

REVIEW

Open Access



Thawing out frozen metabolic accidents

Dario Leister

Abstract

Photosynthesis and nitrogen fixation became evolutionarily immutable as “frozen metabolic accidents” because multiple interactions between the proteins and protein complexes involved led to their co-evolution in modules. This has impeded their adaptation to an oxidizing atmosphere, and reconfiguration now requires modification or replacement of whole modules, using either natural modules from exotic species or non-natural proteins with similar interaction potential. Ultimately, the relevant complexes might be reconstructed (almost) from scratch, starting either from appropriate precursor processes or by designing alternative pathways. These approaches will require advances in synthetic biology, laboratory evolution, and a better understanding of module functions.

Frozen accidents: evolution of the genetic code

A key challenge in genetic engineering and synthetic biology is to *change the unchangeable*—i.e., to thaw the so-called frozen metabolic accidents (FMAs). The term FMA refers to processes that are thought to be immutable, because their modification requires altering multiple intertwined components at the same time. Historically, the term “frozen accident” was coined to explain certain characteristics of the standard genetic code (SGC), which is outlined in the following. In fact, the assignment of the 20 canonical amino acids to the 64 codons of the SGC is clearly non-random, i.e., related amino acids typically occupy contiguous areas in the codon table [1]. Several possible reasons for this have been proposed. Chemical interactions between amino acids and the tertiary structures of RNA-binding sites of codons (or anti-codons; “stereochemical theory” [2], co-evolution of amino-acid biosynthesis and code structure (“coevolution theory”) [3], and selection for robustness (“error minimization theory”) [4] could all have contributed to the evolution of the SGC.

Furthermore, the “frozen accident” perspective introduced by Francis Crick 50 years ago [5] explains the universality of the SGC by its effective fixation in the earliest life forms, such that any major change would be strongly selected against because it would immediately affect large numbers of proteins. This scenario does not require that the original assignment of codons occurred entirely by chance. In fact, the SGC displays clear signs of optimization: it is very robust, albeit not the most robust possible [6, 7], and the 20 canonical amino acids are thought to be virtually ideal for building soluble protein structures with close-packed cores [8]. But the frozen accident perspective also emphasizes that once the SGC assignment had been made, it became essentially immutable.

Among the deleterious effects of codon reassignment, it is thought that inhibition of horizontal gene transfer (HGT) might have played a major role [1]. In fact, HGT remains a key factor in microbial evolution [9] and, assuming that the capacity to undergo HGT was essential for the early prokaryotes [10], even small alterations of the code would have had detrimental effects on HGT, genetically isolating the species affected and dooming them to extinction. HGT itself also explains the universality of the code. Advanced simulations of evolution starting with several genetic codes and extensive HGT lead to a single universal code [11–13] (Fig. 1a).

The “frozen metabolic accident” concept

The term FMA was introduced by Paul Falkowski and coworkers to explain the evolutionary dynamics of photosynthetic genes [14]. Maximal co-evolution among photosynthetic genes occurs when their protein products physically interact with each other, and in prokaryotes, such genes are often clustered at the genomic level [14]. Thus, co-evolved photosynthetic proteins are found in thylakoid multiprotein complexes (photosystems I (PSI) and II (PSII), cytochrome *b₆f* complex, and ATPase) and in soluble enzyme complexes like the tetrapyrrole biosynthesis enzyme magnesium protoporphyrin IX chelatase and the Calvin cycle enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) [14] (Fig. 1b). The rate of amino acid substitution for photosynthetic proteins in

Correspondence: leister@lmu.de

Faculty of Biology, Ludwig-Maximilians-University Munich, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany



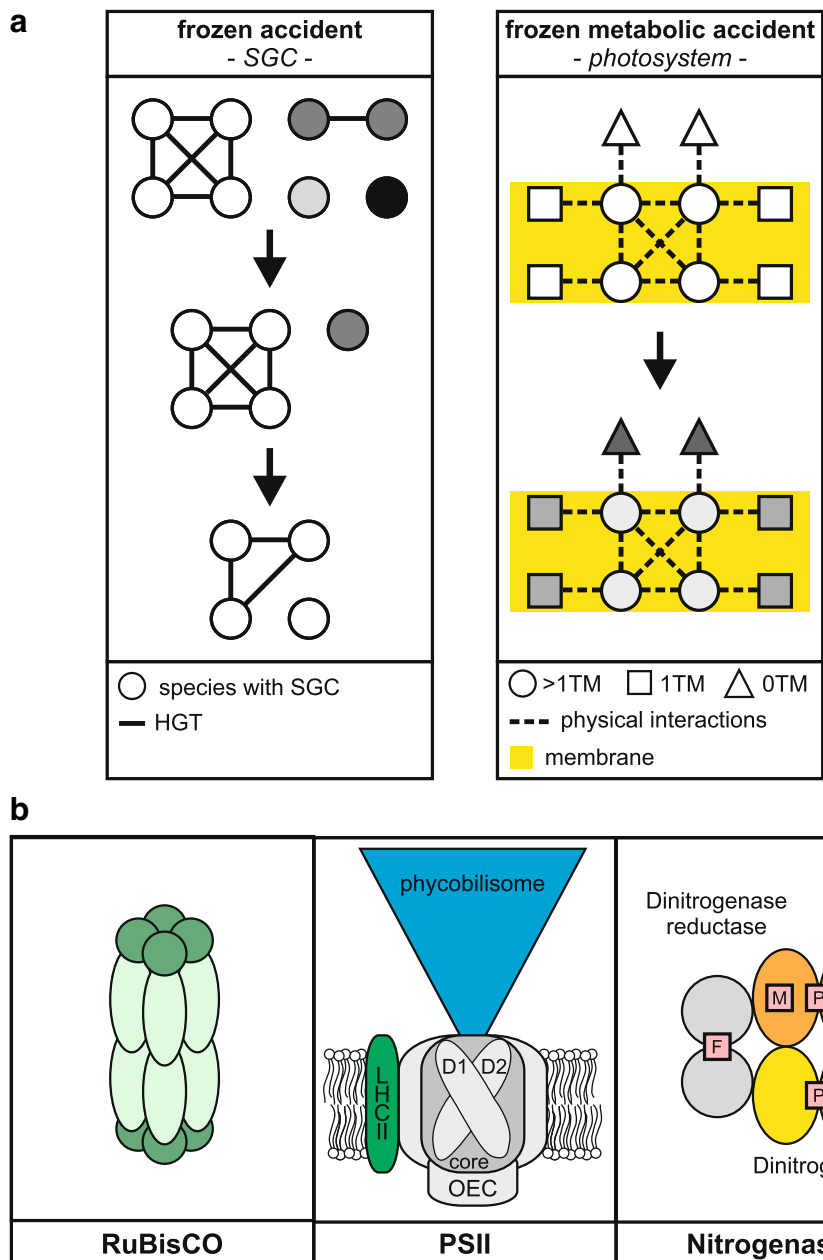


Fig. 1. Evolution and instances of “frozen (metabolic) accidents.” **a** The standard genetic code (SGC) is a frozen genetic accident and became universal during early prokaryote evolution because of the eminent role of horizontal gene transfer (HGT) (*left side*). Each *circle* represents a species and different genetic codes are indicated by different fillings. Species with rare codes became extinct because they were less amenable to stable HGT. The code that remained became the SGC. Once it was universally adopted, it persisted even when HGT became less important. On the *right*, the evolution of photosystem subunits is schematically shown as an example of a FMA. In the thylakoid membrane (*yellow*), the core photosystem subunits (*circles*), having more than one transmembrane domain (TM), multiply interact with each other (*dotted lines*) and with their co-factors. They also interact to a lesser extent with other transmembrane photosystem proteins (with one TM, *rectangles*), as well as with peripheral photosystem proteins and soluble proteins (*triangles*). This leads to different rates and modes of evolution and sequence diversification (highlighted in *gray*, with increased variation being indicated by darker shading). The sequences of photosystem core subunits changed least during evolution and represent FMAs. **b** RuBisCO, PSII, and the molybdenum (Mo)-dependent cyanobacterial nitrogenase. RuBisCO consists of eight large (*light green*) and eight small (*dark green*) subunits and catalyzes carboxylation of ribulose-1,5-bisphosphate during the Calvin cycle. PSII contains a core with several subunits, of which D1 and D2 are shown. Other PSII subunits, including the oxygen-evolving complex (OEC), surround the core. In cyanobacteria, phycobilisomes act as antenna complexes, whereas light-harvesting complex II (LHCII) serves this function in eukaryotes. The Mo-nitrogenase is a two-component system comprising a heterotetrameric dinitrogenase complex with a P-cluster (P; Fe₈S₇) and an Fe-Mo cofactor (M), and a homodimeric dinitrogenase reductase with a Fe₄S₄ cluster (F)

the cores of the photosynthetic multiprotein complexes (with multiple transmembrane domains (TMs)) are indeed markedly lower than those of the small subunits surrounding the cores (with 1 TM), or peripheral or soluble photosynthetic proteins that lack TMs (Fig. 1a). This implies that interactions of proteins with other proteins, lipids, or cofactors constrain their evolution in the core photosynthetic apparatus of cyanobacteria and photosynthetic eukaryotes [14]. Interestingly, proteins involved in different multiprotein complexes also co-evolved, e.g., core proteins of PSI and PSII, as well as cytochrome *b₆f* complex and NADH dehydrogenase proteins, suggesting that their functional linkage in thylakoid electron transport was responsible for their co-adaptation. Moreover, the stoichiometry of the output of the light reactions in terms of ATP and NADPH almost perfectly matches the needs of the Calvin cycle, indicating substantial co-evolution of entire functionally linked processes.

When environmental conditions changed several hundred million years ago, the functions of some of these conserved core components of photosynthesis were compromised. This holds in particular for their performance in the presence of oxygen and high light intensities—conditions that directly (oxygen) or indirectly (the ozone layer that enabled colonization of terrestrial habitats by blocking UV radiation) resulted from oxygenic photosynthesis. The two most prominently affected proteins are the D1 protein of PSII and the RuBisCO enzyme (Fig. 1b). Reactive oxygen species (ROS) directly damage D1 under high light intensities and inhibit its continuous replacement by newly synthesized copies [15–19], while RuBisCO's propensity to employ oxygen instead of CO₂ as substrate essentially wastes light energy [20]. Falkowski and coworkers estimated that these features of PSII and RuBisCO together reduce the potential overall efficiency of photosynthesis by at least 50% [14].

Like D1 and RuBisCO, nitrogenase (Fig. 1b)—the enzyme complex that fixes atmospheric nitrogen in prokaryotes—evolved before oxygen was freely available in the atmosphere, and the core proteins of cyanobacterial nitrogenases remained virtually unchanged following the transition to an oxidizing atmosphere [21, 22]. Hence, nitrogenase represents a third instance of a FMA [23]. Because of the oxygen sensitivity of the iron-sulfur clusters in these nitrogenases (Fig. 1b), an estimated 20 to 30% of marine nitrogenase activity is inhibited by O₂. It is tempting to speculate that this evolutionary inflexibility might also explain why cyanobacterial nitrogenases were not endosymbiotically acquired by eukaryotes.

Redressing the effects of FMAs

The functional shortcomings of RuBisCO, nitrogenase, and the D1 subunit of PSII have stimulated attempts to enhance their efficiency, aiming to improve crop yield

and biomass production. These efforts have encountered many obstacles, which revealed additional facets of the consequences of FMAs. Exploiting interspecific diversity in RuBisCO activity, improving the activity of its auxiliary protein RuBisCO activase, and altering the levels of regulatory metabolites have been identified as promising ways to enhance the enzyme's function in crop plants [20]. While these approaches are still at an early stage and are complicated by the fact that in crop plants the small and large subunits of the enzyme are encoded in the nucleus and chloroplast, respectively, they have already uncovered additional constraints on the evolution of RuBisCO. For instance, the enzyme's subunits (Fig. 1b) co-evolve not only with each other [14], but also with its assembly factors [24]. This highlights why the modification of FMAs will often require the exchange of entire modules of co-evolved components, including structural proteins, auxiliary factors, and the genetic elements necessary for their efficient expression [25]. The number of auxiliary proteins implicated can even exceed the number of structural proteins present in the mature complex, in the case of RuBisCO prompting the coining of the term “rubiscosome,” describing its evolution from a stand-alone enzyme into an enzyme complex that involves various auxiliary factors [26, 27]. Indeed, it has only recently become possible to express a functional plant RuBisCO in *Escherichia coli* by co-expressing the large and small subunits of RuBisCO together with its five assembly factors [28]. This heterologous plant RuBisCO expression system promises to provide a way to test variants of the enzyme for enhanced function in the genetic workhorse *E. coli*. In fact, a non-native Calvin cycle or parts of it have previously been functionally reconstituted in *E. coli* [29, 30], yeast [31], or *Rhodobacter capsulatus* [32], allowing its optimization by directed evolution [33] or laboratory evolution [30, 34, 35]. Because several natural and synthetic autotrophic pathways besides the Calvin cycle exist, it is possible to design alternative carbon fixation cycles in plants, and such efforts are now underway [36–41].

With regard to nitrogenases, three strategies are available for enhancing biological nitrogen fixation in crops: (i) boosting the process in naturally plant-associated bacteria, (ii) inducing formation of the root nodules that permit symbiosis between crop plants and N₂-fixing bacteria, and (iii) directly transferring prokaryotic nitrogenase genes into plant genomes [42, 43]. An important initial step was taken more than 40 years ago, when recombinant *E. coli* strains with nitrogenase activity were constructed by genetic engineering [44]. However, these transgenic strains exhibited much lower nitrogenase activity than the original host and could not support diazotrophic growth on nitrogen-free medium. *E. coli*-based systems were also

used more recently to combine the nitrogenase with electron transport components from plant organelles as power supplies for future engineering of diazotrophy in cereal crops [45]. However, overcoming the FMA of nitrogenases by mitigating their oxygen sensitivity has not yet been accomplished.

While recombinant *E. coli* strains are available for RuBisCO and nitrogenase that provide a platform for modifying these suboptimal enzymes by genetic engineering and laboratory evolution, the situation for the PSII core proteins, including the D1 protein, is less favorable. Reconstitution of the photosynthetic light reactions in *E. coli* has proved to be a challenging task and has not yet been achieved. In fact, owing to their slow tempo of evolution (as expected for products of a FMA), the PSII core proteins are largely conserved between cyanobacteria and plants. Nevertheless, attempts to replace the cyanobacterial PSII core proteins D1, CP43, CP47, and PsbH with their counterparts from flowering plants (with which they share between 78 and 86% identity at the amino acid level) have been only partially successful [25] (Fig. 2). The resulting strains were either non-photoautotrophic or severely impaired in photosynthesis [46–49], demonstrating that the exchange of one component of this module of co-adapted PSII core proteins can markedly perturb the multiple interactions and the functionality of the entire PSII core.

