Stand-alone lipoylated H-protein of the glycine cleavage system enables glycine cleavage and the synthesis of glycine from one-carbon compounds *in vitro*

Yingying Xu[1], Yuchen Li[1], Han Zhang[1], Jinglei Nie[1], Jie Ren[3], Wei Wang[2] and An-Ping Zeng[1,2]*

[1]Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, North Third Ring Road 15, Chaoyang District, 100029, Beijing, China

[2]Institute of Bioprocess and Biosystems Engineering, Hamburg University of Technology Denickestrasse 15, D-21073 Hamburg, Germany

[3]State Key Laboratory for Biology of Plant Diseases and Insect Pests/Key Laboratory of Control of Biological Hazard Factors (Plant Origin) for Agri-product Quality and Safety, Ministry of Agriculture, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100081, China

* Corresponding author: Prof. An-Ping Zeng

Abstract

H-protein, one of the four component proteins (H, T, P and L) of glycine cleavage system (GCS), is generally considered a shuttle protein interacting with the other three GCS-proteins via a lipoyl swinging arm. We report that without P-, T- and L-proteins, lipoylated H-protein (H_{lip}) enables GCS reactions in both glycine cleavage and synthesis directions *in vitro*. This apparent catalytic activity is closely related to the cavity on the H-protein surface where the lipoyl arm is attached. Heating or mutation of selected residues in the cavity destroys or reduces the stand-alone activity of H_{lip}, which can be restored by adding the other three GCS-proteins. Systematic study of the H_{lip}-catalyzed overall GCS reactions and the individual

reaction steps provides a first step towards understanding the stand-alone function of H_{lip}. The results in this work provide some inspiration for further understanding the mechanism of the GCS and give some interesting implications on the evolution of the GCS.

Keywords: H-protein, glycine cleavage system, glycine synthesis, One-carbon metabolism

Significance statement

Glycine cleavage system (GCS) plays central roles in C1 and amino acids metabolisms and the biosynthesis of purines and nucleotides. Manipulations of GCS are desired to promote plant growth or to treat serious pathophysiological processes such as aging, obesity and cancers. Reversed GCS reactions form the core of the reductive glycine pathway (rGP), one of the most promising pathway for the assimilation of formate and CO_2 in the emerging C1-synthetic biology. H-protein, one of the four GCS component proteins (H, T, P and L) is generally considered a shuttle protein interacting with the other three proteins via a lipoyl swinging arm. Here, we discovered that without P-, T- and L-proteins, H-protein alone can catalyze GCS reactions in both glycine cleavage and synthesis directions *in vitro*. The surprising catalytic activities are related to a structural region of H-protein which can be manipulated. The results have impacts on engineering GCS to treat related diseases, to improve photorespiration, and to efficiently use C1-carbon for biosynthesis.

Introduction

In the mitochondria of plant and animal cells as well as in the cytosol of many bacteria, the glycine cleavage system (GCS) comprising four (H, T, P and L) proteins catalyzes the reversible decarboxylation and deamination of glycine to yield CO_2, NH_3 and provide a methylene group for the conversion of tetrahydrofolate (THF) to N^5,N^{10}-methylene-tetrahydrofolate (5,10-CH_2-THF)[1, 2, 3]. The overall reaction cycle catalyzed by GCS comprises three steps as illustrated in **Figure 1a** (hereinafter GCS is used to refer the four-enzyme system regardless of reaction direction). The reaction

is first catalyzed by P-protein (glycine decarboxylase; EC 1.4.4.2) to yield CO_2 from glycine and methylamine-loaded H-protein (H_{int}) from the oxidized form (H_{ox}) of the lipoylated H-protein (H_{lip}). T-protein (aminomethyltransferase; EC 2.1.2.10) then catalyzes the release of NH_3 and transfer the methylene group from H_{int} to THF to form 5,10-CH_2-THF, leaving dihydrolipoyl H-protein (H_{red}). Finally, L-protein (dihydrolipoyl dehydrogenase; EC 1.8.1.4) catalyzes the oxidation of H_{red} to regenerate H_{ox} in the presence of NAD^+. H-protein as a shuttle protein interacts with the other three GCS-proteins via a lipoyl swinging arm and plays central role in the GCS.

The physiological roles of GCS in various organisms have been well studied. In human and most vertebrates, GCS is part of the serine and glycine metabolism pathway. Serine is catalyzed by serine hydroxymethyltransferase (SHMT) to form glycine and 5,10-CH_2-THF, and then the product glycine is degraded by GCS. The final product, 5,10-CH_2-THF, whose methylene group derived from the β-carbon of serine or the α-carbon of glycine, is one of the few C1 donors in the biosynthesis process, such as the biosynthesis of purine and methionine[2]. Decrease or loss in the activity of GCS will lead to glycine accumulation in human body, which is linked to glycine encephalopathy[4] (also known as nonketotic hyperglycinemia). Relevant studies have shown that most patients with glycine encephalopathy have a P-protein deficiency, and the rest are caused by T-protein or H-protein deficiency[5]. Moreover, recent studies have shown that glycine metabolism is associated with tumorigenesis, and P-protein as a key factor regulates glycolysis and methylglyoxal production in cancer cells[6, 7]. In C3 plants, GCS is the key enzymatic system that deals with a large amount of glycine in mitochondria during the photorespiration, and the activity of GCS directly determines the growth rate of plants. Knockout of GCS gene is lethal to plants, which is relevant to impaired one-carbon metabolism[8], whereas overexpression of L-protein[9] or H-protein[10] have been shown to improve photorespiration rates for further increasing biomass yield.

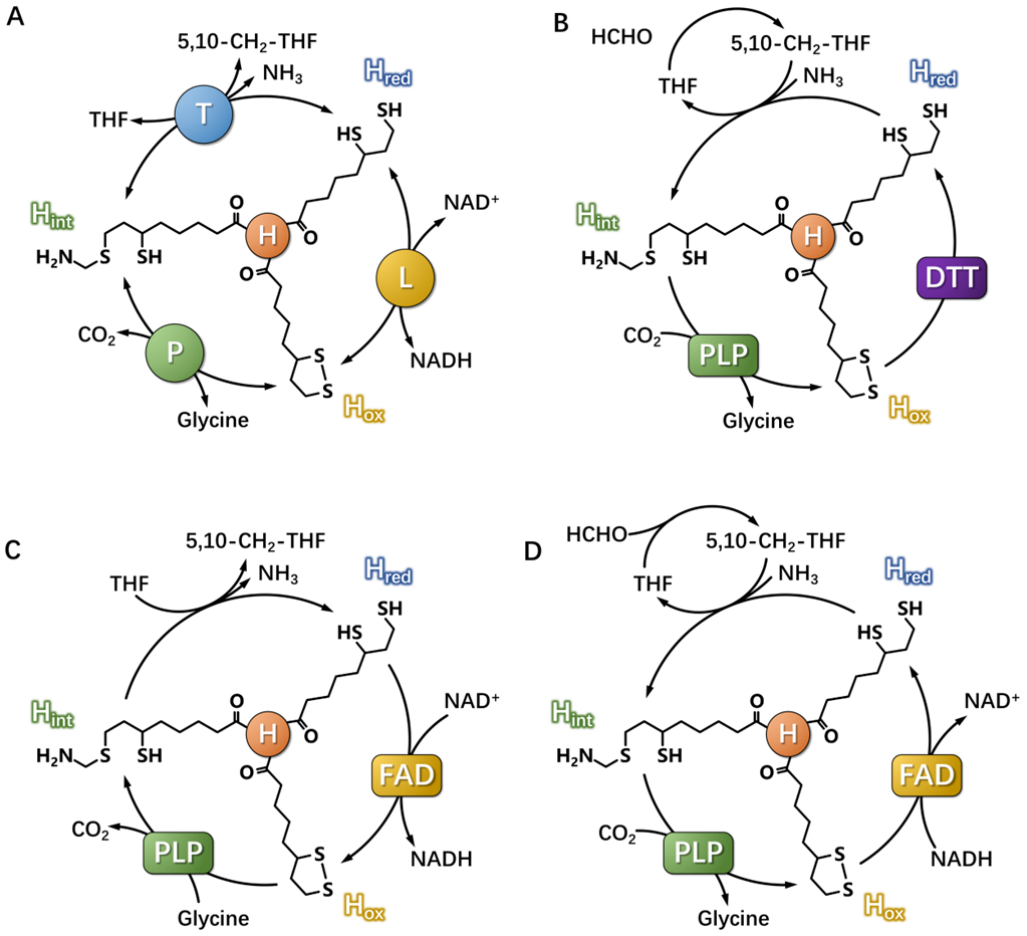


Figure. 1. Schematic diagrams of the reversible glycine cleavage reaction catalyzed by GCS or stand-alone H-protein (H_{lip}). (a) The glycine cleavage and synthesis reactions catalyzed by GCS with the complete set of enzymes; (b) Glycine synthesis reactions catalyzed by H_{lip} alone under the presence of PLP, THF and DTT (DTT as a reductant replacing the functions of FAD and NADH); (c) Glycine cleavage and (d) synthesis reactions catalyzed by H_{lip} alone under the presence of PLP, THF, FAD and NADH.

Although GCS in most organisms runs mostly in the direction of glycine cleavage, it catalyzes glycine synthesis in a few anaerobic bacteria such as *Clostridium acidiurici*[11], *Eubacterium acidaminophilum*[12] and *Arthrobacter globiformis*[13, 14]. The reversibility of GCS was first discovered in mitochondrial extract of rat liver[15], in *A. globiformis*[16] and in cock liver mitochondria[17], and most of these studies were already carried out in 1960-1980s. Today, attributing to the reversibility of GCS, GCS gains renewed attention of researchers, because reductive glycine pathway (rGP), with GCS as its key component pathway, is considered to be the most promising synthetic pathway for the assimilation of formate and CO_2 to produce pyruvate[18], a key

precursor that enters the central metabolic pathway for cell growth and biosynthesis. Sánchez-Andrea *et al.*[19] discovered that rGP functions in an anaerobic sulphate-reducing bacterium *Desulfovibrio desulfuricans*, and stated that it represents the seventh natural CO_2 fixation pathway. Recently, rGP has been successfully introduced into *E.coli*[20, 21, 22, 23] for autotrophic growth on formate and CO_2. At the same time, part of this pathway was successfully transferred into *Saccharomyces cerevisiae*[24], *Cupriavidus necator*[25] and *Clostridium pasteurianum*[26]. However, the flux of rGP is quite low, which limits the growth of microorganism. It has been pointed out that the reaction catalyzed by GCS is the rate-limiting step in rGP[21]. Therefore, it is particularly important to understand the catalytic mechanism of GCS for increasing the flux of rGP. Substantial progress has been made in understanding the catalytic properties of GCS, and H-protein is so far considered to function merely as a shuttle protein of the cofactor lipoic acid. Lipoic acid is attached by an amide linkage to the conserved lysine residue of H-protein at the 64th position, and the lipoylated H-protein (H_{lip}) plays a pivotal role acting as a mobile substrate which undergoes a cycle of reductive methylamination, methylamine transfer and electron transfer in the enzymatic cycle of GCS[27].

In this work, we discovered that H_{lip} alone can enable the GCS reaction cycle in both glycine cleavage (**Figure 1b**) and synthesis directions (**Figures 1c and 1d**) in the absence of P-, T- and L-proteins. The formation of glycine from C1 compounds in the presence of suitable cofactors was demonstrated by choosing HCHO as the source of α-carbon of glycine. More detailed analyses led to the striking finding that H_{lip} can apparently “catalyze” all the GCS reaction steps previously believed to be solely catalyzed by P, T and L-proteins, respectively. These findings not only shed new light into the functions of H-protein, but also provide useful hints for engineering H-protein and GCS, either for treating diseases such as hyperglycinemia, for enhancing biomass yield in plants, or for developing synthetic pathways for technical use of C1-carbons. The fact that stand-alone H_{lip} can catalyze the synthesis of the basic amino acid glycine from inorganic compounds may also have important implications

for the evolution of life.

Results

Effects of components of the GCS reaction system on glycine cleavage and synthesis

On the basis of previous studies[8, 28], we successfully constructed GCS catalyzed glycine cleavage and synthesis reactions *in vitro*. Normally, all the four GCS enzymes are included in the reaction system. During kinetic studies, we found that the reactions of glycine cleavage and glycine synthesis can also occur in the absence of certain GCS enzymes and reaction components. This triggered us to systematically examine the effects of missing a certain component or enzyme in the reaction mixture on the reaction rate of both reaction directions. As shown in **Table 1**, the lack of a certain component or enzyme can cause very different changes of the reaction rate. As expected, the reaction did not occur in the absence of essential substrates (glycine in the cleavage direction and NH_4HCO_3 in the synthesis direction). The presence of the H_{ox} was also vital, as no reaction was observed in the absence of H_{ox}. However, varied reaction rates (10-76 % of the reference values) were observed when only one of the P-, T- and L-proteins was missing. Compared to the effects of GCS proteins, the missing of substrates and cofactors (THF, PLP, NAD or NADH) showed often stronger effects on the cleavage and synthesis of glycine. In this context, the effects of PLP were surprising: (1) missing of both P-protein and PLP resulted in neither cleavage nor synthesis of glycine; (2) while missing PLP alone resulted in strongly impaired glycine synthesis, it had, however, no negative effect on glycine cleavage. This might be partially explained by the fact that PLP is covalently bound to P-protein[29, 30, 31]. Therefore, P-protein expressed in *E. coli* might have PLP covalently bound to it during its expression. The purified P-protein might still have PLP attached to it and can therefore function well in decarboxylation without externally adding PLP. On the contrary, the effect of PLP absence was even worse than the absence of P-protein for glycine synthesis which implied the importance of PLP for the stand-alone catalytic activity of H_{lip}.

Table 1. Effects of missing a certain component of the GCS reaction system on the rates of glycine cleavage (determined as HCHO formation from the degradation of 5,10-CH_2-THF) and glycine synthesis.

Missing component	Glycine cleavage reaction		Glycine synthesis reaction	
	(μM HCHO • min^{-1})	GCS/ %	(μM glycine • min^{-1})	rGCS/ %
None (Reference)	22.48±3.47	100.00	5.95±0.13	100.00
P-protein	2.32±0.52	10.34	2.03±0.20	34.07
T-protein	11.67±0.42	51.91	4.55±0.16	76.53
L-protein	8.44±0.57	37.55	4.45±0.15	74.78
H_{ox}	0.00	0.00	0.00	0.00
P-protein+PLP	0.00	0.00	0.00	0.00
PLP	24.92±2.67	110.86	0.98±0.13	16.42
T-protein+THF	1.11±0.33	4.93	0.43±0.08	7.20
THF	0.87±0.15	3.88	0.53±0.04	8.94
NAD^+/NADH	4.88±1.54	21.73	5.76±0.05	96.72
Glycine	0.00	0.00	-	-
NH_4HCO_3	-	-	0.00	0.00
HCHO	-	-	0.98±0.02	16.42

H_{lip} alone enables glycine synthesis and glycine cleavage reactions

The results in **Table 1** suggested that P-protein, T-protein and L-protein are not essential for the functionality of GCS both in glycine cleavage and glycine synthesis directions. This led us to the question if H_{lip} alone can “catalyze” glycine formation from NH_4HCO_3 and HCHO, or glycine cleavage in the opposite direction.

