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 Xu¹, Yuchen Li¹, Han Zhang¹, Jinglei Nie¹, Jie Ren³, Wei Wang² and An-Ping Zeng^{1,2*}

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

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

³State 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

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

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

35 **Significance statement**

36 Glycine cleavage system (GCS) plays central roles in C1 and amino acids
37 metabolisms and the biosynthesis of purines and nucleotides. Manipulations of GCS
38 are desired to promote plant growth or to treat serious pathophysiological processes
39 such as aging, obesity and cancers. Reversed GCS reactions form the core of the
40 reductive glycine pathway (rGP), one of the most promising pathway for the
41 assimilation of formate and CO₂ in the emerging C1-synthetic biology. H-protein, one
42 of the four GCS component proteins (H, T, P and L) is generally considered a shuttle
43 protein interacting with the other three proteins via a lipoyl swinging arm. Here, we
44 discovered that without P-, T- and L-proteins, H-protein alone can catalyze GCS
45 reactions in both glycine cleavage and synthesis directions *in vitro*. The surprising
46 catalytic activities are related to a structural region of H-protein which can be
47 manipulated. The results have impacts on engineering GCS to treat related diseases,
48 to improve photorespiration, and to efficiently use C1-carbon for biosynthesis.

49 **Introduction**

50 In the mitochondria of plant and animal cells as well as in the cytosol of many
51 bacteria, the glycine cleavage system (GCS) comprising four (H, T, P and L) proteins
52 catalyzes the reversible decarboxylation and deamination of glycine to yield CO₂, NH₃
53 and provide a methylene group for the conversion of tetrahydrofolate (THF) to
54 N⁵,N¹⁰-methylene-tetrahydrofolate (5,10-CH₂-THF)^{1, 2, 3}. The overall reaction cycle
55 catalyzed by GCS comprises three steps as illustrated in **Figure 1a** (hereinafter GCS
56 is used to refer the four-enzyme system regardless of reaction direction). The reaction

57 is first catalyzed by P-protein (glycine decarboxylase; EC 1.4.4.2) to yield CO₂ from
58 glycine and methylamine-loaded H-protein (H_{int}) from the oxidized form (H_{ox}) of the
59 lipoylated H-protein (H_{lip}). T-protein (aminomethyltransferase; EC 2.1.2.10) then
60 catalyzes the release of NH₃ and transfer the methylene group from H_{int} to THF to
61 form 5,10-CH₂-THF, leaving dihydrolipoyl H-protein (H_{red}). Finally, L-protein
62 (dihydrolipoyl dehydrogenase; EC 1.8.1.4) catalyzes the oxidation of H_{red} to
63 regenerate H_{ox} in the presence of NAD⁺. H-protein as a shuttle protein interacts with
64 the other three GCS-proteins via a lipoyl swinging arm and plays central role in the
65 GCS.

66 The physiological roles of GCS in various organisms have been well studied. In
67 human and most vertebrates, GCS is part of the serine and glycine metabolism
68 pathway. Serine is catalyzed by serine hydroxymethyltransferase (SHMT) to form
69 glycine and 5,10-CH₂-THF, and then the product glycine is degraded by GCS. The
70 final product, 5,10-CH₂-THF, whose methylene group derived from the β-carbon of
71 serine or the α-carbon of glycine, is one of the few C1 donors in the biosynthesis
72 process, such as the biosynthesis of purine and methionine². Decrease or loss in the
73 activity of GCS will lead to glycine accumulation in human body, which is linked to
74 glycine encephalopathy⁴ (also known as nonketotic hyperglycinemia). Relevant
75 studies have shown that most patients with glycine encephalopathy have a P-protein
76 deficiency, and the rest are caused by T-protein or H-protein deficiency⁵. Moreover,
77 recent studies have shown that glycine metabolism is associated with tumorigenesis,
78 and P-protein as a key factor regulates glycolysis and methylglyoxal production in
79 cancer cells^{6,7}. In C3 plants, GCS is the key enzymatic system that deals with a large
80 amount of glycine in mitochondria during the photorespiration, and the activity of GCS
81 directly determines the growth rate of plants. Knockout of GCS gene is lethal to plants,
82 which is relevant to impaired one-carbon metabolism⁸, whereas overexpression of
83 L-protein⁹ or H-protein¹⁰ have been shown to improve photorespiration rates for
84 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 acidurici*¹¹, *Eubacterium acidaminophilum*¹² and *Arthrobacter globiformis*^{13, 14}. The reversibility of GCS was first discovered in mitochondrial extract of rat liver¹⁵, in *A. globiformis*¹⁶ and in cock liver mitochondria¹⁷, 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 CO₂ to produce pyruvate¹⁸, 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 CO₂ fixation pathway. Recently, rGP has been successfully introduced into *E.coli*^{20, 21, 22, 23} for autotrophic growth on formate and CO₂. 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 rGP²¹. 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 methylation, methylamine transfer and electron transfer in the enzymatic cycle of GCS²⁷.

