A new-to-nature carboxylation module to improve natural and synthetic CO₂ fixation

Marieke Scheffen¹, Daniel G. Marchal¹, Thomas Beneyton², Sandra K. Schuller³, Melanie Klose¹, Christoph Diehl¹, Jessica Lehmann¹, Pascal Pfister¹, Martina Carrillo¹, Hai He⁵, Selçuk Aslan⁵, Niña S. Cortina¹, Peter Claus¹, Daniel Bollschweiler¹, Jean-Christophe Baret²,⁷, Jan M. Schuller³,⁴,⁸, Jan Zarzycki¹, Arren Bar-Even⁵ and Tobias J. Erb⁵,¹⁴,✉

The capture of CO₂ by carboxylases is key to sustainable biocatalysis and a carbon-neutral bio-economy, yet currently limited to few naturally existing enzymes. Here, we developed glycolyl-CoA carboxylase (GCC), a new-to-nature enzyme, by combining rational design, high-throughput microfluidics and microplate screens. During this process, GCC’s catalytic efficiency improved by three orders of magnitude to match the properties of natural CO₂-fixing enzymes. We verified our active-site redesign with an atomic-resolution, 1.96-Å cryo-electron microscopy structure and engineered two more enzymes that, together with GCC, form a carboxylation module for the conversion of glycylate (C₂) to glycerate (C₃). We demonstrate how this module can be interfaced with natural photorespiration, ethylene glycol conversion and synthetic CO₂ fixation. Based on stoichiometrical calculations, GCC is predicted to increase the carbon efficiency of all of these processes by up to 150% while reducing their theoretical energy demand, showcasing how expanding the solution space of natural metabolism provides new opportunities for biotechnology and agriculture.

The conversion of the C₂ compounds glycolate and glyoxylate into C₃ metabolites plays a central role in many carbon metabolic processes, such as photorespiration and fatty acid assimilation, as well as formato- and methylotrophy. Moreover, glyoxylate is the product of several synthetic CO₂ fixation pathways, where it serves as a hub metabolite that connects to central carbon metabolism. However, there are only very few natural metabolic routes that allow the direct conversion of these C₂ intermediates into C₃ metabolites, and all of them result in the loss of carbon. The glyoxylate cycle¹ and the recently described β-hydroxyaspartate cycle² arrive at C₃ compounds, which need to be decarboxylated to generate C₃ metabolites. Similarly, photorespiration and the glycerate pathway convert two glyoxylate molecules into glycerate through the release of CO₂.³,⁴ The inevitable loss of CO₂ in all of these pathways strongly limits their carbon efficiency, which is especially apparent for photorespiration. It was estimated that in hot and dry climates, agricultural crop yield is diminished by up to 50% due to photorespiratory carbon losses.¹ Therefore, circumventing energy and carbon loss during glycolate assimilation through synthetic pathways is expected to enhance productivity substantially.

Recently, the tartronyl-CoA (TaCo) pathway was proposed as a direct route for the assimilation of glycolate into central carbon metabolism.¹ This hypothetical pathway was designed to fix CO₂ instead of releasing it, and is expected to outperform all naturally evolved glycolate assimilation routes (Fig. 1). However, the pathway remained theoretical until now, as neither tartronyl-CoA nor the individual enzymatic reactions of the TaCo pathway are known to occur in nature. This presented a challenge for its realization, with the engineering of glycolyl-CoA carboxylase (GCC)—the key enzyme of the TaCo pathway—as the biggest obstacle.

Here, we show the successful reconstitution and in vitro implementation of the complete TaCo pathway, applying rational design and high-throughput evolution of enzymes. Furthermore, we interface the TaCo pathway with several biotechnologically and agriculturally relevant processes (in particular, photorespiration, ethylene glycol conversion and synthetic CO₂ fixation) as proof of principle.

Results

Identifying the enzymatic framework for the TaCo pathway. To establish a biosynthetic route to the non-native carbohydrate substrate glycolyl-CoA, we investigated two possible options: a coenzyme A (CoA) transfer from another acyl-CoA donor by CoA transferases, as well as the direct ligation of glycolate to CoA by acyl-CoA synthetases. We screened 11 different native and engineered enzymes (Supplementary Table 1). All of the tested enzymes catalysed the activation of glycolate. The best transferase from Clostridium aminobutyricum (AbfT) showed a catalytic efficiency of 120 M⁻¹ s⁻¹ for the activation of glycolate with acetyl-CoA as the CoA donor (Supplementary Table 1).

The best CoA synthetase was an acetyl-CoA synthetase (ACS) homologue from Erythrobacter sp. NAP1 (EryACS1), which showed a catalytic efficiency of 20 M⁻¹ s⁻¹ with glycolate. It is well known that ACSs are post-translationally regulated through lysine acetylation in vivo.⁶ To suppress post-translational inactivation of EryACS1 during protein production, we changed surface loop residue Leu641 into a proline, which has been reported to prevent post-translational inactivation of EryACS1 during protein production, we changed surface loop residue Leu641 into a proline, which has been reported to prevent
Acetylation of ACS from *Salmonella enterica*<sup>11</sup>. Indeed, *EryACS1* Leu641Pro showed a twofold higher specific activity, but also a tenfold increased apparent *K<sub>c</sub>* for glycolate (Supplementary Table 1). Therefore, we decided to create a lysine acetylase knockdown strain for protein expression (*Escherichia coli* BL21 (DE3) Al Δ*apT*Z). When we produced *EryACS1* in this knockdown strain, the catalytic efficiency of the enzyme was increased almost 30-fold to 540 M<sup>−1</sup>s<sup>−1</sup> (Supplementary Table 1), which is sevenfold higher than the catalytic efficiency of a previously reported engineered *E. coli* ACS for the activation of glycolate (82 M<sup>−1</sup>s<sup>−1</sup>; deacetylated enzyme)<sup>7</sup>. Based on a homology model created with ACS of *S. enterica* (Protein Data Bank (PDB) ID: 2P2R; 62% sequence identity)<sup>12,13</sup>, we identified Val379 as a directed for target mutagenesis to open up the active site for accommodation of the slightly larger glycolate (Supplementary Fig. 1). A substitution of Val379 by alanine (Val379Ala) had previously been reported to enhance the activity of the enzyme with the slightly larger propionate in *S. enterica* ACS<sup>7</sup>. We tested three different Val379 substitutions (Supplementary Table 1), isolating the variant Val379Ala (glycolyl-CoA synthetase (GCS)), which showed an improved apparent *K<sub>c</sub>* for glycolate (13 ± 3 mM) and a catalytic efficiency of 853 M<sup>−1</sup>s<sup>−1</sup> (Table 1). Based on the favourable kinetic parameters and thermodynamic considerations (that is, the irreversibility of the reaction because of immediate hydrolysis of inorganic pyrophosphate in vivo), as well as the independence from other acyl-CoA pools in vivo compared with transferases, we decided to further rely on the engineered GCS for glycolyl-CoA synthesis (Table 1 and Supplementary Fig. 1).