For glycine synthesis, the experimental results with H_{lip} as the only GCS protein in an array of reaction mixtures are presented in **Figure 2a**. Compared with the glycine

synthesis catalyzed by the all four GCS proteins (GCS as control), the reaction rate catalyzed by H_{lip} alone at the same concentration of 10 μM (H-10) was somewhat lower, but glycine formation was well detected. With the increase of H_{ox} concentration the initial reaction rate increased and the final concentration of glycine synthesized was higher than that of the control. The above results prove that H_{lip} alone can apparently catalyze the synthesis of glycine from NH_4HCO_3 and HCHO in the presence of THF, PLP and DTT. To get the optimum reaction conditions for the glycine synthesis, the reaction rate of glycine synthesis catalyzed by H_{lip} alone was investigated under conditions of using different buffers, temperatures and pHs. **Figure 2b** shows the effect of different types of buffer, in which the order of the catalytic ability of H_{lip} was as follows: Tris-HCl≈Mops > HEPES > PBS. The effect of temperature and pH on the activity were studied by changing temperature from 4 °C to 95 °C (**Figure 2c**), and pH from 6.0 to 10.0 (**Figure 2d**). The reaction rate decreased sharply when the temperature was higher than 42 °C or pH was higher than 8.0. The optimum temperature and pH were at 37-42 °C and 7.5-8.0, respectively.

For glycine cleavage, the reaction could not be observed even at high H_{lip} concentrations (up to 80 μM) using the same reaction mixture as used for the GCS-catalyzed glycine cleavage reaction but without P-, T- and L-proteins. Later, we found out that when FAD, the coenzyme of L-protein, was added, H_{lip} alone was indeed able to activate the glycine cleavage, and the reaction rate increased with the increase of H_{lip} concentration, as shown by the time-courses of NADH formation (**Figure 2e**) and initial rates of glycine cleavage (**Figure 2f**). The essentiality of FAD for the glycine cleavage but not for glycine synthesis catalyzed by stand-alone H_{lip} is due to the presence of DTT which can convert H_{ox} to H_{red} required in the direction of glycine synthesis (details see below).

H-protein is a small heat-stable protein, so heating does not lead to precipitation. In literature, thermal stability of H-protein is therefore used to terminate the lipoylation of H-protein catalyzed by the enzyme lipoate-protein ligase A (LplA), in which LplA is completely denatured and precipitated [32, 33]. We have tried to use heated H_{ox} (at 95 °C

for 5 min) to catalyze the reactions of glycine synthesis and cleavage, but no reaction in either direction was observed (the details are discussed in a later section). We therefore speculate that the structure of H_{lip} was altered by heating at high temperature, which made it lose its catalytic activity shown above for glycine synthesis and cleavage.

In order to explore the reasons behind the function of the stand-alone H_{lip} observed, we further studied the effect of H_{lip} alone on the three GCS reaction steps, i.e., the glycine decarboxylation reaction (accompanied by the reductive aminomethylation of H_{ox} to H_{int}) in the absence of P-protein, the aminomethyl transfer reaction in the absence of T-protein, and the electron transfer reaction without the presence of L-protein, respectively.

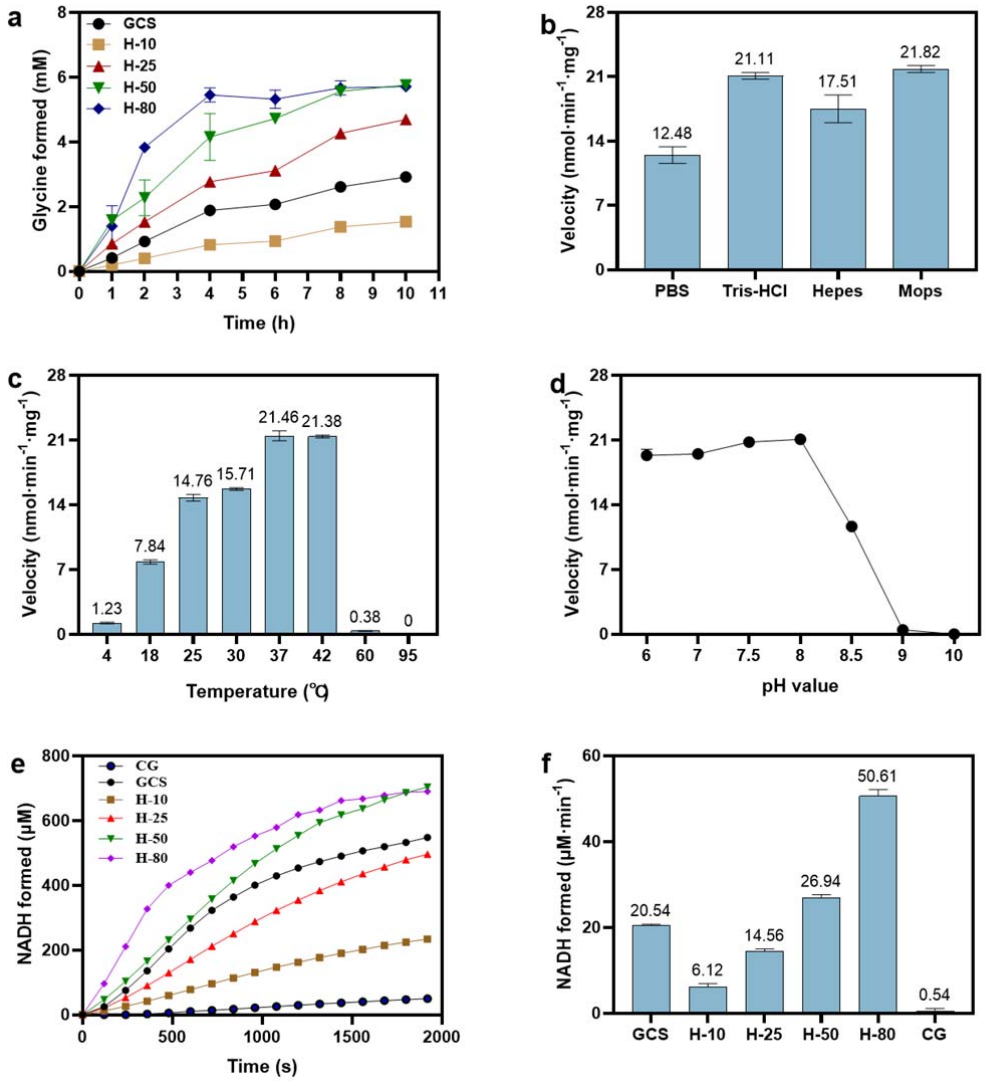


Figure 2. H_{lip} alone enabled glycine synthesis and glycine cleavage. Effects of H_{ox} concentration (a), buffer (b), temperature (c) and pH (d) on glycine synthesis. "GCS" refers to a reaction mixture for glycine synthesis as specified in "Materials and Methods" without missing any reaction components and enzymes; "H-10", "H-25", "H-50" and "H-80" were the same reaction mixture containing no P-, T- and L-proteins but only H_{ox} at 10 μM, 25 μM, 50 μM and 80 μM, respectively. (e) Effects of H_{ox} concentration on glycine cleavage. "CG" refers to no GCS enzymes in the reaction mixture, "GCS" refers to a reaction mixture for glycine cleavage as specified in "Materials and Methods" without missing any reaction components and enzymes; "H-10", "H-25", "H-50" and "H-80" were the same reaction mixture containing no P-, T- and L-proteins but only H_{ox} at 10 μM, 25 μM, 50 μM and 80 μM, respectively. (f) NADH formation rate in glycine cleavage catalyzed by different concentrations of H_{ox}.

Decarboxylation and carboxylation reactions in the absence of P-protein

The results in **Table 1** revealed that no activity could be measured for either the

cleavage or the synthesis of glycine, when both P-protein and PLP were absent in the reaction mixtures. However, activities were observed when only P-protein was missing. We therefore speculate that the presence of PLP alone might be sufficient to enable H_{lip} to catalyze the decarboxylation/carboxylation reaction normally catalyzed by P-protein. This was confirmed by the HPLC results shown in **Figure 3a** for the decarboxylation reaction in the glycine cleavage direction. H_{int} was formed from H_{ox} without the presence of P-protein. This astonishing result suggests that glycine decarboxylation activated by H_{lip} alone can occur independent of P-protein, as long as PLP is present (**Figure 3a**) under the experimental conditions of this study.

For the carboxylation reaction in the glycine synthesis direction, in the absence of P-protein, glycine formation could be still detected and the reaction rate showed a nearly linear increase with the PLP concentration in the low PLP concentration range, albeit that the reaction rate was lower than those determined in the presence of P-protein (**Figure 3b**). This is understood that CO_2 fixation (carboxylation) has a higher energy barrier and P-protein is needed for its activation.

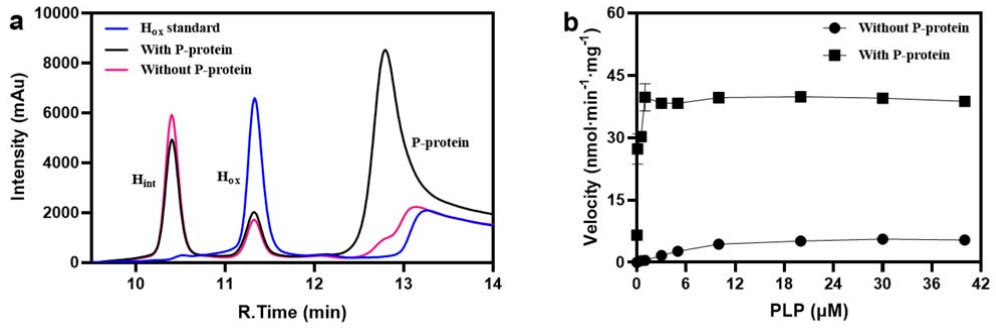


Figure 3. Examination of H_{lip} (H_{ox}) for the function of P-protein. (**a**) Formation of H_{int} from H_{ox} during the glycine decarboxylation reaction was determined using HPLC to show that H_{lip} functions as decarboxylase in the presence of PLP. Chromatograph "With P-protein" refers to a reaction mixture containing 50 mM glycine, 50 μM H_{ox}, 25 μM PLP, and 5 μM P-protein; Chromatograph "Without P-protein" refers to a reaction mixture similar to that "with P-protein" but without the addition of P-protein; Chromatograph "H_{ox} standard" refers to a test solution containing only H_{ox}; (**b**) PLP-dependent glycine formation was determined to show that carboxylation can take place in the absence of P-protein, albeit at lower reaction rate. "With P-protein" refers to a reaction mixture for glycine synthesis as specified in "Materials and Methods" without missing any reaction components and enzymes, "Without P-protein" refers

to a reaction mixture containing all reaction components and enzymes except for P-protein.

Aminomethyl transfer reaction in the absence of T-protein

According to the results in **Table 1**, the overall GCS reaction in both glycine cleavage and glycine synthesis directions could still precede reasonably well in the absence of T-protein in the reaction mixtures. In comparison, it is obvious that the absence of THF had a more significant negative effect on the reaction rate regardless of the reaction directions, i.e. with a reduction of the reaction rate for over 96 % in glycine cleavage and 91 % in glycine synthesis. **Figure 4a** shows the change of NADH production with time in the direction of glycine cleavage under different conditions. The initial rate without adding T-protein was still more than half of that with adding T-protein, whereas in the experimental group "without THF" the formation of NADH could be hardly detected. For the aminomethyl transfer reaction in the direction of glycine synthesis, **Figure 4b** shows that H_{int} could still be generated without adding T-protein in the reaction mixture. It was found by Kochi *et al*[14] that when dihydrolipoic acid, HCHO and NH_4^+ were mixed together, compounds in the form of -S-CH_2NH_2 could be obtained. The question then arises: is the aminomethylation of H_{red} to H_{int} in the absence of T-protein the result of a complete non-enzymatic reaction due to the presence of HCHO and NH_4Cl in which H-protein acts only as the shuttle protein?

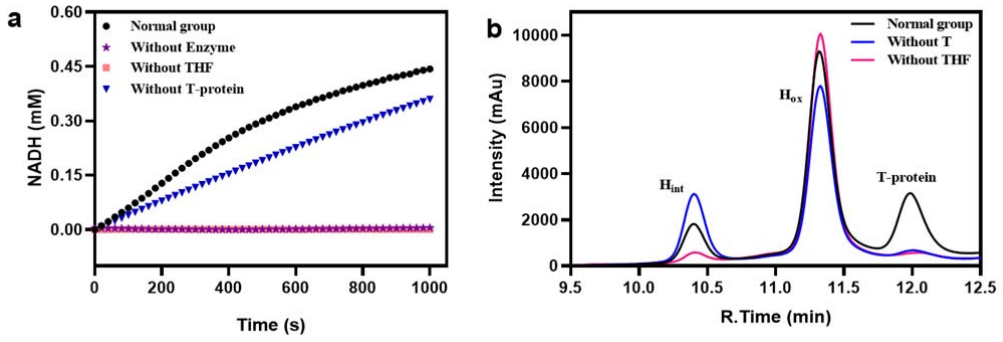


Figure 4. Examination of H_{lip} for the function of T-protein. (**a**) Effects of THF or T-protein absence on the overall glycine cleavage reaction rate. "Normal group" refers to a reaction mixture containing all reaction components and enzymes required; "Without Enzyme" refers to a reaction mixture containing all reaction components but no GCS enzymes; "Without THF" refers to a reaction mixture containing all reaction components and enzymes except for THF;

"Without T-protein" refers to a reaction mixture containing all reaction components and enzymes except for T-protein. (**b**) Detection of H_{int} by HPLC proved that the aminomethyl transfer reaction can take place spontaneously. "Normal group" refers to a reaction mixture containing 50 μM H_{ox}, 0.5 mM THF, 20 mM DTT, 50 mM NH4Cl, 10 mM HCHO, and 5 μM T-protein; The reaction mixtures of "Without T" and "Without THF" were the same as that of "Normal group", but no T-protein or THF was added, respectively.

