Supplementary Materials:

Biocatalytic cascade of sebacic acid production with *in situ* co-factor regeneration enabled by engineering of an alcohol dehydrogenase

Jie Lu [1,2], Dong Lu [1,2], Qiuyang Wu [1,2], Shuming Jin [1,2], Junfeng Liu [1], Meng Qin [1], Li Deng [1,2], Fang Wang [1,2] and Kaili Nie [1,2] *

[1] College of Life Science and Technology, Beijing University of Chemical Technology, Beijing, 100029, PR China

[2] Beijing Key Lab of Bioprocess, National Energy R&D Center for Biorefinery, Beijing University of Chemical Technology, Beijing, 100029, PR China

* Correspondence: niekl@mail.buct.edu.cn

Contents

1. Molecular modeling

The template of a diketoreductase crystalstructure (Protein Data Bank code: 4E12, 283 residues) was obtained through searching the database according to the gene of *Ml*ADH (ADH from *Micrococcus luteus*, NCBI Reference Sequence: ADD83022.1, 310 residues) in SWISS-MODEL (https://swissmodel.expasy.org). The sequence identity with *Ml*ADH is 49.12%, and the quality of the homologous model of *Ml*ADH was evaluated.

As shown in Figure S1, the global quality estimation score (GMQE) of this model was 0.75. GMQE varied between 0 and 1, of which higher values indicating better global quality. The local composite scoring function (QMEAN) was -0.53, which varied between -4 and 0, with values closer to 0 indicating higher local matching degree, such as pairwise atomic distances, torsion angles, solvent accessibility and so on. Based on the evaluation results, the homology model is available for further docking and molecular dynamics (MD) simulations.
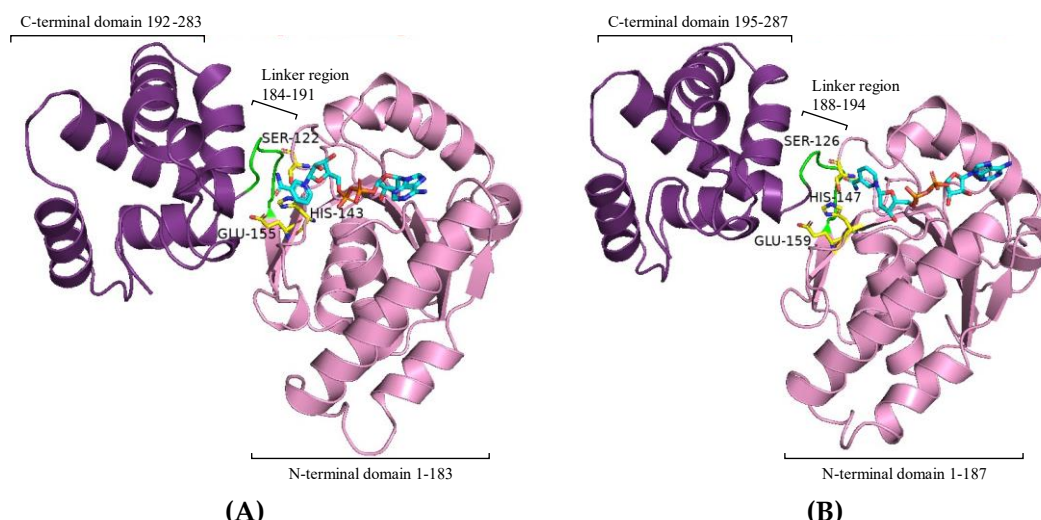

Figure S1. Homology modeling results based on 4E12

2. Molecular docking simulation results

The molecular docking simulations of complexes were performed with the YASARA software package (version 21.8.26). Docking simulation for 50 times were carried out at 298K and atmospheric pressure under AMBER 03 force field with the *Ml*ADH homology model as the rigid receptor and the cofactor as the flexible ligand.

