ChemBioChem

Research Article
doi.org/10.1002/cbic.202300390

Chemistry Europe
European Chemical Societies Publishing

www.chembiochem.org

Special Collection

Engineering a Formate Dehydrogenase for NADPH Regeneration**

Wei Ma,[a] Qiang Geng,[a] Cheng Chen,[a] Yu-Cong Zheng,[a] Hui-Lei Yu,[a] and Jian-He Xu*[a]

Nicotinamide adenine dinucleotide (NADH) and nicotinamide adenine dinucleotide phosphate (NADPH) constitute major hydrogen donors for oxidative/reductive bio-transformations. NAD(P)H regeneration systems coupled with formate dehydrogenases (FDHs) represent a dreamful method. However, most of the native FDHs are NAD^+-dependent and suffer from insufficient reactivity compared to other enzymatic tools, such as glucose dehydrogenase. An efficient and competitive $NADP^+$-utilizing FDH necessitates the availability and robustness of NADPH regeneration systems. Herein, we report the engineering of a new FDH from *Candida dubliniensis* (*Cd*FDH), which showed no strict NAD^+ preference by a structure-guided rational/semi-rational design. A combinatorial mutant *Cd*FDH-M4 (D197Q/Y198R/Q199N/A372S/K371T/△Q375/K167R/H16L/K159R) exhibited 75-fold intensification of catalytic efficiency (k_{cat}/K_m). Moreover, *Cd*FDH-M4 has been successfully employed in diverse asymmetric oxidative/reductive processes with cofactor total turnover numbers (TTNs) ranging from 135 to 986, making it potentially useful for NADPH-required biocatalytic transformations.

Introduction

Enzymatic oxidative/reductive reactions are a vital part of the current pharmaceutical industry.[1] As the most commonly used hydrogen donors, nicotinamide cofactors (NADPH & NADH) account for almost 90% of the biocatalytic redox process,[2] while the rest includes flavin,[3] PQQ,[4] CoQ[5] or other rare cofactors. However, the high cost of these cofactors prevents their stoichiometric supply during industrial applications. Therefore, economical cofactor regeneration systems are usually required.[6]

Over the past two decades, significant efforts have been devoted to the development of organometallic chemical,[7] electrochemical,[8] and photochemical methods[9] for the regeneration of nicotinamide cofactors, however, compared to enzymatic regeneration, few methods have reached high efficiency in regenerating NAD(P)H.[10] Moreover, the beauty of enzymatic NAD(P)H regeneration methods relies on their easy availability, mild reaction conditions, and even the freedom from the extra addition of cofactors when whole-cell catalysts are used.[11] Therefore, great success has been achieved for biocatalytic NAD(P)H regeneration methods, with examples such as glucose dehydrogenases (GDHs)[12] and isopropanol dehydrogenases (ADHs).[13] Nevertheless, the well-known GDH coupled cofactor regeneration systems suffered from the formation of by-product gluconic acid, which led to the demand for additional pH adjustment to avoid acidification of the system.[14] For ADHs, the co-substrate acetone may challenge the stability issue of the enzymes.[15] Therefore, more alternative enzymatic systems are required.

Among the available enzymatic tools for nicotinamide cofactor regeneration, formate dehydrogenases (FDHs) catalyze the transfer of reducing power from formate to $NAD(P)^+$, which leads to the release of nontoxic CO_2.[16] The beauty of this approach relies on the easy expulsion of the by-product CO_2 as well as the stability of pH value during the reaction. Furthermore, a broad pH range (6.0–9.0)[17] or good thermostability (T_m: 50–70 °C)[18] is usually observed in the majority of FDHs, which expands the feasibility of FDHs for more biocatalytic processes. As the metal-containing (molybdenum or tungsten) FDHs are still suffered from lability towards oxygens[19] and availabilities in terms of lacking of the biosynthesis system for the metal cofactor in industrial preferred strains (such as *E. coli* or yeasts),[20] metal-independent FDHs have been popularly applied in multiple bio-processes.[21] Unlike most of the identified native oxidoreductases or monooxygenases,[22] metal-independent FDHs exhibited higher NAD^+ preference than $NADP^+$.[23] Even though Arnold group had developed the Cofactor Specificity Reversal-Structural Analysis and Library Design (CSR-SALAD) method for NADP(H)/NAD(H) preference switching in 2017,[24] the engineering of cofactor switching was not always successful when expanded to some useful but extremely NADPH-dependent biocatalytic tools such as ene-reductase,[25] monooxygenases,[26] imine reductases[27] or reductive aminases,[28] and some examples could show sluggish reactivities when NADPH was employed.[29] Therefore, the generation of an efficient $NADP^+$ utilizing FDH is of great interest, with considerable efforts dedicated to either the identification or engineering of $NADP^+$-dependent FDHs. Although some of the examples were quite competitive,[18,30]

[a] W. Ma, Q. Geng, C. Chen, Dr. Y.-C. Zheng, Prof. Dr. H.-L. Yu, Prof. Dr. J.-H. Xu
State Key Laboratory of Bioreactor Engineering
Shanghai Collaborative Innovation Centre for Biomanufacturing
East China University of Science and Technology
Meilong Road 130, Shanghai, 200237 (China)
E-mail: jianhexu@ecust.edu.cn

[**] NADPH = nicotinamide adenine dinucleotide phosphate.

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202300390

Special Collection This article is part of a joint Special Collection with ChemCatChem dedicated to the Biotrans 2023 conference.

the available FDHs still showed rather low catalytic efficiency (< 10 U mg^{-1}) or poor expression level.

In our previous work, an engineered $NADP^+$-dependent FDH mutant from *Burkholderia stabilis* 15516 (*Bst*FDH-G146M/A287G) was generated by rational design, showing the highest catalytic constant (k_{cat}/K_m: 93.3 $s^{-1}mM^{-1}$) toward $NADP^+$ reduction.[31] However, further random mutagenesis failed to identify any mutants with obviously increased activity. As *Bst*FDH is one of few native FDH members which is strictly $NADP^+$-dependent, it appears to be highly evolved for $NADP^+$. Therefore, we address the challenge to other FDHs which should be a better starting point for protein engineering. In the present study, we chose a new FDH originating from *Candida dubliniensis* (*Cd*FDH) which showed the highest k_{cat}/K_m toward NAD^+ but with abnormally relatively high $NADP^+$ reduction ability as the starting point. By combining the CSR-SALAD with the semi-rational design method, the $NADP^+$ reduction reactivity of the resultant mutant was dramatically intensified, making it competent for various biocatalytic redox processes.

Results and Discussion

Structure and active site of CdFDH.

