DESIGN AND EVOLUTION FOR NEW METABOLISM WITH SPEED AND SCALE

JAMES C. LIAO
Academia Sinica
128 Academia Road, Section 2, Taipei 115, Taiwan

Biology plays a major role in the carbon cycle on earth. It was estimated that terrestrial photosynthesis draws down 120 giga tons of carbon per year, half of which is released back to the atmosphere by respiration. The remaining carbon is slowly converted to CO_2 and CH_4 due to microbial metabolism, with about 10% stored in soil. The carbon cycle reached an equilibrium to establish a relatively stable level of atmospheric CO_2, until the industrial revolution. To date, burning of fossil fuel emits more than 10 giga tons of carbon annually, which upsets the natural balance of the carbon cycle and caused a rapid increase of atmospheric greenhouse gas (GHG) in the past 100 years or so. The rise in GHG in the air has caused global warming and more frequent occurrence of extreme weather events.

The 2018 IPCC report "Global Warming of 1.5^{o}C" recommends that the CO_2 emission must be reduced by 50% by 2030, and reaches net zero by 2050 in order to limit the global warming to 1.5 ^{o}C. In addition, methane emission also needs to be reduced by 50% by 2030. To meet this target, the current fossil fuel-based economy must be changed to renewable-based economy to avoid carbon emission. Thus, conversion of C1 compounds (CO_2, methanol, methane) to useful chemicals is essential to achieve "net zero" by 2050. In this regard, biology becomes an important player because of the specificity and diversity of biological chemistry. However, the biological chemistry nature evolves does not address the problems that we face today which must be solved with speed and scale. Thus, new metabolism is needed to meet our needs.

Redesign and re-evolution of metabolism

The initial stage of metabolic redesign is focusing on demonstrating the plasticity of metabolic systems by converting a carbon substrate to a specific chemical in an organism that does not natively do so. At this stage, the common strategy is the identification of enzymes from various organisms and express them properly in a desirable organism that is amenable to genetic manipulation or possesses specific phenotype of interest. An example is the isobutanol production using *E. coli* [1], which is a non-native yet toxic product for the organism. The redesign of metabolic pathway may be straightforward, but the rewiring of regulatory systems for optimal production often presents a major challenge.

Another example is the production of 1-butanol in *E. coli* [2]. The pathway requires only a handful of heterologous genes, which should be relatively easy to transfer from various organisms to *E. coli* to fit in the existing biotechnology and industrial platform. However, despite many trials by multiple groups, *E. coli* was unable to match the level of production demonstrated by the native butanol producer – *Clostridium*. We then developed an evolution strategy to couple production with growth and showed that by manipulating the driving force of metabolism, *E. coli* can produce to a level matching the *Clotridium* performance. This example demonstrates that rational design coupled with in vivo evolution is a powerful way to develop biological chemistry with speed and efficiency. Up to date, numerous chemicals of various complexity have been engineered

to multiple platform organisms, demonstrating the plasticity of downstream metabolism after substrate assimilation. However, few are focusing on the speed and scale required to combat the climate problem.

To do so, the following problems must be addressed: 1) the redesign of the metabolic infrastructure, including CO_2 fixation, photorespiration, and sugar metabolism, 2) the strategy for re-evolution in vivo to increase the speed and scale (efficiency).

Preventing carbon loss in photosynthesis

Rubisco is responsible for the majority of CO_2 fixation on earth. Rubisco itself is not sensitive to oxygen. However, Rubisco cannot distinguish CO_2 from O_2, causing oxygenase activity as a side reaction. The result of this activity is the oxygenation of the Rubisco substrate ribulose 1,5-bisphosphate to 3-phosphoglycerate and 2-phosphoglycolate. The latter is then oxidized to CO_2 in peroxisome and mitochondria, resulting in the loss of carbon. This phenomenon is called "photorespiration", and is particularly severe for C3 plants which accounts for the majority of the plants on earth. It is thought that Rubisco in nature has been evolved to a limit that is balanced between specificity and activity.