As shown in Figure S2, the optimal docking conformation of 4E12-NAD$^+$(A) and *Ml*ADH-NAD$^+$ (B) was screened according to the binding energy and binding conformation. Three residues of His143-Ser122-Glu155 and the active hydrogen from NAD$^+$ formed the catalytic triad of the enzyme in 4E12[1]. From the comparison of docking results, the possible active center of *Ml*ADH should be His 147-Ser 126-Glu 159.
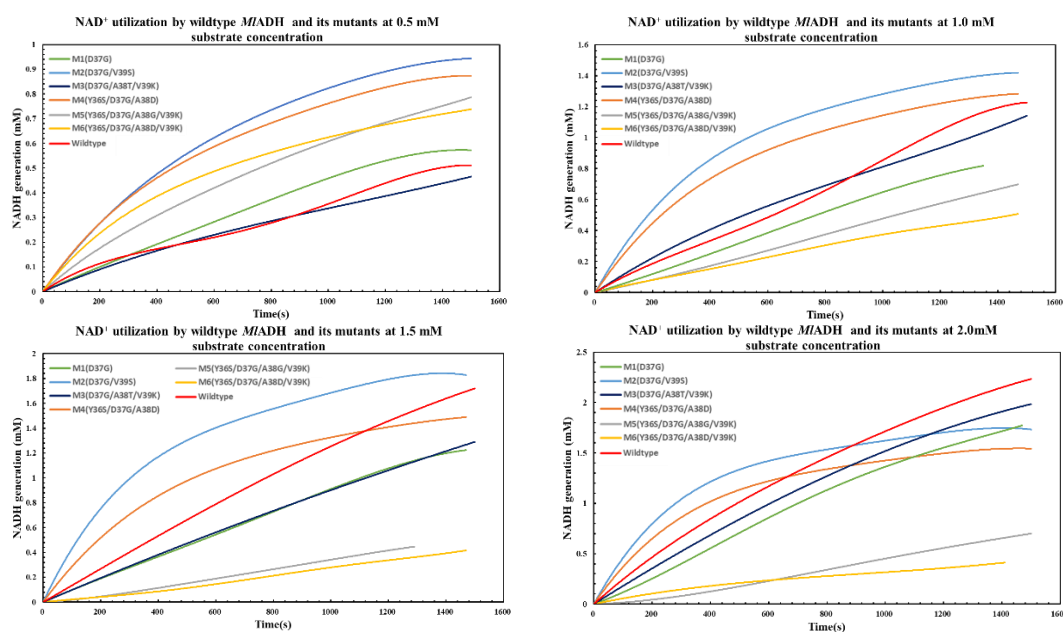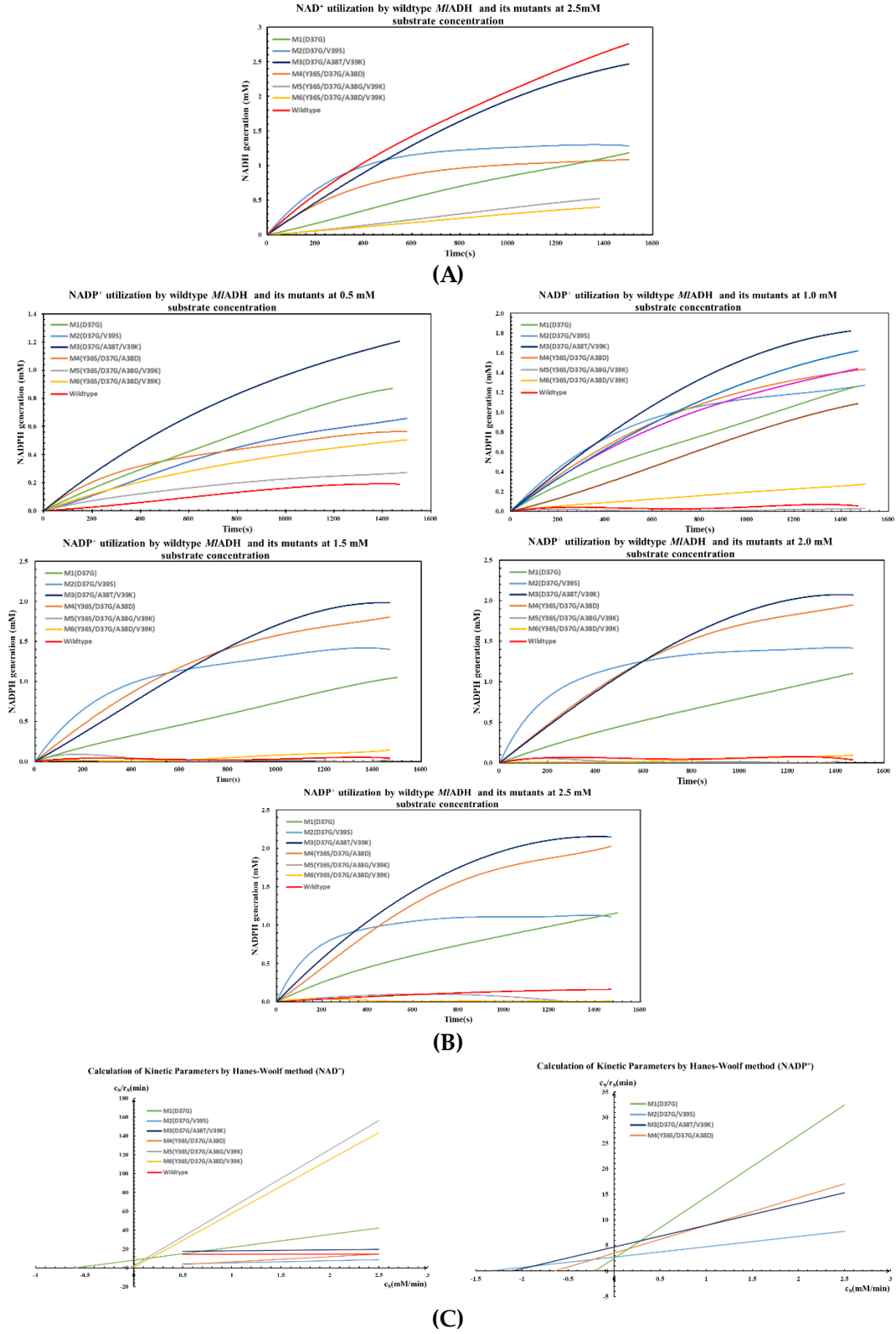
Figure S2. The structure comparison of *Ml*ADH-NAD$^+$ with 4E12-NAD$^+$. (A)reported conformation of 4E12-NAD$^+$ [1]; (B) the simulation results of *Ml*ADH-NAD$^+$