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

130 for the evolution of life.

131 **Results**

132 ***Effects of components of the GCS reaction system on glycine cleavage and*** 133 ***synthesis***

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

159

160 **Table 1.** Effects of missing a certain component of the GCS reaction system on the
161 rates of glycine cleavage (determined as HCHO formation from the degradation of
162 5,10-CH₂-THF) and glycine synthesis.

Missing component	Glycine cleavage reaction		Glycine synthesis reaction	
	(μM HCHO $\cdot \text{min}^{-1}$)	GCS/ %	(μM glycine $\cdot \text{min}^{-1}$)	rGCS/ %
None (Reference)	22.48 $\pm$ 3.47	100.00	5.95 $\pm$ 0.13	100.00
P-protein	2.32 $\pm$ 0.52	10.34	2.03 $\pm$ 0.20	34.07
T-protein	11.67 $\pm$ 0.42	51.91	4.55 $\pm$ 0.16	76.53
L-protein	8.44 $\pm$ 0.57	37.55	4.45 $\pm$ 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 $\pm$ 2.67	110.86	0.98 $\pm$ 0.13	16.42
T-protein+THF	1.11 $\pm$ 0.33	4.93	0.43 $\pm$ 0.08	7.20
THF	0.87 $\pm$ 0.15	3.88	0.53 $\pm$ 0.04	8.94
NAD ⁺ /NADH	4.88 $\pm$ 1.54	21.73	5.76 $\pm$ 0.05	96.72
Glycine	0.00	0.00	-	-
NH ₄ HCO ₃	-	-	0.00	0.00
HCHO	-	-	0.98 $\pm$ 0.02	16.42

163

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

165 The results in **Table 1** suggested that P-protein, T-protein and L-protein are not
166 essential for the functionality of GCS both in glycine cleavage and glycine synthesis
167 directions. This led us to the question if H_{lip} alone can “catalyze” glycine formation
168 from NH₄HCO₃ and HCHO, or glycine cleavage in the opposite direction.

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

171 synthesis catalyzed by the all four GCS proteins (GCS as control), the reaction rate
172 catalyzed by H_{lip} alone at the same concentration of 10 μ M (H-10) was somewhat
173 lower, but glycine formation was well detected. With the increase of H_{ox} concentration
174 the initial reaction rate increased and the final concentration of glycine synthesized
175 was higher than that of the control. The above results prove that H_{lip} alone can
176 apparently catalyze the synthesis of glycine from NH₄HCO₃ and HCHO in the
177 presence of THF, PLP and DTT. To get the optimum reaction conditions for the glycine
178 synthesis, the reaction rate of glycine synthesis catalyzed by H_{lip} alone was
179 investigated under conditions of using different buffers, temperatures and pHs. **Figure**
180 **2b** shows the effect of different types of buffer, in which the order of the catalytic ability
181 of H_{lip} was as follows: Tris-HCl $\approx$ Mops > HEPES > PBS. The effect of temperature and
182 pH on the activity were studied by changing temperature from 4 $^{\circ}$ C to 95 $^{\circ}$ C (**Figure**
183 **2c**), and pH from 6.0 to 10.0 (**Figure 2d**). The reaction rate decreased sharply when
184 the temperature was higher than 42 $^{\circ}$ C or pH was higher than 8.0. The optimum
185 temperature and pH were at 37-42 $^{\circ}$ C and 7.5-8.0, respectively.

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

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

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

205 In order to explore the reasons behind the function of the stand-alone H_{lip} observed,
206 we further studied the effect of H_{lip} alone on the three GCS reaction steps, i.e., the
207 glycine decarboxylation reaction (accompanied by the reductive aminomethylation of
208 H_{ox} to H_{int}) in the absence of P-protein, the aminomethyl transfer reaction in the
209 absence of T-protein, and the electron transfer reaction without the presence of
210 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 μ M, 25 μ M, 50 μ M and 80 μ M, 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 μ M, 25 μ M, 50 μ M and 80 μ M, 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

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

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

240

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

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

252 **Aminomethyl transfer reaction in the absence of T-protein**

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

269

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

275 “Without T-protein” refers to a reaction mixture containing all reaction components and
276 enzymes except for T-protein. (b) Detection of H_{int} by HPLC proved that the aminomethyl
277 transfer reaction can take place spontaneously. “Normal group” refers to a reaction mixture
278 containing 50 μ M H_{ox} , 0.5 mM THF, 20 mM DTT, 50 mM NH_4Cl , 10 mM HCHO, and 5 μ M
279 T-protein; The reaction mixtures of “Without T” and “Without THF” were the same as that of
280 “Normal group”, but no T-protein or THF was added, respectively.

