Electronic Supplementary Material (ESI) for Green Chemistry.

Supporting Information

One-pot chemoenzymatic synthesis of glycolic acid from formaldehyde

Tianzhen Li,[‡ab] Zijian Tan,[‡b] Zijing Tang,[b] Pi Liu,[b] Haifeng Liu,[*c] Leilei Zhu,[*b] and Yanhe Ma[b]

[a] University of Chinese Academy of Sciences, Beijing, 100049, China

[b] Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences; National Technology Innovation Centre of Synthetic Biology, Tianjin, 300308, China

[c] Jiangsu Collaborative Innovation Centre of Chinese Medicinal Resources Industrialization, School of Pharmacy, Nanjing University of Chinese Medicine, 210023, China

[‡]These authors contributed equally to this work.

[*]Corresponding Author: haifeng.liu@njucm.edu.cn; zhu_ll@tib.cas.cn.

Table of Contents

1. Experimental Procedures

1.1 Materials and methods

Formaldehyde solution was purchased from Aladdin (shanghai, China); glycolaldehyde (GA) and 1,3-dihydroxyacetone (DHA) were purchased from Bidepharm (shanghai, China); diphenylamine, glycolic acid, sodium chlorite ($NaClO_2$), sodium hypochlorite (NaClO), hydrogen peroxide (H_2O_2), and sodium persulfate ($Na_2S_2O_8$) were purchased from Macklin (shanghai, China). The other analytical-reagent grades or higher quality chemicals were purchased from Sinopharm (Beijing, China), except the resins (GE Healthcare, Boston, USA) for purifications. EasyTaq DHA polymerase was purchased from TransGen Biotech (Beijing, China); galactose oxidase (GalOD) was purchased from Yuanye Biological (Shanghai, China) and Aladdin (Shanghai, China); horseradish peroxidase (HRP) was purchased from GENVIEW (USA); 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonate) (ABTS) was purchased from Coolaber (Beijing, China). Oligonucleotide syntheses and DNA sequencing analyses were carried out by Genewiz (Tianjin, China). The amount of DNA in cloning experiments was quantified by TECAN microplate reader (Infinite M200, TECAN, Switzerland). The amount of protein was quantified by GenStar BCA protein assay kit (Beijing, China).

1.2 Conversion of commercial GA/DHA to glycolate by different oxidants

140 mM GA or DHA solution was prepared by dissolving commercial GA or DHA in phosphate buffer (50 mM KPi, 5 mM $MgSO_4$, pH 7.0). 200 mM $NaClO_2$ solution, 200 mM H_2O_2 solution, and 200 mM $Na_2S_2O_8$ solution were prepared by dissolving $NaClO_2$, H_2O_2, $Na_2S_2O_8$ in phosphate buffer (50 mM KPi, 5 mM $MgSO_4$, pH 7.0), respectively. Sodium hypochlorite (NaClO) solution was prepared by dissolving it in phosphate buffer (500 mM NaH_2PO_4) to adjust the pH to 7.0, and the available chlorine of sodium hypochlorite was about 3.2% detected by A-1 available chlorine test paper (ANNJET, Shandong, China). 500 μL GA solution (140 mM) or DHA solution (140 mM) was added into 1.5 mL microcentrifuge tube containing 500 μL $NaClO_2$ solution (200 mM), or NaClO solution (3.2% available chlorine), or H_2O_2 solution (200 mM), or $Na_2S_2O_8$ solution (200 mM). The reaction was conducted at 30°C, 1000 rpm for 24 h. The generated glycolate was detected by HPLC. The HPLC system was carried out on an Agilent 1260 system equipped with a RI detector and UV detector (210 nm), and fitted with chromatographic column (DEVOTE Sugar 10H column). Mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 35°C

1.3 Conversion of commercial GA/DHA to glycolate at different pH values by $NaClO_2$

Phosphate buffer (50 mM KPi, 5 mM $MgSO_4$) with different pH values (5.0, 7.0 and 8.0) and carbonate buffer (50 mM Na_2CO_3 and $NaHCO_3$, 5 mM $MgSO_4$) with different pH values (9.0 and 10.0) were prepared. GA solution (140 mM), DHA solution (140 mM) and $NaClO_2$ solution (200 mM) were prepared with the above-mentioned buffer solutions with varied pH value. The reaction was initiated by adding GA solution (500 μL, 140 mM) or DHA solution (500 μL, 140 mM) to 1.5 mL microcentrifuge tube containing $NaClO_2$ solution (500 μL, 200 mM) and conducted for 24 h (30°C, 1000 rpm). The generated glycolate was detected by HPLC. The HPLC system was carried out on an Agilent 1260 system equipped with a RI detector and UV detector (210 nm), and fitted with chromatographic column (DEVOTE Sugar 10H column). Mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 35°C.

1.4 Conversion of commercial GA/DHA to glycolate in presence of different concentration of $NaClO_2$

Different concentration of $NaClO_2$ solution were prepared by dissolving $NaClO_2$ in phosphate buffer (50 mM NaPi, 5 mM $MgSO_4$, pH 8.0). The reactions with GA as substrate were executed in 1.5 mL microcentrifuge tube containing GA (70 mM), different concentration of $NaClO_2$ (100, 140, 210, 280, 350 mM) and phosphate buffer (50 mM NaPi, 5 mM $MgSO_4$, pH 8.0). The reactions with DHA as substrate were executed in 1.5 mL microcentrifuge tube containing DHA (70 mM), different concentration of $NaClO_2$ (100, 140, 210, 280, 350 mM) and phosphate buffer (50 mM NaPi, 5 mM $MgSO_4$, pH 8.0). All reactions were conducted at 30°C, 1000 rpm for 24 h. The generated glycolate was detected by HPLC. The HPLC system was carried out on an Agilent 1260 system equipped with a RI detector and UV detector (210 nm), and fitted with chromatographic column (DEVOTE Sugar 10H column). Mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 35°C.