3. Determination of kinetic parameters of *Ml*ADH and its mutants by Hanes-Woolf method

A series of 200 μL reaction system were set up in a 96-well plate, and continuously monitored at 340nm absorbance on a UV-spectrophotometer (Enspize, PerkinElmer) at 25 °C (Figure S3A and Figure S3B) for the change of NAD(P)H generation. The assay mixture contained a final concentration of 0.5-2.5 mM of 10-hydroxyoctadecanoic acid, 0.1 mM NAD(P)$^+$, 50 mM Tris-HCl buffer (pH 7.5) and an appropriate amount of purified enzyme by His-tag through Ni-NTA affinity chromatograph.

According to the data fitting, the single substrate catalytic reaction conformed to the Michaelis-Menten equation. The slope is obtained by making the tangent of the curve where the reaction time is 0. Then the kinetic parameters were obtained through Hanes-Woolf method [2], and the slope was $1/V_{max}$, the intercept with the ordinate is K_m/V_{max}, the intercept with the abscissa is $-K_m$, and k_{cat} was obtained by dividing V_{max} by the initial substrate concentration (Figure S3C).

(A)

(B)

(C)

Figure S3. Determination of kinetic parameters of MlADH and its mutants by Hanes-Woolf method.

(A) Fitting curve of time and NADH generation at different substrate concentrations;
(B) Fitting curve of time and NADPH generation at different substrate concentrations;
(C) H-W graphic method.