Among the known FDHs, FDH from *Candida dubliniensis* (*Cd*FDH) which was newly identified in our laboratory, harboring the FDH enzyme pool, showed the highest catalytic specificity ($k_{cat}/K_{m,\ NAD+} = 439\ mM^{-1}s^{-1}$) toward NAD^+, as well as the extremely low reactivity toward $NADP^+$ ($k_{cat}/K_{m,\ NADP+} = 0.63\ mM^{-1}s^{-1}$) (see also Table S1), which promises great potential for protein engineering. Therefore, we selected this enzyme as the starting point for further experiments. To obtain exact structural information for protein engineering guidance, the structure of the wild-type enzyme *Cd*FDH complexing with NADH was determined in 2.65 Å resolution (PDB code: **8J3O**), by co-crystallization of purified *Cd*FDH with NADH as well as sodium formate. The space group was the P_1 group, with 4 asymmetric units in each cell. Similar to other FDHs, *Cd*FDH functions as a homodimer.[32] The NADH cofactor was buried in the Rossman fold with the adenosine ribose interacting with Asp197, Tyr198, and Gln199 by hydrogen bonding networks. For the known NAD^+-dependent FDHs, this binding motif of ribose was highly dependent on Asp197, with position 198 occupied by bulky residues (*e.g.* Tyr, Arg, etc.) (Figure S1). However, the analysis of the binding environment of NAD(H) in *Cd*FDH indicated that the residues interacting with 2′-OH of adenosine ribose showed diversities. Besides, the electron density of the *C*-terminal (chain A: from Gly374 to Lys379; chain B: from Gln375 to Lys379; chain C: from Thr370 to Lys379; chain D: from Ala369 to Lys379) of all the subunits was disordered due to high flexibility.

Engineering of CdFDH

First-round cofactor switching aided by CSR-SALAD

Beyond the position Asp197/Tyr198/Gln199 which were targeted as the highest priority by CSR-SALAD prediction, however, as previously reported, successful examples of FDH mutagenesis that alter the cofactor preference did not exist in the suggested mutant library, especially for the position Gln199.[33] To get a better combination of the known mutagenesis candidates with the suggested mutants by CSR-SALAD, the alphabet of mutant Gln199 was further expanded to K, P, Q, R, T, N, and H. Screening of the resulting combination libraries identified the triple mutant D197Q/Y198R/Q199N (M1), with the same mutagenesis with that found in *Cbo*FDH for switching the cofactor preference to $NADP^+$.[34] Interestingly, unlike *Cbo*FDH, the turnover frequency (k_{cat}) of mutant M1 was almost unchanged (WT: $3.4 \pm 0.4\ s^{-1}$ *v.s.* M1: $3.5 \pm 0.2\ s^{-1}$). Combined with the fact that the K_m value of M1 toward $NADP^+$ was decreased from 5.4 ± 0.4 mM to 0.30 ± 0.07 mM, these results indicated that the mutagenesis only led to a better accommodation of the phosphate group of $NADP^+$.

Second-round saturation mutagenesis surrounding the binding pocket of nicotinamide cofactor and the presumed formate

One common feature usually observed in the protein engineering for nicotinamide cofactor preference switching is that the initial cofactor preference relevant mutagenesis could deteriorate the catalytic performance of the mutant as compared to the catalytic process of the native cofactor.[24,35] In our case, the specific activity of *Cd*FDH-M1 toward $NADP^+$ was only 0.5 U/mg. As the residues lining in the cofactor/formate binding pocket were more likely to have a profound effect on the recovery of activity. A total of 8 residues with ≤ 4 Å distance surrounding the binding site of the cofactor, excluding those residues relevant for catalytic binding (Figure S2), were subjected to site-saturation mutagenesis, using *Cd*FDH-M1 as the parent. Screening of the mutant library led to the identification of an additional single mutant in which Ala372 was exchanged with a polar side chain (Ser372), resulting in an approximately 1.5-fold increase in specific activity.

Compared to the reported structures of FDHs, the adenylyl moiety of the nicotinamide cofactor was blocked by their *C*-terminal (Figure S3). Motivated by the mutant A372S, further site-saturation mutagenesis was performed on the *C*-terminal residues adjacent to the adenosine part (from Ala369 to Lys379). As expected, two mutants (K371T and △Q375) were identified, respectively. Further combination of these two mutations led to the identification of *Cd*FDH-M2, with the specific activity recovered to the same magnitude (1.5 U/mg toward $NADP^+$) when compared with wild-type enzyme for NAD^+ reduction (6 U/mg). On the other hand, a combination of K371T and the truncation mutagenesis △Q375 led to the rescue of the NAD^+ reactivity, which was attributed to the dramatically improved K_m value of NAD^+ (Table 1).

Table 1. Kinetic behavior of *Cd*FDH mutants toward different nicotinamide cofactors.

Enzyme	Mutation site	NAD^+			$NADP^+$		
		k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($s^{-1}mM^{-1}$)	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($s^{-1}mM^{-1}$)
WT	–	30.7 ± 0.7	(7.0 ± 0.7)×10^{-2}	439	3.4 ± 0.4	5.4 ± 1.4	0.63
1 (M1)	D197Q/Y198R/Q199N	2.5 ± 0.2	5.1 ± 1.2	0.48	3.5 ± 0.2	0.30 ± 0.07	12
2 ($M1^{I}$)	D197Q/Y198R/Q199N/A372S	3.7 ± 0.3	4.2 ± 1.0	0.90	7.2 ± 0.2	0.60 ± 0.06	12
3 ($M1^{II}$)	D197Q/Y198R/Q199N/A372S/K371T	4.1 ± 0.2	3.8 ± 0.7	1.1	9.7 ± 0.4	0.55 ± 0.09	18
4 ($M1^{III}$)	D197Q/Y198R/Q199N/A372S/△Q375	3.0 ± 0.2	6.1 ± 1.1	0.49	11.3 ± 0.5	0.90 ± 0.13	13
5 (M2)	D197Q/Y198R/Q199N/A372S/K371T/△Q375	4.5 ± 0.2	0.21 ± 0.05	21	14.7 ± 0.3	0.69 ± 0.05	21
6 (M3)	D197Q/Y198R/Q199N/A372S/K371T/△Q375/K167R	6.7 ± 0.3	4.9 ± 0.61	1.37	28.3 ± 0.44	0.76 ± 0.04	37
7 (M4)	D197Q/Y198R/Q199N/A372S/K371T/△Q375/K167R/H16 L/K159R	7.3 ± 0.5	8.0 ± 1.5	0.91	33.6 ± 0.34	0.71 ± 0.03	47

Third-round saturation mutagenesis targeting the interface between the homo-dimers

Inspired by the successful examples of the beneficial hot spots located at the interface between each monomer when the enzyme functions as homo-dimers,[36] all residues showing polar interactions between the two monomers were selected for the third round of site saturation mutagenesis (Figure 1A), using the *Cd*FDH-M2 as the starting parent. Positive hits were found in the library of position 167, resulting in the identification of K167R (*Cd*FDH-M3). Analysis of kinetic behavior indicated the mutagenesis slightly increased the K_m value of $NADP^+$ from 0.69 mM to 0.76 mM (Table 1, Entry 6), while the corresponding k_{cat} was increased by 1.9-fold. Interestingly, the rescued NAD^+ affinity deteriorated again to 4.9 ± 0.61 mM. It was surprising that a residue on the FDH surface can directly affect the binding affinity of the cofactor which was far from this position. Furthermore, the introduction of two additional combination mutations at position 16 (H16 L) and position 159 (K159R) finally resulted in a mutant *Cd*FDH-M4, with the specific activity towards $NADP^+$ increased to 3.2 U/mg. The overall catalytic