In the TaCo pathway, the hypothetical product of glycolyl-CoA carboxylation, tartronyl-CoA, is reduced to glycerate via trartronic semialdehyde. We tested different reductases and found that the bifunctional malonyl-CoA reductase from *Chloroflexus aurantiacus*<sup>14</sup> (CaMCR; tartronyl-CoA reductase (TCR)) was able to convert tartronyl-CoA directly to glycerate in two steps with a *k<sub>c</sub>* of 1.4 s<sup>−1</sup> and a very favourable apparent *K<sub>c</sub>* of 26 μM for tartronyl-CoA (Table 1 and Supplementary Table 1).

**Identification and engineering of GCC.** With both GCS and TCR in hand, we focused our efforts on identifying a suitable enzyme candidate for the key reaction of the TaCo pathway, the carboxylation of glycolyl-CoA to tartronyl-CoA. We screened for promiscuity in four different biotin-dependent propionyl-CoA carboxylases (PCCs), because of the structural similarity between glycolyl-CoA and propionyl-CoA (Supplementary Table 1). Of the tested enzymes in this study, only PCC from *Methylorubrum extorquens* (MePCC) had minuscule activity with glycolyl-CoA. Co-expression of the cognate biotin ligase gene of *M. extorquens* resulted in very low but measurable activity (*k<sub>c</sub>* = 0.01 s<sup>−1</sup>; Fig. 2 and Supplementary Table 1) accompanied by a high ratio of futile ATP hydrolysis compared with tartronyl-CoA formation (~100:1).

Next, we sought to alter the substrate preference of MePCC. Such engineering of biotin-dependent acyl-CoA carboxylases had only been attempted in a few studies<sup>15–17</sup>. For structure-guided rational design, we built a structural model of MePCC, for which we obtained a 3.48-Å cryo-electron microscopy (cryo-EM) dataset (Fig. 2, Supplementary Fig. 2 and Supplementary Table 2). We then targeted several residues in the first and second shell of the active site of the carboxytransferase subunit, which form the substrate binding pocket. The selected residues were assumed to be involved in direct or indirect binding of the natural substrate propionyl-CoA, which differs in the Cα position from glycolyl-CoA by a hydroxyl group (Fig. 2, Supplementary Fig. 3 and Supplementary Table 1). To accommodate this hydroxyl group of glycolyl-CoA in the active site, we created variant GCC M2, in which we introduced a Tyr143His substitution to enable direct hydrogen bonding between enzyme and substrate, and further directed the hydroxyl group towards His143 via a Asp407Ile substitution, which had been reported to enhance substrate promiscuity in a PCC from *Streptomyces coelicolor*<sup>15</sup>. We introduced a third substitution in the active site’s second shell (Leu100Ser), to facilitate the formation of a hydrogen-bonding network and further strengthen the interaction of His143 with glycolyl-CoA. Compared with the wild type, the corresponding triple mutant (GCC M3) exhibited a more than 50-fold increased catalytic efficiency and a more than 15-fold decreased futile ATP hydrolysis (Fig. 2a–c and Supplementary Table 3).

Next, we applied directed evolution using high-throughput approaches to screen for mutants with further decreased futile ATP hydrolysis. We diversified the carboxytransferase subunit of GCC M3 using error-prone PCR. To assess whether the libraries contained enzymes with decreased futile ATP hydrolysis, we established a high-throughput microfluidics screen (Supplementary Fig. 4). In this screen, the formation of tartronyl-CoA was monitored in a TCR-coupled assay under ATP-limiting conditions within picolitre droplets. We used previously described microfluidic workflows<sup>18</sup> to develop short- and long-term multiplexed endpoint assays with single *E. coli* cells expressing individual GCC variants (Supplementary Figs. 5–7). Analysis of the pooled libraries showed that 4.7% of the variants exhibited more favourable ATP hydrolysis ratios compared with GCC M3 (Supplementary Fig. 8). This fraction of positive hits made it feasible to switch to microplate screens, which had the additional benefit that time-dependent kinetics of single variants could be directly obtained. We assessed each library separately to identify and characterize potentially improved enzyme variants. The best variant obtained from two rounds of subsequent random mutagenesis (GCC M5) possessed two additional substitutions (Ile450Val and Trp502Arg; Fig. 2e), a 560-fold increased *k<sub>c</sub>* of 5.6 s<sup>−1</sup> and a more than 25-fold lower ATP hydrolysis compared with the MePCC wild type (Table 1 and Fig. 2a–c). Overall, GCC M5 had a catalytic efficiency of 3.6 × 10<sup>4</sup>M<sup>−1</sup>s<sup>−1</sup> for glycolyl-CoA carboxylation, which is three orders of magnitude higher compared with the wild type (Supplementary Table 3) and similar to the catalytic efficiencies of naturally occurring biotin-dependent acyl-CoA carboxylases (Supplementary Fig. 9).

To exclude incomplete biotinylation of our GCC variants during protein production, we performed avidin-gel shift assays, which confirmed full biotinylation of all characterized enzymes (Supplementary Fig. 10). Furthermore, we assessed the thermostability of all...
enzymes by circular dichroism spectroscopy. All GCC variants exhibited melting temperatures similar to the wild type, indicating that enzyme stability was not affected through our mutations (Supplementary Fig. 11). Finally, we determined the catalytic efficiency of the enzymes with the natural substrate, propionyl-CoA. The catalytic efficiency varied among the different variants (Supplementary Table 3) and was decreased about twofold in GCC M5 compared with the wild type. Taken together, all of these results

---

**Diagram:**

- **Central carbon metabolism** (glycolysis, TCA cycle)
  - CBB cycle
  - RuBisCO
  - EG module
  - CETCH cycle
  - TCR
  - GCS
  - GCC
  - HCO3
  - ATP
  - CoA
  - ADP + Pi
  - NAD+
  - NADH
  - NADP+
  - NADPH
  - CO2
  - C3-SCoA
  - CoA
  - SCoA
  - ATP
  - ADP + Pi
  - NAD+
  - NADH
  - NADP+
  - NADPH
  - CO2
  - C3-SCoA
  - CoA
  - SCoA
  - ATP
  - ADP + Pi
  - NAD+
  - NADH
  - NADP+
  - NADPH
  - CO2
  - C3-SCoA
  - CoA
  - SCoA
  - ATP
  - ADP + Pi
  - NAD+
  - NADH
  - NADP+
  - NADPH
  - CO2
  - C3-SCoA
  - CoA

---

**Graphs:**

- **Graph a:**
  - Glycolyl-CoA
  - Ion count
  - Time (min)

- **Graph b:**
  - Glycolyl-CoA
  - Ion count
  - Time (min)

- **Graph c:**
  - 13C-tartronyl-CoA
  - Ion count
  - Time (min)

- **Graph d:**
  - 13C-glycerate
  - Ion count
  - Time (min)

- **Graph e:**
  - Malate yield (%)
  - Time (min)

- **Graph f:**
  - Glycerate (µM)
  - Time (min)

- **Graph g:**
  - Glycerate (µM)
  - Time (min)
show that we successfully created a GCC that behaves like a naturally evolved enzyme.