Interconversion of H_{ox} and H_{red} in the absence of L-protein

The function of L-protein is to catalyze the interconversion between the H_{ox} and its reduced form dihydrolipoyl H-protein (H_{red}) involving NAD^+/NADH and FAD/$FADH_2$. The experimental results in **Table 1** show that in the absence of L-protein both the glycine cleavage and synthesis systems can still work. Previous studies[34, 35] have shown that the disulfide bond-reducing agent TCEP allows the reduction of the disulfide bond of the lipoyl group associated with the H-protein during the course of the reaction catalyzed by L-protein. Since a certain amount of DTT was added to prevent the oxidation of THF during the preparation of THF stock solution, we expected that DTT should have the same reduction function as TCEP. **Figure 5a** schematically shows the interconversion between H_{ox} and H_{red} through the combination of the reduction of H_{ox} by TCEP (Left) or DTT (Right) with the re-oxidation of H_{red} catalyzed by L-protein, yielding thereby NADH from added NAD^+. The functionality of this combination was demonstrated by the relevant results presented in **Figure 5b**. Next, we compared the effects of three different reducing agents (TCEP, DTT and β-ME) on the reducibility of H_{ox} to H_{red} in different buffers (Tris-HCl, PBS, Mops and HEPES) (**Figure 5c**). It was found that the reducibility of H_{ox} by DTT in four different buffers is PBS > Mops > Tris-HCl > HEPES; TCEP had no reducing effect on H_{ox} in Tris-HCl but showed the strongest reducing power in PBS, even much better than DTT. β-ME was obviously not suitable to reduce H_{ox} in these four buffers. Thus, for glycine synthesis, the functionality of L-protein can be well replaced by DTT or TCEP, as also evidenced by the result shown in **Table 1** that the absence of NAD^+/NADH did not affect glycine synthesis.

In order to verify that H_{lip} has the catalytic function of L-protein, we used FAD, the redox coenzyme of L-protein, instead of L-protein in combination with DTT to observe whether the redox reaction can still occur. To this end, NADH formation as the result of the redox reaction of H_{lip} (**Figure 5a**) by either L-protein or FAD was measured. Considering the results shown in **Figure 5d** and **Figure 5e**, the following conclusions can be drawn: (1) As long as L-protein is present, the redox reaction of the disulfide bond on the lipoyl group is comparable, with the lipoyl group bound to H-protein (H_{lip}) showing a slightly higher activity than that not bound (in lipoic acid). (2) In the absence of L-protein, the redox reaction of the lipoyl group bound to H-protein still occurred with the help of FAD, though to a lesser extent, but FAD was not able to replace L-protein in the reoxidation of dihydrolipoic acid, indicating that binding of the lipoyl group on H-protein is the prerequisite for the function of FAD in the absence of L-protein. To find out whether this is simply due to “fixation effect” of lipoyl group bound to H-protein which enables an easier approach of FAD to the lipoyl group, or this is facilitated through an unknown interaction of FAD with H-protein, we also examined to use heat-treated (95 °C for 5 min) H_{ox}, which still had the lipoyl arm linked to it. Interestingly, in the presence of L-protein, there was nearly no difference in the redox reaction of the lipoyl group between the heated H_{ox} and the unheated H_{ox}, however, FAD completely lost its function on the heated H_{ox}, clearly suggesting that an interaction of FAD with H-protein is required for its function in the absence of L-protein, and heating-induced structural change of H-protein destroyed the possibility of such interaction.

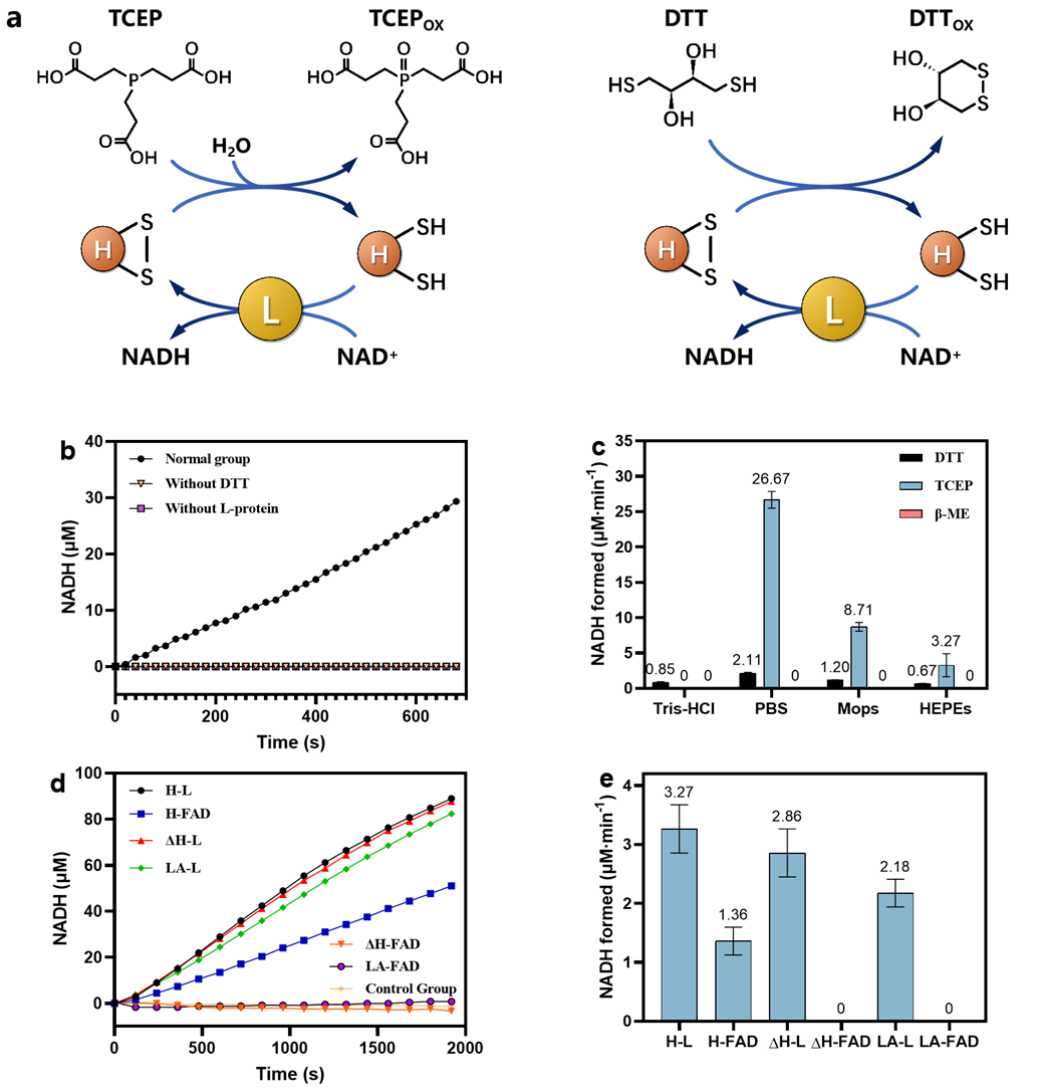


Figure 5. Examination of H_{lip} for the function of L-protein. (**a**) Interconversion between H_{ox} and H_{red} by combining the reduction of H_{ox} to H_{red} by TCEP (Left) or DTT (Right) with the re-oxidation of H_{red} to H_{ox} catalyzed by L-protein. (**b**) Reduction of the lipoamide group of H_{ox} by DTT. "Normal group" refers to a reaction mixture for electron transfer reaction as specified in "Materials and Methods" without missing any reaction components and enzymes. "Without DTT" refers to a reaction mixture similar to "Normal group" but without adding DTT. "Without L-protein" refers to a reaction mixture similar to "Normal group" but without adding L-protein; (**c**) Comparison of different disulfide reductants on the reduction of H_{ox} in different buffer solutions. (**d**) Time courses of NADH formation as the result of the redox reaction of H_{lip} (H) or heated H_{olip} (ΔH) or lipoic acid (LA) in the presence of either L-protein (L) or FAD. (**e**) NADH formation rates in the redox reaction of H_{ox}, heated H_{ox} and lipoic acid in the presence of either L-protein or FAD.

Possible reasons of apparent catalytic functions of H_{lip} in glycine cleavage and glycine synthesis

The above results show that for *in vitro* GCS reactions, H_{lip} alone enables both the glycine synthesis and the glycine cleavage without the presence of P-, T-, and L-proteins. It seems that H_{lip} might functionally replace at least P- and L-proteins and acts as glycine carboxylase and dihydrolipoyl dehydrogenase with the help of PLP and FAD, respectively. It is also suggested by the experiment results shown in **Figure 5d** and **5e** that heated H_{ox} lost the catalytic function of L-protein, obviously because heated H_{ox} cannot interact with FAD for the redox reaction of the lipoyl group bound to H-protein.

We therefore further systematically studied the effects of heating (95 °C for 5 min) on the catalytic activity of H_{lip} for glycine synthesis, in comparison to unheated H_{ox} as well as to lipoic acid. By either using H_{lip} alone or combined with other GCS enzymes, the overall reaction rate of glycine synthesis was measured. From the results in **Table 2,** we can ascertain several interesting observations and conclusions. First, the overall glycine synthesis reaction could be catalyzed by the unheated H_{ox} alone; adding P-protein significantly enhanced glycine synthesis, showing the importance of P-protein; the addition of either T- or L-protein has no positive effect. In fact, the *in vitro* glycine synthesis could run even better without T- and L-proteins, suggesting that in the presence of H_{lip} the two individual steps, namely aminomethyl transfer and electron transfer catalyzed by the two proteins respectively, could take place through spontaneous aminomethylation of H_{red} to H_{int} in the presence of HCHO and NH_4^+ and reduction of H_{lip} to H_{red} by DTT. Second, compared with the unheated H_{ox}, the heated H_{ox} alone could not catalyze the reaction of glycine synthesis; adding P-protein and L-protein partially revived glycine synthesis to different extents, indicating that heated H-protein lost the catalytic ability but was still functional as the shuttle protein of lipoyl group; the addition of T-protein did not bring any effect. Furthermore, to our surprise, when the heated H_{ox} was added with the other three GCS proteins to form a complete GCS, the rate of glycine synthesis was even higher than the GCS containing the unheated H_{ox}, indicating that heated H_{ox} loses its catalytic function but the heat-induced change is even beneficial for H-protein to exert its role as a

lipoyl-carrying protein to work together with the other three GCS proteins. Finally, in the presence of P-, T- and L-proteins, free lipoic acid can act as intermediate substrate to sustain the reversed GCS reaction towards gylcine synthesis under the given experimental conditions, even though to a much lower extent than that observed in the case of lipoyl group bound to H-protein.

Table 2. Synthesis of glycine using unheated H_{ox} or heated H_{ox} or lipoic acid at the same concentration of 10 μM, either alone or in varied combination with other GCS enzymes

Other component(s) added	H_{ox} ($\mu M \cdot min^{-1}$)	heated H_{ox} ($\mu M \cdot min^{-1}$)	Lipoate acid ($\mu M \cdot min^{-1}$)
No other GCS proteins	3.41± 0.18	0.00	0.00
P-protein	6.58± 0.19	1.54± 0.02	0.06± 0.01
T-protein	2.05± 0.38	0.00	0.00
L-protein	3.03± 0.11	0.98± 0.03	0.00
P-, T- and L-proteins	6.07± 0.04	8.45± 0.46	0.21± 0.01

Heating H_{lip} led to the loss of its catalytic activity regarding glycine synthesis. Although we found no change of the HPLC retention time of H_{lip} in the reversed-phase HPLC chromatograph after heating, which indicates no obvious change in the overall polarity and size of H_{lip}, heating may induce structural changes that are vital for the catalytic activity of H_{lip}. We therefore additionally performed HPLC analysis of H_{int} and H_{ox} to further determine the catalytic activity of H_{lip} in the two individual reaction steps normally catalyzed by P-protein and T-protein, respectively. As shown in **Figure 6a**, in the group of unheated H_{ox} the formation of H_{int} from H_{ox} clearly demonstrated that H_{lip} was not only a lipoyl-carring protein but could also replace P-protein in catalyzing the glycine decarboxylation reaction. H_{int} was not detected in the group of heated H_{ox}, indicating that heated H_{ox} lost the catalytic activity of P-protein. **Figure 6b** shows the reaction results of aminomethyl transfer from H_{red} (generated *in situ* from H_{lip}) to H_{int}. The group of unheated H_{ox} clearly exhibited the catalytic activity of T-protein; in the

group of heated H_{ox} there was a small amount of H_{int} formation found, however, this low level of aminomethyl transfer activity was most likely not due to a catalytic activity of T-protein manifested by heated H_{ox}, but due to a spontaneously occurring aminomethyl transfer reaction. Since heated H_{ox} still preserves its function as shuttle protein, the difference between the unheated and heated H_{ox} in the aminomethyl transfer reaction may provide an answer to the question in section 3.4 that in the absence of T-protein, the participation of H_{lip} in the aminomethyl transfer reaction is not only as shuttle protein.

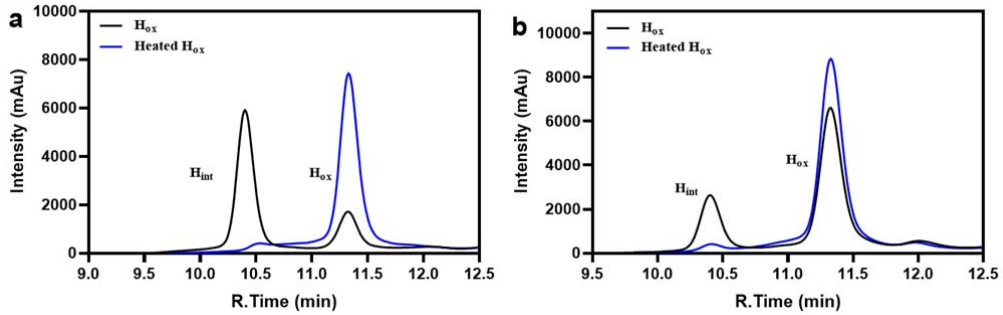


Figure 6. Effects of heating on the catalytic activity of H_{lip} as decarboxylase or aminomethyltransferase. (**a**) Decarboxylation reaction of glycine. "H_{ox}" refers to a reaction mixture containing 50 mM glycine, 50 μM H_{ox}, and 25 μM PLP; "heated H_{ox}" was the same as "H_{ox}" except for using heated H_{ox}. (**b**) Aminomethyl transfer reaction (started from H_{ox}, which was reduced to H_{red} by DTT, before H_{red} is converted to H_{int} by aminomethyl transfer reaction). "H_{ox}" refers to a reaction mixture containing 50 μM H_{ox}, 0.5 mM THF, 20 mM DTT, 50 mM NH_4Cl, and 10 mM HCHO; "heated H_{ox}" was the same as "H_{ox}" except for using heated H-protein.