In light of this limitation, Stephen Mayfield and co-workers proposed to replace the entire PSII core with its counterpart from another species, thereby maintaining the multiple intrinsic interactions within the PSII core that have evolved over millions of years in each photosynthetic species [50] (Fig. 2). To test this hypothesis, the six original core PSII genes (*psbA*, *B*, *C*, *D*, *E*, and *F*) from the chloroplast genome of the green alga *Chlamydomonas reinhardtii* were deleted and replaced by a single synthetic

construct that contained the orthologous genes from *Volvox carteri* or *Scenedesmus obliquus* (two other green algal species) or, as a control, from *C. reinhardtii* [50]. In addition, the effect of replacing only subsets of the six genes was investigated. These experiments showed that (i) the strains reconstituted with the *C. reinhardtii* PSII gene sets showed the best photosynthetic performance, albeit lower than in the starting WT strain, and (ii) in the strains with replacements from *V. carteri* and *S. obliquus*, photosynthetic performance declined with increasing numbers of exchanged genes. These results imply that both the organization of the substitute genes in a synthetic construct (that might lack some *cis*-acting elements and their original operon structures, leading to suboptimal gene expression) and off-target effects of the PSII gene deletions (that might affect other operons and adjacent tRNA genes) in the *C. reinhardtii* chloroplast could decrease the photosynthetic performance of the transgenic strains. Moreover, the experiment could not clarify whether the roadblocks presented by FMAs can be truly removed by exchanging the entire ensemble (or module) of interacting and co-evolved proteins. In fact, the substituted green algal proteins were so similar with respect to their sequence—with average identities of between 93% (*S. obliquus*) and 98% (*V. carteri*) relative to the deleted original *C. reinhardtii* genes—that the strong negative effect on PSII function resulting from perturbations in gene expression in the host system used might have masked any subtle positive effect produced by the retention of the intrinsic interactions between the six co-adapted proteins. In consequence, a more appropriate experimental approach would be to replace entire cyanobacterial photosystem cores by their functional equivalents from higher plants [25], for which single subunit exchanges have proven to be clearly detrimental (see above). In such experiments, the exchange of multiple subunits

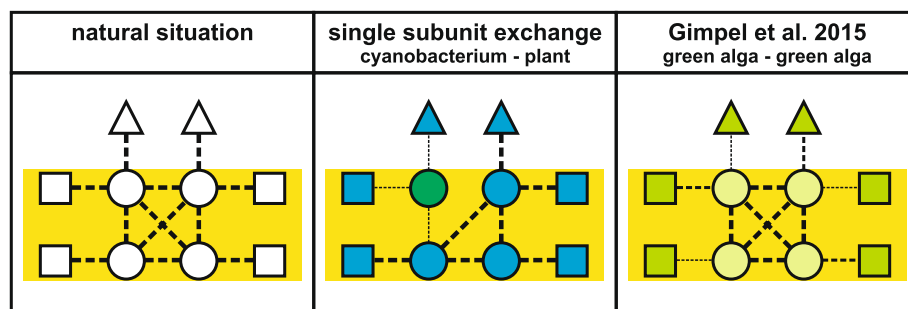


Fig. 2. Current approaches to modifying the PSII core. *Left panel:* In PSII from cyanobacteria, algae and plants, the core subunits (circles) multiply interact (dotted lines) with each other and to a lesser extent with other transmembrane (rectangles) and peripheral (triangles) proteins. *Middle panel:* If core subunits in cyanobacteria (blue) are replaced by their plant pendants (green), their sequence differences may weaken or disrupt these interactions (thinner dotted lines), severely affecting the function of PSII. *Right panel:* Stephen Mayfield and co-workers [50] replaced an entire green algal PSII core with variants from other green algal species (indicated by a lighter shade of green). This would be expected to maintain the interactions between PSII core subunits, but might have affected interactions with other PSII proteins

might alleviate the strong negative effects of single subunit substitution, given that the experimental setup can achieve efficient expression of the introduced genes and efficient assembly of the corresponding proteins—the latter one might even require introduction of plant-specific assembly factors. Moreover, it might be necessary to replace not only the six core subunits, but also additional PSII proteins that physically and/or functionally interact with the PSII core.

Novel approaches to bypassing roadblocks caused by FMAs

Functional resolution of FMAs by replacing the evolutionarily immutable protein (complexes) with more efficient variants is an enormous challenge. But the issue is not only of interest for evolutionary biologists, but would have significant implications for plant productivity, as it might be the only way to enhance photosynthetic performance and nitrogen fixation rates. With respect to photosynthesis, this endeavor involves the realization of an evolutionary development that never actually took place. One needs to create a primordial version of photosynthesis under aerobic conditions at high light intensities—the very conditions produced by the advent of oxygenic photosynthesis (see above)—and let it evolve. The outcome of this evolution might be to reduce the light sensitivity of the PSII reaction center and the current susceptibility of the carbon fixation cycle to the level of oxygen in the atmosphere.

How can this challenge be tackled? A “conservative” solution to create novel proteins would be to fully exploit the potential of the 20 canonical amino acids encoded by the SGC. Indeed, concepts for systematically designing entire libraries of “non-natural proteins” (based on the canonical

amino acid repertoire) that can be employed *in vivo* to replace natural proteins have been presented, and proof of their practicability has been provided [51–53]. A second possibility is to expand or alter the genetic code. In fact, some organisms have succeeded in co-opting the two non-canonical amino acids selenocysteine and pyrrolysine into the code [1], and nonstandard amino acids with residues that have unusual chemical properties could be employed for the design of novel protein functions hitherto not possible with canonical amino acids. Thus, concepts for recoding genomes by re-assigning or deleting codons have been developed with the aim of enabling multivirus resistance, enhanced incorporation of nonstandard amino acids, or biocontainment by synthetic auxotrophy [54]. Such strains have been designed, generated, and subsequently streamlined by adaptive evolution [55–57].

These two concepts for the design of new proteins to counteract the repercussions of FMAs may be applied either “conservatively,” *i.e.*, exploiting the opportunities offered by such scenarios (including the possibility of using non-natural proteins), or “unconventionally,” *i.e.*, employing entirely new components and interactions. In addition, the recognition that all interacting/co-evolving components in a module may need to be altered must be taken into account. In consequence, three approaches for redressing the aftereffects of FMAs can be envisioned: exploiting (i) natural or (ii) non-natural variants of the components of the module concerned, and (iii) re-designing the corresponding biological process (almost) from scratch (Fig. 3). All three approaches can be expected to require extensive laboratory evolution [58] to streamline the novel processes, *i.e.*, exploiting the high rate of evolution in microbial systems to optimize the expression, interactions, and performance of the introduced proteins for maximal