**Atomic-resolution cryo-EM structure of GCC.** To investigate the structural changes in GCC M5, we obtained a detailed 1.96-Å cryo-EM structure of the enzyme that validated our (rational) design strategy at atomic resolution (Fig. 2d,e, Supplementary Figs. 3 and 12 and Supplementary Table 2). For the cryo-EM structures of both MePCC and GCC M5, the central core comprised six β subunits, with well resolved and could be modelled without any gaps. In contrast, only the biotin carboxyl carrier protein domain and an anchoring domain could be modelled for the α subunits. Both of these domains are located right on top of their respective neighbouring β subunits and comprise about 200 amino acids of the C-termini. The actual catalytic domain of the α subunits, the biotin carboxylase domain, could not be modelled due to a much lower resolution and/or mostly disconnected weak electron densities (Supplementary Fig. 12e). The strong discrepancy in resolution and electron density between the β subunit core and the α subunits may hint at extreme flexibility within α subunits. It is therefore possible that, during catalysis, it is not just the biotin carboxyl carrier protein domain that undergoes conformational changes to bridge the distance between the active sites of the biotin carboxylase and the carboxyltransferase, as proposed by the swinging domain model¹⁹. Instead, also, the biotin carboxylase domains themselves may move or bend towards the β subunit core. In both structures, we also observed the biotin cofactor—interestingly, in a position analogous to what was reported for the crystal structure of a chimeric PCC holoenzyme⁴¹. The biotin is notably not located in a catalytically relevant position, as the nitrogen atom that accepts the carboxyl group is in close coordination contact with a backbone carbonyl oxygen (Supplementary Fig. 13d). Moreover, although the biotin is inserted into the neighbouring β subunit, it is still about 10 Å away from the site of the carboxyl transfer to the CoA thioester substrate. The fact that this exact position of the biotin was observed in our cryo-EM structures, as well as in the crystal structure of a homologous enzyme from another organism, suggests that this location is a possible parking position for the cofactor. Unfortunately, we were not able to observe the supplied substrate glycolyl-CoA in the active site of the carboxyltransferase β subunits, but only free CoA (Supplementary Fig. 13bc), probably due to spontaneous hydrolysis of the CoA thioester bond. Nevertheless, the resolution of the β subunit core of GCC M5 was high enough to model almost 900 ordered solvent water molecules with very low B factors. This allowed us to determine the actual distances between important active-site residues, which indicated that the Leu100Ser substitution did not directly hydrogen bond to His143 as initially assumed (Fig. 2e), but probably provided space for a new hydrophilic interaction between His143 and Asp171 in GCC M5. The Tyr143His substitution that was introduced to accommodate the hydroxyl group of glycolyl-CoA was actually held in a more favourable rotamer conformation by Asp171 through an almost perfect H bond of 2.8 Å (Fig. 2e and Supplementary Fig. 3). During random mutagenesis, two additional substitutions were introduced: Ile450Val and Trp502Arg. The Ile450Val substitution is located close to the active site in an α helix (Fig. 2e and Supplementary Fig. 3). Compared with the wild type, the Ile450Val substitution may slightly influence interatomic distances during catalysis. In contrast, the Trp502Arg substitution is far away from the active site, close to the rotational symmetry centre of the β subunit core, where it only affects the position of the loop region between Lys496 and Lys503. Its impact on enzyme activity cannot be simply rationalized by our structures, which represent a single non-catalytic state of the enzyme.

**In vitro reconstitution of the TaCo pathway.** Having established and engineered all of the enzymes, we next calculated the thermodynamic profile of the TaCo pathway using component contribution⁷. Overall, the max-min driving force (MDF)—representing the minimum thermodynamic driving force via the pathway reactions after optimizing metabolite concentrations within a physiological range—was above 7 kJ mol⁻¹ for the TaCo pathway (Methods and Supplementary Fig. 14). This high MDF predicts that all pathway reactions could work with minimal backward flux (that is, <20% of the total flux) and hence close to the maximum rate. In fact, the tartronyl-CoA module seems to be almost as thermodynamically favourable as the Calvin–Benson–Bassham (CBB) cycle itself (MDF approaching 8 kJ mol⁻¹; Supplementary Fig. 14) and more thermodynamically favourable than most central metabolism pathways, including Embden–Meyerhof–Parnas glycolysis (MDF < 2 kJ mol⁻¹)²². Next, we aimed to confirm the functionality of the full TaCo pathway by sequentially reconstituting its reaction sequence and isotopic labelling with ¹³C-bicarbonate (Fig. 1b–d). The reconstituted TaCo pathway converted glycolate into glycerate at a rate of 27 ± 1 nmol min⁻¹ mg⁻¹ total protein (Supplementary Fig. 15). To continuously operate and optimize the TaCo pathway, we coupled it to different ATP regeneration modules²²,²³. While the use of a polyphosphate kinase-based system was limited due to the precipitation of polyphosphate at concentrations above 20 mM, a phosphocreatine-based system proved three times more effective compared with the polyphosphate system (Supplementary Fig. 16). With these optimized conditions, we next aimed to test the TaCo pathway for three different potential applications.

**The TaCo pathway as photorespiratory bypass.** As proof of concept, we first interfaced the TaCo pathway with photorespiration. Natural photorespiration yields 2-phosphoglycerate (2-PG), which is recycled back into the CBB cycle through a complicated reaction sequence of 11 enzymes releasing NH₃ and CO₂ (Supplementary Fig. 17a). According to theoretical calculations, replacing natural photorespiration with the TaCo pathway would allow the direct conversion of 2-PG into the CBB cycle intermediate 3-phosphoglycerate (3-PGA) with only five enzymes, while fixing an additional carbon, instead of releasing it, thereby increasing the carbon efficiency from 75% to 150% (Supplementary Table 4). Furthermore, flux balance analysis showed that the combination of the CBB cycle and the TaCo pathway requires 21% less ATP and 29% less reducing equivalents for the net formation of one 3-PGA molecule from three CO₂ molecules, compared with the CBB cycle coupled with natural photorespiration (Fig. 3 and Supplementary Table 4).

### Table 1 | Kinetic parameters of the enzymes of the TaCo pathway

| Enzyme                  | Abbreviation | Substrate | kcat (s⁻¹) | Apparent Km (mM) | kcat/Km (M⁻¹ s⁻¹) |
|-------------------------|--------------|-----------|------------|-----------------|------------------|
| Glycolyl-CoA synthetase | GCS⁺         | Glycolate | 11.1 ± 0.6 | 13 ± 3          | 853              |
| Glycolyl-CoA carboxylase | GCC M5⁺     | Glycolyl-CoA | 5.6 ± 0.3 | 0.15 ± 0.03 | 3.6 × 10⁴        |
| Tartronyl-CoA reductase  | TCR          | Tartronyl-CoA | 1.4 ± 0.0 | 0.03 ± 0.00 | 5.4 × 10⁴        |

The data represent means ± s.d., as determined from n=18 independent measurements using nonlinear regression. *EryACS1 (Vaj379Aa), *MnPCC (Asp407Ile, Tyr143His, Leu100Ser, Ile450Val and Trp502Arg). *CoMCR.
Notably, TaCo-based photorespiration excels not only natural photorespiration regarding carbon and energy efficiencies (that is, cofactor requirements), but also other synthetic bypasses\textsuperscript{[8,26–30]} that were recently proposed and/or realized (Fig. 3, Supplementary Fig. 17 and Supplementary Table 4).