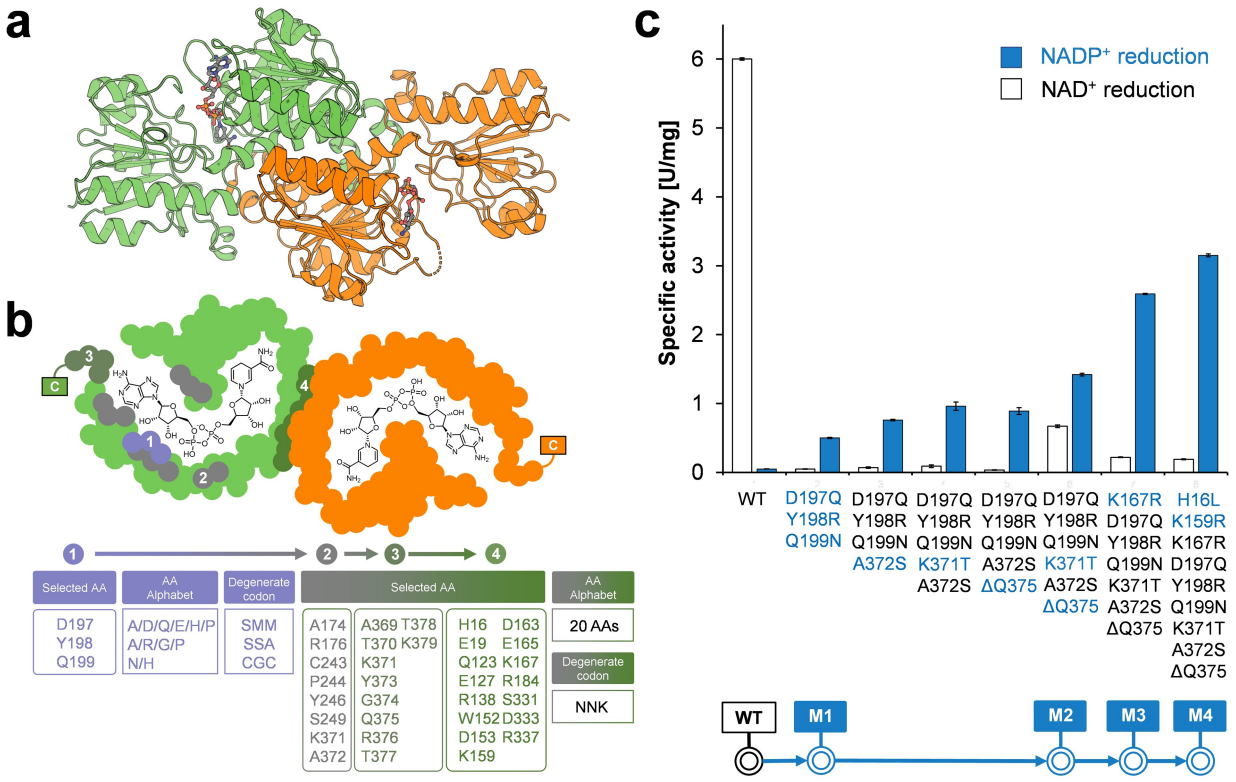


Figure 1. a) Overall structure of a functional homo-dimer of *Cd*FDH. b) Evolution strategy of this work for improving the $NADP^+$ reduction activity of *Cd*FDH. c) Summary of the evolutionary pathways of *Cd*FDH for improving the $NADP^+$ reduction activity. The newly introduced mutagenesis via each round of evolution is highlighted in blue.

constant of the best mutant *Cd*FDH-M4 showed a 75-fold increase as compared with the wild-type enzyme.

Structural insights into the improved mutants

To get better insights into the mechanical improvements of the mutants, the best mutant *Cd*FDH-M4 (PDB code: **8J3P**) was successfully co-crystallized with $NADP^+$. *Cd*FDH-M4 was determined in a different space group compared to the wild-type enzyme, giving a 2.3 Å resolution for the best datasets. The asymmetric unit contained 3 molecules of mutant M4, with 2 molecules forming homo-dimer. The dimers as well as the overall fold of each chain were conserved in the mutant enzyme, as indicated by the overall root mean square deviation (RMSD, for Cα atoms) between the two structures not exceeding 0.40 Å when dimers were superimposed (Figure S4). The binding environment of the nicotinamide cofactor, however, demonstrated profound changes that correlate with the better accommodation of $NADP^+$ in mutant M4. The mutagenesis D197Q, Y198R, and Q199 N allowed larger space near the 2′-OH of the ribose moiety of NAD(H) to accommodate the 2′-*O*-phosphate group of NAD(P)H, where new hydrogen bonding networks could be found near the phosphate group between the mutated Gln197, Arg198, and Asn199 residues. Besides, the new hydrogen bonding networks were introduced by the mutagenesis A372S (Figure 2). Unexpectedly, the Lys3 from another subunit supported additional polar contact with the 2′-*O*-phosphate group of NAD(P)H (Figure S5), which was not observed in the wild-type enzyme.

In addition to the mutagenesis of residues within a layer of 4 Å surrounding the nicotinamide cofactor, the mutagenesis located at the *C*-terminal (K371T & △Q375) or the board between two subunits of the dimer (K167R, H16L, and K159R) were remote from both the nicotinamide and formate binding sites. Combined with the fact that attempts to obtain the complex structure of *Cd*FDH-NADP(H) failed due to the relatively low affinity of *Cd*FDH-WT with NADP(H). Further molecular insights into the machinery of the mutagenesis were obtained by molecular dynamics (MD) simulations. As experimentally determined FDH/$NADP^+$/formate complexes were not available, the ternary complexes of the wild-type enzyme and mutant M4 were modeled and then employed as the starting input structures. Considering the *C*-terminal of the wild-type enzyme was missing in the crystal structure, where a truncation mutagenesis △Q375 was located, this part was supplemented by fusing the crystal structure with the *C*-terminal model predicted by AlphaFold.[37] A general increase in both correlated and anticorrelated motions was observed over the molecular evolution of *Cd*FDH (Figures 3 and S6). Anti-/correlated motions were particularly different between mutant M4 and the wild-type enzyme. The loop from Thr94 to Ser99, together with the loop from Asn242 to Lys255, which consisted of the formate binding pocket[38] and was closed to the adenyl moiety of NADP(H), played a very important role for the correlated or anti-correlated motions in *Cd*FDH. In addition, analysis of the MD trajectories indicated the residue Arg176, which usually showed hydrogen binding interaction with the phosphate group, could block the 2′-*O*-phosphate group of NADP(H). This unfavorable binding conformation thus caused the nicotinamide ring to move away from the canonical binding conformation, moreover, losing the pre-reaction state (PRS) with the small formate molecule (PRS state: WT: 2% v.s. M4: 0.07%, see also Figure S7). In contrast, the binding conformation of $NADP^+$ in mutant M4 showed high stability, which was almost unchanged compared to the crystal structure. Therefore, these mutants might contribute to the correlated nature of residue motion that facilitated the overall catalytic cycles of FDH.