4. Determination of product and substrate by GC-MS

Substrate and target products were identified by GC-MS (Agilent 7890A-5975C). GC analysis program: the temperature of the injector was set at 290 °C. The initial temperature of the column was 50 °C for 1 min, and gradually increased to 290 °C at a rate of 6 °C/min, then maintained for 10 min. The MS was operated with an ion-source of 280 °C and 70 eV. The chromatography (with peak identification) and mass spectrum were provided in Fig S4.
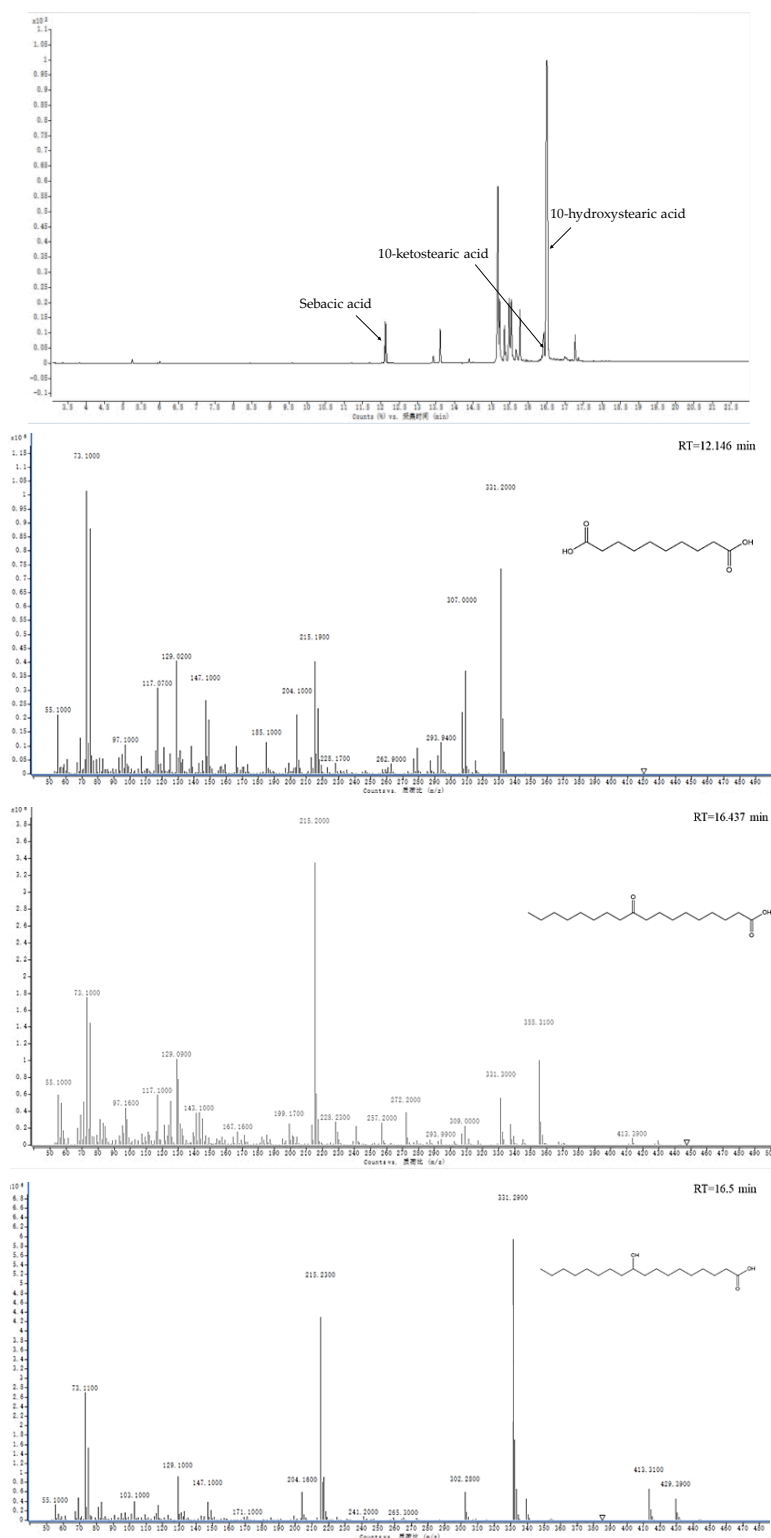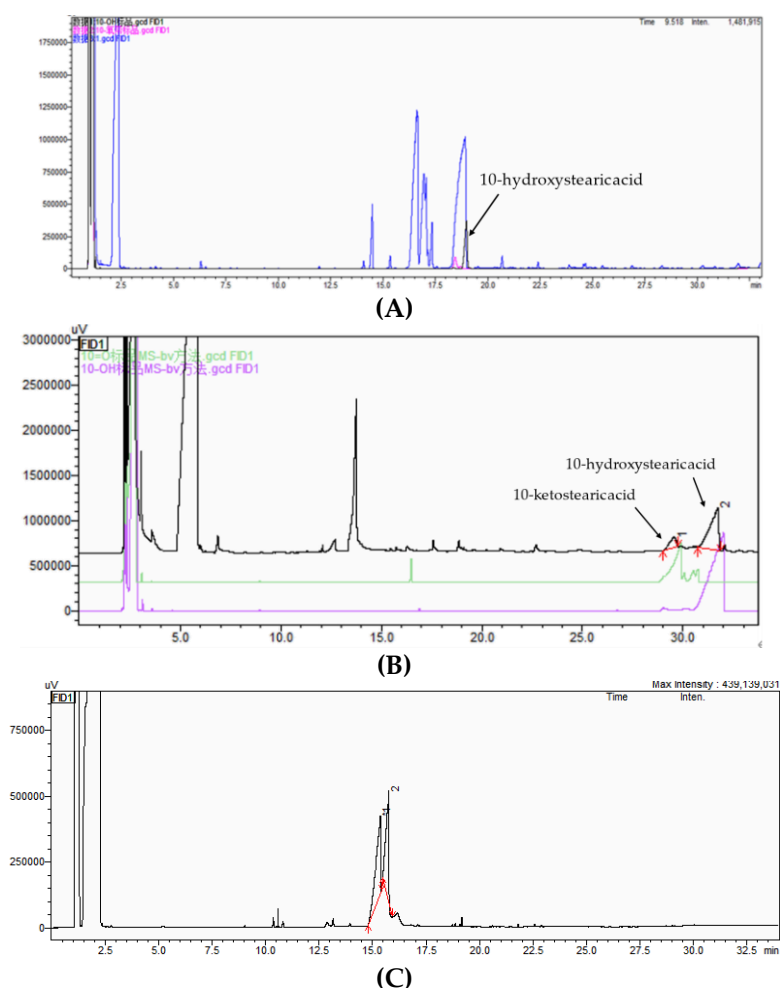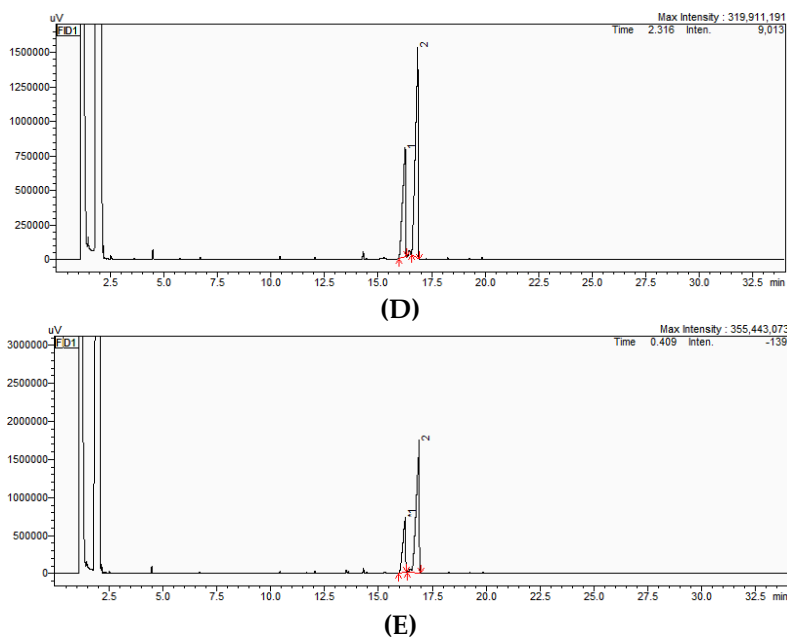