Starting from the photorespiration product 2-PG, we tested the TaCo pathway together with 2-PG phosphatase and glycerate kinase (GlxK) using \( ^{13} \text{C} \)-labelled bicarbonate. Our synthetic pathway produced \((R)\)-glycerate, which was further converted into labelled phosphoglycerate at a rate of 12.3 nmol min\(^{-1}\) mg\(^{-1}\), showing that the TaCo pathway can be successfully interfaced with photorespiration (Supplementary Fig. 18).

We further aimed at testing the TaCo pathway under conditions mimicking 100% RuBisCO oxygenation by the addition of equimolar amounts of 3-PGA and 2-PG. For quantification, we developed a malate read-out module based on isotopic labelling. The read-out module converts 3-PGA into malate, while 2-PG is only converted to malate when the TaCo pathway is active (Fig. 4a). Isotopic labelling with \( ^{13} \text{C} \)-bicarbonate allowed us to distinguish the fraction of malate derived from 3-PGA from that derived through the TaCo
pathway from 2-PG. 3-PGA is converted into malate via one carboxylation step (that is, the carboxylation of phosphoenolpyruvate (PEP) to oxaloacetate by PEP carboxylase; Fig. 4). This carboxylation introduces one \(^{13}\)C label, which leads to a single-labelled malate. The conversion of 2-PG into malate only takes place in the presence of the TaCo pathway. This requires two carboxylation reactions (and therefore double incorporation of \(^{13}\)CO\(_3\)\(^{-}\)); first, the carboxylation of glycolyl-CoA to tartronyl-CoA by GCC; and second, the carboxylation of PEP to oxaloacetate by PEP carboxylase (Fig. 4). Thus, the malate formed via the TaCo pathway is double labelled while malate formed from 3-PGA is only single labelled. Malate formation increased in the presence of the TaCo pathway to drive CO\(_2\) fixation, notably even in a 100% RuBisCO oxygenation reaction (that is, maximum photorespiration in the CBB cycle).

**Ethylene glycol conversion via the TaCo pathway.** Next, we tested the TaCo pathway in the context of ethylene glycol conversion. Ethylene glycol, a constituent of polyethylene terephthalate, is an environmental pollutant that is degraded by aerobic microbes via the glycerate pathway under CO\(_2\) release\(^{33,34}\). We conceived of a TaCo pathway-based route for the conversion of ethylene glycol into the central metabolite glycerate that would increase the carbon efficiency of ethylene glycol assimilation from 75\% to 150\% (Fig. 1a). We combined GCC and TCR of the TaCo pathway with the (L)-lactaldehyde dehydrogenase of *E. coli* (FucO\(^{35}\), which oxidizes ethylene glycol into glycolaldehyde) and an aldehyde dehydrogenase of *Rhodopseudomonas palustris* (BisB18 (PduP\(^{36,37}\) to convert glycolaldehyde into glycolyl-CoA). A first version of the synthetic pathway produced 77\(\mu\)M glycerate over 2h with an initial rate of 0.4 nmol min\(^{-1}\) mg\(^{-1}\) in vitro (Supplementary Fig. 19). To improve the initial ethylene glycol oxidation, we replaced FucO with Gox0313 from *Gluconobacter oxydans* \(^2\) (Supplementary Table 1) and introduced a water-forming NADH oxidase \(^3\) to maintain a high NAD\(^{+}\) concentration. Together with the integration of an efficient ATP regeneration system, glycerate production was increased to 485\(\mu\)M at a fivefold higher rate of 2.1 nmol min\(^{-1}\) mg\(^{-1}\), demonstrating efficient CO\(_2\)-dependent conversion of the environmental pollutant and plastic component ethylene glycol into a central carbon metabolite through the TaCo pathway (Fig. 1f and Supplementary Fig. 19).

**Synthetic CO\(_2\) fixation by the TaCo pathway.** Finally, we tested the TaCo pathway in the context of the crotonyl-CoA/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle, a recently developed synthetic CO\(_2\) fixation pathway\(^2\). The primary CO\(_2\) fixation product of the CETCH cycle is glyoxylate, which can be converted into glycolate using glyoxylate reductase (Fig. 1a). Extending the CETCH cycle with the TaCo pathway would add another CO\(_2\) fixation step and allow direct formation of the central carbon metabolite glycerate from three CO\(_2\) molecules. Thus, the use of GCC directly increases carbon efficiency of the CETCH cycle by 25\% (Supplementary Table 5). For proof of principle, we coupled the CETCH cycle and TaCo pathway and optimized their interplay in several rounds. When we initially combined the 17 enzymes of CETCH (version 5.4)\(^2\) with a semialdehyde reductase for glyoxylate reduction (Gox1801 from *G. oxydans* and the TaCo pathway, the TaCo pathway produced glycerate at rates of 0.1 and 0.4 nmol min\(^{-1}\) mg\(^{-1}\) in continuous and stepwise assays, respectively (Supplementary Fig. 20). Further analysis identified two potential bottlenecks of the overall system. First, the use of formate and formate dehydrogenase to regenerate NADPH led to production of the dead-end metabolite formyl-CoA due to a side reaction of GCS, trapping CoA and consuming ATP. Second, succinyl-CoA reductase of *Clostridium kluyveri* (CkSucD), which was used in the CETCH cycle, had a side reactivity with glycolyl-CoA of 311\(\mu\)M, corresponding to a surplus of 33\%. \(^{13}\)C labelling confirmed that the malate surplus was provided by the TaCo pathway (Figs. 1e and 4). In a control experiment, in which the TaCo pathway was disconnected from the read-out module by omission of GlxK, no additional malate was formed, but 139\(\mu\)M glycerate accumulated (Fig. 4d).

Notably, GlxK from *E. coli* only accepts the (R)-stereoisomer of glycerate\(^{15,17}\), thus confirming that the physiologically relevant stereoisomer (R)-glycerate is the product of the TaCo pathway. Overall, these results showcase the potential of the TaCo pathway as an efficient ATP regeneration system, glycerate production was increased to 485\(\mu\)M at a fivefold higher rate of 2.1 nmol min\(^{-1}\) mg\(^{-1}\), demonstrating efficient CO\(_2\)-dependent conversion of the environmental pollutant and plastic component ethylene glycol into a central carbon metabolite through the TaCo pathway (Fig. 1f and Supplementary Fig. 19).
Fig. 20), which is comparable to CO₂ fixation rates of the CETCH cycle alone (5 nmol min⁻¹ mg⁻¹). These results demonstrate successful combination of the TaCo pathway with the synthetic CETCH cycle into a more carbon- and energy-efficient autotrophic pathway.