1.5 Conversion of commercial GA/DHA to glycolate by $NaClO_2$ at different temperature

DHA solution, GA solution and $NaClO_2$ solution were prepared by dissolving DHA, GA and $NaClO_2$ in phosphate buffer (50 mM NaPi, 5 mM $MgSO_4$, pH 8.0). The reaction was executed in 1.5 mL microcentrifuge tube containing DHA (70 mM), $NaClO_2$ (210 mM) and phosphate buffer (50 mM NaPi, 5 mM $MgSO_4$, pH 8.0), and conducted at different temperatures (35, 40, 45, 50°C), 1000 rpm for 24 h. The reaction was executed in 1.5 mL microcentrifuge tube containing GA (70 mM), $NaClO_2$ (210 mM) and phosphate buffer (50 mM NaPi, 5 mM $MgSO_4$, pH 8.0) at different temperature (30, 50°C) with shaking (1000 rpm) for 24 h. Then glycolate was detected by HPLC. The HPLC system was carried out on an Agilent 1260 system equipped with a RI detector and UV detector (210 nm), and fitted with chromatographic column (DEVOTE Sugar 10H column). Mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 35°C.

1.6 Conversion of commercial DHA to glycolate in presence of different concentration of phosphate buffer by $NaClO_2$

The DHA solution (140 mM) and $NaClO_2$ solution (420 mM) in the varied concentration of phosphate buffer (100 mM, 150 mM, 200 mM NaPi, 5 mM $MgSO_4$, pH 8.0; 250 mM, 300 mM KPi, 5 mM $MgSO_4$, pH 8.0) were prepared. The reaction was initiated by mixing 500 μL DHA solution (140 mM) and equal volume of $NaClO_2$ solution (420 mM) in same concentration of phosphate buffer in 1.5 mL microcentrifuge tube. The reaction was conducted for 24 h (50°C, 1000 rpm). The generated glycolate was detected by HPLC. The HPLC system was carried out on an Agilent 1260 system equipped with a RI detector and UV detector (210 nm), and fitted with chromatographic column (DEVOTE Sugar 10H column). Mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 35°C.

1.7 Construction of formolase error-prone PCR (epPCR) libraries

epPCR libraries were generated by the standard epPCR method using parent formolase M4 (W86R/N87T/L109G/L110E/A460M/H281Y)[1] as template. For the mutagenic PCR (95°C for 2 min, 1 cycle; 95°C, 20 s/60°C, 20 s/72°C, 3 min, 25 cycles; 72°C for 5 min, 1 cycle). EasyTaq DHA polymerase (2.5 U), dNTP mix (0.20 mM), template (25 ng, pET28a harboring

the M4 gene), $MnCl_2$ (0.05 and 0.08 mM), and primers (forward: 5'-CCGCGCGGCAGCCATATG-3', reverse: 5'-GGTGGTGGTGGTGGTGCTCGAGTTATT-3', 10 pmol each) were used. The epPCR products were purified by using a TransGen PCR purification kit (TransGen Biotech, Beijing, China). The purified epPCR products were cloned into expression plasmid pET28a and transformed into *Escherichia coli* BL21-Gold (DE3) cells for further expression and screening.

1.8 Cultivation and expression of formolase variants in 96-well plates

Clones grown on LB agar plates (50 μg/mL kanamycin) were transferred into 96-well microtiter plates (V-bottomed, polystyrene plates; Beyotime, Shanghai, China) containing 150 μL LB liquid medium (50 μg/mL kanamycin). After overnight cultivation in a microtiter plate shaker (37°C, 800 rpm; Zhichu Instrument, Shanghai, China), each well was replicated by a replicator into a second series of 96-well microtiter plates containing 200 μL LB liquid medium (50 μg/mL kanamycin) and 0.1 mM Isopropyl β-D-Thiogalactoside (IPTG). All the 96-well microtiter plates were cultivated in the microtiter plate shaker at 800 rpm at 37°C for the first 4 h and then cultivated at 30°C for next 24 h. After expression, the pellets were harvested in V-bottom 96-well microtiter plates after centrifugation (4°C, 3400 rpm, 15 min; Eppendorf Centrifuge 5810 R; Eppendorf Corporate, Hamburg, German) and then resuspended in 200 μL phosphate buffer (50 mM NaPi, 5 mM $MgSO_4$, pH 7.4). After centrifugation, 96-well microtiter plates were subjected for following screening.

1.9 Library screening for improved activity

For activity screening, the cell pellets in 96-well microtiter plates were resuspended in phosphate buffer (50 mM NaPi, 5 mM $MgSO_4$, pH 7.4, 75 μL), then 75 μL formaldehyde solution (phosphate buffer containing different concentrations of formaldehyde and 1mM thiamine diphosphate (TPP), pH 7.4) was added in above plates. The plates were incubated for 3 h at 30°C. After incubation, the plates were centrifuged (4°C, 3400 rpm, 10 min) and 30 μL supernatant was transferred into 96-well microtiter plates (flat-bottomed, polystyrene plates; Beyotime, Shanghai, China) and 90 μL supernatant was transferred into another 96-well microtiter plates.