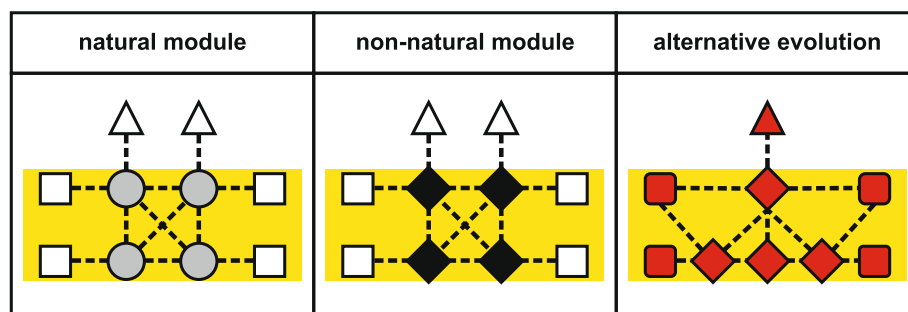


Fig. 3. Overview of the three general strategies for the resolution of FMAs. *Left panel:* The module responsible for the FMA is replaced by a natural variant of the module. This might be effective if the alternative module has evolved beneficial characteristics due to high selective pressure. Possible trade-offs and interactions with proteins extraneous to the transferred module might need to be considered, but could be mitigated by laboratory evolution (see Fig. 4 for more details). *Middle panel:* The module is replaced by non-natural proteins that have the same interaction potentials but different sequences. This will very likely require a step-wise strategy, with successive replacement of individual proteins via a complementation-based approach, combined with streamlining of protein functions by laboratory evolution (see Fig. 5 for more details). *Right panel:* Alternative evolution of a primordial version of the FMA or its entire *de novo* design might be the only solution if the general design of the FMA is fundamentally flawed and cannot be corrected by the first two approaches (see Fig. 6 for more details)

implementation of the desired function. The third approach must follow an “alternative evolutionary trajectory,” either beginning with a primordial progenitor version of the process or re-designing the process from scratch. These three general concepts will now be discussed in the context of the re-engineering of the PSII core in order to bypass a FMA. In this instance, the difficulties arise from the special biophysical properties and pigment-binding sites of its components, as well as their multiple interactions and complex biogenesis [59–62].

Exploiting natural sequence variations to enhance PSII functions

Although FMAs, by definition, are largely refractory to natural evolution, screening of existing biodiversity might nevertheless provide variants that mitigate their negative effects. For PSII, and in particular the D1 protein, the isolation of the non-model green alga *Chlorella ohadii* (from desert soil crusts) and its physiological characterization may lead to a breakthrough in enhancing the light resistance of plant PSII [63–66]. Even under extremely high irradiation intensities, growth of *C. ohadii* remains unimpaired and its photosynthetic O₂ evolution actually increases at high light levels [63]. In contrast to other photosynthetic eukaryotes, very little of its D1 is degraded under such conditions, indicating that (i) PSII in *C. ohadii* might generate fewer ROS than do canonical PSII complexes [63] or (ii) PSII is less accessible to or better protected against ROS. Because green algae and flowering plants employ similar PSII complexes, with light-harvesting proteins as antennas (Fig. 1b), transfer of the entire *C. ohadii* PSII core module into a flowering plant should be more straightforward than that between cyanobacteria and flowering plants (see above), assuming that efficient expression of all components can be achieved, and biogenesis of the complex does not require

(too many) algal-specific auxiliary proteins. However, although the molecular mechanisms of the high light tolerance of *C. ohadii* are still elusive [63], it appears to be unlikely that a modified PSII core is responsible for the enhanced light tolerance, with a combination of special antenna complexes and auxiliary factors providing a more plausible explanation. Moreover, the specialized ecological niche occupied by *C. ohadii* also raises the possibility of an as-yet unknown trade-off connected with the alga’s high light resistance, although enhanced growth of *C. ohadii* is observed even at low light intensities [65], which argues against the explanation that solely special antenna complexes (in terms of a reduced size) are responsible for the high light tolerance of this organism. But even if transfer of the PSII core of *C. ohadii* is not a promising approach to enhance light stress resistance, the question arises whether it is practicable to exchange entire PSII cores between green algae and flowering plants, thereby extending the approach of Gimpel et al. [50] to the next level of complexity. In fact, dysfunctional interactions between the transferred module and extrinsic components might affect the overall function of the process (Fig. 4). The latter can in principle be rectified by laboratory evolution (although this is not yet practicable in flowering plants) or by replacing the extrinsic components. A third possibility would be to streamline the plant PSII complex with a green algal core in a microbial test system that is amenable to genetic manipulation and laboratory evolution.

Exploiting non-natural sequence variations to enhance PSII functions

While the exact number of natural protein sequences is unknown, it clearly represents a minuscule fraction of all possible proteins. Indeed, for a protein containing 100 amino acids, 20¹⁰⁰ possible sequences exist, enough to fill a volume larger than that of Avogadro’s

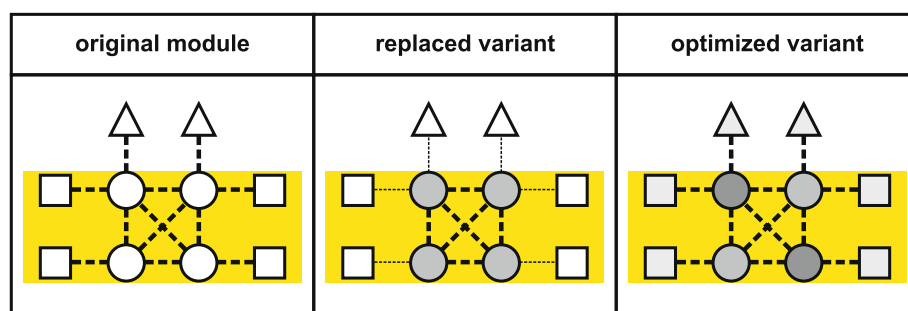


Fig. 4. Replacing modules resulting from FMAs by natural variants. The original variant (left) is replaced by another natural variant with advantageous characteristics (middle). While the intrinsic interactions within the transferred module should already be optimal (thick dotted lines), perturbation of interactions (thin dotted lines) with extrinsic components might affect the overall function of the process. In that case, these extrinsic components must also be exchanged or suitably modified by laboratory evolution (right) to accommodate the chimeric multiprotein complex, perhaps also involving adaptive sequence variation (indicated by altered shading) of contact sites in the complex itself. Note that this approach is not likely to succeed for PSII (see main text) but might be feasible for other instances of FMAs

number of universes [67]. Certainly, only a small fraction of all possible proteins is compatible with biological systems, because many sequences would simply result in intrinsically disordered proteins. Of the compatible fraction, only a very tiny part has been exposed to evolutionary pressures on Earth, leaving a vast number of non-natural proteins capable of replacing existing natural proteins or mediating entirely novel functions. This concept is not only plausible; it has already been shown to be practicable [51]. Libraries of non-natural proteins with the potential to sustain the growth of living cells can be constructed in a systematic and knowledge-based manner, and non-natural proteins that perform specific functions have been identified by phenotypically complementing mutations in natural proteins [51]. This approach has resulted in the identification of non-natural proteins that rescued deletion mutants lacking natural proteins with various specific activities, including phosphoserine phosphatase, citrate synthase, threonine deaminase, and enterobactin esterase [68]. Moreover, established engineering parameters were used to generate simple non-natural four- α -helix bundles and build basic oxidoreductase activities into these scaffolds to create completely artificial redox proteins (“maquettes”) that can plug into natural biochemical pathways and are functionalized in vivo [52, 53, 69].

