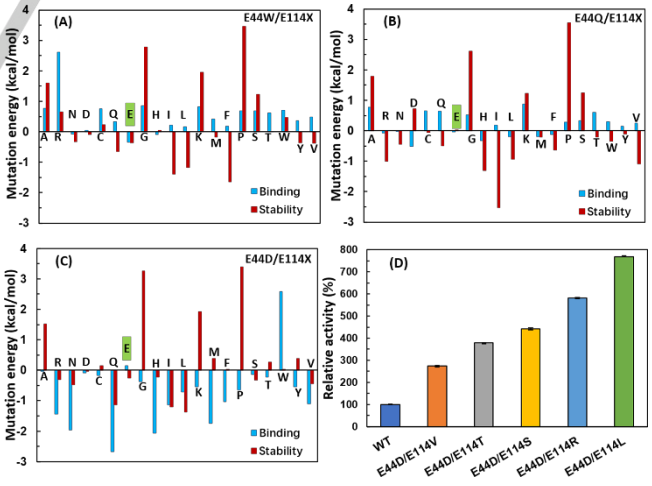


Figure 4. Computational mutation energies of saturation mutagenesis libraries of E114 using (A)E44D, (B) E44W, (C)E44Q as templates. (D) Relative activity of several *Ss*GDH double mutants.

Unlike MEB values, a similar trend in MES change was observed in all three virtual libraries. Calculations show that a significant decrease in protein stability could be caused by alanine, glycine, lysine, or proline substitutions at site 114 whereas leucine and isoleucine substitutions have positive contributions to the protein stability. Taken together, there is a certain correlation

between the simulated mutation energies and wet experimental results, although not completely. For example, both MEB and MES values of five double mutants in Figure 4D imply a trend of stabilization. Glutamine, asparagine, isoleucine and leucine substitutions at E114 also exhibit a favorable reduction in both MEB and MES, in actual assays however, these mutants did not show higher activity than the template protein. Consequently, the mutation energy could provide some guidance, whereas wet experimental screening is still necessary for the engineering of *Ss*GDH toward BANA$^+$ utilization.

Kinetic parameters of SsGDH and its variants

Apparent kinetic parameters of WT *Ss*GDH and six variants toward BANA$^+$ were determined. For comparison, binding energies were calculated based on the optimized structures after *in situ* ligand minimization (Table 1). A quite low affinity of WT toward BANA$^+$ is indicated by a high K_m value (19.16 mM) and extremely high binding energy (21.43 kcal/mol). The K_m value of E44D is half of the WT, suggesting that the removal of negative-negative interactions could improve enzyme-cofactor binding. This result is also consistent with the interaction analysis in Figure 3C, and the significantly decreased binding energy (-90.43 kcal/mol). All double mutants except E44D/E114S, have lower K_m values than that of E44D. The lowest K_m and binding energy were observed in E44D/E114R. The best mutant E44D/E114L did not show obvious advantage of cofactor affinity among all double mutants, while displays the highest v_{max} value. Regarding catalytic efficiency, all the double mutants demonstrate 4–11-fold improvements in k_{cat}/K_m values.

Table 1 Kinetic parameters of *Ss*GDH and its variants toward BANA$^+$.

	K_m [mM]	v_{max} [μmol·min^{-1}·mg^{-1}·10^{-3}]	k_{cat} [s^{-1}]	k_{cat}/K_m [s^{-1}·mM]	Binding energy [kcal/mol]
WT	19.16	1.37	0.959	0.050	21.43
E44D	11.06	1.84	1.288	0.116	-90.43
E44D/E114V	9.73	2.83	1.981	0.204	-123.69
E44D/E114T	9.46	4.19	2.933	0.310	-122.65
E44D/E114S	12.15	4.33	3.031	0.249	-120.35
E44D/E114R	8.51	5.19	3.633	0.427	-214.00
E44D/E114L	9.08	6.87	4.809	0.530	-123.09

Each data was collected in triplicate and averaged.