For GA activity screening, 150 μL assay buffer A (1.5 g diphenylamine was dissolved in 100 mL acetic acid and then added 1.5 mL H_2SO_4) was added in the plate containing 30 μL supernatant. The plate was incubated at 90°C for 30 min, and subsequently was detected at 650 nm using a microtiter plate reader (VersaMax; Molecular Devices, Sunnyvale, USA).

For DHA activity screening, assay buffer B (0.2 mg/mL GalOD, 24 U/mL HRP, 50 mM NaPi, 5 mM $MgSO_4$, pH 7.4, 60 μL) and assay buffer C (4 mM ABTS, 50 mM NaPi, 5 mM $MgSO_4$, pH 7.4, 50 μL) were added in the plates. The plates were immediately measured at 410 nm using a microtiter plate reader.

1.10 Expression of formolase variants in shaking flask and purification

Single clone was inoculated into 20 mL LB liquid medium (50 μg/mL kanamycin) and grown overnight. The cell cultures were transferred into 600 mL LB liquid medium (50 μg/mL kanamycin) with the ratio of 1:100. All the flasks were shaken at 200 rpm at 37°C. When the optical density at 600 nm (OD_{600}) reached 0.6-0.8, gene expressions were induced using 0.1 mM IPTG at 20°C for 24 h. Subsequently, *E. coli* cells were harvested by centrifugation (4°C, 5000 rpm, 15 min) and resuspended in 25 mL phosphate buffer (50 mM NaPi, 5 mM $MgSO_4$, pH 7.4), and subsequently followed by disruption with a low temperature ultra-high pressure coutionuous flow cell disrupter (LN-3000plus; JNBIO, Guangzhou, China). The disrupted cells were centrifuged (4°C, 16000 rpm, 40 min), and the supernatants were further filtered through millipore express membrane filters (0.22 μm; Sterile Millex Filter Unit; Merck kGaA, Darmstadt, Germany).

The purification procedure was carried out as following: the supernatants were applied to a Ni-NTA column fast protein liquid chromatography (FPLC) system (ÄKTA go; Cytiva, America) with a His-Spin protein column, the target proteins eluted at ~150 mM imidazole when using a 20-250 mM imidazole gradient. The purified protein fractions were run on SDS-PAGE to ascertain the size and purity. After concentrated in a centrifugal filter unit (Amicon Ultra-15, 10 kDa Centrifugal Filter Devices; Merck kGaA, Darmstadt, Germany), purified formolase variant was desalted by FPLC system with a desalt column. The concentration of protein was determined using a protein assay kit (GenStar BCA Protein Assay Kit; GenStar, Beijing, China) with BSA as the standard. The homogeneity of the purified sample was controlled by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) by using standard molecular biology techniques.

1.11 Kinetic characterization of formolase and improved variant

The k_{cat} and K_m /K_{half} values were determined from initial-velocity data measured as a function of substrate concentration. First, 70 μL different concentration of substrate solution (5 to 1000 mM formaldehyde, 1 mM TPP in phosphate buffer) was added in 96-well microtiter plates. Enzyme reactions were initiated by the addition of equal volume of purified formolase variants. Samples were incubated (at 30°C, 800 rpm) and taken out at different time points.

The produced GA was determined as followings: 30 μL reaction solution in the 96-well microtiter plates, 150 μL assay buffer A (1.5 g diphenylamine was dissolved in 100 mL acetic acid and then added 1.5 mL H_2SO_4) for the detection of GA was added in the plate. After incubation at 90°C for 30 min, the absorbance at 650 nm was measured.

The produced DHA was determined as followings: 90 μL reaction solution in the 96-well microtiter plate, 60 μL assay buffer B (0.3 mg/mL GalOD, 36 U/mL HRP, 50 mM NaPi, 5 mM $MgSO_4$, pH 7.4) was added in the plate, and then 50 μL assay buffer C (3.2 mM ABTS, 50 mM NaPi, 5 mM $MgSO_4$, pH 7.4) was added in the plates. The plates were immediately measured at 410 nm for 20 min by using a microtiter plate reader.

The initial-velocity data for GA was fitted to Michaelis-Menten equation ($v=V_{max}\times S/(K_m+S)$), where v is the initial enzyme velocity, S is substrate concentration, V_{max} is the maximum enzyme velocity, K_m is the substrate concentration needed to achieve a half-maximum enzyme velocity. The initial-velocity data for DHA was fitted to allosteric sigmoidal equation ($v=V_{max}\times S^n/(K_{half}^n+S^n)$), where v is the initial enzyme velocity, S is substrate concentration, V_{max} is the maximum enzyme velocity, K_{half} is the ligand concentration at which half of the active sites are occupied (concentration of half saturation). All experiments were conducted in triplicate.

1.12 Formaldehyde resistance of parent M4 and the improved variant

Purified M4 (parent) and M4V2 in phosphate buffer (50 mM NaPi, 5 mM $MgSO_4$, pH 7.4) were quantified. 0.027 mM of purified M4 or M4V2 were incubated in presence of different concentration of formaldehyde (200 mM, 500 mM, and 1 M) and 0.5 mM TPP with shaking (1000 rpm, 30°C) for a range of time (0.5, 1, 2, 4, 6, 8 h). GA and DHA were detected by HPLC. The HPLC system was carried

out on an Agilent 1260 system equipped with a RI detector and UV detector (210 nm), and fitted with chromatographic column (DEVOTE Sugar 10H column). Mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 65°C.