The central question in the context of PSII is whether a protein with the same tertiary structure and the same complex interaction pattern and cofactor-binding sites as a natural PSII core protein might be capable of substituting for that PSII protein even though their primary sequences might be entirely different. Indeed, analogously to the first test cases for non-natural proteins in *E. coli* described above, such an approach might employ libraries of non-natural proteins expressly designed to have features similar to those of the target to be replaced, e.g., the same number of transmembrane domains and the same cofactor-binding sites. Rescuing a mutant that lacks a specific natural protein with the aid of a suitable non-natural protein could then be accomplished by a complementation-based strategy, provided that the recipient can also be propagated heterotrophically (as in cyanobacterial and algal model systems). Alternatively, the heterotrophic step can be avoided if the natural protein is replaced by the non-natural one via homologous recombination, which is feasible in cyanobacteria and the chloroplasts of some eukaryotes. The resulting lines might then contain a non-natural protein that has replaced a natural co-adapted protein whose evolution had been constrained by a FMA (Fig. 5). The next step would be to optimize the introduced non-natural proteins by laboratory evolution, either by optimizing

each protein before the next one is introduced or by introducing and optimizing them as a batch. In the former case, one might create a system comprised of components that display interactions and features similar to those of the original, but differ markedly from them in their primary sequences, whereas in the latter case interactions and features might differ substantially from the original. This concept is intriguing, but will be very difficult to implement. First, it requires detailed knowledge of the structure–function relationships of the proteins concerned. Even for the best-studied protein complexes—like the photosystems, RuBisCO, or ribosomes—our current understanding might not be sufficient. Secondly, a non-natural protein or maquette with similar structure and interactions to its natural “predecessor” but completely different in sequence might be refractory to the natural assembly process, making novel assembly factors necessary. Thirdly, as laboratory evolution is feasible only in microbial systems, if specific aspects of plant photosynthesis need to be re-designed, the corresponding components must first be transferred into a suitable microbial host. Furthermore, if a new system consisting of non-natural proteins displays the same intrinsic interactions, it might also represent a soon-to-be FMA. Finally, replacing the original proteins with non-natural proteins or maquettes with similar functions will not enhance the process if the basic structure of the process is inherently suboptimal, for instance if any heterodimerically organized reaction center of PSII inevitably leads to D1-based photoinhibition [70]. In the latter case, its re-design with non-natural proteins with similar characteristics will also be in vain, and the only solution is to redesign the process from scratch as outlined below.

Redesign from (almost) scratch to enhance PSII functions

How can the deleterious consequences of a FMA be redressed if the process itself is already intrinsically flawed? Or, with respect to PSII, what if all designs of water-splitting PSII found in extant organisms share the same inevitable flaw of photosensitivity and D1 damage? In fact, it can be argued that the energetics of the reaction mediated by PSII necessitates PSII being susceptible to ROS generation. In this scenario, re-designing PSII such that it circumvents ROS production is impossible, although it might well be engineered so that less ROS are produced, ROS are better scavenged, or that central parts of the complex are less accessible to ROS. Another thought is that a certain level of photosensitivity of PSII has been positively selected for during evolution because this serves

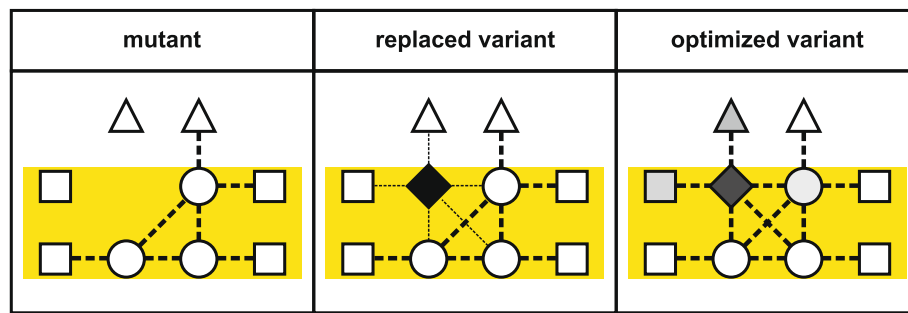


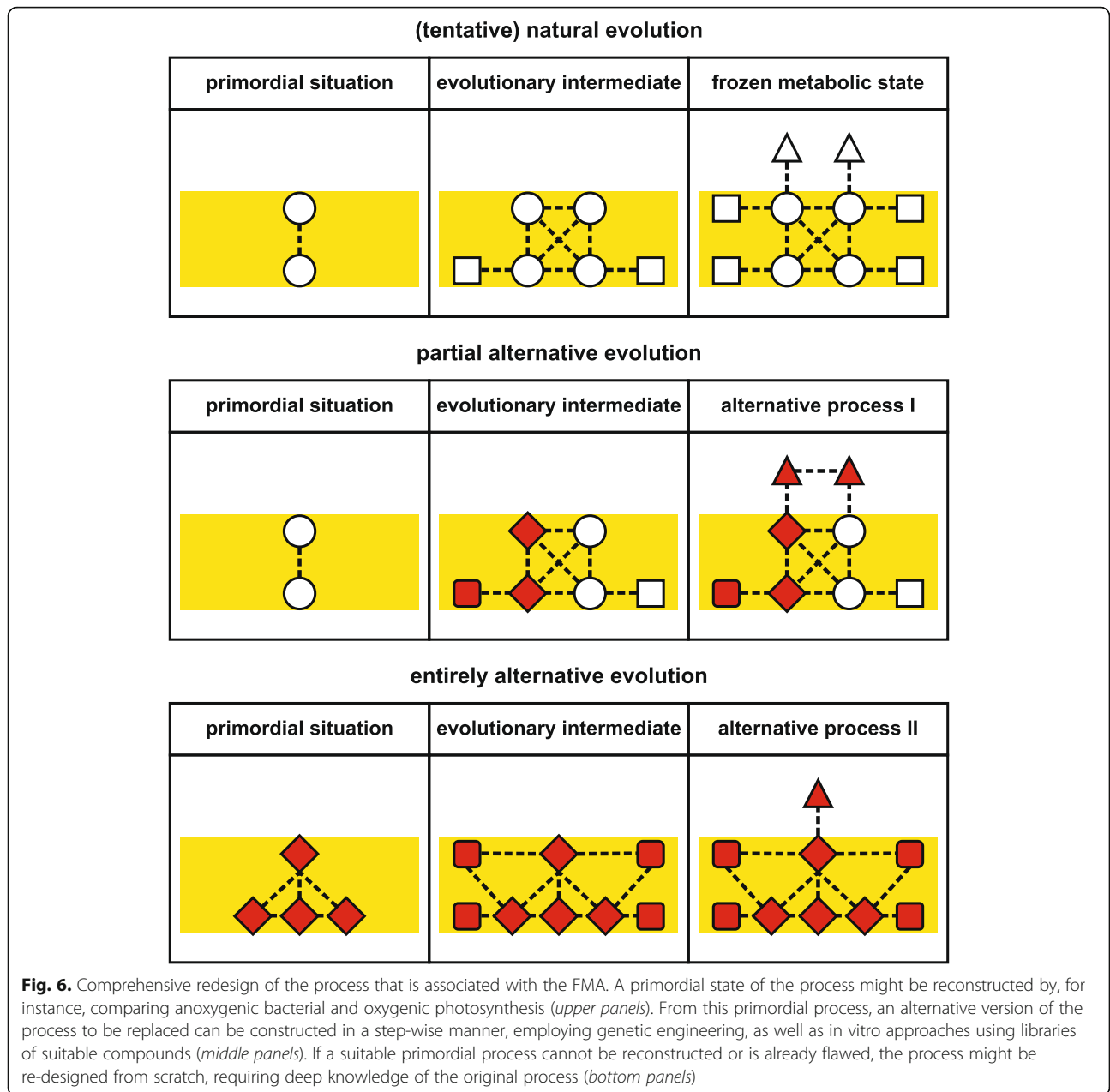
Fig. 5. Resolving the consequences of FMAs using non-natural components. A similar workflow as for replacing the original variant by another natural variant (Fig. 4) is employed, except that, in the first step, the natural protein (*left*) is replaced by a functional non-natural protein from a library specifically designed for this purpose (*middle*). This can be achieved via homologous recombination (in the case of cyanobacterial or plastid transformation) or by complementation of a mutant that is grown heterotrophically (shown here). The function of the protein can be enhanced by laboratory evolution (*right*), optimizing interactions (from *thin dotted lines* to *thick dotted lines*) and involving adaptive sequence variation (indicated by altered shading) of the introduced non-natural protein and its binding partners