Figure S4. GC-MS determination of combined catalysis of *Ml*ADH and *Pa*BVMO

The enzymatic cascade reaction contains 1.0 mM 10-hydroxystearic acid, 1 mM NAD(P)$^+$, 10 μL Tween 20 and an appropriate amount of purified *Ml*ADH or its mutants. The total volume of reaction was 20 mL (50 mM Tris-HCl buffer, pH 7.5). The reactions were carried out at 25 °C for 24 hours. Extraction was carried out with ethyl acetate as solvent, which was removed by blowing nitrogen. Then, the products were collected after hydrolyzation (KOH (1.0 M) at 80 °C for 2 h), acidification (to pH 2.0 with diluted HCl), and extraction (ethyl acetate,1:1(v/v)). Finally, the solvent was dried under nitrogen, and derived by *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide (TMS).

The dehydrogenation conversion performance of the wild-type *Ml*ADH with NAD$^+$ or NADP$^+$ as co-factor were verified (Figure S.5A-5B). The conversion rate of 10-hydroxyoctadecanoic acid to 10-oxooctadecanoic acid was over 98% when NAD$^+$ was applied, whereas the yields of 10-oxooctadecanoic acid was only 15% when the co-factor changed to NADP$^+$. The results indicated that the NADP$^+$ is not a suitable co-factor for the wild-type *Ml*ADH.

According to the kinetic parameters, enzyme activity of M1-M3 with NADP$^+$ preference was amplified and verified (Figure S.5C-E). When the NADP$^+$ was utilized as the only source of co-factors, the yields of 10-oxooctadecanoic acid with the mutants M3, M1, M2 as catalyst was 56%, 36%, and 27%, respectively.
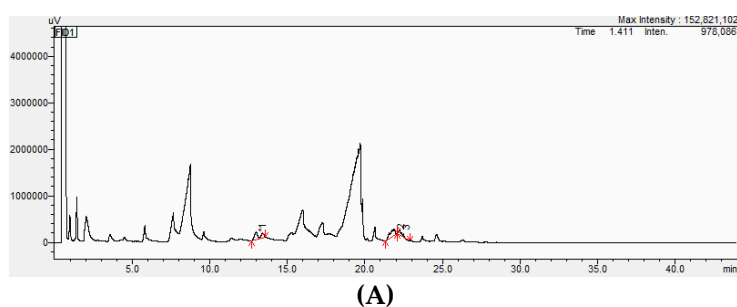

(A)

(B)

(C)

(D)

(E)