1.13 Molecular docking and generation of the variants in silico

Molecular docking was performed using Schrödinger2018 software. Crystal structure of BFD-M3 (PDB ID: 6M2Z) without TPP was used as receptor protein.[1] TPP -FA (complex of thiamine diphosphate and one-carbon formaldehyde) and TPP-GA (complex of thiamine diphosphate and two-carbon glycolaldehyde) were used as ligands. The receptor was prepared with protein preparation wizard module of Schrodinger software. Receptor grid was set to 30 X 30 X 30 Å and centered to the mass center of TPP. Glides SP of docking algorithm was used to perform docking. Then, M3-TPP-FA and M3-TPP-GA complex structures were mutated to M4-TPP-FA, M4-TPP-GA, M4V2-TPP-FA and M4V2-TPP-GA with mutate module of Schrodinger software.

1.14 Molecular dynamic (MD) simulation

MD simulation of M4-TPP-FA, M4-TPP-GA, M4V2-TPP-FA and M4V2-TPP-GA was performed with software Amber20. The ff19SB and gaff2 force field parameters were used to set up simulation systems. 100 ns MD simulation were performed for all the systems. All simulations were carried out according to the following procedure: initially, we performed a series of energy minimization steps to eliminate any bad contacts in the initially built structures. During the minimization, protein backbone was restrained with harmonic force constant of 20 kcal/mol/Å. The minimization step included 2000 steps steepest descent followed by 200 steps of conjugate gradient method. After the energy minimization, the system was slowly heated up to 310 K in 500 ps using 2 fs integration time step. After this, we performed 500 ps NPT equilibration of the structures with same restraints. Finally, NPT production simulations were performed at 310 K and 1 atm pressure with 2 fs integration time step with restraint maintained. We have implemented periodic boundary condition across the system using a TIP3P water box.[2] We used Particle Mesh Ewald (PME) techniques integrated with AMBER package to account for the long-range part of the electrostatic interactions.[3] During the MD simulation, all the bonds involving hydrogen were restrained using the SHAKE algorithm.[4] Langevin thermostat with collision frequency of 2 ps^{-1} was used to maintain constant temperature while the pressure was controlled by anisotropic Monte-Carlo barostat.[5] The accelerated GPU version of PMEMD was performed on NVIDIA GeForce RTX 30 Series cards.[6]

1.15 Calculation of binding free energy

MM-GBSA (molecular mechanics/generalized born surface area) was used to calculate the binding energy between TPP intermediate ligands (TPP-FA and TPP-GA) and formolase M4 and M4V2. 100 snapshots of the last 10 ns MD simulation were fetched out to calculate the binding free energy. The entropy contributions were neglected because the same receptor was used and that the normal mode analysis calculations are computationally expensive and cause a large margin of error which introduces significant uncertainty in the result.

$$\Delta G = \Delta E_{MM} + \Delta G_{solv} - T \cdot \Delta S$$
$$= \Delta E_{bat} + \Delta E_{vdw} + \Delta E_{coul} + \Delta G_{solv.p} + \Delta G_{solv.np} - T \cdot \Delta S$$

The free energy for each species (ligand, receptor, and complex) was decomposed into a gas-phase MM energy, polar, and nonpolar solvation terms, and an entropy term, as shown in the following equation: E_{MM} is composed of E_{bat} (the sum of bond, angle, and torsion terms in the force field), a Van der Waals term E_{vdW} and a coulombic term E_{coul}. $G_{solv,p}$ is the polar contribution to the solvation free energy, often computed via the Generalized-Born (GB) approximation. $G_{solv,np}$ is the nonpolar solvation free energy, usually computed as a linear function of the solvent-accessible surface area (SASA).

1.16 The influence of the concentration of TPP on the conversion of formaldehyde catalyzed by M4V2

0.0886 mM purified M4V2 were incubated with 600 mM formaldehyde and in presence of different concentration of TPP (0-0.5 mM) with shaking (1000 rpm, 30°C) for 3 h. The generated GA and DHA were detected by HPLC. The HPLC system was carried out on an Agilent 1260 system equipped with a RI detector and UV detector (210 nm), and fitted with chromatographic column (DEVOTE Sugar 10H column). Mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 65°C.

1.17 The influence of the concentration of M4V2 on the conversion of formaldehyde catalyzed by M4V2

Different concentration of M4V2 (0.0886 mM, 0.0709 mM, 0.0532 mM, 0.0355 mM) were incubated with 600 mM formaldehyde, 0.5 mM TPP and with shaking (1000 rpm, 30°C) for 3 h. The generated GA and DHA were detected by HPLC. The HPLC system was carried out on an Agilent 1260 system equipped with a RI detector and UV detector (210 nm), and fitted with chromatographic column (DEVOTE Sugar 10H column). Mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 65°C.

1.18 Chemoenzymatic synthesis of glycolate from formaldehyde by M4V2 and $NaClO_2$

Enzymatic conversion of formaldehyde to GA and DHA (total volume 2 mL) containing fomolase variant M4V2 (0.0886 mM), formaldehyde (600 mM), TPP (0.5 mM), $MgSO_4$ (5 mM) and phosphate buffer (50 mM NaPi, pH 7.4). The reaction was conducted at 30°C, 1000 rpm for 3 h. The generated GA and DHA were detected by HPLC. Subsequently, 0.5 mL enzymatic reaction mixture was added in a 1.5 mL microcentrifuge tube which containing 0.5 mL $NaClO_2$ solution (633 mM) which was dissolved in 250 mM phosphate buffer (250 mM NaPi, 5 mM $MgSO_4$, pH 8.0). The reaction was conducted at 50°C for 24 h. The generated glycolate was detected by HPLC. The HPLC system was carried out on an Agilent 1260 system equipped with a refractory index (RI) detector and UV detector (210 nm), and fitted with chromatographic column (DEVOTE Sugar 10H column). (A) For GA and DHA: mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 65°C; (B) For glycolate: mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 35°C.