Catalytic performance of CdFDH mutants for NADPH regeneration

In order to examine the catalytic performance of the *Cd*FDH mutants under real biocatalytic processes, different cofactor regeneration systems were constructed comparatively by using the wild-type enzyme and mutant M4 that showed the highest $NADP^+$ reactivity. Biocatalytic asymmetric reduction of 50 mM 2-(2,5-difluorophenyl) pyrrolidine (**1a**) catalyzed by a strict

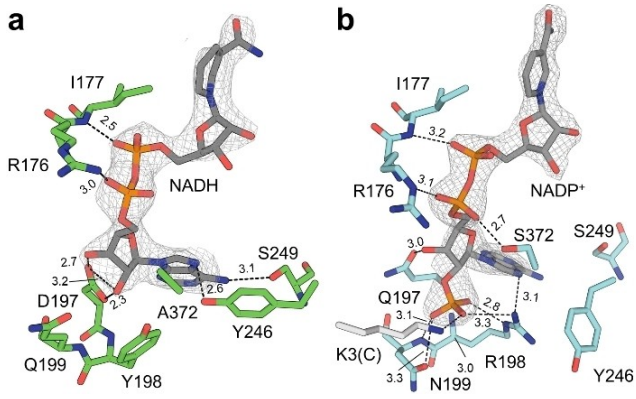


Figure 2. Comparison of the nicotinamide cofactor binding environment in the crystal structure of wild-type enzyme (a) with that of mutant M4 (b). The 2*Fo-Fc* maps of NADH and $NADP^+$ were contoured at 1.0 σ and 1.2 σ, respectively. The distance of the hydrogen bonding networks is given in Å.

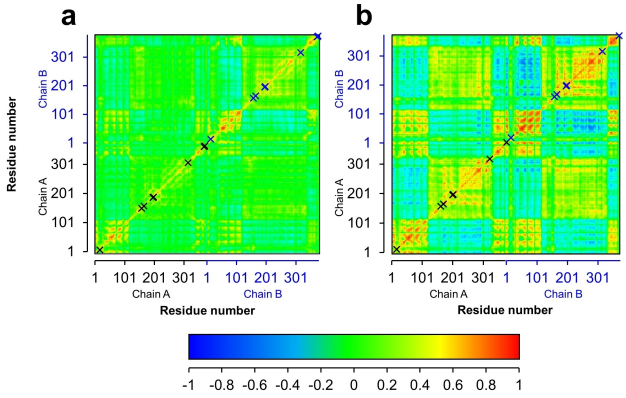


Figure 3. Dynamic cross-correlation maps (DCCMs) of *Cd*FDH-WT (a) and mutant M4 (b). The positions of the mutations were represented with the symbol ×.

NADP(H)-dependent imine reductase mutant from *Streptomyces clavuligerus* (*Sc*IRED-R3-V4)[39] was used as a model reaction (Figure 4). As expected, the reaction afforded 90% conversion of **1a** after 24 h when *Cd*FDH-M4 was employed. In contrast, only 6% conversion was observed when the bio-transformation was coupled with the wild-type enzyme, which was consistent with the fact that the wild-type enzyme showed little $NADP^+$ reactivity. However, it was too low to satisfy the requirement of reaction when high substrate loading was employed.

With the best $NADP^+$ reactive *Cd*FDH at hand, the cofactor regeneration ability for adapting this *Cd*FDH-M4 in more redox bio-transformations under high substrate loading was therefore questioned. Besides the transformation of substrate **1a**, several typical NADPH requiring bio-transformations for preparation of either pharmacological precursors or drugs were carried out by employing *Cd*FDH-M4, involving the asymmetric reductive amination of 3-(3-(trifluoromethyl) phenyl) propanal (**1b**) via IRED from *Penicillium camemberti* (*Pc*IRED-M3),[40] asymmetric reduction of ethyl 2-oxo-4-phenylbutanoate (**1c**) and *tert*-butyl 3-oxopiperidine-1-carboxylate (**1d**) catalyzed by carbonyl reductase from *Candida glabrata*,[41] and monooxygenation of cyclohexanone (**1e**) and pyrmetazole (**1f**) catalyzed by Baeyer-Villiger monooxygenases,[42] respectively (Scheme 1). Moreover, to better examine the NADPH regeneration capacity of *Cd*FDH-M4, the same bio-transformations were also carried out by coupling the well-known industrially applied GDH mutant from *Bacillus megaterium* (*Bm*GDH-M6)[12] with the same initial activity loading, but without any pH control. Except for **1a**, all systems coupled with FDH exhibited a higher cofactor total turnover number (TTN: from 135 to 986) compared to those systems coupled with GDH, which afforded a moderate to complete conversion (from 45% to 99% conversion). For GDH coupled reactions (**1b**–**1f**), the reaction appeared to be terminated when 22% to 63% of the substrates were converted in the target transformation. It was understandable that **2a** was an organic amine, which could partially neutralize the system during gluconic acid regeneration. However, since GDH produced stoichiometric gluconic acid during NADPH regeneration when the substrate loading was higher than the buffer capacity of the system, the GDH coupled systems would be deteriorated by the acidification.

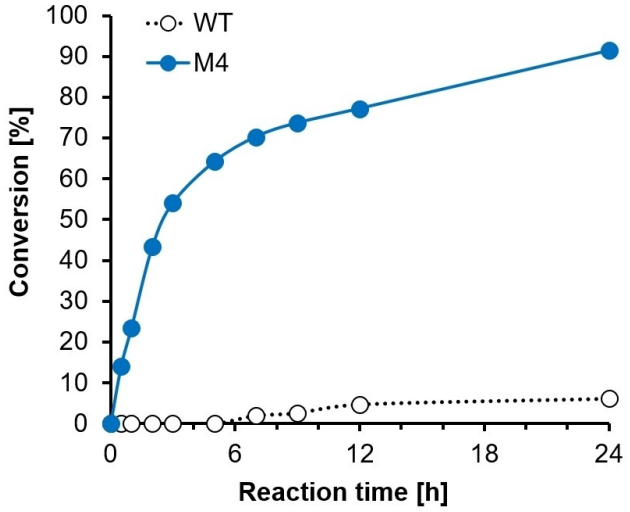


Figure 4. Comparison between *Cd*FDH-WT and *Cd*FDH-M4 for the NADPH regeneration driving the asymmetric biocatalytic reduction of 2-(2,5-difluorophenyl) pyrrolidine.