To alleviate this problem, several authors have engineered plants to recycle 2-phosphoglycolate to productive pathways. The first approach [3] is to construct an oxidation pathway to break down 2-phosphoglycolate to two CO_2 in chloroplasts, instead of mitochondria. This would allow Rubisco to re-assimilate CO_2 and partially reduce the carbon loss. The second approach [4] is to convert two molecules of 2-phosphoglycolate to glycerate with a loss of CO_2. The former can then be assimilated in the CBB cycle. The third approach [5] is to convert 2-phosphoglycolate to a common biosynthetic precursor, acetyl-CoA without carbon loss. This would allow the plant to synthesize compounds such as fatty acids that are useful as fuel and chemical. These approaches utilize enzymes in different organisms and express them in the chloroplasts of plant cells. The last approach [6] evolves a new-to-nature enzyme, glycolyl-CoA carboxylase (GCC) to fix another CO_2 to glycolate and eventually convert it to glycerate for re-assimilation in the CBB cycle.

These strategies highlight the different approaches for establishing "new metabolism" in cells. Taking advantage of enzymes nature already evolve would accelerate the progress, but creating new enzymes may also have advantages in the future. In either case, the construction of new metabolism by rational design must be followed by in vivo evolution to adapt to the cellular environment.

Rubisco-independent CO_2 fixation

Another approach to solve the carbon loss problem in photorespiration is to avoid using Rubisco in carbon fixation. Nature has evolved seven pathways for CO_2 fixation, of which only the CBB pathway uses Rubisco. Since it is oxygen-tolerant, it is widely used by phototrophic organisms that use light to split water. The other CO_2 fixation pathways are all oxygen-sensitive, and involve enzymes that depend heavily on oxygen-sensitive iron-sulfur clusters. However, there are some oxygen-insensitive carboxylases that are used for carboxylation reactions in the cell, and can be repurposed to form CO_2 fixation pathways. Many of the theoretical CO_2 fixation pathways have been designed based on kinetics and thermodynamics considerations [7]. However, most of them involve oxygen sensitive enzymes and are difficult to implement. So far, two CO_2 fixation pathways

have been constructed using cell-free systems to demonstrate the feasibility of the pathway in vitro. The first one (CETCH cycle) [8] uses only an enoyl-CoA carboxylases/reductases (ECR) to fix a carbon on crotonyl-CoA (C4) and acrylyl-CoA (C3), respectively, and form a synthetic cycle. The cycle produces glyoxylate, a C2 compound. The second synthetic cycle (rPS-MCG)[9] (Fig. 1A) uses a PEP carboxylase (PPC) and an ECR to fix two CO_2, and produces acetyl-CoA, which is also an intermediate in the cycle. The cycle can also output any C2, C3, and C4 intermediate, forming a self-replenishing cycle. The rPS-MCG cycle also features a possibility to recover some energy from NADPH in the form of FADH2, analogous to the TCA cycle. These cycles are demonstrated in "one-pot" in vitro systems to prove the chemical principle and demonstrate the efficiency. The in vitro systems tested protein-protein interactions, stabilities, and cofactor recycling controls, which must be carefully balanced, even in in vitro systems. These results demonstrate the potential challenges in implementing a fundamental pathway in cells and highlight the importance of in vivo evolution in implementing new metabolism, which we will address shortly.

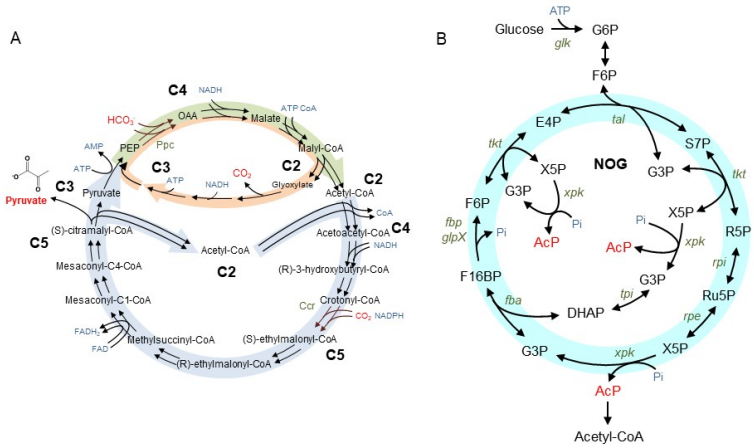

Fig. 1. A) The synthetic rPS-MCG cycle[9] for CO2 fixation without Rubisco. B) The non-oxidative glycolysis (NOG) pathway[10].