281

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

283 The function of L-protein is to catalyze the interconversion between the H_{ox} and its
284 reduced form dihydrolipoyl H-protein (H_{red}) involving $NAD^+/NADH$ and $FAD/FADH_2$.
285 The experimental results in **Table 1** show that in the absence of L-protein both the
286 glycine cleavage and synthesis systems can still work. Previous studies^{34, 35} have
287 shown that the disulfide bond-reducing agent TCEP allows the reduction of the
288 disulfide bond of the lipoyl group associated with the H-protein during the course of
289 the reaction catalyzed by L-protein. Since a certain amount of DTT was added to
290 prevent the oxidation of THF during the preparation of THF stock solution, we
291 expected that DTT should have the same reduction function as TCEP. **Figure 5a**
292 schematically shows the interconversion between H_{ox} and H_{red} through the
293 combination of the reduction of H_{ox} by TCEP (Left) or DTT (Right) with the re-oxidation
294 of H_{red} catalyzed by L-protein, yielding thereby $NADH$ from added NAD^+ . The
295 functionality of this combination was demonstrated by the relevant results presented
296 in **Figure 5b**. Next, we compared the effects of three different reducing agents (TCEP,
297 DTT and β -ME) on the reducibility of H_{ox} to H_{red} in different buffers (Tris-HCl, PBS,
298 Mops and HEPES) (**Figure 5c**). It was found that the reducibility of H_{ox} by DTT in four
299 different buffers is PBS > Mops > Tris-HCl > HEPES; TCEP had no reducing effect on
300 H_{ox} in Tris-HCl but showed the strongest reducing power in PBS, even much better
301 than DTT. β -ME was obviously not suitable to reduce H_{ox} in these four buffers. Thus,
302 for glycine synthesis, the functionality of L-protein can be well replaced by DTT or
303 TCEP, as also evidenced by the result shown in **Table 1** that the absence of
304 $NAD^+/NADH$ did not affect glycine synthesis.

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

342 The above results show that for *in vitro* GCS reactions, H_{lip} alone enables both the
343 glycine synthesis and the glycine cleavage without the presence of P-, T-, and
344 L-proteins. It seems that H_{lip} might functionally replace at least P- and L-proteins and
345 acts as glycine carboxylase and dihydrolipoyl dehydrogenase with the help of PLP
346 and FAD, respectively. It is also suggested by the experiment results shown in **Figure**
347 **5d** and **5e** that heated H_{ox} lost the catalytic function of L-protein, obviously because
348 heated H_{ox} cannot interact with FAD for the redox reaction of the lipoyl group bound to
349 H-protein.

350 We therefore further systematically studied the effects of heating (95 °C for 5 min)
351 on the catalytic activity of H_{lip} for glycine synthesis, in comparison to unheated H_{ox} as
352 well as to lipoic acid. By either using H_{lip} alone or combined with other GCS enzymes,
353 the overall reaction rate of glycine synthesis was measured. From the results in **Table**
354 **2**, we can ascertain several interesting observations and conclusions. First, the
355 overall glycine synthesis reaction could be catalyzed by the unheated H_{ox} alone;
356 adding P-protein significantly enhanced glycine synthesis, showing the importance of
357 P-protein; the addition of either T- or L-protein has no positive effect. In fact, the *in*
358 *vitro* glycine synthesis could run even better without T- and L-proteins, suggesting that
359 in the presence of H_{lip} the two individual steps, namely aminomethyl transfer and
360 electron transfer catalyzed by the two proteins respectively, could take place through
361 spontaneous aminomethylation of H_{red} to H_{int} in the presence of HCHO and NH₄⁺ and
362 reduction of H_{lip} to H_{red} by DTT. Second, compared with the unheated H_{ox}, the heated
363 H_{ox} alone could not catalyze the reaction of glycine synthesis; adding P-protein and
364 L-protein partially revived glycine synthesis to different extents, indicating that heated
365 H-protein lost the catalytic ability but was still functional as the shuttle protein of lipoyl
366 group; the addition of T-protein did not bring any effect. Furthermore, to our surprise,
367 when the heated H_{ox} was added with the other three GCS proteins to form a complete
368 GCS, the rate of glycine synthesis was even higher than the GCS containing the
369 unheated H_{ox}, indicating that heated H_{ox} loses its catalytic function but the
370 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 glycine 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} (μM • min ⁻¹)	heated H _{ox} (μM • min ⁻¹)	Lipoate acid (μM • 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-carrying 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 μ M H_{ox} , and 25 μ M 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 μ M 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 lipamide-methylamine arm protected in the cavity (ecH_{int}). (b) Docking of the lipamide-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.