Conclusions
We have constructed a new-to-nature metabolic reaction sequence for the CO₂-dependent assimilation of glycolate and demonstrated its use for different applications. To develop the TaCo pathway, it was necessary to identify and engineer enzymes able to catalyse reactions, which to the best of our knowledge are not known from any natural pathways. A crucial step was the engineering of a new-to-nature CO₂-fixing enzyme, GCC, which shows great potential to improve photosynthetic yield in natural and synthetic carbon fixation, even under very low CO₂ conditions and with 100% oxygenation of RuBisCO.

Our engineered enzymes of the TaCo pathway show turnover numbers (1.4–11.1 s⁻¹) that are comparable to naturally evolved enzymes, which show on average a kcat of 10 s⁻¹ (ref. 39), while most RubisCOs notably fall in the range of 1–10 s⁻¹ (ref. 40). Nevertheless, to unlock the full potential of the TaCo pathway, future efforts might involve directed evolution to further improve enzyme...
activities, enhance flux through the pathway and potentially eliminate the unfruitful ATP hydrolysis of GCC.

Overall, our results showcase how hitherto unknown but theoretically feasible enzyme reactions can be developed on the scaffold of naturally existing proteins to extend the solution space of natural metabolism\textsuperscript{41–45}. We expect similar future approaches to allow access to novel routes with carbon and energy efficiencies superior to naturally evolved pathways\textsuperscript{46}, which may greatly impact current efforts in biotechnology and metabolic engineering.

Methods

Materials. Chemicals were obtained from Sigma–Aldrich, Carl Roth, Santa Cruz Biotechnology and Merck. NaH\textsubscript{13}CO\textsubscript{3} was obtained from Cambridge Isotope Laboratories. Biochemicals and materials for cloning and protein expression were obtained from Thermo Fisher Scientific, New England Biolabs and Macherey–Nagel. Coenzyme A was bought from Roche Diagnostics. Materials and equipment for protein purification were obtained from GE Healthcare, Bio-Rad and Merck Millipore. Pyruvate kinase/lactic dehydrogenase, malic dehydrogenase, glucose-6-phosphate dehydrogenase, glucose dehydrogenase and PEP carboxylase were bought from Sigma–Aldrich.

Library generation of GCC variants. Plasmid libraries of randomly mutagenized GCC were created by megaprimed-based whole-plasmid PCR (MEGAHWOP)\textsuperscript{[17]}. To generate randomized fragments of the \(e\) subunit of GCC M3 (\(\text{pET1412}\)) for the first round of mutagenesis, error-prone PCR\textsuperscript{[18]} was performed using 2.5 U Taq-polymerase with Mg-free buffer (New England Biolabs; M03205), 7 mM MgCl\textsubscript{2}, 0.4 mM dGTP and dATP each, 2 mM dCTP and dTTP each, 0.4 µM primer PccB\textsubscript{FW_P1}, 0.4 µM primer PccB\textsubscript{REV_P2} (Supplementary Table 6), 10% (vol/vol) dimethyl sulfoxide, 50 ng template DNA of pET1412 and 200–500 µM MnCl\textsubscript{2} in a 50 µl reaction. The randomized fragments were digested with DpnI, purified again by agarose gel electrophoresis and used as mega primers for a whole-plasmid PCR (MEGAHWOP), as described elsewhere\textsuperscript{[17]}, or subjected to another error-prone PCR reaction to further increase the mutation rate. The MEGAHWOP reaction (50 µl) contained 1× KOD Hot Start reaction buffer (Novagen), 0.2 mM dNTPs, 1.5 mM KOD Hot Start reaction buffer (Novagen), 0.2 mM dNTPs, 1.5 mM MgCl\textsubscript{2} in subsequent error-prone PCR); library 1_2 contained 1.1 mutations per kbp (500 µM MnCl\textsubscript{2}); and library 1_3 contained 0.2 mutations per kbp (200 µM MnCl\textsubscript{2}).

For the second round of random mutagenesis, plasmid pTE3100 (GCC M4) was used as a template for error-free PCR as well as for MEGAHWOP, following the procedure described above. Here, we aimed to create libraries with mutation rates of 0.2 and 1.1 per kbp and used 200 and 500 µM MnCl\textsubscript{2}, respectively.

Screen of GCC libraries in microwell plates. The libraries 1_1, 1_2 and 1_3 were transformed into E. coli BL21 BirA (for the bacterial strains used, see the Supplementary Methods) and colonies were picked into 96-deep-well plates (PlateOne) with lysogeny broth (Miller) containing 100 µg ml\textsuperscript{-1} spectinomycin and 2 µg ml\textsuperscript{-1} tetracycline (for detailed UPLC-MS/MS methods, see the Supplementary Methods).

Ethylene glycol conversion by the TaCo pathway. The conversion of ethylene glycol by the TaCo pathway was determined by measuring glycerate formation. All of the assays were run at 37 °C for 2 h. The samples were withdrawn at different time points and immediately quenched with 1% HCl. The samples were centrifuged at 17,000g for 20 min at 4 °C, derivatized with 3-NP and analyzed via UPLC-MS/MS for malate and glycerate (for detailed UPLC-MS/MS methods, see the Supplementary Methods).

CT experiment 1. The assay for CT experiment 1 was run as described earlier for CETCH (version 5.4)\textsuperscript{[14]} with the following changes. The assay mix did not contain Mcl or Gox and additionally contained 434 µg ml\textsuperscript{-1} TPK2 II. The assay for the no Gox control additionally contained 4.6 mg ml\textsuperscript{-1} GCS, 0.6 mg ml\textsuperscript{-1} GCC M4 and 1.07 mg ml\textsuperscript{-1} TCR. All of the assays were started with the addition of 5 mM ATP, 1 mM 2-PG and 1 mM 3-PGA. Aliquot samples were withdrawn at different time points and immediately quenched with 1% HCl. The samples were centrifuged at 17,000g for 20 min at 4 °C, derivatized with 3-NP and analyzed via UPLC-MS/MS for malate and glycerate (for detailed UPLC-MS/MS methods, see the Supplementary Methods).