Cofactor spectrum of *Ss*GDH and its variants

As shown in Figure 5, four other biomimetic cofactors were synthesized and interrogated. Among them, BNA$^+$, P2NA$^+$, and P3NA$^+$, have elongated alkyl chains between nicotinamide and benzene rings, while without carboxyl substitution. In terms of BNA$^+$, only 44D/114R and 44D/114L are superior to WT, while 44D, 44D/114V, 44D/114T, and 44D/114S showed lower activity. When tested with P2NA$^+$, 44D/114S, 44D/114R, and 44D/114L exhibited 1.5–2 fold increased catalytic activity, while other mutants showed similar catalytic activity to WT. As the carbon chain was lengthened, 44D/114L showed a significant advantage with about 3-fold increase in catalytic activity in experiments with P3NA$^+$, whereas all the other double mutants were only 2-fold higher. Based on their chemical structures, B2NA$^+$ differs from P2NA$^+$ only by a carboxyl group. 44D/114T and 44D/114L showed a more significant increase in the catalytic activity of B2NA$^+$ that containing carboxyl groups. In summary, double mutants showed 1.1–3.3-fold improvements in activity when tested with these four biomimetic cofactors, However, the results are not comparable with those achieved using BANA$^+$ as a cofactor. It is therefore not surpris297ing that these double mutations are initially designed for BANA$^+$, and these differences further confirm that the selection of sites 44 and 114 are well targeted.

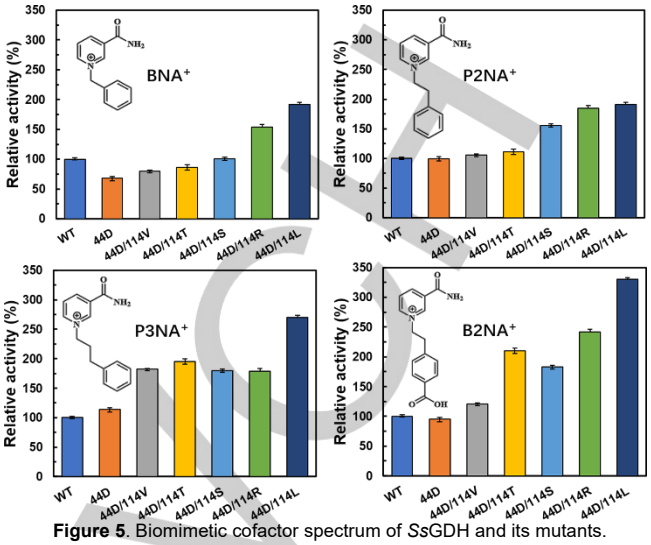


Figure 5. Biomimetic cofactor spectrum of *Ss*GDH and its mutants.

Conclusion

In this study, a semi-rational design method for the improvement of *Ss*GDH activity toward BANA$^+$ was developed. Residues E44 and E114, which exhibit unfavorable negative-negative interactions with the carboxyl moiety of BANA$^+$, were selected as hotspots based on *in situ* ligand minimization. Subsequently, single mutants with improved catalytic activity were identified from saturation mutagenesis libraries. Interaction analysis reveals the importance of stabilization of carboxyl group for the binding and localization of BANA$^+$. To further explore the synergistic effect between the two sites, saturation mutagenesis of E114 was carried out and the energy effect of each mutant were calculated. Wet experiments confirmed that a synergistic effect could achieved only when E44D as template, and the optimal mutant E44D/E114L shows a 7.67- fold increase in activity over WT. Since mutations at sites 44 and 114 are designed specifically for BANA$^+$, the activity improvements are not as significant with other four biomimetic cofactors. These results can provide valuable information for the rational design of nicotinamide cofactor-dependent oxidoreductases as well as biomimetic cofactors.

Experimental Section

Reagents

All reagents used in the NCBs synthesis were purchased from Macklin Inc. (Macklin, Shanghai, China). All other reagents were purchased from Sinopharm Chemical Reagent (Sinopharm, Wuhan, China). The sequence of *Ss*GDH was commercially synthesized (Yixin, Wuxi, China). All plasmids were from our laboratory and expressed in *Escherichia. coli* BL21 (DE3).