To explain the above experimental observations we examined the possibility that the catalytic ability of unheated H_{ox} is due to intermolecular interactions of H_{lip} itself, and such interactions would enable H_{lip} to catalyze the interconversion of the different forms (H_{lip}, H_{red} and H_{int}) to complete cycle by cycle the electron transfer, aminomethyl transfer, and reductive aminomethylation. In order to verify this hypothesis, we developed a concept to distinguish a possible intermolecular interaction of H_{lip} from its function as shuttle protein. On the one hand, the lysine residue at position 64 of H-protein for binding lipoic acid was mutated into alanine to obtain H_{K64A} mutant protein, which is unable to be lipoylated and consequently unable to act as shuttle

protein, but is still expected to be able to uphold the capability of intermolecular interactions of the wild-type H-protein. On the other hand, as discussed above, heated H_{ox} losing its catalytic activity for glycine synthesis well preserves its function as a shuttle protein. Therefore, if the hypothesis holds true, we would expect that mixing heated H_{ox} with H_{K64A} would "restore" the catalytic activity exhibited by unheated H_{ox} towards glycine synthesis. No glycine formation could be detected which rules out the hypothesis and points out that the interconversion of the three forms of lipoylated H-protein, and consequently the stand-alone catalytic activity of H-protein, is not the result of intermolecular interactions of H-protein itself.

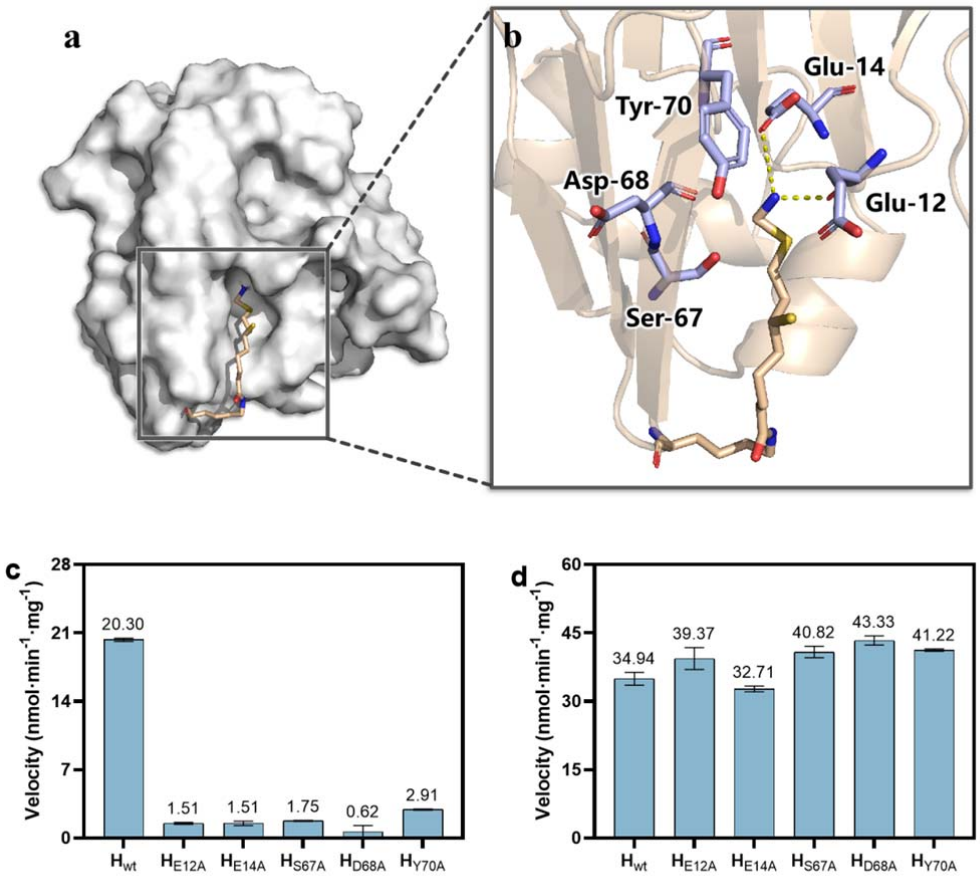


Figure 7. Study of the essential roles of the H-protein cavity. (**a**) Three-dimensional structure of *E. coli* H-protein bearing the lipoamide-methylamine arm protected in the cavity (ecH_{int}). (**b**) Docking of the lipoamide-methylamine arm in the cavity with highlighting of some surrounding amino acid residues selected for mutation. (**c**) The rate of glycine synthesis catalyzed by stand-alone H_{lip} or its mutants. (d) The rate of glycine synthesis catalyzed by GCS comprising H_{lip} or its mutants and the other three GCS proteins.

An answer to the question of why heating destroys the catalytic ability of H-protein but still preserves its function as shuttle protein might lie in the special surface structure of H-protein. As shown by the crystal structure of H_{int} (**Figure 7a**), following methylamine transfer the lipoamide-methylamine arm enters into the hydrophobic cavity on the surface of the H-protein and prevents thereby it from nucleophilic attack by water molecules[28, 36, 37, 38]. Based on our previously molecular dynamic simulation study, the lipoamide-methylamine arm may interact with some amino acid residues in the proximity of the cavity[39], such as Glu-12, Glu-14, Ser-67, Cys-68 and Tyr-70 (**Figure 7b**). We speculated that the catalytic activity found for the stand-alone H_{lip} is related to the structure of the cavity. Therefore, to verify this assumption, the above five residues of the wild-type H-protein were mutated to alanine. Assays of glycine synthesis were then performed with these mutated H_{lip} in comparison with the wild-type H_{lip}. As shown in **Figure 7c**, compared to the wild-type H_{lip} the glycine synthesis rates were strongly reduced in reaction mixtures containing H_{lip} mutants. However, as shown in **Figure 7d**, when these mutants were combined with the other three GCS proteins, the glycine synthesis rates of all the mutants were increased to levels comparable or even better than that of the wild-type H_{lip}. The results are very similar to what observed with heated H_{ox} (**Table 2**). It is therefore confirmed that the cavity on the H-protein surface plays a decisive role in the catalytic functions of H-protein, and alterations of the cavity structure (in size or form) either through mutation or heating will reduce or even destroy the stand-alone catalytic functions of H-protein (**Figure 7c**, **Table 2**) due to yet unclear mechanisms. One possible reason might be that the lipoamide-methylamine arm cannot properly enter the deformed cavity, resulting in the failure or imbalance of the GCS cycle in the presence of heated H_{ox} alone. By adding other GCS proteins, H-protein is mainly required to act as a shuttle protein and, consequently, GCS reactions can be revived (**Table 2**). Moreover, under the given *in vitro* reaction conditions, the lipoamide-methylamine arm may undergo fast and continuous reaction in the GCS cycle that not only minimizes the probability of its hydrolysis, but also maintains or even increases the reaction efficiency by omitting the process of its entry and exit from the cavity (**Figure 7d,**

Table 2).

Discussion

In this work, we show for the first time that stand-alone lipoylated H-protein (H_{lip}) has the catalytic functions so far believed to be carried out by the P-, T- and L-proteins of GCS. It enables glycine cleavage reactions, as well as the reversed reactions towards glycine synthesis with NH_4HCO_3 and HCHO as the substrates. The K_{cat} value for the overall synthesis reaction is about 0.01 s^{-1} for GCS catalyzed reaction and 0.0057 s^{-1} for H-protein alone catalyzed reaction. After purification of the GCS proteins, we used the most commonly used methods SDS-PAGE and HPLC to verify that there was no obvious residual of other GCS proteins in the purified H-protein solutions, though we did not confirm this by using more precise methods like mass spectroscopy. Through calculations, we can state that even if other GCS proteins would exist in the H-protein solution (e.g. up to 10%), it will not qualitatively affect the main conclusions drawn in our work (see Supplementary Materials for detailed explanation). Therefore, the purity of H-protein meets the requirement needed for this study.

The stand-alone catalytic activity of H_{lip} is closely related to the cavity on the H-protein surface, where the lipoyl swing arm is bound. Both heating H_{lip} and mutating cavity-related amino acid residues result in complete loss or strong reduction of the stand-alone catalytic activity of H_{lip}, because they may cause deformation of the cavity, resulting in failure of the lipoamide-methylamine arm to properly enter the cavity and consequently failure of GCS reactions. Cohen-Addad *et al.*[36] suggested that the lipoamide-methylamine arm is locked into a very stable configuration within the hydrophobic cavity and therefore highly stable against the non-enzymatic hydrolysis (which leads to the release NH_3 and HCHO) due to nucleophilic attack by water molecules.

However, our experiments surprisingly show that when heated H_{ox} or a H_{ox} mutant was combined with the other GCS proteins, the rate of glycine synthesis was

recovered or even increased (**Table 2** and **Figure 7**). This implies that under the given *in vitro* reaction conditions, heated H_{ox} truly acts only as shuttle protein in the presence of P-, T- and L-proteins, it is unnecessary and even disadvantageous for the lipoamide-methylamine arm to enter the cavity in order to undergo GCS reactions. Indeed, there was no obvious decrease of the peak area of H_{int} on HPLC even after hours of waiting, indicating that the hydrolysis of H_{int} is very slow (data not shown). This is in consistence with our recent finding that mutations of the key residue Ser-67 which reduce the residence time of the lipoamide-methylamine arm in the cavity can significantly increase the *in vitro* GCS activity[39]. The molecular dynamic simulations of H-protein carried out by our group[39] also implies that the lipoamide-methylamine arm can leave the cavity of unheated H-protein even without the interaction with T-protein. This is confirmed by further molecular dynamic simulations of H-protein (results not published yet). These results therefore raise the question why the lipoamide-methylamine arm should enter the cavity and be protected in the GCS system? Based on the fact that the cavity is closely related to the stand-alone catalytic activity of H-protein, perhaps we may make a bold speculation for the interpretation of these results from a perspective of evolution. *In vivo*, GCS may have evolved from a simple system for glycine cleavage catalyzed by H-protein alone at the early time of evolution to a sophisticated system, in which H-protein is assisted by specialized P-, T-, L-proteins for more effective catalysis of the GCS reactions to meet the growing metabolic requirements of organisms. However, the cavity structure and the stand-alone catalytic functions of H-protein have been retained till now.

Oliver[40] discovered that the H-protein and the small subunit of ribulose-1,5-bisphosphat-carboxylase/-oxygenase (RuBisCo) have obvious similarities in plants. The two proteins are not only about the same size, but also have similar mechanism in terms of transcriptional control of the corresponding genes. It has been reported that there are striking sequence and structure similarity between H-protein and the E2 protein of pyruvate dehydrogenase complex (PDC)[27]. Therefore, further research on the catalytic mechanism of H-protein may give useful hints for

522 understanding the evolution and function of PDC and other 2-oxoacid dehydrogenase
523 multi-enzyme complexes (e.g. alpha-ketoglutarate dehydrogenase complex) which
524 are all of fundamental importance in cellular metabolism, governing the synthesis of
525 C1-C4 metabolites for life.

526 H_{lip} alone catalyzes the glycine cleavage and synthesis *in vitro* with the help of the
527 cofactors PLP, THF and FAD. This seems in contradictory to the results reported in
528 the previous literature that deletion in gcvP or gcvT is lethal for organisms[8, 41]. By
529 comparing the differences between the in *vivo* and *in vitro* conditions, it is conceivable
530 that *in vivo* these cofactors are stoichiometrically linked to P-, T and L-proteins,
531 respectively, to play their dedicated catalytic roles in a concerted action with H-protein.
532 The *in vivo* concentrations of these cofactors are much lower than what we used in
533 our *in vitro* experiments to facilitate the cleavage or synthesis of glycine catalyzed by
534 H-protein alone. Thus, at one hand, it may not be feasible for cells to maintain high
535 concentration pools of these cofactors; at the other hand, it might be beneficial for
536 cells to have the sophisticated complete GCS system for fast and fine tuning and
537 adapting of this important system to the changes in metabolic demands. This may be
538 one reason why P-, T-, L-proteins are necessary *in vivo*. This study shows that for the
539 stand-alone catalytic functions of H_{lip} an interaction of FAD with H_{lip} is required for its
540 function in the absence of L-protein, and heating-induced structural change of H_{lip}
541 destroyed the possibility of such interaction (**Figure 5d**). It is most likely that PLP also
542 needs to interact with H_{lip} to exert its function in the absence of P-protein. This thus
543 raises questions: How do PLP and FAD interact with H_{lip}? Is it similar to the binding of
544 PLP to P-protein and FAD to L-protein? To answer these questions, more experiments
545 will be carried out in the future studies.

546 In addition, the overall reaction rate of GCS with T-protein deficiency was only
547 reduced to 52 % in the direction of glycine cleavage and 76.5 % for glycine synthesis,
548 compared to those of the complete GCS, indicating that T-protein had the least effect
549 on the catalytic activities of GCS. These results are in agreement with the results of
550 Timm *et al.*[8] who showed that knockdown mutants of Arabidopsis containing very low

T-protein expression under physiological conditions were able to grow and propagate in normal air, only showing some minor changes. Meanwhile, their study also found that the knockout mutation without T-protein expression was lethal even in non-photorespiratory environment of air enriched to 1 % CO_2. This result may indicate that THF needs weak binding with T-protein to participate in glycine cleavage and synthesis reaction *in vivo*. When the T-protein is completely knocked out, THF cannot participate in the reaction, resulting in slow reaction rate and plant death. The deficiency of THF is far more detrimental and has the greatest impact on the reaction rate of GCS *in vitro* (**Table 1**).

Although we have identified the cavity of the H-protein as the key structural region that determines the catalytic activities of stand-alone H_{lip}, the specific catalytic mechanism is not explored from the perspective of structure and molecular interaction. By studying the structures and their dynamics of heated and unheated H-proteins, it will certainly help to better understand the mechanism. This in turn will help to better engineer H-protein and GCS, leading to new possibilities to improve the growth and physiology of cells, organisms and plants and to design industrial microbes for utilizing C1 compounds for biosynthesis.

Materials and Methods

Materials

The substrates glycine, nicotinamide adenine dinucleotide (NAD^+, NADH), Tris (2-carboxyethyl) phosphine (TCEP) and the derivatization reagents dansyl chloride were purchased from Yuanye Bio-Technology (Shanghai, China). Dithiothreitol (DTT), β-mercaptoethanol (β-ME), pyridoxal 5'-phosphate monohydrate (PLP) and flavin adenine dinuleotide (FAD) were obtained from Aladdin (Shanghai, China). 6-(RS)-Tetrahydrofolate (THF) was obtained from Sigma-Aldrich (St. Louis, MO, USA). Other chemicals in this study were of analytical grade and purchased from Solarbio (Beijing, China) or Sinopharm (Shanghai, China), unless otherwise noted. *Escherichia coli* Top10 and BL21 (DE3) were used for plasmid construction and the

overexpression of recombinant proteins, respectively. Ni^{2+}-NTA resin was purchased from Genscript (Nanjing, China). Amicon® Ultra-15 filtration devices (molecular size cut-off 10 KDa for H-protein, 30 KDa for T-protein and L-protein and 100 KDa for P-protein) were purchased from Millipore (Billerica, MA, USA). Mut Express II Fast Mutagenesis kit V2 was purchased from Vazyme (Jiangsu, China). BCA protein assay kit and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels were purchased from SolarBio (Beijing, China). Luria-Bertani (LB) medium containing tryptone (10 g/L), yeast extract (5 g/L) and NaCl (10 g/L) were used for cloning and expression, and tryptone and yeast extract were purchased from Oxoid.