Peak1: 10-ketostearic acid; Peak2: 10-hydroxystearic acid

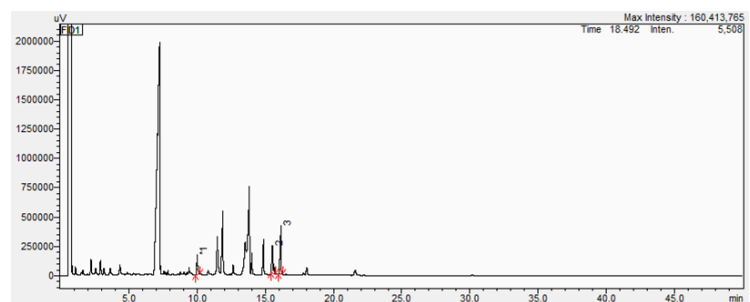**Figure S5.** GC chromatography of enzymatic reaction results of purified *Ml*ADH and its mutants using NAD(P)⁺

(A) wild-type *Ml*ADH using NAD⁺; (B) wild-type *Ml*ADH using NAD⁺
(C) *Ml*ADH mutants of M3(D37G/A38T/V39K) using NADP⁺; (D) *Ml*ADH mutants of
M1(D37G) using NADP⁺; (E) *Ml*ADH mutants of M2(D37G/V39S) using NADP⁺.

The enzymatic cascade reactions were completed to verify sebacic acid production with *in situ* NADP⁺ regeneration system, containing 1.0 mM 10-hydroxystearic acid, 1 mM NAD(P)⁺, 10 μL Tween 20, 0.5 mg/mL *Ml*ADH (or its mutants), 1 mg/mL *Pa*BVMO. The total volume of reaction was 20 mL (50 mM Tris-HCl buffer, pH 7.5). The reactions were carried out at 25°C for 24 hours. After reaction, the reaction solution was treated in the same manner as described above.
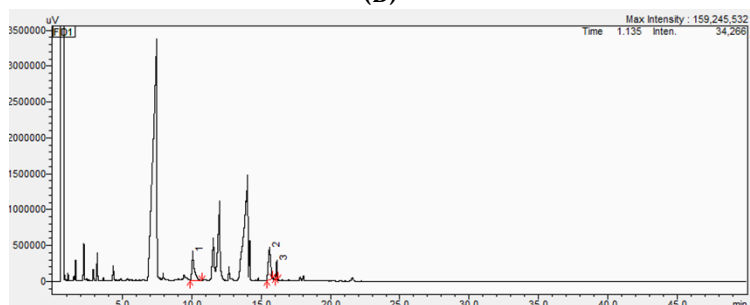
The traditional multi-enzymatic co-factor regeneration system was set as controlled group, in which the NOX and the GDH was used for NAD⁺ and NADPH regeneration, respectively. The final yield of sebacic acid was 41.2%. The combined catalysis of wild-type *Ml*ADH and *Pa*BVMO obtained a sebacic acid yield of 20.1%. The combined catalysis of M1(D37G), M2(D37G/V39S) and M3(D37G/A38T/V39K) with *Pa*BVMO had sebacic acid yields of 42.4%, 25.8% , and 49.3%, respectively.
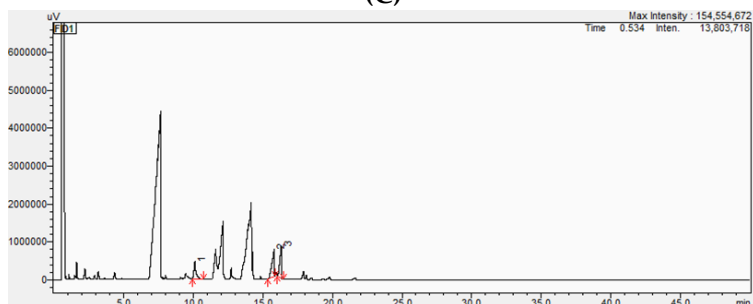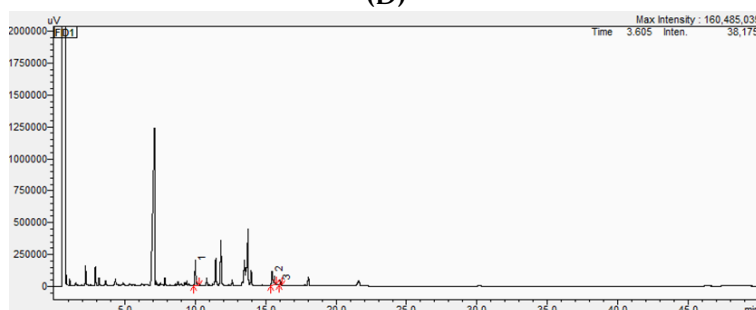

(A)