to protect PSI against photoinhibitory damage, which is considered to be more dangerous than PSII photo-damage [71–74]. In this case, high light tolerance of both PSII and PSI has to be enhanced to have a positive impact on the resistance of the photosynthetic light reactions towards high levels of light, and only two enhancement strategies remain open: (1) re-design of the process, starting from a primordial version that does not exhibit the relevant defect(s), or (2) complete replacement with a synthetic process designed from scratch (Fig. 6). With regard to the former possibility, all natural light-dependent and water-splitting systems in cyanobacteria, algae, and plants are related to each other, but anoxygenic bacterial relatives of PSII have a less complex structure and might serve as models of primordial versions of PSII [70, 75–78]. Starting from these primitive versions of PSII, one needs to recapitulate the evolution from non-oxygenic, bacterial, single-photosystem situations to a cyanobacterial, oxygenic, two-photosystem solution that operates optimally in the presence of oxygen and high light. This option presents various formidable technical and conceptual obstacles, and will require a more detailed understanding of the mechanisms and evolution of photosynthesis than has yet been attained. The second possibility—inventing a synthetic system that is devoid of the flaws of the natural one—is even more challenging. Although it is clear that water splitting capacity will always be associated with the potential to produce ROS and that it is unlikely that a new type of *in vivo* photosynthesis can be constructed from scratch in the near future, the associated opportunities are fascinating, as one could also combine chlorophylls and bacteriochlorophylls in the same organism, or even incorporate novel

pigments that absorb more of the light spectrum than conventional ones can utilize.

Harnessing synthetic biology to resolve frozen metabolic accidents

Due to their immense significance for the biosphere in terms of nitrogen and carbon fixation, efforts to enhance the processes associated with FMAs are of the utmost interest. Elegant manipulations of the genetic code have demonstrated that it is feasible to get around the frozen *genetic* accident, but frozen *metabolic* accidents involve another level of complexity, owing to the intricate biophysical and structural interdependencies implicated in the frozen state, which extend to auxiliary components. To resolve the roadblocks caused by FMAs, a comprehensive knowledge of the process to be modified is vital, and elegant genetic engineering strategies have to be designed. The most promising and technically feasible approaches of resolving FMAs currently are to exploit the few evolutionary opportunities that they have left open by screening extant biological diversity for possible remedies (e.g., the high-light resistant PSII from *C. ohadii*) and applying laboratory evolution in cases where the process can be transferred to a suitable host (as in the case of nitrogenase and the Calvin cycle in *E. coli*). A much more challenging approach will be to test large sets of non-natural proteins or maquettes for their ability to overcome some of the limitations imposed by FMAs, and this will be only possible in suitable microbial systems that are accessible to efficient genetic engineering and laboratory evolution. The ultimate challenge will be to re-design from scratch a process that has been trapped in an evolutionary blind alley and is in principle suboptimally adapted to current conditions, at least in terrestrial environments. However, replacing this process



with a completely different one amounts to discovering an evolutionary path that was never realized in nature, because the appropriate combination of starting conditions and selective forces never materialized. In this context it remains to be shown whether PSII truly qualifies as FMA because its sensitivity to light stress due to ROS production might have evolved regardless of the conditions (aerobic or pre-aerobic) and might therefore be inevitable, although it is clear that at least some organisms apparently have developed mechanisms that can largely overcome this limitation (the *C. ohadii* case). Moreover, the photosensitivity of PSII

also protects PSI against photoinhibition, extending the need for re-design of both photosystems. Whichever explanation(s) for the inherent light sensitivity of PSII will finally turn out to be correct, for enhancing the function of a process assumed to be a FMA one might ideally start from a primordial process that is sufficiently flexible to allow it to be re-designed by employing non-natural proteins and laboratory evolution. For such experiments, *E. coli* is the ideal workhorse, and given that fixation of nitrogen [44] and carbon [30] and the biosynthesis of chlorophyll [79] and carotenoids [80–82] have already been reconstituted in this host, the light reactions

of photosynthesis are an especially attractive target of this approach. Thawing out FMAs will, however, require synthetic biology and laboratory evolution, but also further knowledge of the reactions themselves.

Acknowledgements

The author thanks Paul Hardy and Tatjana Kleine for critical reading of the manuscript.

Funding

The author receives funding for his work on synthetic biology and enhancing photosynthesis from the Deutsche Forschungsgemeinschaft (DFG; GRK 2062 and SFB-TRR 175).

Availability of data and materials

Not applicable.

Author's contributions

DL wrote, read, and approved the final manuscript.

Competing interests

The author declares that he has no competing interests.