Synthesis of totally synthetic biomimetic cofactors

BNA$^+$, P2NA$^+$, P3NA$^+$ were synthesized according to Nowak et al.[9],BANA$^+$ and B2NA$^+$ were synthesized according to Falcone et al.[19]. The NMR spectra were shown in Figure S4-S8.

Synthesis of1-benzyl-3-carbamoylpyridin-1-ium (BNA$^+$ Br$^-$)

Nicotinamide (40 mmol, 4.88 g) was dissolved in 40 mL acetonitrile and heated to reflux. Subsequently, benzyl bromide (40 mmol, 4.8 mL) was added, and the reaction mixture was stirred under reflux for 15 h. Progress of the reaction was monitored by TLC. After cooling to room temperature, 50 mL of diethyl ether was added, and the precipitate was filtered and

washed three times with diethyl ether. Product was obtained as a white solid (7.17 g) after dried with yield of 61 %.
^{1}H NMR (600 MHz, Deuterium Oxide) δ 9.40 (t, *J* = 1.7 Hz, 1H), 9.11 (dt, *J* = 6.2, 1.4 Hz, 1H), 8.94 (dt, *J* = 8.1, 1.5 Hz, 1H), 8.23 (dd, *J* = 8.1, 6.2 Hz, 1H), 7.54 (d, *J* = 1.3 Hz, 5H), 5.94 (s, 2H).
^{13}C NMR (151 MHz, Deuterium Oxide) δ 165.65, 146.50, 144.31, 134.02, 132.19, 130.14, 129.64, 129.25, 128.58, 65.12.

Synthesis of 3-carbamoyl-1-phenethylpyridin-1-ium (P2NA$^+$ Br$^-$)

Nicotinamide (20 mmol, 2.44 g) was dissolved in 40 mL acetonitrile and heated to reflux. Subsequently, (2-Bromoethyl) benzene (40 mmol, 5.45 mL) was added, and the reaction mixture was stirred under reflux (130 °C) for 20 h. Progress of the reaction was monitored by TLC. After cooling to room temperature, 50 mL of diethyl ether was added, and the precipitate was filtered and washed three times with diethyl ether. Product was obtained as a white solid (4.51 g) after dried with yield of 61 %.
^{1}H NMR (400 MHz, DMSO-d_6) δ 9.64 (t, *J* = 1.5 Hz, 1H), 9.25 (dt, *J* = 6.2, 1.3 Hz, 1H), 8.98 (dt, *J* = 8.1, 1.4 Hz, 1H), 8.67 (s, 1H), 8.25 (dd, *J* = 8.1, 6.0 Hz, 1H), 8.20 (s, 1H), 7.30 (d, *J* = 4.4 Hz, 4H), 7.29 – 7.19 (m, 1H), 4.96 (dd, *J* = 8.6, 6.8 Hz, 2H), 3.34 (t, *J* = 7.6 Hz, 2H).
^{13}C NMR (101 MHz, DMSO-d_6) δ 163.22, 146.91, 145.32, 144.11, 136.61, 134.02, 129.49, 129.08, 128.14, 127.55, 62.19, 36.81.

Synthesis of 3-carbamoyl-1-(3-phenylpropyl) pyridin-1-ium (P3NA$^+$ Br-)

Nicotinamide (20 mmol, 2.44 g) was dissolved in 40 mL acetonitrile and heated to reflux. Subsequently, 1-Bromo-3-phenylpropane (40 mmol, 5.45 mL) was added, and the reaction mixture was stirred under reflux (100 °C) for 36 h. Progress of the reaction was monitored by TLC. After cooling to room temperature, 50 mL of diethyl ether was added, and the precipitate was filtered and washed three times with diethyl ether. Product was obtained as a white solid (5.37 g) after dried with yield of 84 %.
^{1}H NMR (400 MHz, DMSO-d_6) δ 9.62 (d, *J* = 1.6 Hz, 1H), 9.32 (dt, *J* = 6.1, 1.3 Hz, 1H), 8.97 (dt, *J* = 8.0, 1.5 Hz, 1H), 8.65 (s, 1H), 8.26 (dd, *J* = 8.1, 6.0 Hz, 1H), 8.19 (s, 1H), 7.31 – 7.12 (m, 5H), 4.77 (t, *J* = 7.4 Hz, 2H), 2.68 (dd, *J* = 9.4, 6.4 Hz, 2H), 2.37 – 2.25 (m, 2H).
^{13}C NMR (101 MHz, DMSO-d_6) δ 163.27, 146.94, 145.26, 143.97, 140.81, 134.11, 128.85, 128.70, 128.30, 126.56, 61.44, 32.43, 32.08.