Conclusions

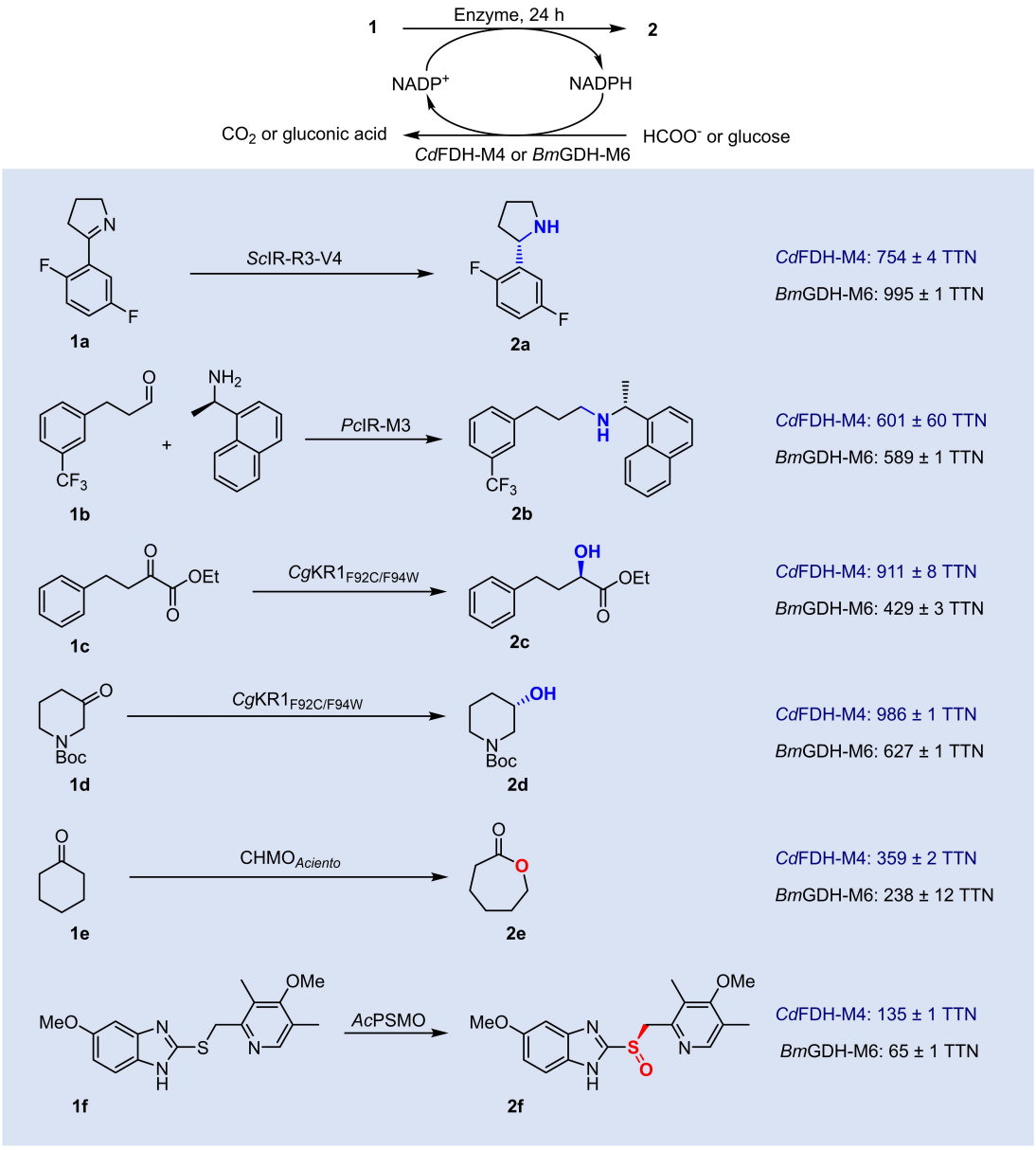
In summary, we have successfully constructed a new FDH mutant (*Cd*FDH-M4) for efficient NADPH recycling via a structure-guided semi-rational design. This *Cd*FDH-M4 had sufficient $NADP^+$ reactivity, which could support different types of NADPH-required biocatalytic reactions. Structural analysis and molecular dynamics revealed the key role of the mutated residues in facilitating the substrate/cofactor binding and the turnover of catalytic cycles of FDHs. This study, therefore, provided a valuable example for expanding the enzymatic NADPH regeneration system, as well as a lesson on how to improve the enzymatic behavior of this FDH superfamily.

Experimental Section

Materials: Primers were chemically synthesized by Tsingke Biotechnology Co., Ltd. (Shanghai, China). Unless specifically mentioned, other chemicals were purchased from commercial suppliers with the highest purity and directly used without purification. GC was performed on a SHIMADZU GC-2014 equipped with a CP-Chirasil-DEX CB column (25 m×0.25 mm×0.25 μm, Agilent, USA), or a DB-1701 column (30 m×0.25 mm×0.25 μm, Agilent, USA) and a flame ionization detector (FID). UPLC was performed on an LC-2010 system equipped with a CHIRALPAK IA-U column (1.6 μm, 3.0 mm×50 mm, Daicel Co., Japan) and a UV detector.

Mutant library construction: Site-directed saturation mutagenesis was performed by PCR using degenerate primers and the mutated position was targeted with NNK codons. The forward and reverse primers were shown in Table S2 in the supplemental material. The PCR mixture (20 mL) contained 100 ng of template DNA, primers (0.01 mM each), and 0.5 U of Prime STAR^{HS}, and the PCR was performed with the following program: 95 °C for 3 min, followed by 25 cycles of 98 °C for 10 s, 55 °C for 15 s, and 72 °C for 6 min 48 s, and then a final extension step at 72 °C for 10 min. The PCR product was digested by *Dpn* I to degrade the template plasmid and then *E. coli* BL21(DE3) was transformed by the PCR product.

Screening of the mutant library: Colonies were randomly picked into the primary 96-deep-well plates containing LB medium (300 μL) with ampicillin (50 μg mL^{-1}) per well. After incubated at 37 °C, 800 rpm overnight, the preculture (20 μL) was transferred from the primary plate to the secondary 96-deep-well plate containing LB medium (430 μL) with ampicillin (50 μg mL^{-1}) and was further incubated at 37 °C, 800 rpm for 2 to 3 h, then the isopropyl-β-*D*-thiogalactopyranoside (IPTG) was added to the culture plates with the final concentration of 0.2 mM to induce protein expression at 16 °C for 24 h. Next, the cells were harvested by centrifugation at 2500 x *g*, 4 °C for 15 min and then were resuspended in 200 μL lysis buffer (lysozyme, 750 mg L^{-1}, DNase, 10 mg L^{-1}, potassium phosphate, 100 mM, pH 7.5), the suspension was further incubated at 30 °C, 800 rpm for 2 h. A clear cell-free extract was obtained by centrifugation at 2500 x *g*, 4 °C for 10 min. An aliquot (100 μL) of the supernatant was used to evaluate the mutant's reactivity with $NADP^+$ by mixing with 100 μL of screening



Scheme 1. Comparison of the *Cd*FDH-M4 against *Bm*GDH-M6 for the cofactor regeneration in NADPH-required bio-transformations.

buffer (sodium formate, 0.2 M; $NADP^+$, 2 mM; potassium phosphate, 100 mM, pH 7.5). The concentration of the $NADP^+$ in the screening buffer and the addition amount of cell-free extract was stepwise decreased to 0.5 mM and 20 μL with the increased activity of the *Cd*FDH mutants, respectively. The $NADP^+$ reduction was monitored at 340 nm, 30 °C for 5 min by using a BioTek Synergy microplate reader (Shanghai, China). Positive hits with higher 340 nm absorbance change were sequenced, cultured by shake flask fermentation, and subjected to Ni^{2+}-affinity chromatography for further characterization.