Publisher's Note

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

Published online: 30 January 2019

References

- Koonin EV, Novozhilov AS. Origin and evolution of the universal genetic code. *Annu Rev Genet.* 2017;51:45–62.
- Yarus M, Widmann JJ, Knight R. RNA-amino acid binding: a stereochemical era for the genetic code. *J Mol Evol.* 2009;69(5):406–29.
- Wong JT, Ng SK, Mat WK, Hu T, Xue H. Coevolution theory of the genetic code at age forty: pathway to translation and synthetic life. *Life (Basel).* 2016;6(1):E12.
- Archetti M. Selection on codon usage for error minimization at the protein level. *J Mol Evol.* 2004;59(3):400–15.
- Crick FH. The origin of the genetic code. *J Mol Biol.* 1968;38(3):367–79.
- Freeland SJ, Hurst LD. The genetic code is one in a million. *J Mol Evol.* 1998;47(3):238–48.
- Koonin EV, Novozhilov AS. Origin and evolution of the genetic code: the universal enigma. *IUBMB Life.* 2009;61(2):99–111.
- Doig AJ. Frozen, but no accident - why the 20 standard amino acids were selected. *FEBS J.* 2017;284(9):1296–305.
- Vos M, Hesselman MC, Te Beek TA, van Passel MWJ, Eyre-Walker A. Rates of lateral gene transfer in prokaryotes: high but why? *Trends Microbiol.* 2015;23(10):598–605.
- Koonin EV. Frozen accident pushing 50: stereochemistry, expansion, and chance in the evolution of the genetic code. *Life (Basel).* 2017;7(2):E22.
- Vetsigian K, Woese C, Goldenfeld N. Collective evolution and the genetic code. *Proc Natl Acad Sci U S A.* 2006;103(28):10696–701.
- Sengupta S, Aggarwal N, Bandhu AV. Two perspectives on the origin of the standard genetic code. *Orig Life Evol Biosph.* 2014;44(4):287–91.
- Aggarwal N, Bandhu AV, Sengupta S. Finite population analysis of the effect of horizontal gene transfer on the origin of an universal and optimal genetic code. *Phys Biol.* 2016;13(3):036007.
- Shi T, Bibby TS, Jiang L, Irwin AJ, Falkowski PG. Protein interactions limit the rate of evolution of photosynthetic genes in cyanobacteria. *Mol Biol Evol.* 2005;22(11):2179–89.
- Miyao M. Involvement of active oxygen species in degradation of the D1 protein under strong illumination in isolated subcomplexes of photosystem II. *Biochemistry.* 1994;33(32):9722–30.
- De Las Rivas J, Shipton CA, Ponticos M, Barber J. Acceptor side mechanism of photoinduced proteolysis of the D1 protein in photosystem II reaction centers. *Biochemistry.* 1993;32(27):6944–50.
- Nishiyama Y, Allakhverdiev SI, Murata N. A new paradigm for the action of reactive oxygen species in the photoinhibition of photosystem II. *Biochim Biophys Acta.* 2006;1757(7):742–9.
- Murata N, Allakhverdiev SI, Nishiyama Y. The mechanism of photoinhibition *in vivo*: re-evaluation of the roles of catalase, alpha-tocopherol, non-photochemical quenching, and electron transport. *Biochim Biophys Acta.* 2012;1817(8):1127–33.
- Ohad I, Adir N, Koike H, Kyle DJ, Inoue Y. Mechanism of photoinhibition *in vivo*. A reversible light-induced conformational change of reaction center II is related to an irreversible modification of the D1 protein. *J Biol Chem.* 1990;265(4):1972–9.
- Parry MA, Andralojc PJ, Scales JC, Salvucci ME, Carmo-Silva AE, Alonso H, et al. Rubisco activity and regulation as targets for crop improvement. *J Exp Bot.* 2013;64(3):717–30.
- Berman-Frank I, Lundgren P, Falkowski P. Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. *Res Microbiol.* 2003;154(3):157–64.
- Mus F, Alleman AB, Pence N, Seefeldt LC, Peters JW. Exploring the alternatives of biological nitrogen fixation. *Metallomics.* 2018;10(4):523–38.
- Berman-Frank I, Chen Y-B, Gao Y, Fennel K, Follows MJ, Milligan AJ, et al. Feedbacks between the nitrogen, carbon and oxygen cycles. In: Capone DG, Bronk DA, Mulholland MR, Carpenter EJ, editors. *Nitrogen in the marine environment*. 2nd ed. Burlington: Elsevier; 2008. p. 1537–63.
- Whitney SM, Birch R, Kelso C, Beck JL, Kapralov MV. Improving recombinant Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1 chaperone. *Proc Natl Acad Sci U S A.* 2015;112(11):3564–9.
- Leister D. Genetic engineering, synthetic biology and the light reactions of photosynthesis. *Plant Physiol.* 2018. <https://doi.org/10.1104/pp.18.00360>.
- Erb TJ, Zarzycki J. A short history of RubisCO: the rise and fall (?) of Nature's predominant CO₂ fixing enzyme. *Curr Opin Biotechnol.* 2018;49:100–7.
- Liu D, Ramya RCS, Mueller-Cajar O. Surveying the expanding prokaryotic Rubisco multiverse. *FEMS Microbiol Lett.* 2017;364(16). <https://doi.org/10.1093/femsle/fnx156>.
- Aigner H, Wilson RH, Bracher A, Calisse L, Bhat JY, Hartl FU, et al. Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD2. *Science.* 2017;358(6368):1272–8.
- Antonovsky N, Gleizer S, Noor E, Zohar Y, Herz E, Barenholz U, et al. Sugar synthesis from CO₂ in *Escherichia coli*. *Cell.* 2016;166(1):115–25.
- Parikh MR, Greene DN, Woods KK, Matsumura I. Directed evolution of RuBisCO hypermorphs through genetic selection in engineered *E. coli*. *Protein Eng Des Sel.* 2006;19(3):113–9.
- Guadalupe-Medina V, Wisselink HW, Luttkik MA, de Hulster E, Daran JM, Pronk JT, et al. Carbon dioxide fixation by Calvin-cycle enzymes improves ethanol yield in yeast. *Biotechnol Biofuels.* 2013;6(1):125.
- Varaljay VA, Satagopan S, North JA, Witte B, Dourado MN, Anantharaman K, et al. Functional metagenomic selection of ribulose 1, 5-bisphosphate carboxylase/oxygenase from uncultivated bacteria. *Environ Microbiol.* 2016;18(4):1187–99.
- Mueller-Cajar O, Morell M, Whitney SM. Directed evolution of rubisco in *Escherichia coli* reveals a specificity-determining hydrogen bond in the form II enzyme. *Biochemistry.* 2007;46(49):14067–74.
- Cai Z, Liu G, Zhang J, Li Y. Development of an activity-directed selection system enabled significant improvement of the carboxylation efficiency of rubisco. *Protein Cell.* 2014;5(7):552–62.
- Greene DN, Whitney SM, Matsumura I. Artificially evolved *Synechococcus* PCC6301 Rubisco variants exhibit improvements in folding and catalytic efficiency. *Biochem J.* 2007;404(3):517–24.
- Schwander T, Schada von Borzyskowski L, Burgener S, Cortina NS, Erb TJ. A synthetic pathway for the fixation of carbon dioxide *in vitro*. *Science.* 2016;354(6314):900–4.
- Mall A, Sobotta J, Huber C, Tschirner C, Kowarschik S, Bacnik K, et al. Reversibility of citrate synthase allows autotrophic growth of a thermophilic bacterium. *Science.* 2018;359(6375):563–7.
- Konneke M, Schubert DM, Brown PC, Hugler M, Standfest S, Schwander T, et al. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc Natl Acad Sci U S A.* 2014;111(22):8239–44.
- Berg IA. Ecological aspects of the distribution of different autotrophic CO₂ fixation pathways. *Appl Environ Microbiol.* 2011;77(6):1925–36.
- Antonovsky N, Gleizer S, Milo R. Engineering carbon fixation in *E. coli*: from heterologous RuBisCO expression to the Calvin-Benson-Bassham cycle. *Curr Opin Biotechnol.* 2017;47:83–91.
- Erb TJ, Zarzycki J. Biochemical and synthetic biology approaches to improve photosynthetic CO₂-fixation. *Curr Opin Chem Biol.* 2016;34:72–9.