435 An answer to the question of why heating destroys the catalytic ability of H-protein
436 but still preserves its function as shuttle protein might lie in the special surface
437 structure of H-protein. As shown by the crystal structure of H_{int} (**Figure 7a**), following
438 methylamine transfer the lipoamide-methylamine arm enters into the hydrophobic
439 cavity on the surface of the H-protein and prevents thereby it from nucleophilic attack
440 by water molecules^{28, 36, 37, 38}. Based on our previously molecular dynamic simulation
441 study, the lipoamide-methylamine arm may interact with some amino acid residues in
442 the proximity of the cavity³⁹, such as Glu-12, Glu-14, Ser-67, Cys-68 and Tyr-70
443 (**Figure 7b**). We speculated that the catalytic activity found for the stand-alone H_{lip} is
444 related to the structure of the cavity. Therefore, to verify this assumption, the above
445 five residues of the wild-type H-protein were mutated to alanine. Assays of glycine
446 synthesis were then performed with these mutated H_{lip} in comparison with the
447 wild-type H_{lip}. As shown in **Figure 7c**, compared to the wild-type H_{lip} the glycine
448 synthesis rates were strongly reduced in reaction mixtures containing H_{lip} mutants.
449 However, as shown in **Figure 7d**, when these mutants were combined with the other
450 three GCS proteins, the glycine synthesis rates of all the mutants were increased to
451 levels comparable or even better than that of the wild-type H_{lip}. The results are very
452 similar to what observed with heated H_{ox} (**Table 2**). It is therefore confirmed that the
453 cavity on the H-protein surface plays a decisive role in the catalytic functions of
454 H-protein, and alterations of the cavity structure (in size or form) either through
455 mutation or heating will reduce or even destroy the stand-alone catalytic functions of
456 H-protein (**Figure 7c, Table 2**) due to yet unclear mechanisms. One possible reason
457 might be that the lipoamide-methylamine arm cannot properly enter the deformed
458 cavity, resulting in the failure or imbalance of the GCS cycle in the presence of heated
459 H_{ox} alone. By adding other GCS proteins, H-protein is mainly required to act as a
460 shuttle protein and, consequently, GCS reactions can be revived (**Table 2**). Moreover,
461 under the given *in vitro* reaction conditions, the lipoamide-methylamine arm may
462 undergo fast and continuous reaction in the GCS cycle that not only minimizes the
463 probability of its hydrolysis, but also maintains or even increases the reaction
464 efficiency by omitting the process of its entry and exit from the cavity (**Figure 7d**,

465 **Table 2).**

466 **Discussion**

467 In this work, we show for the first time that stand-alone lipoylated H-protein (H_{lip})
 468 has the catalytic functions so far believed to be carried out by the P-, T- and L-proteins
 469 of GCS. It enables glycine cleavage reactions, as well as the reversed reactions
 470 towards glycine synthesis with NH_4HCO_3 and HCHO as the substrates. The K_{cat} value
 471 for the overall synthesis reaction is about $0.01\ s^{-1}$ for GCS catalyzed reaction and
 472 $0.0057\ s^{-1}$ for H-protein alone catalyzed reaction. After purification of the GCS
 473 proteins, we used the most commonly used methods SDS-PAGE and HPLC to verify
 474 that there was no obvious residual of other GCS proteins in the purified H-protein
 475 solutions, though we did not confirm this by using more precise methods like mass
 476 spectroscopy. Through calculations, we can state that even if other GCS proteins
 477 would exist in the H-protein solution (e.g. up to 10%), it will not qualitatively affect the
 478 main conclusions drawn in our work (see Supplementary Materials for detailed
 479 explanation). Therefore, the purity of H-protein meets the requirement needed for this
 480 study.

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

491 However, our experiments surprisingly show that when heated H_{ox} or a H_{ox} mutant
 492 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 activity³⁹. The molecular dynamic simulations of H-protein carried out by our group³⁹ 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.