Expression and purification of *Cd*FDH: Individual recombinant *E. coli* BL21(DE3) colonies that harbored the pET-21a(+)-*Cd*FDH or its mutant were picked into Luria Broth (LB, tryptone, 1%, *w/v*; yeast extract, 0.5%, *w/v*; NaCl, 1%, *w/v*) medium containing ampicillin (50 μg mL^{-1}) and was further incubated at 37 °C, 200 rpm until the optical density arrived at 0.6–0.8. Cells were subjected to induction with IPTG to the final concentration of 0.2 mM, the culture broth was further incubated at 16 °C, 200 rpm for 20–24 h. Biomass was harvested by centrifugation at 6000 x *g*, 4 °C for 10 min and washed with NaCl aqueous solution (0.85%, *w/v*). Cells were resuspended in Ni-NTA buffer A (imidazole, 10 mM; NaCl, 0.5 M; β-mercaptoethanol, 5 mM; potassium phosphate, 20 mM, pH 7.4), then disrupted by sonication, the clear cell-free extract was obtained by centrifugation at 10,000 x *g*, 4 °C for 0.5 h, then loaded onto a self-packed Ni Sepharose™ 6 Fast Flow affinity column (4-mL, GE Healthcare, USA) which was equilibrated with Ni-NTA buffer A. The loaded column was washed with 20-column volumes of Ni-NTA buffer A containing 10 mM imidazole, and the proteins were eluted with Ni-NTA buffer A supplied with 10–300 mM imidazole. The pooled fraction containing the desired protein was concentrated and desalted with desalt buffer (NaCl, 0.2 M, potassium phosphate, 20 mM, pH 7.4) three times by using Amicon Ultra-15 (4 °C, MWCO: 10 kDa, Millipore, Germany). Protein samples were further 1:1 diluted with storage buffer (glycerol, 40%, *v/v*; NaCl, 0.2 M, potassium phosphate, 20 mM, pH 7.4), then flash frozen in liquid nitrogen and

stored at −80 °C. The protein concentration was calculated by measuring UV absorbance at 280 nm and using the extinction coefficients (1.05 M^{-1} cm^{-1}) predicted by the online service (https://web.expasy.org/protparam/).

Crystallization and data processing: The concentrated *Cd*FDH enzyme purified by Ni^{2+}-affinity chromatography was applied to a Superdex 75 pre-packed column (4 °C, 120 mL, GE Healthcare, USA), and eluted with Superdex buffer (NaCl, 0.2 M; dithiothreitol, 1 mM; potassium phosphate, 20 mM, pH 7.2), and then concentrated to 15 mg mL^{-1}. Protein samples were incubated at 0 °C for 0.5 h with the addition of NADH/sodium formate or $NADP^+$ (10 mM), respectively. One microliter of the resulting protein solution was mixed with one microliter of the corresponding reservoir in each crystallization screening condition. Initial screening of the crystallization conditions was carried out by using commercial Nextal JCSG-Suite screens, using the sitting drop vapor-diffusion method in 96-well plates at 18 °C. Positive hits were further optimized with pH and/or precipitant gradients, using the sitting drop vapor-diffusion method in 24-well plates at 18 °C. Sheet crystals of *Cd*FDH wild-type enzyme were obtained by mixing 2 μL protein sample with 2 μL reservoir solution (110 mM potassium thiocyanate, PEG2000MME, 33 % (*w*/*v*)). For *Cd*FDH-M4, diamond crystals were obtained by mixing 2 μL protein sample with 2 μL reservoir solution (0.2 M sodium chloride, 0.1 M sodium cacodylate pH 6.5, 2 M ammonium sulfate), after incubation for 1–2 days, then cryo-protected (soaking for 5–30 s) in their corresponding reservoir solution and glycerol (30 %, *v*/*v*).

BL19 U1 (Shanghai Synchrotron Radiation Facility, Shanghai, China) was used to obtain X-ray diffraction data sets, using the wavelength of 0.9879 Å. Data were collected at 100 K. Data were processed and scaled by the XDS package.[43] The structure was solved by Phaser[44] by molecular replacement, in which peptide chain A of formate dehydrogenase from *Candida boidinii* (*Cbo*FDH, PDB code: 5DN9, 58 % sequence identity) was employed as the input model. Iterative cycles of model rebuilding and refinement were carried out manually by using COOT (Version 0.8.9)[45] and phenix,[46] respectively.

Activity and kinetics assays: The activity of FDH was measured in a 1.4-mL quartz cuvette containing 1 mL of the assay mixture at 30 °C by monitoring the change of absorbance at 340 nm within 1 min on a UV-spectrophotometer (UV-1900, Shimadzu, Japan). The reaction mixture (1 mL) contained potassium phosphate buffer (100 mM, pH 7.5), sodium formate (500 mM), $NAD(P)^+$ (0.5 mM), and appropriately diluted purified *Cd*FDH variants. One unit of enzyme activity was defined as the amount of FDH required for the formation of 1 μM NAD(P)H in 1 min.

The concentrations of sodium formate for K_m determination were varied from 10 mM to 800 mM in potassium phosphate buffer (100 mM, pH 7.5) containing 1 mM $NAD(P)^+$. Similarly, the concentrations of $NAD(P)^+$ for the measurement of K_m were varied from 0.1 mM to 10 mM in potassium phosphate buffer (100 mM, pH 7.5) containing 500 mM sodium formate. The reaction mixture was incubated at 30 °C, then monitored the increase of the absorbance at 340 nm.

Computational details: The missing *C*-terminal of the *Cd*FDH wild-type enzyme was extracted from the predicted model by AlphaFold using *Cd*FDH sequence (GenBank Accession No. XP_002420645.1). Then it was manually fused to the crystal structure of the *Cd*FDH wild-type enzyme. NADH was manually exchanged by $NADP^+$ using COOT. The resulting binary complex and the crystal structure of the *Cd*FDH-M4-$NADP^+$ complex were docked with formate by Autodock Vina,[47] respectively. Molecular dynamics were performed using the Amber 20 suite.[48] The proton state of the protein was set at pH 7.0 and predicted by the online server PDB2PQR (https://server.poissonboltzmann.org/pdb2pqr).[49] The protein was described by AMBER ff14SB force field.[50] $NADP^+$ and formate were described by AMBER GAFF force field and parameters were calculated by the restrained electrostatic potential (RESP), respectively. The partial charges of formate were calculated according to the Merz-Singh-Kollman scheme54 using Gaussian 09.[51] Partial charges of $NADP^+$ were directly taken from Bülow's work.[52] Topology and coordinate files were then generated by the AMBER tleap tool in an octahedral box of 12×12×12 Å TIP3P water model and sodium ions. The following system minimization, heating, equilibration, and productive MD simulation were performed based on our previous work.[53]

Biocatalytic redox transformation using *Cd*FDH for NADPH regeneration: Biocatalytic oxidative/reductive transformations were carried out in 2-mL Eppendorf tubes containing 1 mL reaction mixture and agitated at 30 °C for 24 h.

Asymmetric reduction of 2-(2,5-difluorophenyl) pyrrolidine (1 a): The reaction mixtures contained 100 mM **1 a**, 0.1 mM $NADP^+$, 1 % (*v*/*v*) DMSO, 12 U *Sc*IR-R3-V4, 1.8 U *Cd*FDH-M4 or *Bm*GDH-M6, potassium phosphate buffer (100 mM, pH 6.0) and 150 mM sodium formate or glucose. At the scheduled time, the reaction mixtures were extracted with 1 mL methyl tert-butyl ether. The organic layer was dried over anhydrous Na_2SO_4, then analyzed by GC.

Asymmetric reductive amination of 3-(3-(trifluoromethyl) phenyl) propanal (1 b): The reaction mixtures contained 100 mM **1 b**, 100 mM (*R*)-1-(naphthalen-1-yl) ethan-1-amine, 0.1 mM $NADP^+$, 15 % (*v*/*v*) DMSO, 2 U *Pc*IR-M3, 1.8 U *Cd*FDH-M4 or *Bm*GDH-M6, potassium phosphate buffer (100 mM, pH 6.0) and 150 mM sodium formate or glucose. At the scheduled time, the reaction mixtures were extracted with 1 mL methyl tert-butyl ether. The organic layer was dried over anhydrous Na_2SO_4, then analyzed by GC.