1.19 Synthesis of glycolate from the chemoenzymatic conversion of formaldehyde in bigger volume

Enzymatic conversion of formaldehyde to GA and DHA (total volume 50 mL) containing fomolase variant M4V2 (0.0886 mM), formaldehyde (600 mM), TPP (0.5 mM), $MgSO_4$ (5 mM) and phosphate buffer (50 mM NaPi, pH 7.4). The reaction was conducted at 30°C, 1000 rpm for 3 h. The generated GA and DHA were detected by HPLC. Subsequently, 45 mL enzymatic reaction mixture was added in a 500 mL flask which containing 45 mL $NaClO_2$ solution (612 mM) which was dissolved in 250 mM phosphate buffer (250 mM KPi, 5 mM $MgSO_4$, pH 8.0). The reaction was conducted at 50°C for 24 h. The generated glycolate was detected by HPLC.

1.20 Conversion of commercial GA/DHA to glycolate in presence of different concentration of H_2O_2

Different concentration of H_2O_2 solution were prepared by diluting 30% H_2O_2 in phosphate buffer (50 mM KPi, 5 mM $MgSO_4$, pH 8.0). The reactions with GA as substrate were executed in 1.5 mL microcentrifuge tube containing GA (70 mM), different concentration of H_2O_2 (140, 210, 280, 420, 560 mM) and phosphate buffer (50 mM KPi, 5 mM $MgSO_4$, pH 8.0). The reactions with DHA as substrate were executed in 1.5 mL microcentrifuge tube containing DHA (70 mM), different concentration of H_2O_2 (140, 210, 280, 420, 560 mM) and phosphate buffer (50 mM KPi, 5 mM $MgSO_4$, pH 8.0). All reactions were conducted at 30°C, 1000 rpm for 24 h. The reaction solution was added to $Na_2S_2O_3$ solution with the same volume to consume the remaining H_2O_2 in the reaction system. The generated glycolate was detected by HPLC. The HPLC system was carried out on an Agilent 1260 system equipped with a RI detector and UV detector (210 nm), and fitted with chromatographic column (DEVOTE Sugar 10H column). Mobile phase: 5 mM H_2SO_4, flow rate: 0.5 mL/min, sample volume: 20 μL, column temperature: 35°C.

1.21 Conversion of commercial GA/DHA to glycolate by H_2O_2 at different temperature

The DHA solution, GA solution and H_2O_2 solution were prepared by dissolving DHA, GA and 30% H_2O_2 in phosphate buffer (50 mM KPi, 5 mM $MgSO_4$, pH 8.0). The reaction was executed in 1.5 mL microcentrifuge tube containing GA (70 mM) or DHA (70 mM), H_2O_2 (280 mM) and phosphate buffer (50 mM KPi, 5 mM $MgSO_4$, pH 8.0) at different temperatures (35°C, 40°C, 45°C, 50°C) with shaking (1000 rpm) for 24 h. The reaction solution was added to $Na_2S_2O_3$ solution with the same volume to consume the remaining H_2O_2 in the reaction system. The generated glycolate was detected by HPLC.

1.22 Optimization of the concentration of phosphate buffer for the conversion commercial GA or DHA to glycolate in presence of H_2O_2

The GA solution (140 mM), DHA solution (140 mM) and H_2O_2 solution (560 mM) in varied concentration of phosphate buffer (100 mM, 150 mM, 200 mM, 250 mM, 300 mM KPi, 5 mM $MgSO_4$, pH 8.0) were prepared. The reaction was initiated by mixing 500 μL GA solution (140 mM) or DHA solution (140 mM) with equal volume of H_2O_2 solution (560 mM). The reaction was conducted for 24 h (50°C, 1000 rpm). The reaction solution was added in $Na_2S_2O_3$ solution with the same volume to consume the remaining H_2O_2 in the reaction system. The generated glycolate was detected by HPLC.

1.23 Conversion of commercial GA/DHA to glycolate by H_2O_2 at different pH

Phosphate buffer (300 mM KPi, 5 mM $MgSO_4$) with different pH values (5.0, 7.0 and 8.0) and carbonate buffer (300 mM Na_2CO_3 and $NaHCO_3$, 5 mM $MgSO_4$) with different pH values (9.0 and 10.0) were prepared. GA solution (140 mM), DHA solution (140 mM) and H_2O_2 solution (560 mM) were prepared with the above-mentioned buffer solutions with varied pH value. The reaction was initiated by adding GA solution (500 μL, 140 mM) or DHA solution (500 μL,140 mM) to H_2O_2 solution (500 μL, 560 mM) and conducted for 24 h (50°C, 1000 rpm). The reaction solution was added to $Na_2S_2O_3$ solution with the same volume to consume the remaining H_2O_2 in the reaction system. The generated glycolate was detected by HPLC.

1.24 Chemoenzymatic synthesis of glycolate from formaldehyde by M4V2 and H_2O_2

The reaction was executed in 2.0 mL microcentrifuge tube containing 600 mM formaldehyde, 0.0886 mM M4V2 and 0.5 mM TPP with shaking (1000 rpm, 30°C) for 3h. The generated GA and DHA were detected by HPLC. Subsequently, 0.5 mL enzymatic reaction mixture was added in a 1.5 mL microcentrifuge tube which containing 0.5 mL H_2O_2 solution (876 mM) which was dissolved in 550 mM phosphate buffer (550 mM KPi, 5 mM $MgSO_4$, pH 8.0). All reactions were conducted at 50°C for 24 h. The reaction solution by H_2O_2 was added in same-volume $Na_2S_2O_3$ solution to remove H_2O_2 in the reaction system and the generated glycolate was detected by HPLC.