The carbon loss in glycolysis

Once CO_2 is assimilated, it is converted to sugar and enters glycolytic *Embden–Meyerhof–Parnas (EMP) pathway,* the most conserved pathway present in almost all organisms. This pathway partially oxidizes glucose to make the 2-carbon unit acetyl-CoA for the synthesis of fatty acids, amino acids, and most of the bioproducts of interest to industry. However, during this process, one third of the carbon is lost as CO_2, resulting in a great reduction in carbon yield. This was the reason for the relatively low carbon yield of all bioproducts and the significantly reduced competitiveness relative to petroleum-derived products, and the problem is independent of the carbon substrates, as long as the metabolism uses the EMP pathway to produce acetyl-CoA. To address this problem, we designed a synthetic non-oxidative glycolysis (NOG) pathway [10] (Fig. 1B) that avoids losing carbon during the breakdown of sugars. The proof of concept was established by demonstrating NOG's function in vitro and in vivo. We further developed an *E. coli* strain that uses NOG as the sole pathway for sugar catabolism [11] by rational

design followed by step-wise in vivo evolution. The in vivo evolution process turns out to be time consuming and suffers from escape strains that utilize alternative pathways for growth. This result highlights the need for a more systematic in vivo evolution process, and a strategy to align the "design goal" with the organism's goal for efficient growth. Nevertheless, the development of a strain that uses NOG to replace EMP demonstrate that even a fundamental pathway can be altered by a combination of design and evolution strategies.

The methanol utilization problem

Similar efforts in redesigning and re-evolving metabolism for methanol utilization in microorganisms have also met with comparable problems. For example, a methanol condensation cycle (MCC) [12] to ethanol or butanol without carbon loss has been designed and shown in vitro. An *E. coli* that can utilize methanol as the sole carbon source for growth has been designed and evolved [13]. In short, while designing novel pathways has seen repeated success, implementing in microorganisms to a level that can perform new metabolism with speed and at scale is still challenging.

Challenges and opportunities in in vivo evolution

Previous examples have shown that purely relying on rational design could not accomplish the goal of "establishing new metabolism for speed and scale". A few strategies have been developed to move toward this goal. The most common one is to artificially create a library of variants in promoter, ribosome-binding site, gene-knockout, or protein sequence, followed by screening. The process has also been aided by machine-learning and automation. However, the constraint on the library size significantly limits the success. Hence, additional tools are required to establish new metabolism with speed and scale that may impact the climate solution requires additional tools.

In conjunction with rational design, we explored natural evolution and searched for properties that can be exploited in creating the *E. coli* strain that can grow with methanol as a sole carbon source. *E. coli* has all but three enzymes that are required to grow with methanol as a sole carbon source [13]. However, despite years of efforts by many groups world-wide, the task has proved to be more difficult than initially expected. We first used a metabolic model to identify and then removed two potential kinetic traps that may cause a metabolic dead-end. The remaining steps were relying on in vivo laboratory evolution. The final result showed that the cell used a combination of single nucleotide polymorphism (SNP) and copy number variation (CNV) to create diversity, and achieved a delicate balance between detoxification and assimilation. Through evolution, the metabolic pathway for methanol assimilation was established and optimized by balancing the production and utilization of the toxic intermediate, formaldehyde. Through this example, we discovered that *E. coli* can dynamically tune the copy number of a specific fragment of DNA on the chromosome. The organism utilizes two mechanisms to adjust the copy number of a DNA fragment. One depends on IS-elements and the other does not. We have identified the essential elements that are required for the dynamic CNV. Through these mechanisms, the cell can amplify the expression of a large number of genes. Upon selection, strains with appropriately amplified gene sets are enriched. By effectively combining SNPs with CNV, the in vivo evolution can be exploited.