Oliver⁴⁰ 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

551 T-protein expression under physiological conditions were able to grow and propagate
552 in normal air, only showing some minor changes. Meanwhile, their study also found
553 that the knockout mutation without T-protein expression was lethal even in
554 non-photorespiratory environment of air enriched to 1 % CO₂. This result may indicate
555 that THF needs weak binding with T-protein to participate in glycine cleavage and
556 synthesis reaction *in vivo*. When the T-protein is completely knocked out, THF cannot
557 participate in the reaction, resulting in slow reaction rate and plant death. The
558 deficiency of THF is far more detrimental and has the greatest impact on the reaction
559 rate of GCS *in vitro* (**Table 1**).

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

568 **Materials and Methods**

569 **Materials**

570 The substrates glycine, nicotinamide adenine dinucleotide (NAD⁺, NADH), Tris
571 (2-carboxyethyl) phosphine (TCEP) and the derivatization reagents dansyl chloride
572 were purchased from Yuanye Bio-Technology (Shanghai, China). Dithiothreitol (DTT),
573 β-mercaptoethanol (β-ME), pyridoxal 5'-phosphate monohydrate (PLP) and flavin
574 adenine dinucleotide (FAD) were obtained from Aladdin (Shanghai, China).
575 6-(RS)-Tetrahydrofolate (THF) was obtained from Sigma-Aldrich (St. Louis, MO, USA).
576 Other chemicals in this study were of analytical grade and purchased from Solarbio
577 (Beijing, China) or Sinopharm (Shanghai, China), unless otherwise noted. *Escherichia*
578 *coli* Top10 and BL21 (DE3) were used for plasmid construction and the

overexpression of recombinant proteins, respectively. Ni²⁺-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 OD₆₀₀ 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 V2³⁹. 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 previously⁴². 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
<i>E. coli</i> 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 Zhang et al. 2020³⁹
with point mutation of Y70A

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 μ M PLP, 5 mM NAD⁺, 5 μ M P-protein,
618 5 μ M T-protein, 5 μ M L-protein and 10 μ M 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 method¹⁷.

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 μ M PLP, 5 mM NAD⁺, 40 μ M FAD and 10 μ M 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 μ M PLP, 5 mM NADH,
631 5 μ M P-protein, 5 μ M T-protein, 5 μ M L-protein and 10 μ M H_{ox}. The reaction was
632 initiated by adding 50 mM NH₄HCO₃ 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 μ M PLP, 50 mM NH_4HCO_3 , 40 μ M FAD and 5 mM NADH, and different concentration of H_{ox} (10-80 μ M). 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 μ M H_{ox} and 25 μ M PLP. 5 μ M 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 μ M H_{ox} , 0.5 mM THF, 20 mM DTT, 50 mM NH_4Cl , and 10 mM HCHO. 5 μ M 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 μ M H_{ox} and 0.2 μ M L-protein. In order to prove that H_{ox} can still undergo redox reaction without L-protein, 40 μ M 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 lab³³. 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·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 NaHCO₃ and 200 µL of 5.4 mg·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 C₁₈ 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.

693 **Data availability**

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

698

699 **Acknowledgments and funding**

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

702

703 **Author Contributions**

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

711

712 **Reference**

- 713 1. Douce R, Bourguignon J, Neuburger M, Rébeillé F. The glycine decarboxylase system: a
714 fascinating complex. *Trends Plant Sci.* **6**, 167–177 (2001).
- 715 2. Kikuchi G, Motokawa Y, Yoshida T, Hiraga K. Glycine cleavage system: reaction mechanism,
716 physiological significance, and hyperglycinemia. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 246–263
717 (2008).
- 718 3. Hong Y, Ren J, Zhang X, Wang W, Zeng A-P. Quantitative analysis of glycine related
719 metabolic pathways for one-carbon synthetic biology. *Curr. Opin. Biotechnol.* **64**, 70–78
720 (2020).
- 721 4. Pardal-Fernández JM, Carrascosa-Romero MC, Vega CdC-dl, Iniasta-López I,
722 Martínez-Gutiérrez A. A typical glycine encephalopathy in an extremely low birth weight
723 infant: description of a new mutation and clinical and electroencephalographic analysis.
724 *Epileptic Disord.* **11**, 48–53 (2009).
- 725 5. Azize NAA, *et al.* Mutation analysis of glycine decarboxylase, aminomethyltransferase and
726 glycine cleavage system protein-H genes in 13 unrelated families with glycine
727 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-Calcano 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-CO₂ exchange system. *J. Biol. Chem.* **259**, 6085-6089
741 (1984).
- 742 12. Schneeberger A, Frings J, Schink B. Net synthesis of acetate from CO₂ 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, Lee SY. Assimilation of formic acid and CO₂ 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 CO₂
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 mitochondria. *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).