2. Results and Discussion

2.1 The yield of glycolate generated from the oxidation of GA by $NaClO_2$ at different temperature

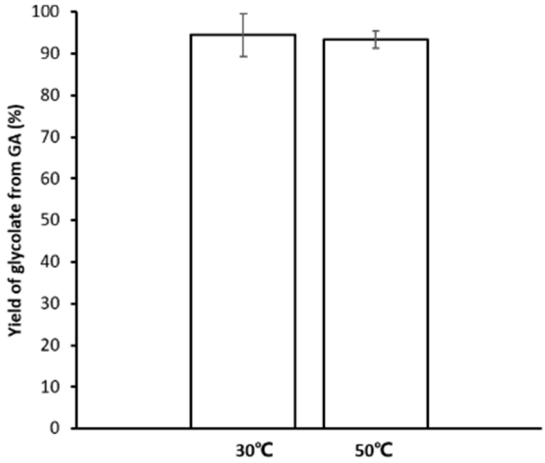

Figure S1. The yield of glycolate generated from the oxidation of GA by $NaClO_2$ at 30℃ and 50℃. The molar ratio between oxidant and substrate was 3:1. The reaction was performed in NaPi buffer (50 mM, pH 8.0) for 24 h

2.2 Structure analysis of the formolase variant M4V2

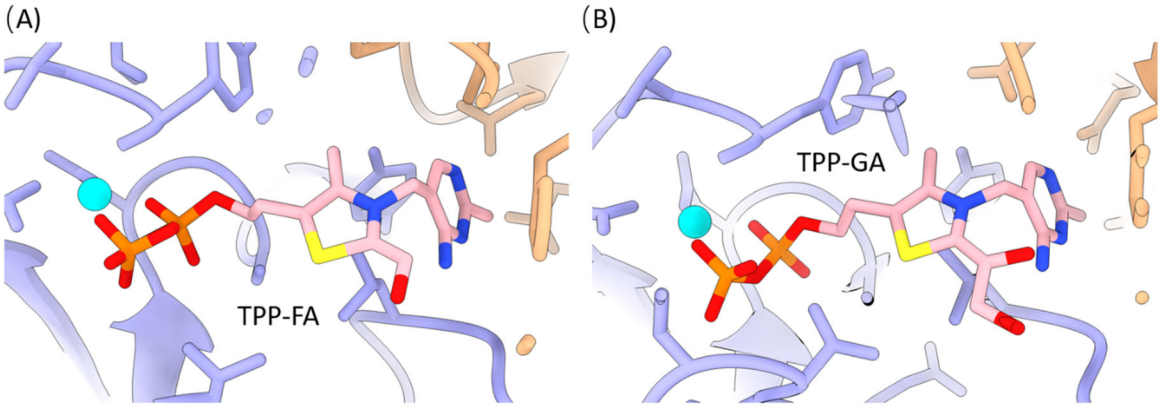

Figure S2. A binding mode of the intermediates TPP-FA (complex of thiamine diphosphate and C1 formaldehyde) (A) and TPP-GA (complex of thiamine diphosphate and C2 glycolaldehyde) (B) in binding pocket. Mg^{2+} was colored in cyan.

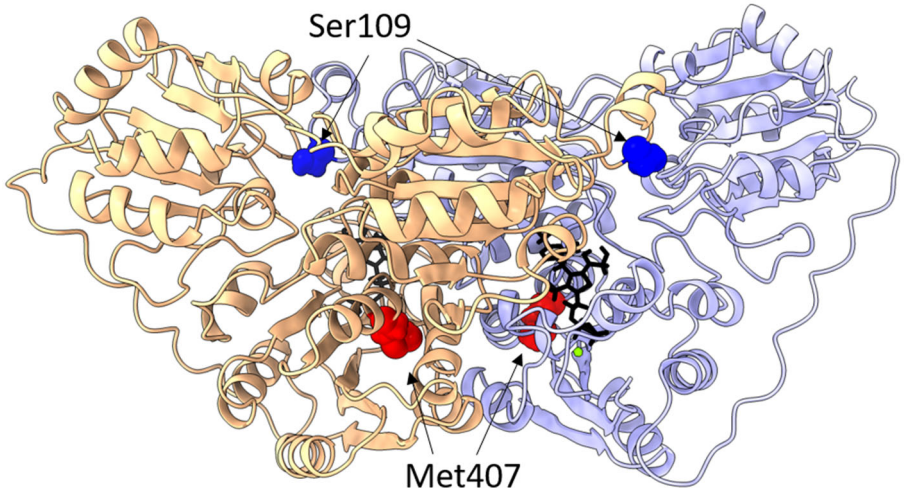

Figure S3. Dimeric structure model of M4V2, the residue Ser109 was colored blue, and residue Met407 was colored in red.

Table S1. The MM/GBSA binding free energy of TPP intermediate with M4 (parent) and M4V2.

Formolase variants	TPP-FA-1	TPP-FA-2	TPP-GA-1	TPP-GA-2
M4	-137.18±7.25	-128.97±8.32	-112.80±6.98	-121.15±8.05
M4V2	-157.68±8.47	-144.83±9.41	-130.52±7.69	-134.25±8.98

Table S2. The MM/GBSA binding free energy between monomers in M4 (parent) and M4V2.