CT experiment 2. The assay for CT experiment 2 was run at 30 °C and 450 r.p.m. and contained 200 mM MOPS/KOH (pH 7.5), 10 mM MgCl\textsubscript{2}, 1 mM coenzyme A, 20 mM ATP, 10 mM NADPH, 40 mM formate, 40 mM polyphosphate, 100 mM NaH\textsubscript{13}CO\textsubscript{3} and 1.6 µg ml\textsuperscript{-1} carbonic anhydrase. We used the following amounts of enzymes as described previously for CETCH (version 5.4)\textsuperscript{[14]}: Instead of GlcB and Mcl, we added 200 µl Gox1801. The reaction was started with the addition of 2 µM pimonidazole-CoA. After 120 min, the TaCo assay mix (final concentrations: 20 mM ATP, 0.5 mM coenzyme A, 2 mg ml\textsuperscript{-1} GCS, 2 mg ml\textsuperscript{-1} GCC M4 and 1 mg ml\textsuperscript{-1} TCR) was added in a 1:2 ratio and the reactions were transferred to 37 °C. For the negative control, buffer was added, corresponding to the concentrations in the TaCo mix. Aliquot samples were withdrawn at different time points, quenched in 4% formic acid and centrifuged at 17,000g for 20 min at 4 °C. The glycerate concentration of the samples was measured using derivatization with 3-NP following analysis via UPLC-MS/MS (for detailed UPLC-MS/MS methods, see the Supplementary Methods).
CT experiment 3. The assay for CT experiment 3 contained 100 mM HEPES (pH 7.5), 5 mM MgCl₂, 0.5 mM CoA, 2 mM ATP, 5 mM NADPH, 20 mM glucose, 20 mM phosphate, 50 mM NaHCO₃, 250 µM propionyl-CoA, 0.8 µg/ml carbonic anhydrase, and the CETCH enzymes in the amounts previously described² (CETCH version 5.4). Instead of GlcB, Mcl and Fdh, we added 100 mM Gox1801 and glucose dehydrogenase (Sigma–Aldrich). The assay was run for 3 h at 30°C and then diluted in a 3:1 ratio with the TaCo master mix (final concentration: 100 mM HEPES–HCl (pH 7.5), 5 mM ATP, 10 mM MgCl₂, 2 mM coenzyme A, 2 mM GTP, 50 mM glucose, 100 mM potassium phosphate, 6.6 mM phosphocreatine, 22 µM imidoadenosine 5’-triphosphate, 2 mM MgCl₂, and 4 mM glycoly-CoA was added to cold-discharged Quantifoil 2/1 grids, blotted with 3.5 s with force 4 in a Vitrobot Mark IV (Thermo Fisher Scientific) at 100% humidity and 4°C and then frozen in liquid ethane that had been cooled by liquid nitrogen.

For MePCC, cryo-EM data were collected on an FEI Glacios transmission electron microscope operated at 200 keV using the SerialEM software package⁴. A total of 2,437 video frames were recorded at a calibrated pixel size of 1.18 Å using a magnification of 35,000×, for which 41% of total frames were damaged or were not particle candidates and damaged particles. Ab initio model generation using the software package. The dose-fractionated videos were gain-normalized, aligned, defocused and higher-order contrast transfer function (CTF) correction in the Refinement_New algorithm, the resolution was improved to 3.48 Å with a B-factor of 86.8 Å².

For GCC M5, cryo-EM data were acquired with an FEI Titan Krios transmission electron microscope using SerialEM software. Video frames were recorded at a nominal magnification of 105,000x (calibrated physical pixel size: 0.8512 Å px⁻¹) using a K3 direct electron detector (Gatan) in correlated double sampling mode and a GIF quantum energy filter (Gatan) at a 20 eV slit width. The total electron dose of 55–58 Å² was distributed over 35 frames. Micrographs were recorded in a defocus range of −0.5 to −3.5 µm.

All subsequent imaging processing steps were performed in the cryoSPARC software package. The dose-fractionated videos were gain-normalized, aligned and dose-weighted using the patch motion algorithm. Defocus values were estimated using the CTFFIND4 implementation, and 168,488 particles were picked template free on 20,057 acquired micrographs. Candidate particles were imported into cryoSPARC, used for ab initio construction of the GCTF 52. Using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/), 5,398,283 particles were extracted with a pixel box size of 340 Å using RELION 3.1 and 168,488 particles were subjected to multiple rounds of heterogeneous refinement to a local resolution of 1.9 Å in the β subunit core of the molecule and a flexible amino-terminal α subunit. Visual inspection of the resulting electron density map revealed protein features that were consistent with the determined high resolution (for example, holes in the aromatic ring systems and the zigzag structure of extended sidechains such as Arg and Lys). Additionally, in the density maps, coordinated water molecules were visible (Supplementary Fig. 12).

Structural modelling and analysis. Homology modelling was performed using SWISS-MODEL⁵. As a template for homology modelling of EryA3C, the structure of an ACS mutant of S. enterica (PDB ID 2P2B; 62% amino acid identity) was used, which had a adenosine-5’-monophosphate-propyl ester (propyl-AMP) bound—an analogue of the reaction intermediate propionyl-AMP. Based on the position of the propyl-AMP in the S. enterica ACS, a glycolyl-AMP was modelled into the active site of the EryA3C homology model.

For initial homology modelling of the β subunit of MePCC, the structure of the β subunit of PCC from Ruegeria pomeroyi (PDB ID 3N6R; 72% amino acid identity) was used.

Initial cryo-EM map fitting was performed in UCSF Chimera version 1.14 (ref. ⁶) using homology models based on PDB ID 3N6R⁷. Automatic refinement of the structure was done using phenix.real_space_refine of the PHENIX 1.17.3–3660 suite⁸. Manual refinements and water picking were performed with COOT 0.8.9.2 (ref. ⁹). All structural depictions were created using PyMOL (the PyMOL Molecular Graphics System; version 1.8; Schrödinger). Modelling of propionyl-CoA and glycolyl-CoA into the active sites of the MePCC and GCC, respectively, was based on the position of methylmalonyl-CoA in the structure of a methylmalonyl-CoA carboxyltransferase from Propionibacterium freudenreichii (PDB ID 1JNO; 32% amino acid identity). Manual fitting and adjustments of the CoA thioesters reflecting differences in active-site architectures were done with COOT and PyMOL.

Flux balance analysis. To compare the natural and synthetic photosresponsive pathways in terms of consumption of ATP, NADP+ and ferredoxin, as well as required terns of RuBiCO, a stoichiometric analysis was performed by applying flux balance analysis with COPRAp⁹. Simplified models were constructed including the reactions of CBB, the specific reactions of each photosresponsive pathway and cofactor regeneration reactions for example, ADP + inorganic phosphate → ATP; NAD+ → NADH and NADH + NADP+ → NAD+ + NADP+, as well as key interconversion reactions, such as adenylate kinase. The ratio between the carbonic anhydrase and oxygenation reactions catalysed by RuBiCO was set to 3:1 (25% oxygenation). The stoichiometric requirement of ATP, NADP+, ferredoxin and RuBiCO turns (that is, the sum of carboxylation and oxygenation reactions) was computed for the production of one 3-PGA molecule. The full code and the list of reactions of each model can be found at https://github.com/he-hai/PubSuppl within the 2020_TaCo directory.

Pathway thermodynamics analysis. MDF analysis was applied to evaluate the thermodynamics feasibility of the taurinyl-CoA module. The Python packages equilibrator_api, equilibrator_assets and equilibrator_pathway were used for the analysis. The change in Gibbs energy of the reactions was estimated using the component contribution method⁴. CO₂ was considered as the substrate for the carboxylation reactions as its concentration is pH independent, unlike that of bicarbonate, thus simplifying the calculations. Metabolite concentrations were constrained to the range 1 µM to 10 mM, as described previously². The pH was assumed to be 7.0, the ionic strength was assumed to be 0.25 M and –log[Mg²⁺] (pMg) was assumed to be 3. The scripts and details can be found at https://github.com/he-hai/PubSuppl within the 2020_TaCo directory.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability. The structural data that support the findings of this study have been deposited in the PDB with the accession codes 6YBP (MePCC wild type) and 6YBQ (GCC M5), and in the Electron Microscopy Data Bank with the accession codes EMD-10770 (MePCC wild type) and EMD-10771 (GCC M5). Other publicly available PDB structures used in this study are 2P2B, 3N6R and 1ON3. Other data supporting the findings of this study are available in the paper and Supplementary Information, or from the corresponding author upon reasonable request.