(B)

(C)

(D)

(E)

Peak1: sebacic acid; Peak2: 10-ketostearic acid; Peak3: 10-hydroxystearic acid

Figure S6. GC chromatography of sebacic acid production by multistep catalytic reactions.
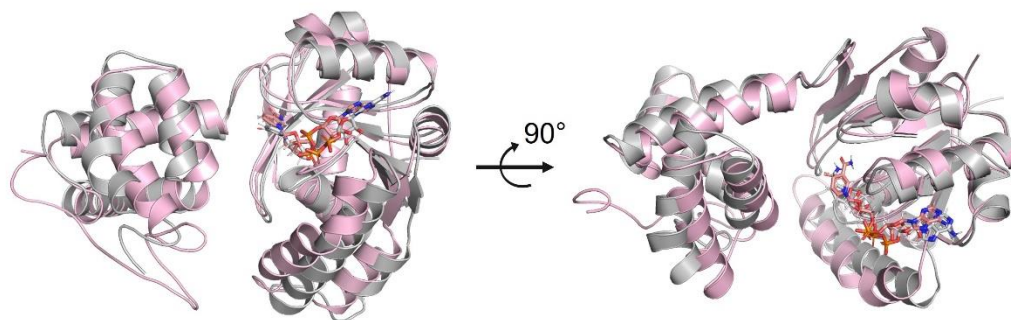(A) Wildtype *Ml*ADH & *Pa*BVMO & GDH & NOX using NADP$^+$ and NAD$^+$;
(B) Wildtype *Ml*ADH & *Pa*BVMO using NADP$^+$; (C) M1(D37G) & *Pa*BVMO using NADP$^+$;
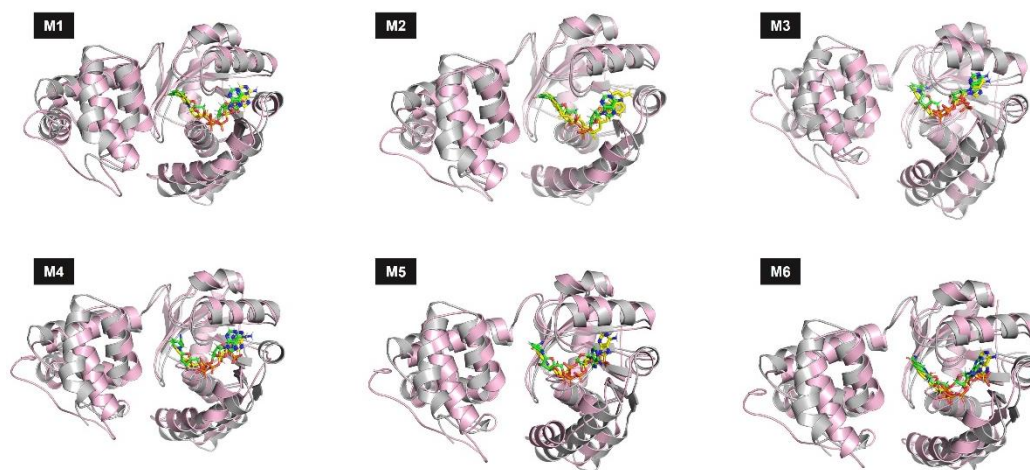(D) M2(D37G/V39S) & *Pa*BVMO using NADP$^+$; (E) M3(D37G/A38T/V39K) & *Pa*BVMO using NADP$^+$.