Formolase variants	TPP-FA	TPP-GA
M4	-193.45±10.16	-187.22±10.28
M4V2	-201.76±9.75	-210.85±12.75

Table S3. The average RMSF of amino acid residues around position 407.

Formolase variants	TPP-FA	TPP-GA
M4	0.6126	0.5874
M4V2	0.5872	0.5554

2.3 The influence of the concentration of TPP in the M4V2 catalytic system

As shown in Figure S2, when the concentration of TPP is 0.3 mM to 0.5 mM, the total yield of GA and DHA are comparable. 0.1-0.2 mM TPP resulted in 84.3% and 90.8% total products of that of 0.5 mM TPP, respectively. When no external TPP was added to the system, the yield of GA and DHA (32.1 mM and 59.3 mM respectively) was 42.9 % of that of 0.5 mM TPP, indicating substantial amount of TPP was produced by the *E.coli* host cell and bounded with the generated enzyme M4V2. Therefore, the concentration of TPP could be reduced from 0.5 mM to 0.2 mM, or even to 0.1 mM for the reaction system utilizing purified M4V2.

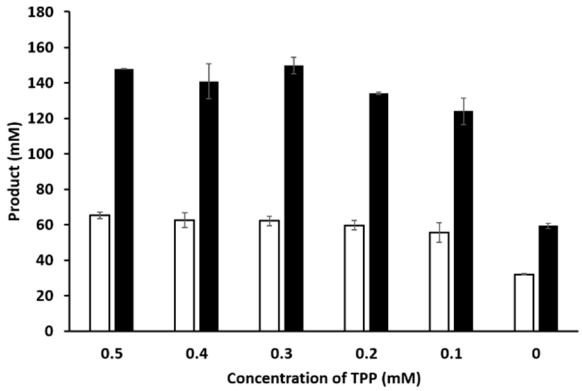

Figure S4. Yield of GA (white bar) and DHA (black bar) produced by M4V2 with different concentration of TPP; 600 mM formaldehyde was used as substrate.

2.4 The influence of the concentration of M4V2 in the M4V2 catalytic system

As shown in Figure S7, the total yield of GA and DHA maintained 95.5% when decreasing the enzyme concentration decreased from 0.0886 mM to 0.0709 mM. The yield of GA and DHA decreased to 78.4% and 71.6% when decreasing the enzyme concentration from 0.0886 mM to 0.0532 mM and 0.0355 mM, respectively.

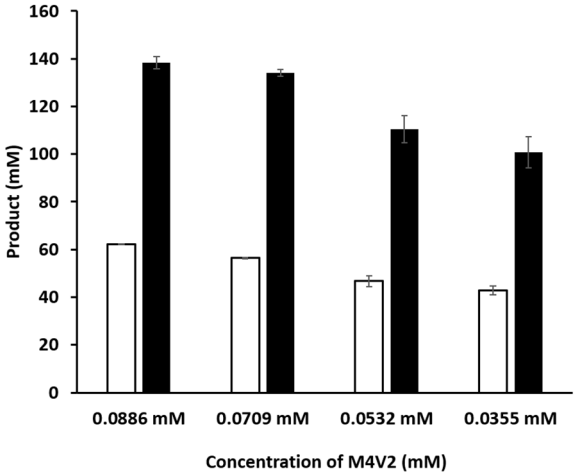

Figure S5. Yield of GA (white bar) and DHA (black bar) produced by different concentration of M4V2. 600 mM formaldehyde was used as substrate.

2.5 Conversion of commercial GA/DHA to glycolate with different H_2O_2: substrate ratio

As shown in Figure S4, the molar ratio between oxidant (H_2O_2) and substrate (GA or DHA) were optimized to achieve highest yield. 4:1 was the best oxidant: substrate ratio. Over all, the yield of glycolate from GA and DHA oxidized by H_2O_2 were lower than that of $NaClO_2$.

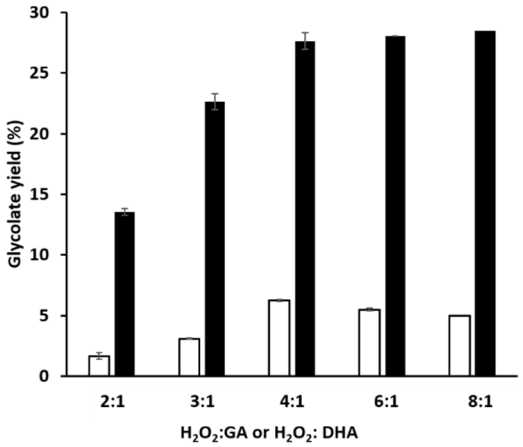

Figure S6. The effect of different molar ratio (H_2O_2/GA and H_2O_2/DHA) on the glycolate yield (phosphate buffer (50 mM, pH 8.0), 5 mM $MgSO_4$ 30°C, 24 h); 70 mM GA (white bar) or 70 mM DHA (black bar) were used as substrates.

2.6 Conversion of commercial GA/DHA to glycolate by H_2O_2 at different temperature

As shown in Figure S5, with the increased of the temperature of reaction system, the yield of glycolate from GA or DHA were gradually increased. Therefore, the H_2O_2 oxidation of GA and DHA were performed at 50°C in the subsequent experiments.