42. Buren S, Rubio LM. State of the art in eukaryotic nitrogenase engineering. *FEMS Microbiol Lett.* 2018;365(2). <https://doi.org/10.1093/femsle/fnx274>.
43. Oldroyd GE, Dixon R. Biotechnological solutions to the nitrogen problem. *Curr Opin Biotechnol.* 2014;26:19–24.
44. Dixon RA, Postgate JR. Genetic transfer of nitrogen fixation from *Klebsiella pneumoniae* to *Escherichia coli*. *Nature.* 1972;237(5350):102–3.
45. Yang J, Xie X, Yang M, Dixon R, Wang YP. Modular electron-transport chains from eukaryotic organelles function to support nitrogenase activity. *Proc Natl Acad Sci U S A.* 2017;114(12):E2460–E5.
46. Carpenter SD, Ohad I, Vermaas WF. Analysis of chimeric spinach/cyanobacterial CP43 mutants of *Synechocystis* sp. PCC 6803: the chlorophyll-protein CP43 affects the water-splitting system of photosystem II. *Biochim Biophys Acta.* 1993;1144(2):204–12.
47. Nixon PJ, Rogner M, Diner BA. Expression of a higher plant psbA gene in *Synechocystis* 6803 yields a functional hybrid photosystem II reaction center complex. *Plant Cell.* 1991;3(4):383–95.
48. Vermaas WF, Shen G, Ohad I. Chimeric CP47 mutants of the cyanobacterium *Synechocystis* sp. PCC 6803 carrying spinach sequences: construction and function. *Photosynth Res.* 1996;48(1–2):147–62.
49. Chiamonte S, Giacometti GM, Bergantino E. Construction and characterization of a functional mutant of *Synechocystis* 6803 harbouring a eukaryotic PSII-H subunit. *Eur J Biochem.* 1999;260(3):833–43.
50. Gimpel JA, Nour-Eldin HH, Scranton MA, Li D, Mayfield SP. Refactoring the six-gene photosystem II core in the chloroplast of the green algae *Chlamydomonas reinhardtii*. *ACS Synth Biol.* 2016;5(7):589–96.
51. Hecht MH, Zarzhitsky S, Karas C, Chari S. Are natural proteins special? Can we do that? *Curr Opin Struct Biol.* 2018;48:124–32.
52. Anderson JLR, Armstrong CT, Kodali G, Lichtenstein BR, Watkins DW, Mancini JA, et al. Constructing a man-made c-type cytochrome maquette *in vivo*: electron transfer, oxygen transport and conversion to a photoactive light harvesting maquette. *Chem Sci.* 2014;5(2):507–14.
53. Lichtenstein BR, Farid TA, Kodali G, Solomon LA, Anderson JL, Sheehan MM, et al. Engineering oxidoreductases: maquette proteins designed from scratch. *Biochem Soc Trans.* 2012;40(3):561–6.
54. Lajoie MJ, Soll D, Church GM. Overcoming challenges in engineering the genetic code. *J Mol Biol.* 2016;428(5 Pt B):1004–21.
55. Ostrov N, Landon M, Guell M, Kuznetsov G, Teramoto J, Cervantes N, et al. Design, synthesis, and testing toward a 57-codon genome. *Science.* 2016;353(6301):819–22.
56. Lajoie MJ, Rovner AJ, Goodman DB, Aerni HR, Haimovich AD, Kuznetsov G, et al. Genomically recoded organisms expand biological functions. *Science.* 2013;342(6156):357–60.
57. Wannier TM, Kunjapur AM, Rice DP, McDonald MJ, Desai MM, Church GM. Adaptive evolution of genomically recoded *Escherichia coli*. *Proc Natl Acad Sci U S A.* 2018;115(12):3090–5.
58. Conrad TM, Lewis NE, Palsson BO. Microbial laboratory evolution in the era of genome-scale science. *Mol Syst Biol.* 2011;7:509.
59. Rast A, Heinz S, Nickelsen J. Biogenesis of thylakoid membranes. *Biochim Biophys Acta.* 2015;1847(9):821–30.
60. Roose JL, Frankel LK, Mummadisetti MP, Bricker TM. The extrinsic proteins of photosystem II: update. *Planta.* 2016;243(4):889–908.
61. Barber J. Photosystem II: its function, structure, and implications for artificial photosynthesis. *Biochemistry (Mosc).* 2014;79(3):185–96.
62. van Amerongen H, Croce R. Light harvesting in photosystem II. *Photosynth Res.* 2013;116(2–3):251–63.
63. Treves H, Raanan H, Kedem I, Murik O, Keren N, Zer H, et al. The mechanisms whereby the green alga *Chlorella ohadii*, isolated from desert soil crust, exhibits unparalleled photodamage resistance. *New Phytol.* 2016;210(4):1229–43.
64. Ananyev G, Gates C, Kaplan A, Dismukes GC. Photosystem II-cyclic electron flow powers exceptional photoprotection and record growth in the microalga *Chlorella ohadii*. *Biochim Biophys Acta.* 2017;1858(11): 873–83.
65. Treves H, Murik O, Kedem I, Eisenstadt D, Meir S, Rogachev I, et al. Metabolic flexibility underpins growth capabilities of the fastest growing alga. *Curr Biol.* 2017;27(16):2559–67 e3.
66. Treves H, Raanan H, Finkel OM, Berkowicz SM, Keren N, Shotland Y, et al. A newly isolated *Chlorella* sp. from desert sand crusts exhibits a unique resistance to excess light intensity. *FEMS Microbiol Ecol.* 2013; 86(3):373–80.
67. Beasley JR, Hecht MH. Protein design: the choice of de novo sequences. *J Biol Chem.* 1997;272(4):2031–4.
68. Fisher MA, McKinley KL, Bradley LH, Viola SR, Hecht MH. De novo designed proteins from a library of artificial sequences function in *Escherichia coli* and enable cell growth. *PLoS One.* 2011;6(1):e15364.
69. Koder RL, Anderson JL, Solomon LA, Reddy KS, Moser CC, Dutton PL. Design and engineering of an O(2) transport protein. *Nature.* 2009;458(7236):305–9.
70. Rutherford AW, Faller P. Photosystem II: evolutionary perspectives. *Philos Trans R Soc Lond Ser B Biol Sci.* 2003;358(1429):245–53.
71. Chauv F, Peltier G, Johnson X. A security network in PSI photoprotection: regulation of photosynthetic control, NPQ and O₂ photoreduction by cyclic electron flow. *Front Plant Sci.* 2015;6:875.
72. Kangasjarvi S, Tikkanen M, Durian G, Aro EM. Photosynthetic light reactions—an adjustable hub in basic production and plant immunity signaling. *Plant Physiol Biochem.* 2014;81:128–34.
73. Scheller HV, Haldrup A. Photoinhibition of photosystem I. *Planta.* 2005; 221(1):5–8.
74. Sonoike K. Photoinhibition of photosystem I. *Physiol Plant.* 2011;142(1):56–64.
75. Raymond J, Blankenship RE. The evolutionary development of the protein complement of photosystem 2. *Biochim Biophys Acta.* 2004;1655(1–3):133–9.
76. Martin WF, Bryant DA, Beatty JT. A physiological perspective on the origin and evolution of photosynthesis. *FEMS Microbiol Rev.* 2018;42(2):205–31.
77. Cardona T. A fresh look at the evolution and diversification of photochemical reaction centers. *Photosynth Res.* 2015;126(1):111–34.
78. Cardona T, Murray JW, Rutherford AW. Origin and evolution of water oxidation before the last common ancestor of the cyanobacteria. *Mol Biol Evol.* 2015;32(5):1310–28.
79. Chen GE, Canniffe DP, Barnett SFH, Hollingshead S, Brindley AA, Vasilev C, et al. Complete enzyme set for chlorophyll biosynthesis in *Escherichia coli*. *Sci Adv.* 2018;4(1):eaq1407.
80. Wang C, Zada B, Wei G, Kim SW. Metabolic engineering and synthetic biology approaches driving isoprenoid production in *Escherichia coli*. *Bioresour Technol.* 2017;241:430–8.
81. Misawa N. Pathway engineering for functional isoprenoids. *Curr Opin Biotechnol.* 2011;22(5):627–33.
82. Harada H, Misawa N. Novel approaches and achievements in biosynthesis of functional isoprenoids in *Escherichia coli*. *Appl Microbiol Biotechnol.* 2009; 84(6):1021–31.