Enzyme preparation

The plasmids and bacterial strains used in this experiment were given in **Table 3**. Oligonucleotide sequences of primers used for cloning target proteins were given in Table 1 of Supplementary Materials. The genes coding for P-protein, T-protein, L-protein and H-protein were amplified from *E. coli* K12 genomic DNA, then cloned into the expression vector pET28a (NdeI and XhoI). *E.coli* BL21 (DE3) harboring the resulting constructs (pET28a-P, pET28a-T, pET28a-L and pET28a-H) were cultured in LB medium supplemented with 50 mg/L of kanamycin at 37°C until the OD_{600} of the culture reached 0.6-0.8, Isopropyl-beta-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.2 mM to induce protein expression for 12 h at 30 °C.

The plasmid pET28a-H was used as a template to generate mutations using Mut Express II Fast Mutagenesis kit V2[39]. Lipoylation of H-protein was performed during its over-expression *in vivo*. To this end, the strain containing the plasmid pET28a-H or a H-protein mutant were added with lipoic acid (200 μM, pH 7.0), prior to starting the cultivation to directly obtain lipoylated H-protein (H_{lip}). Following the overexpression, enzymes were purified as described previously[42]. The purified enzymes were checked by SDS-PAGE (shown as Figure S1 in Supplementary Materials). In the lane of the H protein no residuals of T, P and L proteins were found except for the existence of low amount of the inactive H apo-protein (H_{apo}). In addition, HPLC analysis also confirmed that there was no obvious residual of other GCS proteins in the purified H-protein

608 solutions. H_{lip} obtained in such a way is considered to primarily exist in the oxidized
609 form of H-protein (H_{ox}). Whereas H_{lip} is mostly used to refer to lipoylated H-protein in
610 general, it is numerically equivalent to H_{ox} in this work when concrete reactions
611 involving H_{lip} are referred to and vice versa.

612 **Table 3. Strains and plasmids used in this study**

	Description	Reference/Source
***E. coli* Strains**		
Top 10	Host for cloning plasmids	WEIDI Ltd.
BL21 (DE3)	Host for protein overexpression and purification	WEIDI Ltd.
Plasmids		
pET28a (+)	Plasmid for protein overexpression	Novagen
pET22b (+)	Plasmid for protein overexpression	Novagen
pET28a-P	pET28a vector containing P-protein gene (NCBI No. WP_112929453.1)	This study
pET28a-T	pET28a vector containing T-protein gene (NCBI No. WP_099356926.1)	This study
pET28a-L	pET28a vector containing L-protein gene (NCBI No. WP_110826218.1)	This study
pET28a-H	pET28a vector containing H-protein gene (NCBI No. WP_001295377.1)	This study
pET28a-H-K64A	pET28a-H containing H-protein gene with point mutation of K64A	This study
pET28a-H-E12A	pET28a-H containing H-protein gene with point mutation of E12A	Zhang et al. 2020[39]
pET28a-H-E14A	pET28a-H containing H-protein gene with point mutation of E14A	Zhang et al. 2020[39]
pET28a-H-S67A	pET28a-H containing H-protein gene with point mutation of S67A	Zhang et al. 2020[39]
pET28a-H-D68A	pET28a-H containing H-protein gene with point mutation of D68A	Zhang et al. 2020[39]

pET28a-H-Y70A	pET28a-H containing H-protein gene with point mutation of Y70A	Zhang et al. 2020[39]

613

614 ***Enzyme activity assays***

615 ***The overall reaction of glycine cleavage***

616 **(1) Glycine cleavage catalyzed by GCS.** The reaction mixture contained Tris-HCl
617 (50 mM, pH 7.5), 0.5 mM THF, 20 mM DTT, 25 μM PLP, 5 mM NAD^+, 5 μM P-protein,
618 5 μM T-protein, 5 μM L-protein and 10 μM H_{ox}. After premixing and centrifugation, the
619 reactions were initiated by the addition of 50 mM glycine, and carried out for 30 min at
620 37 °C. The overall GCS activity was determined by measuring either NADH formation
621 at 340 nm using an Enspire multimode plate reader (PerkinElmer, USA) or
622 formaldehyde formation according to our previous reported method[17].

623 **(2) Glycine cleavage enabled by H_{lip} alone.** The glycine cleavage reaction enabled
624 by H_{lip} alone was monitored by determining the formation of NADH at 340 nm. The
625 reaction mixture contained Tris-HCl (50 mM, pH 7.5), 0.5 mM THF, 20 mM DTT, 25
626 μM PLP, 5 mM NAD^+, 40 μM FAD and 10 μM H_{ox}. The reaction was initiated by adding
627 50 mM glycine.

628 ***The overall reaction of glycine synthesis***

629 **(1) Glycine synthesis catalyzed by GCS.** The reaction mixture containing Tris-HCl
630 (50 mM, pH 7.5), 0.5 mM THF, 20 mM DTT, 10 mM HCHO, 25 μM PLP, 5 mM NADH,
631 5 μM P-protein, 5 μM T-protein, 5 μM L-protein and 10 μM H_{ox}. The reaction was
632 initiated by adding 50 mM NH_4HCO_3 to the reaction mixture and carried out for 2 h at
633 37 °C. The amount of glycine formed was determined using HPLC. One unit of glycine
634 synthesis activity was defined as the amount (in mg) of H-protein that catalyzed the
635 formation of 1 nmol of glycine per minute.

636 **(2) Glycine synthesis enabled by H_{lip} alone.** The reaction mixture contained
637 Tris-HCl (50 mM, pH 7.5), 0.5 mM THF (added with β-ME to prevent its oxidative

degradation), 10 mM HCHO, 25 μM PLP, 50 mM NH_4HCO_3, 40 μM FAD and 5 mM NADH, and different concentration of H_{ox} (10-80 μM). Alternatively, 20 mM DTT can be used to replace FAD and NADH for the reduction of H_{ox} to H_{red}. The reaction condition and enzyme activity calculation were the same as stated above in 2.3.2 (1).

Individual GCS reaction steps in the presence of H_{lip} with or without the corresponding enzymes

(1) Glycine dcarboxylation reaction catalyzed by P-protein. The reaction mixture contained Tris-HCl (50 mM, pH 7.5), 50 mM glycine, 50 μM H_{ox} and 25 μM PLP. 5 μM P-protein was added to the reaction mixture as the control group. The reaction was carried out for 2 h at 37 °C. The substrate H_{ox} and the product H_{int} were measured using HPLC.

(2) Aminomethyl transfer reaction catalyzed by T-protein. In this reaction of converting H_{red} to H_{int} through aminomethyl transfer, H_{red} required was generated by reducing H_{ox} with DTT (see 2.3.3 below), and 5,10-CH_2-THF was derives from the condensation of HCHO and THF. Therefore, the reaction mixture contained Tris-HCl (50 mM, pH 7.5), 50 μM H_{ox}, 0.5 mM THF, 20 mM DTT, 50 mM NH_4Cl, and 10 mM HCHO. 5 μM T-protein was added to the reaction mixture as the control group. The reaction was carried out for 2 h at 37 °C. The substrate H_{ox} and the product H_{int} were measured by HPLC.

Electron transfer reaction between H_{ox} and H_{red} with or without* the presence *of L-protein

The interconversion of H_{ox} and H_{red} was performed according to a reported enzymatic assay using an excess amount of a reductant (8 mM TCEP or 20 mM DTT) for the reduction of the H-protein-bound lipoic acid[34, 35], and then the produced H_{red} is re-oxidized by L-protein in the presence of NAD^+. For the assay, the reaction mixture contained different types of buffer (50 mM, pH 7.5), 8 mM TCEP or DTT, 5 μM H_{ox} and 0.2 μM L-protein. In order to prove that H_{ox} can still undergo redox reaction without L-protein, 40 μM FAD is used instead of L-protein. The reactions were initiated by the

addition of 5 mM NAD^+. The rate of NADH formation was determined spectrophotometrically at 340 nm.

Analytical methods using HPLC

H_{ox} and H_{int} proteins were analyzed based on the HPLC method previously developed in our lab[33]. The analysis was performed with a Inertsil WP300 C4 column (5 μm, 4.6×150 mm) and monitored at 280 nm using a diode array detector (DAD).The mobile phase consisted of acetonitrile (A) and 0.1 % trifluoroacetic acid aqueous solution (B). The volumn percentage of buffer B was varied as follows: linearly increased from 30 % to 50 % (0-13.4 min), sharply increased from 50 % to 90 % (13.4-13.41 min), held at 90 % (13.41-14.2 min), and then sharply decreased to 30 % (14.2-14.21 min), held at 30 % to 18 min. The flow rate was 1.0 $mL \cdot min^{-1}$.

Glycine concentration in the reaction mixture was determined by pre-column dansyl chloride derivatization. To this end, 40 μL of a reaction mixture was mixed with 160 μL of 0.2 M $NaHCO_3$ and 200 μL of 5.4 $mg \cdot mL^{-1}$ dansyl chloride in acetonitrile. Derivatization occurred at 30 °C for 30 min. After the reaction, 600 μL of 0.12 M HCl was added to adjust the pH of the sample to weak acidic. After centrifuged at 10,000 rpm the supernatant was filtered with 0.22 μm membrane. The dansyl derivative of glycine was measured using HPLC (Shimadzu LC-2030C system) on a Shim-pack GIST C_{18} column (5 μm, 4.6×150 mm) at 30 °C, with a mobile phase composed of acetonitrile and 20 mM potassium phosphate buffer pH 6.0 (25:75 v/v) at a flow rate of 0.8 mL/min. The effluent was monitored at 254 nm using a diode array detector (DAD). The HPLC results of glycine were given in Supplementary materials Figure 2.

Statistics and Reproducibility. Enzyme activities and reaction rates were measured by three independent experiments and averaged for report. Individual data points are added in the graphs, and error bars are defined by the standard deviation.

Data availability

Major data generated and analyzed during this study are included in the article. The source data underlying the graphs and charts presented in the main figures are available as Supplementary Data. Other datasets generated and analyzed during the study are available from the corresponding author on reasonable request.

Acknowledgments and funding

This work was financially supported by the Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology.

Author Contributions

Y.X. designed and performed the experiments, wrote the initial manuscript. Y.L. assisted in experiments and data analysis. H.Z. provided H-protein mutants and participated in data analysis. J.N. assisted in experiments and preparing the figures. J.R. involved in experimental design, data analysis and drafting the manuscript. W. W. involved in data analysis and revised most of the manuscript content. A.-P.Z. supervised the project, involved in experimental design, data analysis and discussion, reviewed and revised the paper.

Reference

1. Douce R, Bourguignon J, Neuburger M, Rébeillé F. The glycine decarboxylase system: a fascinating complex. *Trends Plant Sci.* **6**, 167-177 (2001).

2. Kikuchi G, Motokawa Y, Yoshida T, Hiraga K. Glycine cleavage system: reaction mechanism, physiological significance, and hyperglycinemia. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 246-263 (2008).

3. Hong Y, Ren J, Zhang X, Wang W, Zeng A-P. Quantitative analysis of glycine related metabolic pathways for one-carbon synthetic biology. *Curr. Opin. Biotechnol.* **64**, 70-78 (2020).

4. Pardal-Fernández JM, Carrascosa-Romero MC, Vega CdC-dl, Iniesta-López I, Martínez-Gutiérrez A. A typical glycine encephalopathy in an extremely low birth weight infant: description of a new mutation and clinical and electroencephalographic analysis. *Epileptic Disord.* **11**, 48-53 (2009).

5. Azize NAA*, et al.* Mutation analysis of glycine decarboxylase, aminomethyltransferase and glycine cleavage system protein-H genes in 13 unrelated families with glycine encephalopathy. *J. Hum. Genet.* **59**, 593-597 (2014).

728 6. Zhang WC, *et al.* Glycine decarboxylase activity drives non-small cell lung cancer
729 tumor-initiating cells and tumorigenesis. *Cell* **148**, 259-272 (2012).

730 7. Kang PJ, *et al.* Glycine decarboxylase regulates the maintenance and induction of
731 pluripotency via metabolic control. *Metab. Eng.* **53**, 35-47 (2019).

732 8. Timm S, *et al.* T-protein is present in large excess over the other proteins of the glycine
733 cleavage system in leaves of Arabidopsis. *Planta* **247**, 41-51 (2018).

734 9. Timm S, *et al.* Mitochondrial dihydrolipoyl dehydrogenase activity shapes photosynthesis and
735 photorespiration of *arabidopsis thaliana*. *Plant Cell* **27**, 1968-1984 (2015).

736 10. E. Lopez-Calcagno P, Fisk S, L. Brown K, E. Bull S, F. South P, A. Raines C.
737 Overexpressing the H-protein of the glycine cleavage system increases biomass yield in
738 glasshouse and field-grown transgenic tobacco plants. *Plant Biotechnol. J.* **17**, 1-11 (2018).

739 11. Gariboldi R, T. , Drake H, L. . Glycine synthase of the purinolytic bacterium, *Clostridium*
740 *acidiurici.* Purification of the glycine-CO_2 exchange system. *J. Biol. Chem.* **259**, 6085-6089
741 (1984).

742 12. Schneeberger A, Frings J, Schink B. Net synthesis of acetate from CO_2 by *Eubacterium*
743 *acidaminophilum* through the glycine reductase pathway. *FEMS Microbiol. Lett.* **177**, 1-6
744 (1999).

745 13. Kechi H, Kikuchi G. Mechanism of the reversible glycine cleavage reaction in arthrobacter
746 globiformis. I. purification and function of protein components required for the reaction. *J.*
747 *Biochem.* **75**, 1113-1127 (1974).

748 14. Hideo K, Goro K. Mechanism of reversible glycine cleavage reaction in Arthrobacter
749 globiformis: function of lipoic acid in the cleavage and synthesis of glycine. *Arch. Biochem.*
750 *Biophys.* **173**, 71-81 (1976).

751 15. Kawasaki H, Sato T, Kikuchi G. A new reaction for glycine biosynthesis. *Biochem. Biophys.*
752 *Res. Commun.* **23**, 227-233 (1966).

753 16. Kochi H, Kikuchi G. Reactions of glycine synthesis and glycine cleavage catalyzed by
754 extracts of *Arthrobacter globiformis* grown on glycine. *Arch. Biochem. Biophys.* **132**,
755 359-369 (1969).