Code availability. For flux balance analysis of different photosresponsive pathways, the full code and list of reactions of each model, as well as the scripts and details of the pathway thermodynamics analysis, can be found at https://github.com/he-hai/PubSuppl within the 2020_TaCo directory.

Received: 30 June 2020; Accepted: 19 November 2020; Published: 04 January 2021

References
1. Bar-Even, A., Noor, E., Lewis, N. E. & Milo, R. Design and analysis of synthetic carbon fixation pathways. Proc. Natl Acad. Sci. USA 107, 8899–8904 (2010).
2. Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N. S. & Erb, T. I. A synthetic pathway for the fixation of carbon dioxide in vitro. Science 354, 900–904 (2016).
3. Kornberg, H. L. & Krebs, H. A. Synthesis of cell constituents from C₃-units by a modified tricarboxylic acid cycle. Nature 179, 988–991 (1957).
4. Schada von Borzyskowski, L. et al. Marine Proteobacteria metabolize glycolate via the β-hydroxyaspartate cycle. Nature 575, 500–504 (2019).
5. Krakov, G., Barkulis, S. S. & Hayashi, J. A. Glyoxylate acid carboxilase: an enzyme present in glycolate-grown Escherichia coli. J. Bacteriol. 81, 509–518 (1961).
27. Yu, H., Li, X., Duchoud, F., Chuang, D. S. & Liao, J. C. Augmenting the synthesis of acetyl-CoA by deacetylating of active lysine. Science 298, 2390–2392 (2002).

28. Starai, V. J. & Escalante-Semerena, J. C. Identification of the protein acetyltransferase (Pat) enzyme that acetylates acetyl-CoA synthetase in Salmonella enterica. J. Mol. Biol. 340, 1005–1012 (2004).

29. Starai, V. J., Gardiner, J. G. & Escalante-Semerena, J. C. Residue Leu-641 of the beta-subunit of acetyl-CoA carboxylase: structure-based engineering of substrate specificity. Biochemistry 43, 14027–14036 (2004).

30. South, P. F., Cavanagh, A. P., Liu, H. W. & Ort, D. R. Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. J. Mol. Biol. 289, 407–412 (2002).

31. Erb, T. J., Jones, P. R. & Bar-Even, A. Synthetic metabolism: metabolic engineering meets enzyme design.Curr. Opin. Chem. Biol. 37, 56–62 (2017).

32. Miyaizaki, K. MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. Methods Enzymol. 498, 399–406 (2011).

33. Chou, A., Clomburg, J. M., Qian, S. & Gonzalez, R. 2-Hydroxyacyl-CoA lyase catalyzes acetoxylation of dihydroxyacetone. J. Biol. Chem. 289, 863–891 (2013).

34. Li, W. J. et al. Laboratory evolution reveals the metabolic and regulatory basis of ethylene glycol metabolism by Pseudomonas putida KT2440. Environ. Microbiol. 21, 3669–3682 (2019).

35. Sridhar, S., Wu, T. T., Choudhury, T. M. & Lin, E. C. F. Ferrous-activated nitroimidazole adenine dinucleotide-linked dehydrogenase from a mutant of Escherichia coli capable of growth on 1,2-propanediol. J. Bacteriol. 98, 87–95 (1969).

36. Zarzycki, J., Sutter, M., Cortina, N. S., Erb, T. J. & Kerfeld, C. A. In vitro characterization and concerted function of three core enzymes of a glyoxyl cycle enzyme-associated bacterial microcompartment. Sci. Rep. 7, 42757 (2017).

37. Zhang, X., Zhang, B., Lin, J. & Wei, D. Oxidation of ethylene glycol to glycolaldehyde using a highly selective alcohol dehydrogenase from Gluconobacter oxydans. J. Mol. Catal. B Enzym. 112, 69–75 (2015).

38. Nakaj, C. et al. A water-forming NADH oxidase from Lactobacillus pentosus pentosus suitable for the regeneration of synthetic biomimetic cofactors. Front. Microbiol. 6, 957 (2015).

39. Bar-Even, A. et al. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. Biochemistry 50, 4042–4410 (2011).

40. Flambolz, A. I. et al. Revisiting trade-offs between Rubisco kinetic parameters. Biochemistry 58, 3365–3376 (2019).

41. Erb, T. J. Back to the future: why we need enzymology to build a synthetic metabolism of the future. Beilstein J. Org. Chem. 15, 551–557 (2019).

42. Zheng, S. Q. et al. Emergent chemoautotrophy during evolution of a highly active artificial aldolase. Nat. Chem. 9, 50–56 (2017).

43. Chou, A., Clomburg, J. M., Qian, S. & Gonzalez, R. 2-Hydroxyacyl-CoA lyase catalyzes acetoxylation of dihydroxyacetone. J. Biol. Chem. 289, 900–906 (2019).

44. Siegel, J. B. et al. Computational protein design enables a novel one-carbon metabolic pathway. Proc. Natl Acad. Sci. USA 112, 3784–3790 (2015).

45. Chen, K. & Arnold, F. H. Engineering new catalytic activities in enzymes. Nat. Catal. 3, 203–213 (2020).

46. Erb, T. J., Jones, P. R. & Bar-Even, A. Synthetic metabolism: metabolic engineering meets enzyme design. Curr. Opin. Chem. Biol. 37, 56–62 (2017).

47. Miyazaki, K. MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. Methods Enzymol. 498, 399–406 (2011).

48. Girino, P. C., Mayer, K. M. & Umeno, D. Generating mutant libraries using error-prone PCR. Methods Mol. Biol. 231, 3–9 (2003).

49. Macrornarte, D. N. Automated electro microscop tomography using robust prediction of specimen movements. J. Struct. Biol. 152, 36–51 (2005).

50. Kyani, E. et al. Focus: the interface between data collection and data processing in cryo-EM. J. Struct. Biol. 198, 124–133 (2017).

51. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332 (2017).

52. Li, X., Grigorieff, N. & Cheng, Y. GPU-enabled FREALIGN: accelerating reconstruction of high-resolution 3D cryo-EM structures. J. Struct. Biol. 172, 407–412 (2010).

53. Scheres, S. H. A bayesian view on cryo-EM structure determination. J. Mol. Biol. 415, 406–418 (2012).

54. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. AcryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296 (2017).

55. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612 (2004).

56. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D. Struct. Biol. 75, 861–877 (2019).

57. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D. Biol. Crystallogr. 66, 486–501 (2010).