Synthesis of 3-carbamoyl-1-(3-phenylpropyl) pyridin-1-ium (BANA$^+$ Br-)

Nicotinamide (15 mmol, 1.83 g) was dissolved in 75 mL mix solution consisting of 1,4-Dioxane and methanol (V: V=2:1). Subsequently, 4-(bromomethyl) benzoic acid (7 mmol) was added, and the reaction mixture was stirred under reflux (50 °C) for 8 h. Progress of the reaction was monitored by TLC. After cooling to room temperature, 50 mL of dichloromethane was added, and the precipitate was filtered and washed three times with diethyl ether. Product was obtained as a white solid (3.70 g) after dried with yield of 78 %.
^{1}H NMR (400 MHz, DMSO-d_6) δ 13.14 (s, 1H), 9.76 (d, *J* = 1.6 Hz, 1H), 9.41 (dt, *J* = 6.2, 1.3 Hz, 1H), 9.05 (dt, *J* = 8.2, 1.4 Hz, 1H), 8.68 (s, 1H), 8.33 (dd, *J* = 8.1, 6.1 Hz, 1H), 8.22 (s, 1H), 8.01 – 7.95 (m, 2H), 7.74 – 7.67 (m, 2H), 6.10 (s, 2H).
^{13}C NMR (101 MHz, DMSO-d_6) δ 167.22, 163.16, 147.08, 145.54, 144.56, 138.94, 134.54, 132.01, 130.44, 129.62, 128.84, 63.29.

Synthesis of 3-carbamoyl-1-(4-carboxyphenethyl)pyridin-1-ium (B2NA$^+$ Br-)

Nicotinamide (15 mmol, 1.83 g) was dissolved in 75 mL mix solution consisting of 1,4-Dioxane and methanol (V: V=2:1). Subsequently, 4-(bromomethyl) benzoic acid (7 mmol) was added, and the reaction mixture was stirred under reflux (150 °C) for 16 h. Progress of the reaction was monitored by TLC. After cooling to room temperature, 50 mL of dichloromethane was added, and the precipitate was filtered and washed three times with diethyl ether. Product was obtained as a white solid (1.66 g) after dried with yield of 64 %.
^{1}H NMR (400 MHz, DMSO-d_6) δ 12.93 (s, 1H), 9.56 (d, *J* = 1.9 Hz, 1H), 9.15 (dd, *J* = 6.1, 1.3 Hz, 1H), 8.95 (dt, *J* = 8.2, 1.5 Hz, 1H), 8.58 (s, 1H), 8.25 (dd, *J* = 8.1, 6.0 Hz, 1H), 8.17 (s, 1H), 7.92 – 7.85 (m, 2H), 7.41 (d, *J* = 8.2 Hz, 2H), 4.96 (t, *J* = 7.6 Hz, 2H), 3.40 (t, *J* = 7.6 Hz, 2H).
^{13}C NMR (101 MHz, DMSO-d_6) δ 167.54, 163.24, 146.87, 145.46, 144.02, 141.80, 134.26, 130.10, 130.07, 129.72, 128.18, 61.84, 36.70.

Site directed mutagenesis library screening

Saturation mutagenesis libraries of E44 and E114 were constructed using the whole-plasmid protocol, The PCR mixture (20 μL) contained plasmid (template, 10 ng), KOD Plus- Neos (0.4 U), Mg2SO4 (30 ng), dNTPs (4 ng), KOD buffer and primers. The PCR procedures were as follows: 95°C for 2 min, followed by 25 cycles of 95°C for 20 s, 55°C for 20 s, 68°C for 3 min and 40 s, then 68°C for 10 min. The PCR product (2 μL) was digested with Dpn I at 37°C for 2 h. Then the digested plasmids were transformed into *E. coli* BL21(DE3) host cells.