Asymmetric reduction of ethyl 2-oxo-4-phenylbutanoate (1 c) and tert-butyl 3-oxopiperidine-1-carboxylate (1 d): The reaction mixtures contained 100 mM **1 c** or **1 d**, 5 % (*v*/*v*) EtOH, 100 U $CgKR1_{F92C/F94W}$, 2.4 U *Cd*FDH-M4 or *Bm*GDH-M6, potassium phosphate buffer (100 mM, pH 6.0) and 150 mM sodium formate or glucose. At the scheduled time, the reaction mixtures were extracted with 1 mL EtOAc, the organic layer was dried over anhydrous Na_2SO_4, analyzed by Chiral GC.

Baeyer-Villiger oxidation of cyclohexanone (1 e): The reaction mixtures contained 50 mM **1 e**, 0.1 mM $NADP^+$, 2 % (*v*/*v*) EtOH, 1.3 U $CHMO_{Acineto}$ wild-type enzyme, 1.3 U *Cd*FDH-M4 or *Bm*GDH-M6, potassium phosphate buffer (100 mM, pH 7.0), 75 mM sodium formate or glucose. At the scheduled time, the reaction mixtures were extracted with 1 mL EtOAc, the organic layer was dried over anhydrous Na_2SO_4, and analyzed by GC.

Asymmetric Kagan oxidation of pyrmetazole (1 f): The reaction mixtures contained 30 mM **1 f**, 0.1 mM $NADP^+$, 10 % (*v*/*v*) MeOH, 20 mU *Ac*PSMO, 1.3 U *Cd*FDH-M4 or *Bm*GDH-M6, potassium phosphate buffer (100 mM, pH 8.0), 45 mM sodium formate or glucose. At the scheduled time, the reaction mixtures were extracted with 1 mL EtOAc, the organic layer was dried over anhydrous Na_2SO_4, and analyzed by chiral UPLC.

Supporting Information

The authors have cited additional references within the Supporting Information.[33a,34,54]

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (grants no. 21871085 and no. 31971380), the National Key Research and Development Program of China (2019YFA0905000), and the Fundamental Research Funds for the Central Universities (222201714026). We thank Dr. Lian Wu (SIOC) for the initial diffraction data processing. We are also grateful for the access to beamline BL19 U1 at Shanghai Synchrotron Radiation Facility and thank the beamline staff for their technical assistance.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

Research data are not shared.

Keywords: cofactor regeneration · formate dehydrogenase · protein engineering · cofactor preference