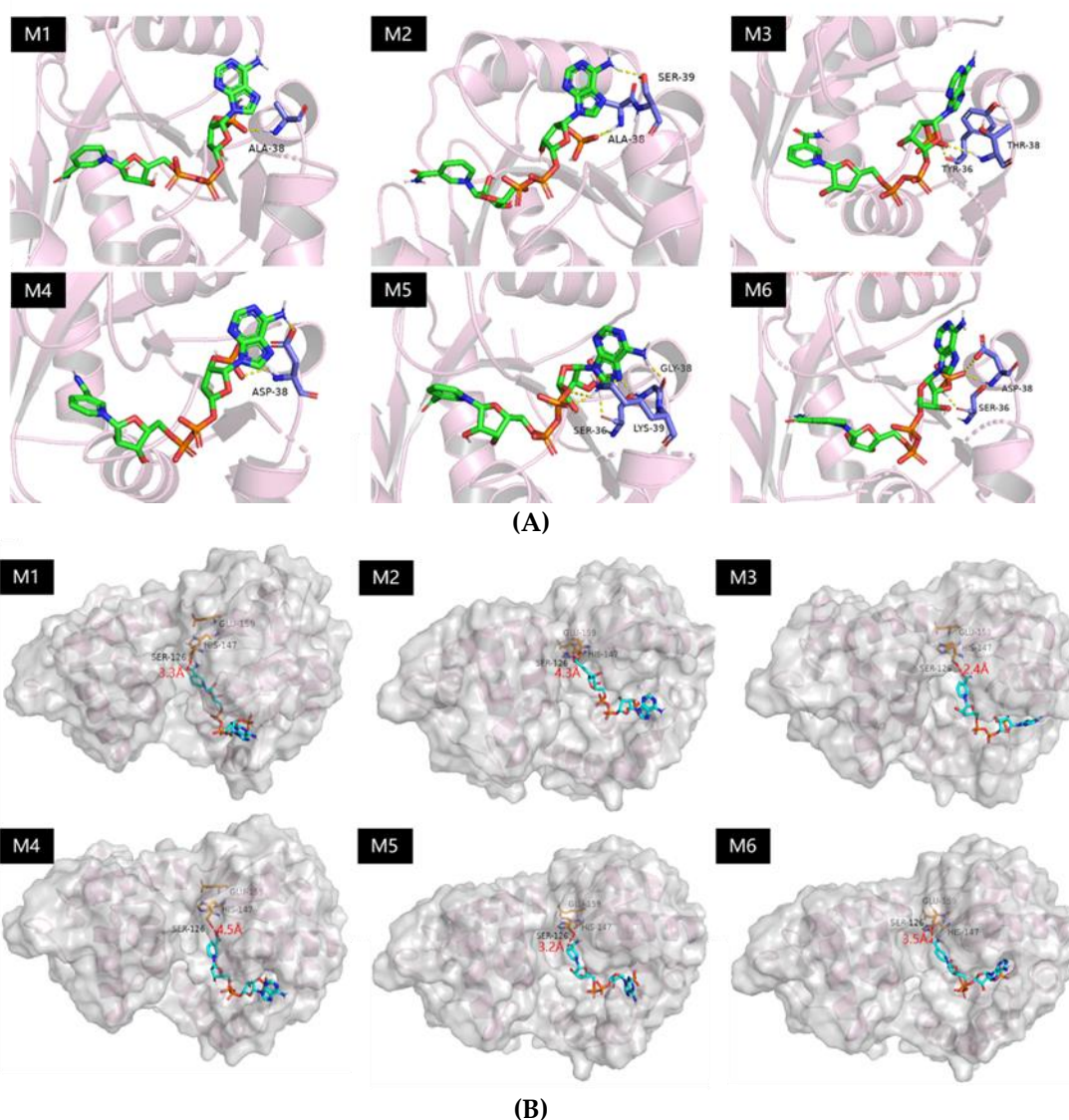**5. Comparison of simulation results of homology model from 4E12 and 6KQ9**

Figure S7. Structure comparison of homology model from PDB: 4E12 and PDB: 6KQ9. Protein structure of 6KQ9 was in gray, homology model from 4E12 was in lightpink.

Figure S8. Comparison of simulation results of homology model from 4E12 and 6KQ9 with reasonable mutation candidates.

(A) The binding conformation of NADP$^+$ in the enzyme; (B) The whole conformation of the ADH with NADP$^+$. Mutants of M1 to M6 was D37G, D37G-V39S, D37G-A38T-V39K, Y36S-D37G-A38D, Y36S-D37G-A38G-V39K, and Y36S-D37G-A38D-V39K, respectively. Protein structure of homology model from 6KQ9 was in gray (with MADP$^+$ in green), homology model from 4E12 was in lightpink (with MADP$^+$ in yellow).

Docking conformation of 6KQ9 - NADP⁺

(A)

(B)

Figure S9. Simulation results of reasonable mutation candidates with PDB 6KQ9 as templet (A) The binding conformation of NADP⁺ in the enzyme; (B) The whole conformation of the ADH with NADP⁺ (with PDB 6KQ9 as templet). Mutants of M1 to M6 was D37G, D37G-V39S, D37G-A38T-V39K, Y36S-D37G-A38D, Y36S-D37G-A38G-V39K, and Y36S-D37G-A38D-V39K, respectively.

References

1. Lu, M.; Huang, Y.; White, M.A.; Wu, X.; Liu, N.; Cheng, X.; Chen, Y. Dual catalysis mode for the dicarbonyl reduction catalyzed by diketoreductase. *Chemical Communications* **2012**, *48 (92)*, 11352-11354.
2. Qin, Y.Z.; Xia, J.; Wang, B.W. *Biological reaction engineering*, 2rd ed.; Chemical Industry Press: Beijing, China, 2009; pp. 9-14.