756 17. Hiraga K, Kechi H, Motokawa Y, Kikuchi G. Enzyme complex nature of reversible glycine
757 cleavage system of cock liver mitochondria. *J. Biol. Chem.* **72**, 1285-1289 (1972).

758 18. Bar-Even A, Noor E, Flamholz A, Milo R. Design and analysis of metabolic pathways
759 supporting formatotrophic growth for electricity-dependent cultivation of microbes. *Biochim.*
760 *Biophys. Acta-Bioenerg.* **1827**, 1039-1047 (2013).

761 19. Sánchez-Andrea I, *et al.* The reductive glycine pathway allows autotrophic growth of
762 Desulfovibrio desulfuricans. *Nat. Commun.* **11**, 5090 (2020).

763 20. Yishai O, Bouzon Madeleine, Döring V, Bar-Even A. In Vivo Assimilation of One-Carbon via
764 a Synthetic Reductive Glycine Pathway in *Escherichia coli*. *ACS Synth. Biol.* **7**, 2023-2028
765 (2018).

766 21. Bang J, Leea SY. Assimilation of formic acid and CO_2 by engineered *Escherichia coli*
767 equipped with reconstructed one-carbon assimilation pathways. *Proc. Natl. Acad. Sci. U. S.*
768 *A.* **115**, E9271-E9279 (2018).

769 22. Kim S, Aslan SNLS, Yishai O, Wenk S, Schann K, Bar-Even A. Growth of *E. coli* on formate
770 and methanol via the reductive glycine pathway. *Nat. Chem. Biol.* **16**, 538-545 (2020).

771 23. Bang J, Hwang CH, Ahn JH, Lee JA, Lee SY. *Escherichia coli* is engineered to grow on CO_2
772 and formic acid. *Nat. Microbiology* **5**, 1459-1463 (2020).

773 24. Cruz JGdl, Machens F, Messerschmidt K, Bar-Even A. Core catalysis of the reductive
774 glycine pathway demonstrated in yeast. *ACS Synth. Biol.* **8**, 911-917 (2019).

775 25. Claassens NJ, Bordanaba-Florit G, Cotton CAR, Maria AD, Bar-Even A. Replacing the
776 Calvin cycle with the reductive glycine pathway in Cupriavidus necator. *Metab. Eng.* **62**,
777 30-41 (2020).

778 26. Hong Y, Arbter P, Wang W, Rojas LN, Zeng A. Introduction of glycine synthase enables
779 uptake of exogenous formate and strongly impacts the metabolism in Clostridium
780 pasteurianum. *Biotechnol. Bioeng.* **112**, doi.org/10.1002/bit.27658 (2020).

781 27. Cronan JE. Assembly of lipoic acid on its cognate enzymes: an extraordinary and essential
782 biosynthetic pathway. *Microbiol. Mol. Biol. Rev.* **80**, 429-450 (2016).

783 28. Kazuko O-I, *et al.* Crystal structure of aminomethyltransferase in complex with
784 dihydrolipoyl-H-protein of the glycine cleavage system. *J. Biol. Chem.* **285**, 18684-18692
785 (2010).

786 29. Nakai T, Nakagawa N, Maoka N, Masui R, Kuramitsu S, Kamiya N. Structure of P-protein of
787 the glycine cleavage system: implications for nonketotic hyperglycinemia. *Embo J.* **24**,
788 1523-1536 (2005).

789 30. Hasse D, *et al.* Structure of the homodimeric glycine decarboxylase P-protein from
790 Synechocystis sp. PCC 6803 suggests a mechanism for redox regulation. *J. Biol. Chem.* **288**,
791 35333-35345 (2013).

792 31. Go MK, Zhang WC, Bing L, Wen SY. Glycine decarboxylase is an unusual amino acid
793 decarboxylase involved in tumorigenesis. *Biochemistry* **53**, 947-956 (2014).

794 32. Fujiwara K, Toma S, Okamura-Ikeda K, Motokawa Y, Nakagawa A, Taniguchi H. Crystal
795 structure of lipoate-protein ligase A from *Escherichia coli*. determination of the lipoic
796 acid-binding site. *J. Biol. Chem.* **280**, 33645-33651 (2005).

797 33. Zhang X, Nie J, Zheng Y, Ren J, Zeng AP. Activation and competition of lipoylation of H
798 protein and its hydrolysis in a reaction cascade catalyzed by the multifunctional enzyme
799 lipoate-protein ligase A. *Biotechnol. Bioeng.* **117**, 3677-3687 (2020).

800 34. Neuburger M*, et al.* Interaction between the lipoamide-containing H-protein and the
801 lipoamide dehydrogenase (L-protein) of the glycine decarboxylase multienzyme system. *Eur.*
802 *J. Biochem.* **267**, 2882-2889 (2000).

803 35. Gueguen V*, et al.* Structural and functional characterization of H protein mutants of the
804 glycine decarboxylase complex. *J. Biol. Chem.* **274**, 26344-26352 (1999).

805 36. Claudine C-A, Serge P, Larry S, Michel N, Roland D. The lipoamide arm in the glycine
806 decarboxylase complex is not freely swinging. *Nat. Struct. Biol.* **2**, 63-68 (1995).

807 37. Guilhaudis L*, et al.* Combined structural and biochemical analysis of the H-T complex in the
808 glycine decarboxylase cycle: evidence for a destabilization mechanism of the H-protein.
809 *Biochemistry* **39**, 4259-4266 (2000).

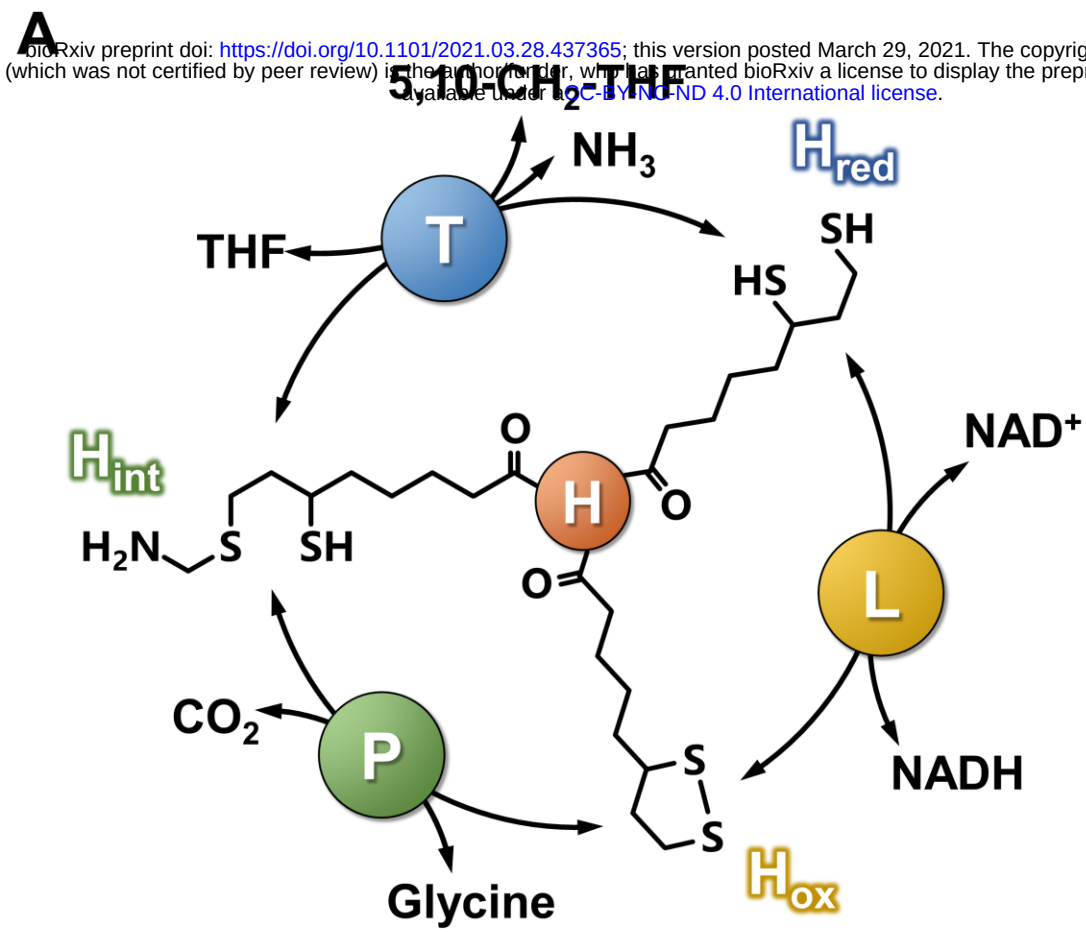
810 38. Okamura-Ikeda K*, et al.* Crystal structure of human T-protein of glycine cleavage system at
811 2.0 Å resolution and its implication for understanding non-ketotic hyperglycinemia. *J. Mol.*
812 *Biol.* **351**, 1146-1159 (2005).

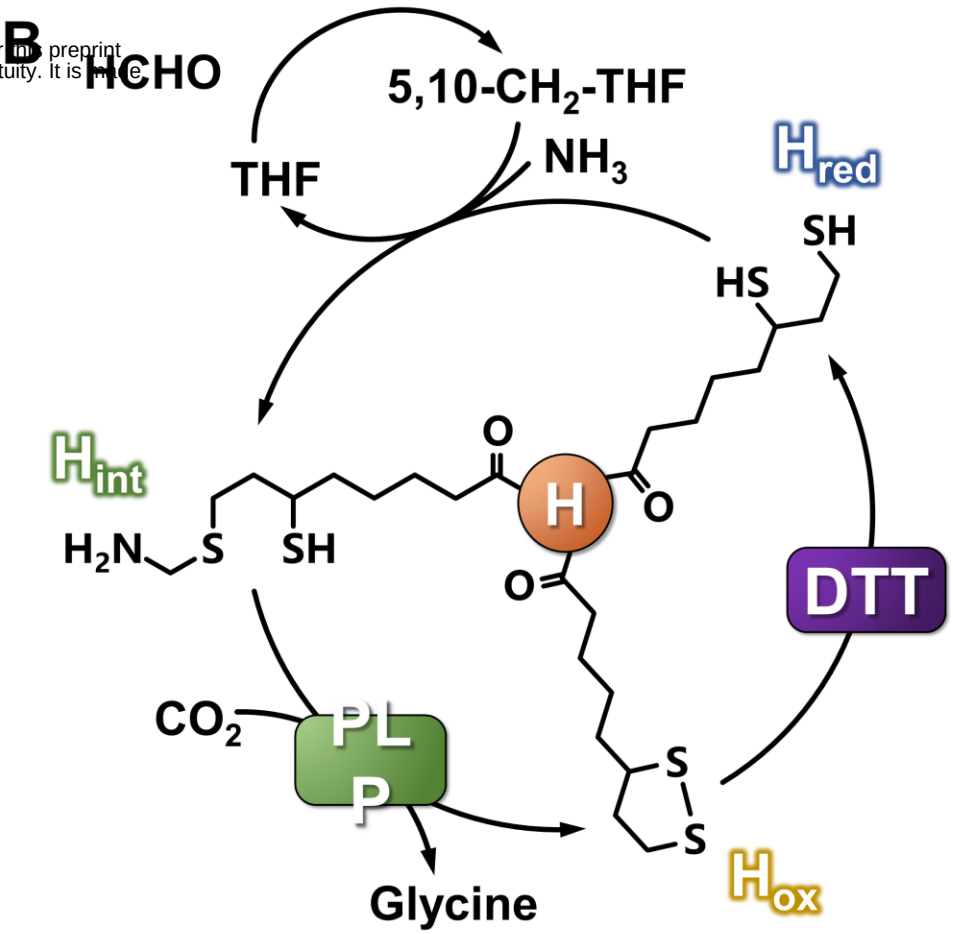
813 39. Zhang H, Li Y, Nie J, Jie R, Zeng A-P. Structure-based dynamic analyses of the glycine
814 cleavage system suggests key residues for control of a key reaction step. *Commun. Biol.* **3**,
815 756 (2020).

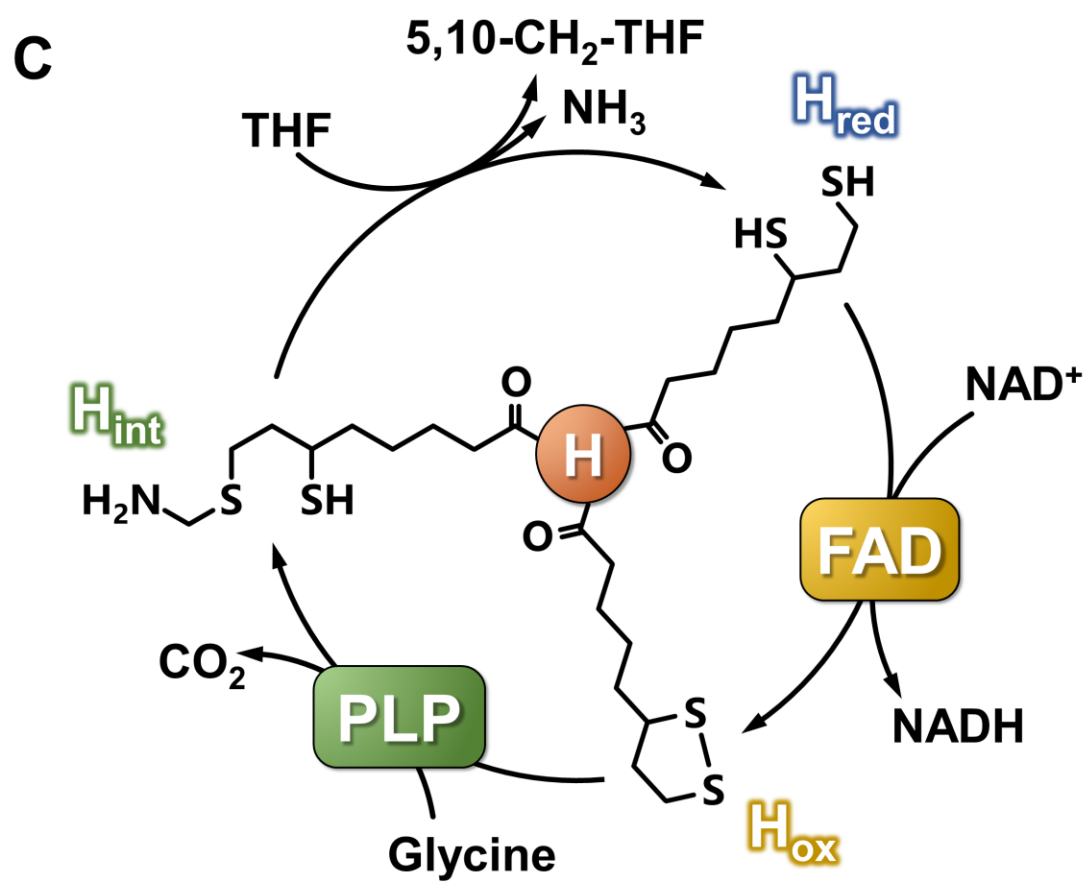
816 40. Oliver D, J. The glycine decarboxylase complex from plant mitochandria. *Annu. Rev. Plant*
817 *Physiol Plant Mol. Biol.* **45**, 323-337 (1994).