58. Hall, S. P. R. et al. Transcrysylase 125 structural crystal symmetry assembly and substrate binding to a multienzyme complex. EMBO J. 22, 2334–2347 (2003).

59. Ebrahimi, A., Lerman, J. A., Palsson, B. O. & Hyduke, D. R. COBRAPP: constraints-based reconstruction and analysis for Python. BMC Syst. Biol. 7, 74 (2013).

60. Shankar, T. D. Estimating the rate of photorespiration in leaves. Physiol. Plant. 73, 147–152 (1988).
programme, part of the University of Bordeaux Initiative of Excellence (IDEX Bordeaux; ANR-10-IDEX-03-02). J.M.S. is grateful to the DFG for an Emmy Noether grant (SCHU 3364/1-1). We thank E. Noor (ETH Zürich) for help with the thermodynamic calculations. We are thankful to the Core Facility for Protein Spectroscopy at the Institute of Cytobiology (Philipps University of Marburg), and especially to S.-A. Freibert and V. Schulz for the circular dichroism measurements of GCC variants.

Author contributions
M.S., J.Z., S.A., A.B.-E. and T.J.E. conceived of the project. M.S., D.G.M. and J.Z. identified and tested the GCS, GCC and TCR candidates for the TaCo pathway, and Gox0313 and PduP for ethylene glycol conversion. P.P. tested the PCC homologues for GCC activity and the influence of biotin ligase on carboxylation activity. S.K.S., D.B. and J.M.S. performed the cryo-EM sample preparation and analysis. J.Z. refined the cryo-EM structures. J.L. constructed the GCS expression strain E. coli BL21 (DE3) AI ΔpatZ and performed the initial characterization of GCS. M.C. identified and tested CdSucD and Gox1801 for the CETCh cycle. J.Z., M.S. and T.J.E. designed active-site variants for GCC and GCS. M.S. and M.K. designed error-prone libraries for GCC. M.S., T.B. and J.-C.B. developed the microfluidics platform to screen the GCC libraries. M.S. and D.G.M. developed and performed microtitre plate assays to screen the GCC libraries, and isolated and characterized the GCC variants. M.S. performed all of the in vitro reconstruction experiments and optimizations. C.D. and M.S. performed the CETCh in vitro reconstruction experiments. N.C.S., C.D. and P.C. performed the mass spectrometric analyses. H.H. performed the flux balance analysis and thermodynamic calculations. The manuscript was written by M.S., J.Z. and T.J.E. with contributions from all other authors.

Funding
Open access funding provided by Max Planck Society

Competing interests
T.J.E. and A.B.-E. have filed a patent covering the TaCo pathway. The other authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41929-020-00557-y.

Correspondence and requests for materials should be addressed to T.J.E.

Peer review information Nature Catalysis thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed

☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

☐ The statistical test(s) used AND whether they are one- or two-sided

Only common tests should be described solely by name; describe more complex techniques in the Methods section.

☐ A description of all covariates tested

☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

☐ For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted

Give P values as exact values wherever suitable.

☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

☐ Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Data were collected using Agilent Technologies Cary WinUV Kinetics Application V.5.0.0.999, Agilent Technologies MassHunter Workstation Software Data Acquisition for 6500 Series QTOF Version B.06.01 Build 6.0.1.6157, Agilent Technologies MassHunter Workstation Software Data Acquisition for 6400 Series Triple Quadrupole Version B.08.02 Build B.2.8260.0, Tecan Austria GmbH Tecan l-control Version 3.9.1.0, Microfluidic data were collected using a field programmable gate array board (cNIO; National Instruments) using a program written in LabView (National Instruments). Cryo-EM data were collected using the SerialEM software package [SerialEM_3-8-0beta8_64 & SerialEM_3-8-0beta11_64]. CD data were collected using Jasco Spectra Manager 2.06.

Data analysis

Data were analyzed using Microsoft Excel 2013 (15.0.5153.1000) MSO (15.0.5153.1000) 32-bit, GraphPad Prism Version 8.0.0, Agilent Technologies MassHunter Workstation Software Quantitative Analysis for TOF Version B.09.00 Build 9.0.647.0, Agilent Technologies MassHunter Workstation Software Quantitative Analysis for QQQ Version B.09.00 Build 9.0.647.0, Agilent Technologies MassHunter Workstation Software Qualitative Analysis Navigator Version 9.08.00 Build 9.0.8208.0, Schrodinger LLC. The PyMOL Molecular Graphics System Version 1.8.0, SWISS-MODEL online Server (2019), UCSF-Chimera 1.14, PHENIX 1.17-3660 suite, Coot 0.8.9.2. Microfluidic data were analyzed using a routine programmed in Sclab. Cryo-EM data were processed using the cryoSPARC software package v2.08-r.2.12 (constantly updated) including the Patchmotion algorithm and Refinement_New algorithm, the Focus software package v1.1.0 (revision 6c7e4683), MotionCor2 v 1.1.0, CTF v 1.06, Gautomatch v 0.56, CTFIND v4.1.9and RELION 3.1. CD data were analyzed using the Jasco Spectra Manager 2.06 software. FBA was done using COBRApy 0.15.4. The full code and the list of reactions of each model can be found at https://github.com/he-ha/PubData within the "2020_Taco" directory. For pathway thermodynamic analysis, Python 3.7.3 and the Python packages equilibrator_api 0.3.2, equilibrator_assets 0.1.untaggered.59.g966e15, equilibrator_pathway 0.3.2 + 5.1622c00c and equilibrator_cache 0.3.2 were used. The scripts and details can be found at https://github.com/he-ha/PubData within the "2020_Taco" directory.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy Information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Structural data that support the findings of this study have been deposited in the Protein Data Bank (PDB) with the accession codes 6YBP (MePCC WT) and 6YBQ (GCC MS) and in the Electron Microscopy Data Bank (EMDB) with the accession codes EMD-10770 (MePCC WT) and EMD-10771 (GCC MS). Other publicly available PDB structures used in this study are 2P2B, 3N6R and 1ON3.

Field-specific reporting
Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design
All studies must disclose on these points even when the disclosure is negative.

| Sample size | No statistical methods were used to predetermine sample size. At least two or three independent experiments were performed. Sample sizes were determined on the basis of previous experiments by considering biological/technical variability. |
| Data exclusions | No data were excluded from analyses. |
| Replication | All experiments were performed at least two or three times. All attempts of replication were successful. |
| Randomization | Randomization was not applicable for the experiments performed, because no statistical comparisons were performed. |
| Blinding | Blinding was not relevant, because no statistical comparisons were performed. |

Reporting for specific materials, systems and methods
We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

| Materials & experimental systems | Methods |
|---------------------------------|---------|
| n/a | n/a | Involved in the study |
| x | | Antibodies |
| x | | Eukaryotic cell lines |
| x | | Palaeontology and archaeology |
| x | | Animals and other organisms |
| x | | Human research participants |
| x | | Clinical data |
| x | | Dual use research of concern |
| x | | CHIP-seq |
| x | | Flow cytometry |
| x | | MRI-based neuroimaging |

1/23/20