Library screening

Recombinant plasmids were constructed through whole-plasmid PCR using NNK codon substitution at target sites of 44 and 114. For each library, 84 colonies were screened using two 48-well plates, each with 42 mutant colonies and 6 WT colonies as controls. All the colonies were inoculated into 300 μL of LB medium containing 50 μg/mL kanamycin and grown at 37 °C for 12 h. Then 100 μL of cell culture was transferred into 1 mL of fresh LB medium containing 50 μg/mL kanamycin and incubated for another 2 h. Enzyme expression was induced by addition of IPTG to a final concentration of 0.2 mM. After 14 h of incubation at 16 °C, cells were harvested by centrifugation, then lysed by addition of 1000 U of lysozyme and shaking at 37 °C for 1 h. The plates were centrifuged at 4000 × g and 4 °C for 15 min, and 80 μL of supernatant was transferred to a 96-well plate. The activity was determined by adding 120 μL of enzymatic assay solution comprised of 5 mM D-glucose and 25 mM BANA$^+$, 100 mM HEPES buffer (pH 7.4), and the reaction was monitored at 358 nm (Biotek, USA). Extinction coefficients of BANAH (ε=2.33 L·mmol^{-1}·cm^{-1}) was referenced to the studies of Falcone et al.

Protein expression and purification

An Erlenmeyer flask containing Luria-Bertani (LB) added with 50 μg·mL-1 kanamycin (kan) and a colony from E. coli harboring WT and variants of SsGDH, then cultivated at 37°C and 180 rpm. When the optical density at 600 nm (OD600) reached 0.8, isopropyl-β-Dthiogalactopyranoside added to a final concentration of 0.2 mM, and the temperature decreased to 16°C for protein overexpression. After collecting cells by centrifugation and performing ultrasonication in HEPES buffer (pH 7.4). Then cell lysates were centrifuged at 8000 rpm for 30 min and the supernatant of SsGDH was purified by Ni-NTA Purose 6 Fast Flow column (Qianchun Bio, Jiaxing, China)). After washing with binding buffer containing 20 mM imidazole, the target protein was eluted with 100 mM HEPES containing 500 mM imidazole pH 7.4. The eluted protein was desalted using a 30-kD ultrafiltration tube (Millipore, USA).

Activity assay

The activity of WT SsGDH and its variants toward biomimetic cofactors was determined by monitoring the increase in the absorbance of reduced biomimetics at 360 nm at 45°C. The assay mixture (200 μL) consisted of 5 mM D-glucose and 25 mM oxidized biomimetic cofactors and 30 μL enzyme solution. One unit of enzyme activity was defined as the amount of enzyme required for catalyzing the reduction of 1 μmol of oxidized biomimetic cofactors per minute. All experiments were conducted in triplicate. Extinction coefficients were referenced to the studies of Falcone et al. and Nowak et al.: BANAH, ε=2.33 L·mmol^{-1}·cm^{-1}; BNAH, ε=7.4 L·mmol^{-1}·cm^{-1}; P2NAH, ε=3.9 L·mmol^{-1}·cm^{-1}; and P3NAH, ε= 8.3 L·mmol^{-1}·cm^{-1}.

Kinetic parameters

Kinetic measurements of WT *Ss*GDH and variants were performed in 100 mM HEPES buffer (pH 7.4) with 5 mM D-glucose and varying cofactor concentrations at 45 °C for 60 min. Kinetics were calculated by fitting enzyme kinetic curves using the Origin Pro software (Version 8.5). All measurements were performed in triplicate.

Molecular docking

The crystal structure of WT *Ss*GDH was docked with P2NA$^+$ using CDOCKER module of Discovery studio 2021 by applying a CHARMm force field. Prior to docking, SsGDH and P2NA$^+$ were prepared using the prepare protein and prepare ligand protocols, respectively. The active pocket was defined as volume based on NADPH in the *Ss*GDH-glucose-NADP$^+$ complex crystal (PDB: 2CDB). The number of hits ligand was set as 100, the number of starting random conformations generated from equilibration and minimization of the starting ligand structure was set as 10, Maximum bad orientations was set as 800, and orientation van der Waals energy threshold was set as 300. The generated docking conformations were evaluated with the primary requisite of being consistent with the catalytic mechanism, i.e., the distance between C1 of the glucose and C4 of the nicotinamide ring needs to be around 3.7 Å, while the distance from C1 of the glucose to catalytic zinc and zinc-coordinated water molecule needs to be around 3.7 Å and 3.4 Å, respectively. Based on above settings, the pose with the highest CDOCKER score is selected as the template for the manually modification from P2NA$^+$ to BANA$^+$.