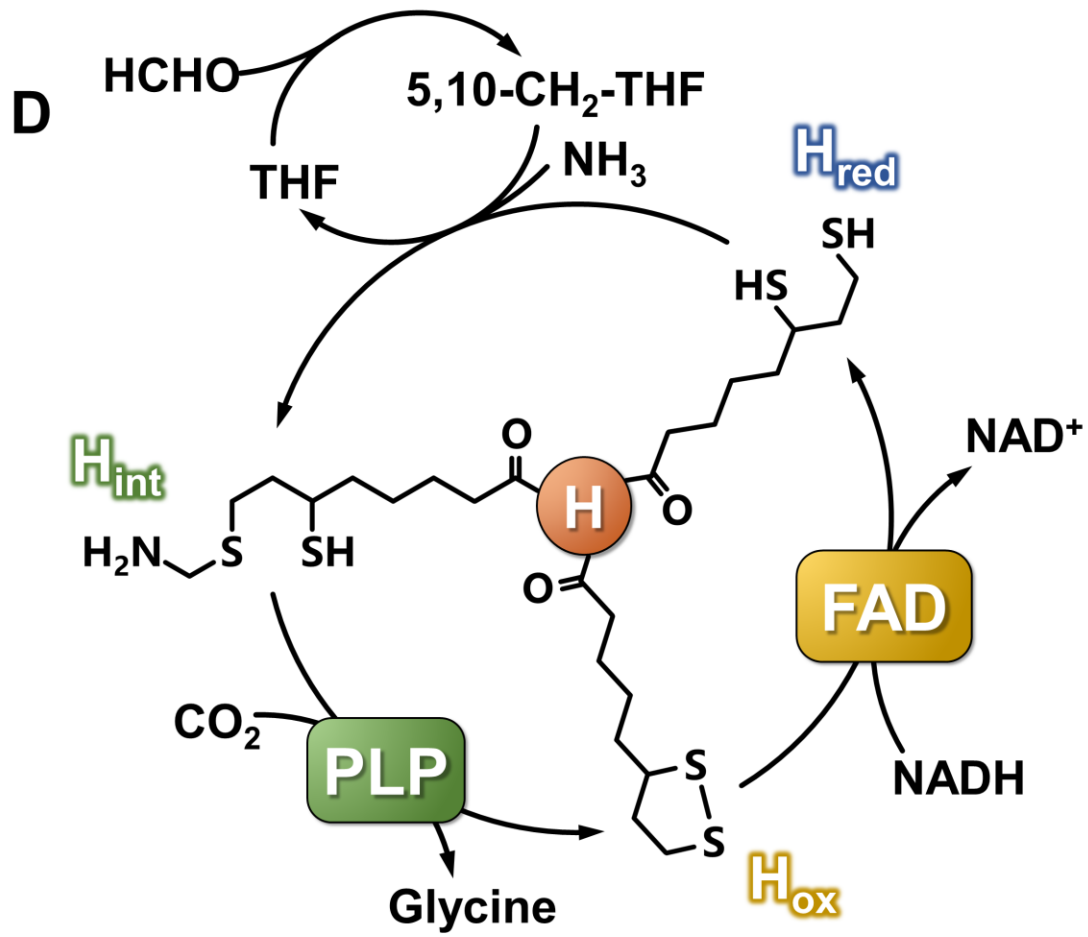
818 41. Nadja E*, et al.* Deletion of glycine decarboxylase in Arabidopsis is lethal under
819 nonphotorespiratory conditions. *Plant Physiol.* **144**, 1328-1335 (2007).

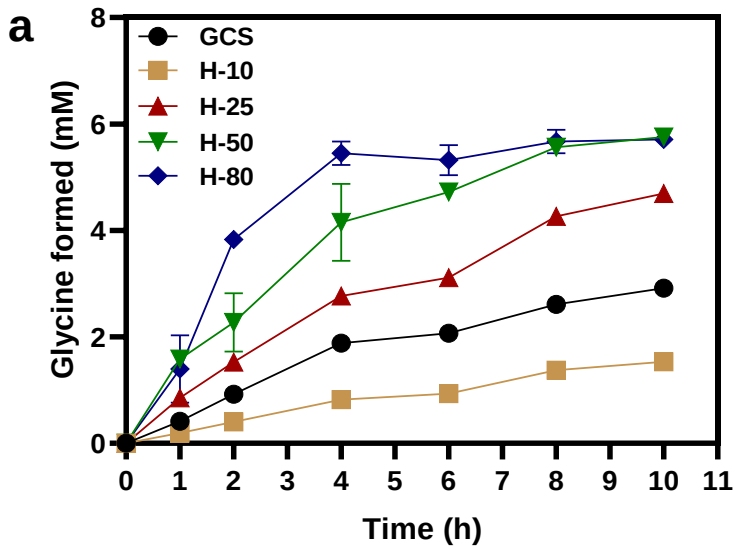
820 42. Xu Y, Meng H, Ren J, Zeng A-P. Formaldehyde formation in the glycine cleavage system and
821 its use for an aldolase-based biosynthesis of 1,3-propanediol. *J. Biol. Eng.* **14**, 15 (2020).