[1] a) U. T. Bornscheuer, G. Huisman, R. Kazlauskas, S. Lutz, J. Moore, K. Robins, *Nature* **2012**, *485*, 185–194; b) D. J. Pollard, J. M. Woodley, *Trends Biotechnol.* **2007**, *25*, 66–73; c) F. Hollmann, I. W. Arends, D. Holtmann, *Green Chem.* **2011**, *13*, 2285–2314.
[2] H. Wu, C. Tian, X. Song, C. Liu, D. Yang, Z. Jiang, *Green Chem.* **2013**, *15*, 1773–1789.
[3] P. Macheroux, B. Kappes, S. E. Ealick, *FEBS J.* **2011**, *278*, 2625–2634.
[4] K. Takeda, K. Umezawa, A. Várnai, V. G. Eijsink, K. Igarashi, M. Yoshida, N. Nakamura, *Curr. Opin. Chem. Biol.* **2019**, *49*, 113–121.
[5] R. Banerjee, J. Purhonen, J. Kallijärvi, *FEBS J.* **2022**, *289*, 6936–6958.
[6] S. Mordhorst, J. N. Andexer, *Nat. Prod. Rep.* **2020**, *37*, 1316–1333.
[7] Y. Maenaka, T. Suenobu, S. Fukuzumi, *J. Am. Chem. Soc.* **2012**, *134*, 9417–9427.
[8] a) Y. S. Lee, R. Gerulskis, S. D. Minteer, *Curr. Opin. Biotechnol.* **2022**, *73*, 14–21; b) Y. H. Kim, Y. J. Yoo, *Enzyme Microb. Technol.* **2009**, *44*, 129–134.
[9] a) I. Willner, D. Mandler, *Enzyme Microb. Technol.* **1989**, *11*, 467–483; b) S. H. Lee, Y. C. Kwon, D. M. Kim, C. B. Park, *Biotechnol. Bioeng.* **2013**, *110*, 383–390.
[10] W. Liu, P. Wang, *Biotechnol. Adv.* **2007**, *25*, 369–384.
[11] K. Bachosz, J. Zdarta, M. Bilal, A. S. Meyer, T. Jesionowski, *Sci. Total Environ.* **2023**, 161630.
[12] W. Z. Qian, L. Ou, C. X. Li, J. Pan, J. H. Xu, Q. Chen, G. W. Zheng, *ChemBioChem* **2020**, *21*, 2680–2688.
[13] Q. Jia, Y.-C. Zheng, H.-P. Li, X.-L. Qian, Z.-J. Zhang, J.-H. Xu, *Appl. Environ. Microbiol.* **2022**, *88*, e00341–00322.
[14] X. Wang, T. Saba, H. H. Yiu, R. F. Howe, J. A. Anderson, J. Shi, *Chem.* **2017**, *2*, 621–654.
[15] Z. Yang, H. Fu, W. Ye, Y. Xie, Q. Liu, H. Wang, D. Wei, *Catal. Sci. Technol.* **2020**, *10*, 70–78.
[16] A. Zhang, X. Zhuang, J. Liu, J. Huang, L. Lin, Y. Tang, S. Zhao, R. Li, B. Wang, B. Fang, *Nat. Catal.* **2023**, 1–10.
[17] A. V. Mesentsev, V. S. Lamzin, V. I. Tishkov, T. B. Ustinnikova, V. O. Popov, *Biochem. J.* **1997**, *321*, 475–480.
[18] V. I. Tishkov, V. O. Popov, *Biomol. Eng.* **2006**, *23*, 89–110.
[19] M. Moon, G. W. Park, J.-p. Lee, J.-S. Lee, K. Min, *J. CO_2 Util.* **2020**, *42*, 101353.
[20] a) R. R. Mendel, S. Leimkühler, *JBIC J. Biol. Inorg. Chem.* **2015**, *20*, 337–347; b) J.-i. Yang, S. H. Lee, J.-Y. Ryu, H. S. Lee, S. G. Kang, *Front. Microbiol.* **2022**, *13*.
[21] a) C. Wandrey, A. Liese, D. Kihumbu, *Org. Process Res. Dev.* **2000**, *4*, 286–290; b) N. Xu, J. Zhu, Y. Wu, Y. Zhang, J. H. Xu, *Org. Process Res. Dev.* **2020**; c) J. Zhu, Q. Geng, Y.-Y. Liu, J. Pan, H. L. Yu, J.-H. Xu, *Org. Process Res. Dev.* **2022**, *26*, 1978–1983; d) Y. Peng, Z. Chen, J. Xu, Q. Wu, *Org. Process Res. Dev.* **2022**, *26*, 1900–1913.
[22] L. S. Vidal, C. L. Kelly, P. M. Mordaka, J. T. Heap, *Biochim. Biophys. Acta.* **2018**, *1866*, 327–347.
[23] V. Tishkov, V. Popov, *Biochemistry* **2004**, *69*, 1252–1267.
[24] J. K. Cahn, C. A. Werlang, A. Baumschlager, S. Brinkmann-Chen, S. L. Mayo, F. H. Arnold, *ACS Synth. Biol.* **2017**, *6*, 326–333.
[25] C. Mähler, F. Kratzl, M. Vogel, S. Vinnenberg, D. Weuster-Botz, K. Castiglione, *Adv. Synth. Catal.* **2019**, *361*, 2505–2513.
[26] a) K. Balke, M. Kadow, H. Mallin, S. Saß, U. T. Bornscheuer, *Org. Biomol. Chem.* **2012**, *10*, 6249–6265; b) O. Döhr, M. J. Paine, T. Friedberg, G. C. Roberts, C. R. Wolf, *Proc. Nat. Acad. Sci.* **2001**, *98*, 81–86.
[27] N. Borlinghaus, B. M. Nestl, *ChemCatChem* **2018**, *10*, 183–187.
[28] M. Lenz, N. Borlinghaus, L. Weinmann, B. M. Nestl, *World J. Microbiol. Biotechnol.* **2017**, *33*, 1–10.
[29] a) A. Beier, S. Bordewick, M. Genz, S. Schmidt, T. Van Den Bergh, C. Peters, H. J. Joosten, U. T. Bornscheuer, *ChemBioChem* **2016**, *17*, 2312–2315; b) M. Gand, C. Thöle, H. Müller, H. Brundiek, G. Bashiri, M. Höhne, *J. Biotechnol.* **2016**, *230*, 11–18; c) N. M. Kamerbeek, M. W. Fraaije, D. B. Janssen, *Eur. J. Biochem.* **2004**, *271*, 2107–2116.
[30] S. Alpdagtas, B. Binay, *Biocatal. Biotransform.* **2020**, *39*, 260–268.
[31] H.-W. Jiang, Q. Chen, J. Pan, G.-W. Zheng, J.-H. Xu, *Appl. Biochem. Biotechnol.* **2020**, *192*, 530–543.
[32] V. S. Lamzin, Z. Dauter, V. O. Popov, E. H. Harutyunyan, K. S. Wilson, *J. Mol. Biol.* **1994**, *236*, 759–785.
[33] a) A. E. Serov, A. S. Popova, V. V. Fedorchuk, V. I. Tishkov, *Biochem. J.* **2002**, *367*, 841–847; b) A. Andreadeli, D. Platis, V. Tishkov, V. Popov, N. E. Labrou, *FEBS J.* **2008**, *275*, 3859–3869; c) K. Hoelsch, I. Sührer, M. Heusel, D. Weuster-Botz, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 2473–2481; d) L. Calzadiaz-Ramirez, C. Calvó-Tusell, G. M. Stoffel, S. N. Lindner, S. Osuna, T. J. Erb, M. Garcia-Borràs, A. Bar-Even, C. G. Acevedo-Rocha, *ACS Catal.* **2020**, *10*, 7512–7525.
[34] W. Wu, D. Zhu, L. Hua, *J. Mol. Catal. B* **2009**, *61*, 157–161.
[35] A. M. Chánique, L. P. Parra, *Front. Microbiol.* **2018**, *9*, 194.
[36] a) T. Geppert, B. Hoy, S. Wessler, G. Schneider, *Chem. Biol.* **2011**, *18*, 344–353; b) E. Y. Hong, S. G. Lee, B. J. Park, J. M. Lee, H. Yun, B. G. Kim, *Biotechnol. J.* **2017**, *12*, 1700278; c) A. Bosshart, S. Panke, M. Bechtold, *Angew. Chem.* **2013**, *125*, 9855–9858.
[37] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, *Nature* **2021**, *596*, 583–589.
[38] V. O. Popov, V. S. Lamzin, *Biochem. J.* **1994**, *301*, 625–643.
[39] Q. Chen, B.-B. Li, L. Zhang, X.-R. Chen, X.-X. Zhu, F.-F. Chen, M. Shi, C.-C. Chen, Y. Yang, R.-T. Guo, *ACS Catal.* **2022**, *12*, 14795–14803.
[40] F.-F. Chen, X.-F. He, X.-X. Zhu, Z. Zhang, X.-Y. Shen, Q. Chen, J.-H. Xu, N. J. Turner, G.-W. Zheng, *J. Am. Chem. Soc.* **2023**, *145*, 4015–4025.
[41] G.-W. Zheng, Y.-Y. Liu, Q. Chen, L. Huang, H.-L. Yu, W.-Y. Lou, C.-X. Li, Y.-P. Bai, A.-T. Li, J.-H. Xu, *ACS Catal.* **2017**, *7*, 7174–7181.
[42] a) N. A. Donoghue, D. B. Norris, P. W. Trudgill, *Eur. J. Biochem.* **1976**, *63*, 175–192; b) S. M. Ren, F. Liu, Y. Q. Wu, Q. Chen, Z. J. Zhang, H. L. Yu, J. H. Xu, *Biotechnol. Bioeng.* **2021**, *118*, 737–744.
[43] W. Kabsch, *Acta Crystallogr. Sect. D* **2010**, *66*, 125–132.
[44] A. J. McCoy, *Acta Crystallogr. Sect. D* **2007**, *63*, 32–41.
[45] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126–2132.
[46] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, *Acta Crystallogr. Sect. D* **2010**, *66*, 213–221.
[47] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455–461.
[48] R. Salomon-Ferrer, D. A. Case, R. C. Walker, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3*, 198–210.
[49] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen, K. Liles, *Protein Sci.* **2018**, *27*, 112–128.
[50] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

[51] U. C. Singh, P. A. Kollman, *J. Comput. Chem.* **1984**, *5*, 129–145.
[52] N. Holmberg, U. Ryde, L. Bülow, *Protein Eng.* **1999**, *12*, 851–856.
[53] Y.-K. Qi, Y.-C. Zheng, Q. Chen, Y. He, Z.-J. Zhang, J.-H. Xu, *ACS Sustainable Chem. Eng.* **2021**, *9*, 12514–12519.
[54] a) X. Robert, P. Gouet, *Nucleic Acids Res.* **2014**, *42*, W320-W324; b) A. Boldt, M. B. Ansorge-Schumacher, *Adv. Synth. Catal.* **2020**, *362*, 4109–4118; c) V. Tishkov, A. Galkin, A. Egorov, *Dokl. Akad. Nauk SSSR.* **1991**, *317*, 745–748; d) H. Nanba, Y. Takaoka, J. Hasegawa, *Bioscience, Biotechnol., Biochem.* **2003**, *67*, 720–728; e) N. Gul-Karaguler, R. B. Sessions, A. R. Clarke, J. J. Holbrook, *Biotechnol. Lett.* **2001**, *23*, 283–287; f) R. Hatrongjit, K. Packdibamrung, *Enzyme Microb. Technol.* **2010**, *46*, 557–561; g) S. Alpdağtaş, S. Yücel, H. A. Kapkaç, S. Liu, B. Binay, *Biotechnol. Lett.* **2018**, *40*, 1135–1147; h) S. Fogal, E. Beneventi, L. Cendron, E. Bergantino, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 9541–9554.

Manuscript received: May 24, 2023
Revised manuscript received: July 9, 2023
Accepted manuscript online: July 16, 2023
Version of record online: August 28, 2023