***in situ* ligand minimization**

Ligand minimization was performed using the *in situ* ligand minimization tool based on the complex of *Ss*GDH-P2NA$^+$ obtained by molecular docking. By using the "Finding sites as volume of selected ligand" tool (Discovery studio 2021), BANA$^+$ was selected to define a sphere inside which atoms are free to move in the receptor. The coordinates of the sphere center are (23.5291, 16.3031, 12.135), and radius of the sphere is

8.7 Å. Residues within 6 Å around BANA$^+$ were allowed to move, including C39, G40, T41, E44, I45, L50, H66, E114, I117, Q150, P151, D154, I155, K157, S158, I192, L276, F277, G278, F279, L305, V306, N307 and G308. CHARMm was chosen for atom typing, the acceptor hydrogen atom near the ligand is allowed to be flexible, and substructure constraints was set as false. Smart minimizer algorithm was chosen for performing the minimization and minimization steps was set as 1000. The RMS gradient, indicating the tolerance (kcal/(mol-Å)) that should be applied to the average gradient over a minimization period, was set to 0.001. Minimization Energy Change, representing a tolerance that should be applied to the change in total energy during a cycle of minimization, was set to 0. After calculation, the simulation results were converged to a single structural output for force analysis. For *Ss*GDH mutants, before performing *in situ* ligand minimization, model of each mutant was constructed using "Built Mutants" protocol. The model with the lowest probability density function (PDF) Total Energy was selected for the following in situ ligand minimization and parameters were set the same as those of WT SsGDH.

Calculation of mutation energy of binding (MEB) and stability (MES)
After *in situ* minimization, converged single poses of E44W-BANA$^+$, E44Q-BANA$^+$ and E44D-BANA$^+$ were used as the input for mutation energy calculation, respectively. Site 114 was mutated into 20 standard amino acids, and the temperature of the solvent was set as 45°C according to wet experiments. For mutation energy of binding (MEB), the differences in the free energy of binding between the WT and mutant structures is calculated. For mutation energy of stability (MES), the difference between the folding free energy of WT and mutant structures is calculated. The values of mutation energy (both MEB and MES) were calculated as a sum of scaled van der Waals, electrostatic, non-polar, and entropy terms.

Acknowledgements

This work was supported by the National Key R&D Program (2021YFC2102700), the National Natural Science Foundation of China (22077054, 21907040), and the National First-Class Discipline Program of Light Industry Technology and Engineering (LITE2018-07).

Keywords: biomimetic cofactors, glucose dehydrogenase, semi-rational design, virtual screening