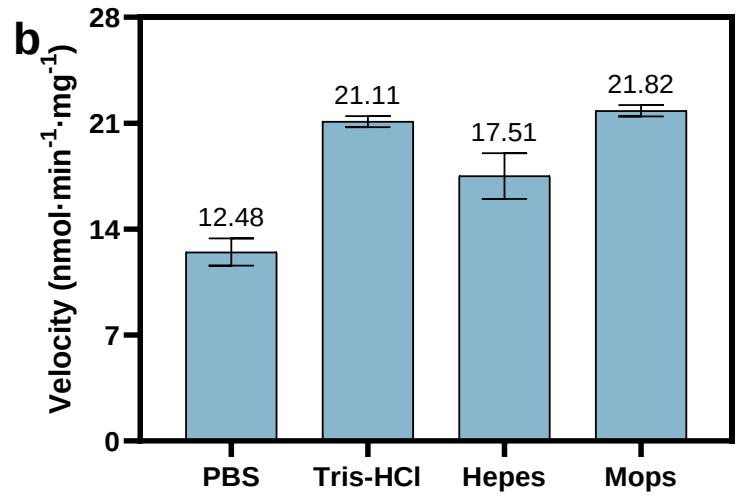


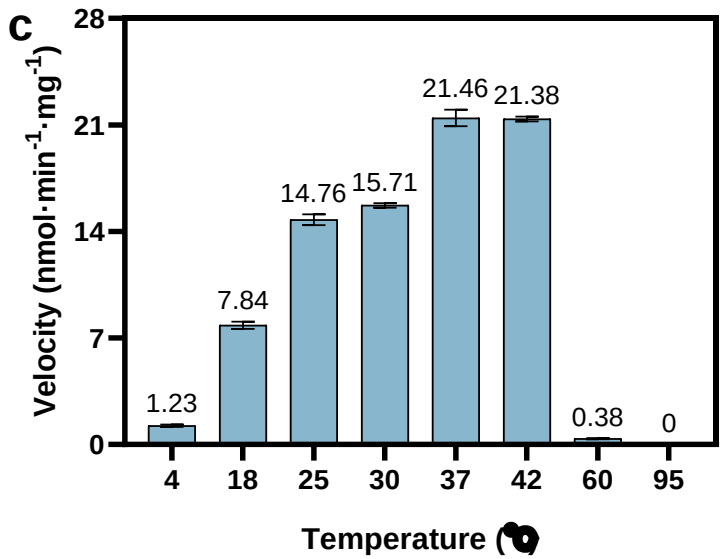


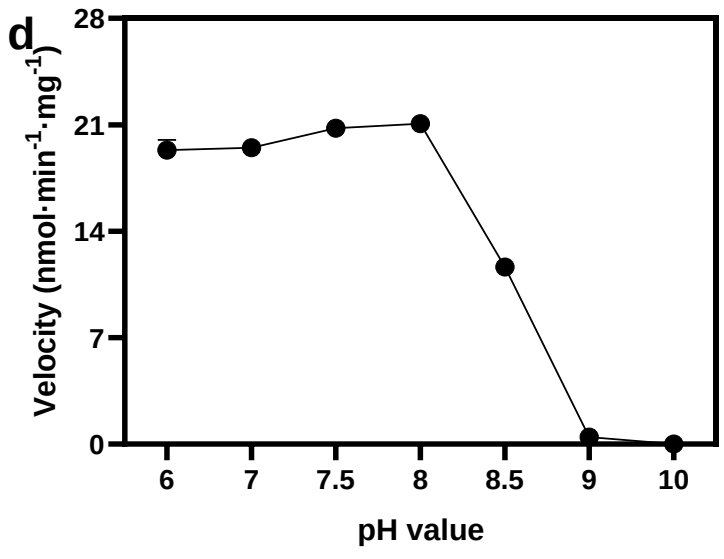


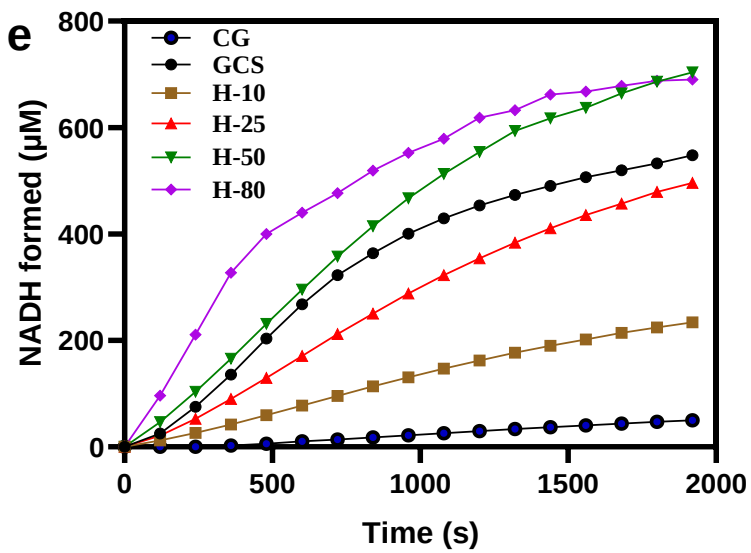


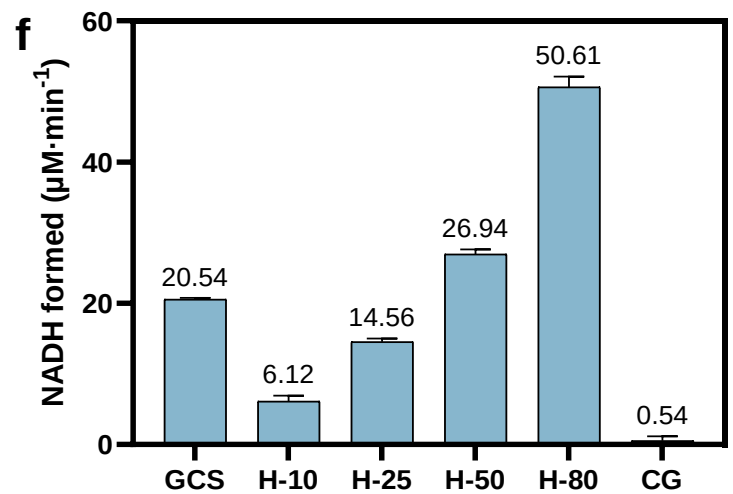


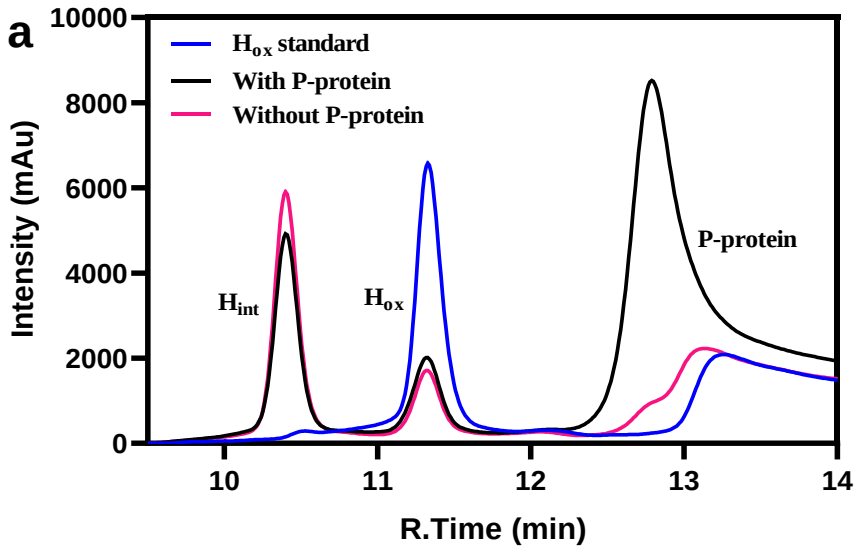


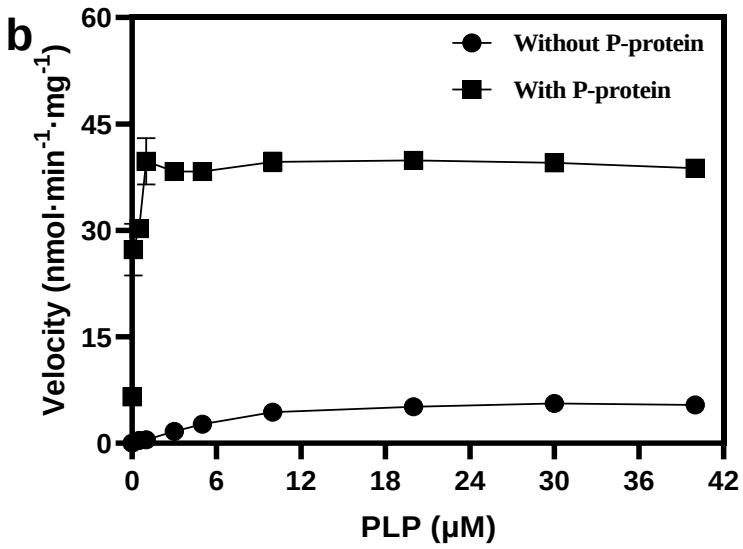


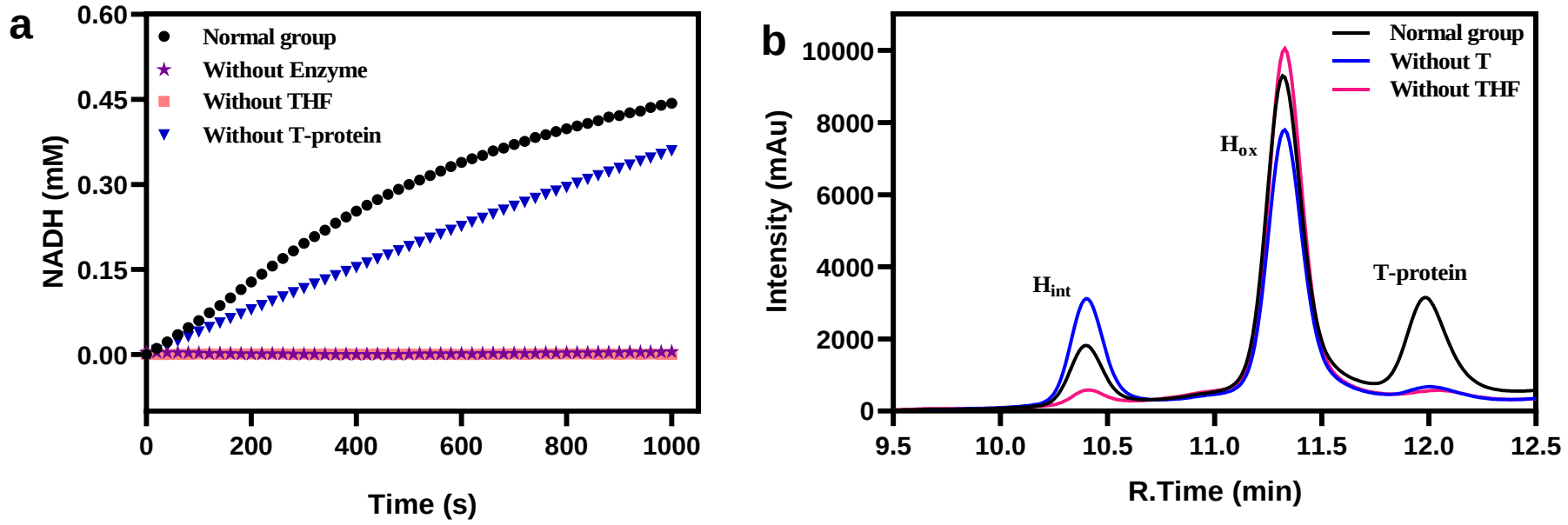


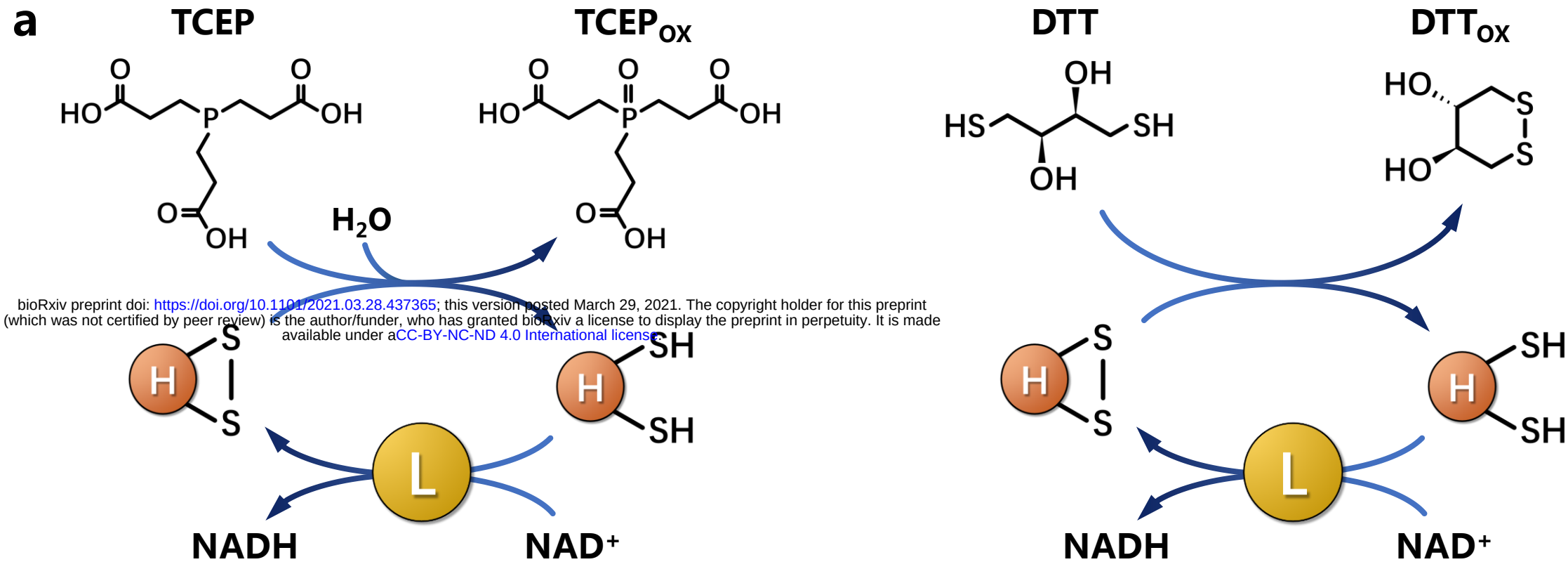


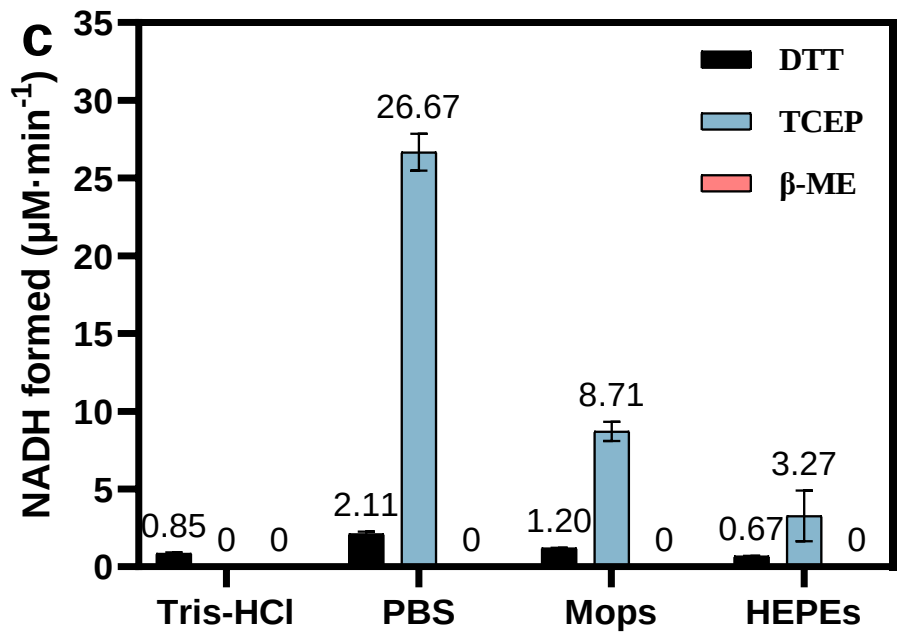


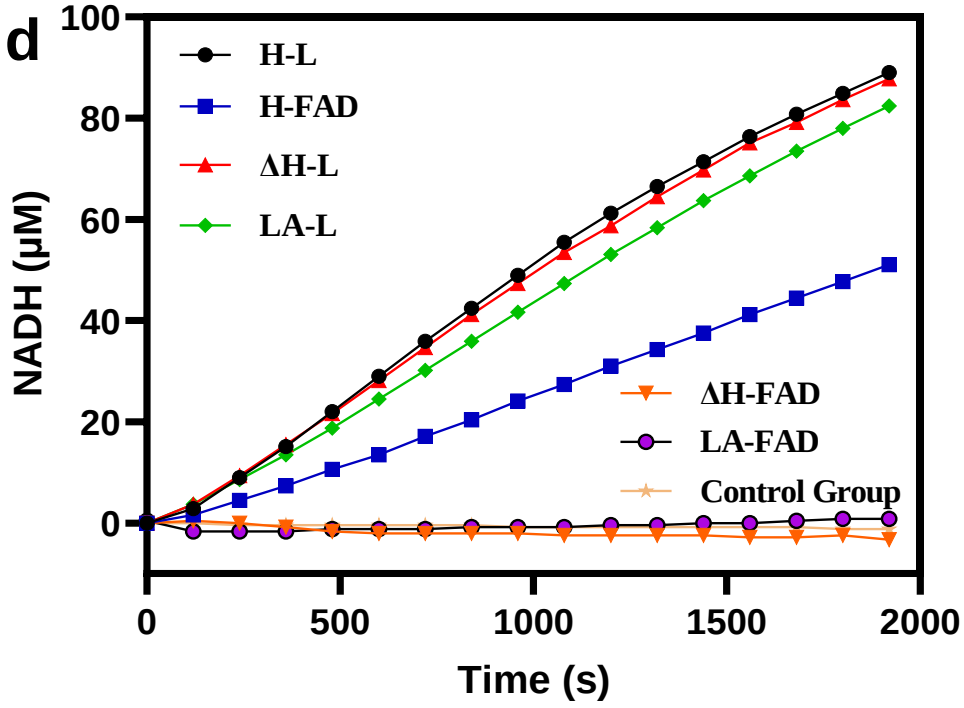

a
10000
8000
6000
4000
2000
0
Intensity (mAu)
H_{ox} standard
With P-protein
Without P-protein
H_{int}
H_{ox}
P-protein
10
11
12
13
14
R.Time (min)

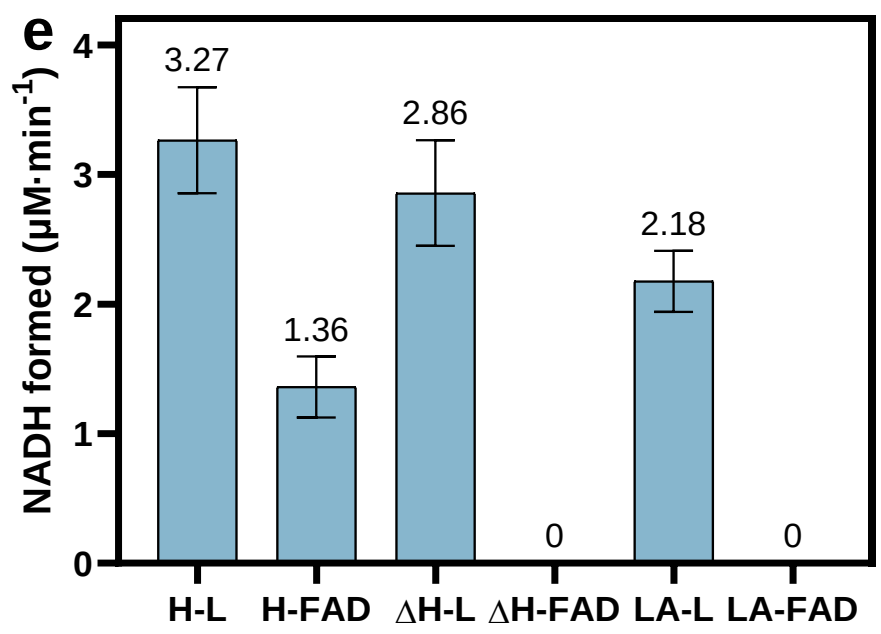

b
60
45
30
15
0
Velocity (nmol·min^{-1}·mg^{-1})
Without P-protein
With P-protein
0
6
12
18
24
30
36
42
PLP (μM)

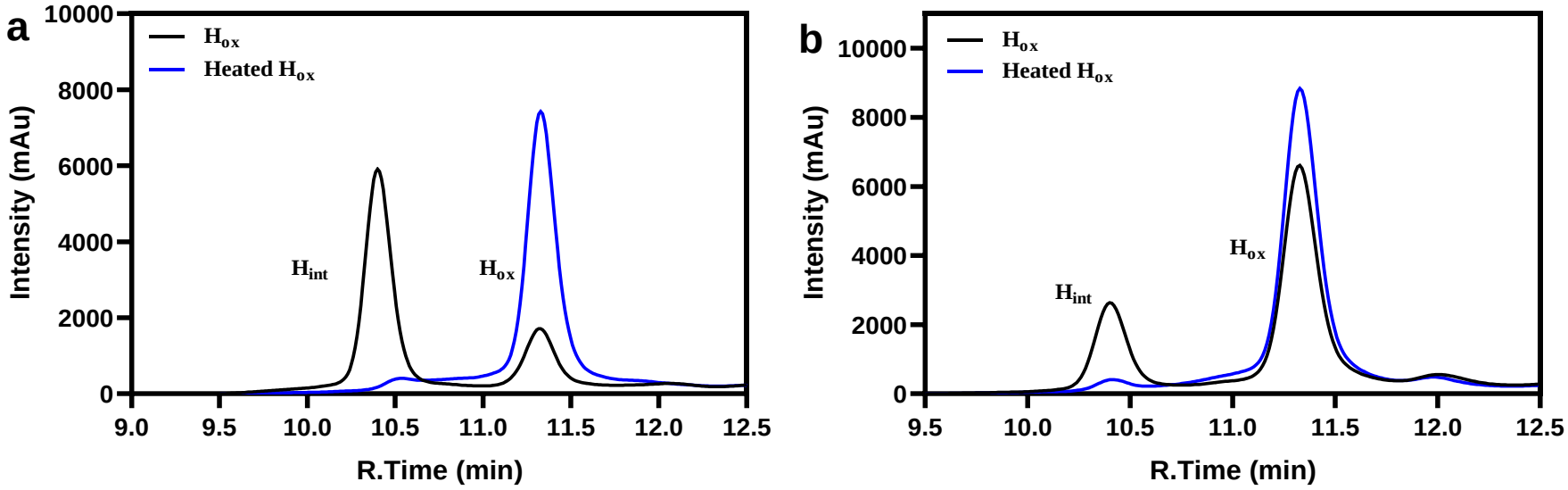

a
0.60
0.45
0.30
0.15
0.00
NADH (mM)
Normal group
Without Enzyme
Without THF
Without T-protein
0
200
400
600
800
1000
Time (s)
b
10000
8000
6000
4000
2000
0
Intensity (mAu)
Normal group
Without T
Without THF
H_ox
H_int
T-protein
9.5
10.0
10.5
11.0
11.5
12.0
12.5
R.Time (min)










a
b
H_{ox}
Heated H_{ox}
H_{int}
H_{ox}
Intensity (mAu)
R.Time (min)
0
2000
4000
6000
8000
10000
9.0
9.5
10.0
10.5
11.0
11.5
12.0
12.5