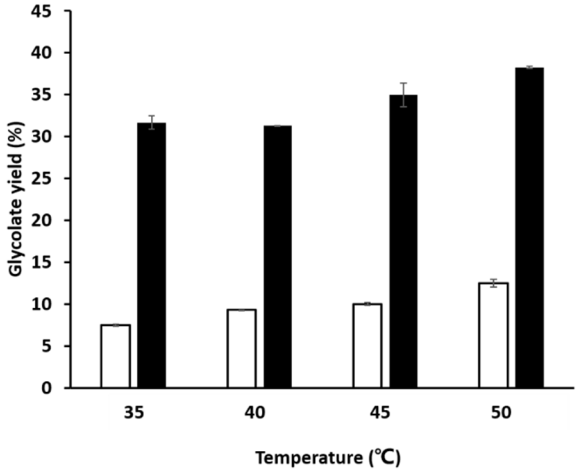

Figure S7. The effect of different temperature on the glycolate yield using 70 mM GA (white bar) or 70 mM DHA (black bar) as substrate (280 mM H_2O_2, phosphate buffer (50 mM, pH 8.0), 5 mM $MgSO_4$, 24 h).

2.7 Optimization of the concentration of phosphate buffer for the conversion of commercial GA or DHA to glycolate in presence of H_2O_2

As shown in Figure S6, with the increased concentration of phosphate buffer, the glycolate yield from GA or DHA were gradually increased. The highest glycolate yield from GA (40.12%) or DHA (64.18%) was achieved in 300 mM KPi buffer after 24 h reaction. Thus, the following oxidation reaction was performed at 50°C in 300 mM KPi buffer.

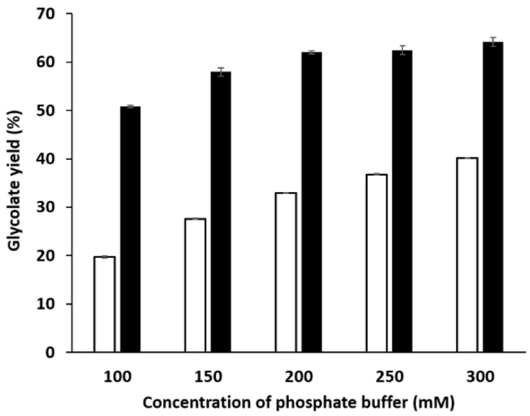

Figure S8. The effect of the concentration of phosphate buffer on the glycolate yield using 70 mM GA (white bar) and 70 mM DHA (black bar) as substrate (280 mM H_2O_2, phosphate buffer (pH 8, 100-300 mM), 5 mM $MgSO_4$, 50°C, 24 h).

2.8 Conversion of commercial GA/DHA to glycolate by H_2O_2 at different pH

As shown in Figure S7, the total yield of GA or DHA to glycolate were gradually increased with the pH elevated from 5 to 8, but the yield of glycolate from GA dramatically decreased at pH 9 and 10. Therefore, the following oxidation reaction was performed at 50°C in 300 mM KPi (pH 8) for 24 h where the molar ratio of H_2O_2/substrate was 4:1.

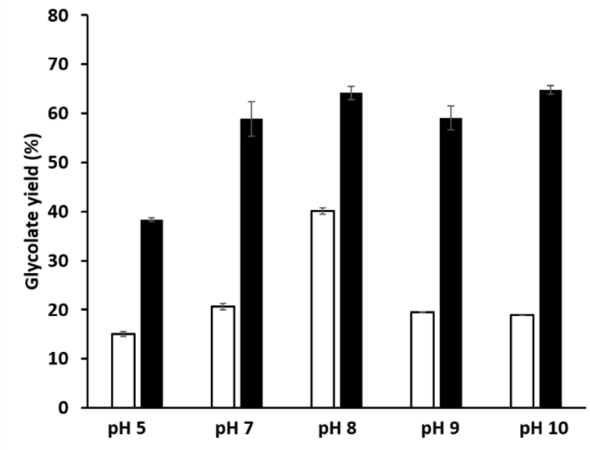

Figure S9. Oxidation of GA (70 mM, white bar) or DHA (70 mM, black bar) into glycolate with H_2O_2 (280 mM) at different pH (pH 5.0-8.0: 300 mM phosphate buffer; pH 9-10: 300 mM carbonate buffer) supplemented with $MgSO_4$ (5 mM) at 50°C for 24 h.

2.9 Comparison of the conversion of GA and DHA into glycolate by using $NaClO_2$ and H_2O_2

We compared the yield of the reaction catalyzed with $NaClO_2$ and H_2O_2 under their optimal reaction conditions. As shown in Table S1, the optimal oxidant: substrate ratio for H_2O_2 was 4:1, while for $NaClO_2$ it was 3:1. The glycolate yield of the reaction catalyzed by H_2O_2 was 67.7% of $NaClO_2$, suggesting that H_2O_2 could be an alternative oxide to produce glycolate from GA and DHA.

Table S4. Comparison of the reaction catalyzed with $NaClO_2$ and H_2O_2 under their optimal reaction conditions. 600 mM formaldehyde was used as substrate in the enzymatic step.

Oxidants	Phosphate buffer (mM)	oxidant: substrate ratio (molar ratio)	Glycolate yield (mM)
$NaClO_2$	150	3:1	94.80±6.52
H_2O_2	300	4:1	64.25±0.65

3. References

1. T. Li, Z. Tang, H. Wei, Z. Tan, P. Liu, J. Li, Y. Zheng, J. Lin, W. Liu, H. Jiang, H. Liu, L. Zhu and Y. Ma, *Green Chem.*, 2020, **22**, 6809-6814.
2. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
3. D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel and P. Kollman, *Comput. Phys. Commun.*, 1995, **91**, 1-41.
4. H. C. Andersen, *J. Comput. Phys.*, 1983, **52**, 24-34.
5. K.-H. Chow and D. M. Ferguson, *Comput. Phys. Commun.*, 1995, **91**, 283-289.
6. S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374-380.