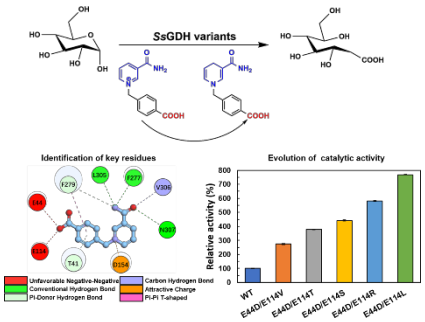
[1] L. S. Vidal, C. L. Kelly, P. M. Mordaka, J. T. Heap. *Biochimica. Et Biophysica. Acta. Proteins and Proteomics*. **2018**,*1866*, 327-347.
[2] A. A. Wu, Y. J. Bai, T. P. Fan, X. H. Zheng. *Syst. Microbiol. and Biomanuf.* **2022**, 2, 473-486.
[3] D. Hofmann, A. Wirtz, B. Santiago-Schübel, U. Disko, M. P. Hofmann, *Anal. Bioanal. Chem*. **2010**, *398*, 2803-2811.
[4] P. Karrer, F. J. Stare, *Helv. Chim. Acta*. **1937**, *20*, 418-423.
[5] I. Zachos, C. Nowak, V. Sieber, *Curr. Opin. Chem. Biol*. **2019**, *49*, 59-66.
[6] S. A. Lçw, I. M. Lçw, M. J. Weissenborn, B. Hauer. *ChemCatChem*, **2016**, *8*, 911-915.
[7] J. R. Sunderland, X. Tao, E. E. Butrick, L. C. Keilich, C. E. Villa, J. R. Miecznikowski, S. S. Jain, *Polyhedron*, **2016**, *114*, 145-151.
[8] E. King, S. Maxel, H. Li, *Curr. Opin. Biotechnol*. 2020,66:217-226.
[9] A. Geddes, C. E. Paul, S. Hay, F. Hollmann, N. S. Scrutton, *J. Am. Chem. Soc.* **2016**, *138*, 11089-11092.
[10]C. Nowak, A. Pick, P. Lommes, V. Sieber, *ACS Catal*. 2017,7(8):5202-5208.
[11] C. Nowak, A. Pick, L. I. Csepei, V. Sieber. *ChemBioChem*. **2017**, *18*, 1944-1949.
[12] C. You, R. Huang, X. L. Wei, Z. G. Zhu, Y. H. P. Zhang, *Synth. Syst. Biotechnol.* **2017**, *2*, 208-218.
[13] Y. Liu, Q. Li, L. Wang, X. J. Guo, J. T. Wang, *ChemBioChem*. **2021**, *21*, 1972-1975.
[14] M. Basle, H. Padley, F. L. Martins, G. S. Winkler, C. M. Jäger, A. Pordea, *J. Inorg. Biochem*. **2021**, *220*, 111446.
[15] H. C. Lo, J. D. Ryan, J. B. Kerr, D. S. Clark, R. H. Fish, *J. Organomet. Chem.* **2017**, *839*, 38-52.
[16] I. Zachos, S. Guner, A. Essert, P. Lommesa, V. Sieber, *Chem. Commun. (Camb)*. **2022**, *58*, 11945-11948.
[17] C. E. Sequeiros-Borja, B. Surpeta, J. Brezovsky, *Brief Bioinform*. **2021**, *22*, 1-15.
[18] D. N. Woolfson, *J Mol Biol*. **2021**, *433*, 167160-167178.
[19] W. B. Black, L. Zhang, W. Shun, S. Maxel, Y. Cui, E. Edward, B. Fong, A. S. Martinez, J. B. Siegel, H. Li, *Nat. Chem. Biol.* **2020**, *16*, 87-94.
[20] N. Falcone, Z. She, J. Syed, A. Lough, H. B. Kraatz, *ChemBioChem*. **2019**, *20*, 838-845.
[21] C. C. Milburn, H. J. Lamble, A. Theodossis, S. D. Bull, D. W. Hough, M. J. Danson, G. L. Taylor. *Biol. Chem.* **2006**, *281*, 14796-14804.
[22] V. Z. Spassov, L. Yan, S. Szalma, *J. Phys. Chem. B.* **2002**,*106*, 8726-8738.
[23] V. Z. Spassov, L. Yan, J, *Comput. Chem.* **2016**, *37*, 2573-2587.
[24] V. Z. Spassov, L. Yan, *Protein Sci.* **2008**, *17*, 1955-1970.
[25] V. Z. Spassov, L. Yan, *Proteins.* **2013**, *81*, 704-714.

10.1002/cbic.202300066

WILEY-VCH

RESEARCH ARTICLE

Entry for the Table of Contents



Glucose dehydrogenase from *Sulfolobus solfataricus* (*Ss*GDH) was engineered to favor a synthesized unnatural cofactor 3-carbamoyl-1-(4-carboxybenzyl) pyridin-1-ium (BANA$^+$). Mutation hotspots E44 and E114 were identified by *in situ* ligand minimization program. Double mutant E44D/E114L with 7.7-fold increased activity and 10.6-fold increased catalytic efficiency toward BANA$^+$ was obtained.