In another example, we knocked out an essential gene, *panD*, coding for aspartate 1-decarboxylase which is essential for β-alanine synthesis in *E. coli* for growth in glucose

minimal medium [14]. We performed an in vivo evolution experiment and select cells that can grow without any nutritional supplement. The organism was able to "repair" this metabolic block by rerouting metabolism in different ways. The first pathway is to repurpose the uracil degradation pathway, coupled with necessary auxiliary functions for detoxification. The second is to mutate ornithine decarboxylase (SpeC) to gain a simultaneous decarboxylation and oxidative deamination function. This novel activity coupled with other transamination reaction and dehydrogenase reaction also "repaired" the *panD* mutation. This example demonstrated that in vivo evolution may have the ability to evolve novel activities that is not present in nature.

Outlook to future developments

Currently, using biological approaches to solve climate problems depend on a few fundamental metabolic pathways that were not evolved for such a purpose. To make a practical impact, speed and scale of biological reactions must be improved. Both rational design and evolution approaches are needed in this endeavor. In particular, in vivo evolution that is responsible for establishing natural biological reactions must be further explored. While the basic principles for DNA replication, repair, and recombination are reasonably established, new mechanistic details are required to explore the full capacity of in vivo evolution. Together with rational designs, these approaches are expected to play a major role in building new metabolism that can achieve the speed and scale required to convert single carbon green house gases to chemicals to alleviate the dependence on fossil sources.

References

1. S. Atsumi; T. Hanai and J.C. Liao. *Nature* **451**, 86 (2008)
2. C.R. Shen; E.I. Lan; Y. Dekishima; A. Baez; K.M. Cho *et al*., *Appl Environ Microbiol* **77**, 2905 (2011)
3. P.F. South, A.P. Cavanagh, J.W. Liu, and D.R. Ort, *Science* **363,** 6422 (2019)
4. R. Kebeish, M. Niessen, K. Thiruveedhi, R. Bari, H.-J.Hirsch *et al., Nat Biotechnol* **5,** 593 (2007)
5. H. Yu, X. Li, F. Duchoud, D.S. Chuang, and J.C. Liao *Nat Commun* **9**, 2008 (2018)
6. M. Scheffen, D.G. Marchal, T. Beneyton, S. K. Schuller, M. Klose *et al. Nat Catal* **4**, 105 (2021)
7. A. Bar-Even, E. Noor, N.E. Lewis, and R. Milo *Proc Natl Acad Sci USA* **107**, 8889 (2010)
8. T. Schwander, L.S. von Borzyskowski, S. Burgener, S., N.S. Cortina, & T.J. Erb *Science* **354**, 900 (2016)
9. S. Luo, P.P. Lin, L.Y. Nieh, G.B. Liao, B.W. Tang *et al., Nat Catal* **5**, 154 (2022)
10. I.W. Bogorad, T.S. Lin, and J.C. Liao *Nature* **502**, 693 (2013)
11. P.P. Lin, A.J. Jaeger, T.Y. Wu, S.C. Xu, A.S. Lee *et al., Proc Natl Acad Sci USA* **115**, 3538 (2018)
12. I.W. Bogorad, C.T. Chen, M.K. Theisen, T.Y. Wu, A.R. Schlenz *et al*., *Proc Natl Acad Sci USA*. **111**,15928 (2014)
13. F.Y-H. Chen, H.-W. Jung, C.Y. Tsuei, and J.C. Liao *Cell* **182**, 933 (2020)
14. S. Pontrelli, R. Fricke, S.T. Teoh, S.T. W.A. Laviña, S.P. Putri *et al., Nat Chem Biol* **14**,1005 (2018)