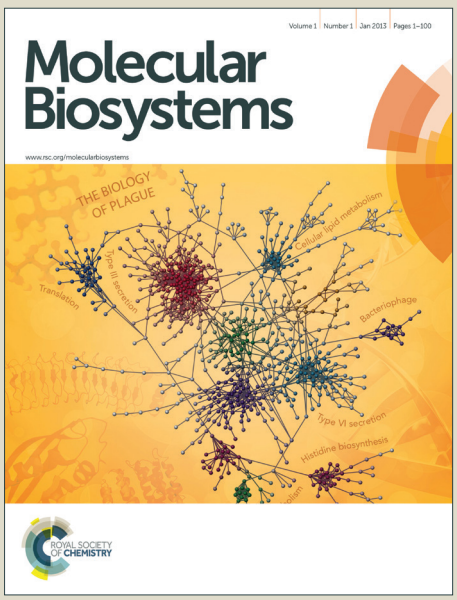
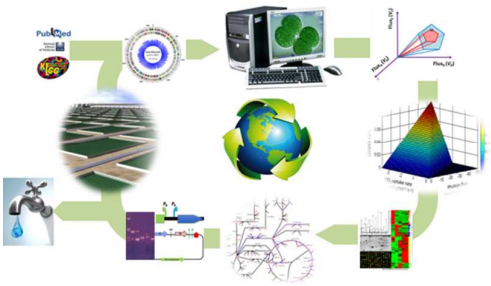
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems



A review of cyanobacterial biocatalysts highlighting their metabolic features that argues for the need for systems-level metabolic engineering.

Cyanobacteria as photosynthetic biocatalysts: A systems biology perspective

Steinn Gudmundsson[1] and Juan Nogales[2‡]

[1]Center for Systems Biology, University of Iceland, Sturlugata 8, 101 Reykjavik, Iceland.
[2]Department of Environmental Biology, Centro de Investigaciones Biológicas-CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

‡To whom correspondence should be addressed: Department of Environmental Biology, Centro de Investigaciones Biológicas (CSIC), Madrid, Spain. E-mail: jnogales@cib.csic.es.

Keywords: Cyanobacteria, genome-scale modeling, metabolic engineering, chemical building blocks, synthetic biology, systems biology, biofuels

Abstract

The increasing need to replace oil-based products and to address global climate change concerns, has triggered a considerable interest in photosynthetic microorganisms. Cyanobacteria, in particular, have great potential as biocatalysts for fuels and fine-chemicals. During the last few years the biotechnological applications of cyanobacteria have experienced an unprecedented increase and the use of these photosynthetic organisms for chemical production is becoming a tangible reality. However, the field is still immature and many concerns about the economic feasibility of the biotechnological potential of cyanobacteria remain. In this review we describe recent successes in biofuel and fine-chemical production using cyanobacteria. We discuss the role of the photosynthetic metabolism and highlight the need for systems-level metabolic optimization in order to achieve the true potential of cyanobacterial biocatalysts.

Introduction

The recycling of CO_2 into usable fuels, chemical building blocks and fine chemicals by photosynthetic organisms has received considerable interest in recent years due to the ever increasing demand for energy, the depletion of fossil fuels and climate change. First generation biodiesel and chemicals derived from crops and biomass are increasingly being questioned over concerns such as high production costs and the competition with edible crops over land use.[1, 2] Microalgae are an alternative source of biodiesel and selected chemicals such as carotenoids, however, the costs of biomass harvesting and downstream processing are still far too high to make them an economically feasible source of fuels.[3] In an attempt to reduce the costs, a new direct approach has been proposed. This approach uses photosynthetic organisms such as eukaryotic microalgae and cyanobacteria which have been engineered to convert CO_2 directly into biofuel or high-value chemicals, without the need to synthesize and process high amounts of biomass. This approach implies continuous production and secretion of target metabolites from the culture while minimizing the production of biomass and undesirable byproducts.

Cyanobacteria have been used as food sources and biofertilizers for centuries,[4] they produce a broad spectrum of high-value compounds,[5] have minimal nutritional requirements, [6] and higher photosynthetic efficiency in term of sunlight conversion into biomass and growth rates than all other photosynthetic organisms. [7-9] These properties have led to a great interest in using cyanobacteria as photobiocatalysts for chemical production. In addition, their primary metabolic capabilities have been modeled at genome-scale.[10-14] enabling quantitative predictions of cellular behavior, and they are amenable to genetic manipulation.[15] Building upon these features and establishing a proof of concept of the direct production of chemicals from oxygenic photosynthesis, cyanobacteria have been successfully engineered to produce high value and biofuel-like compounds in the last few years (Table S1). However, only a limited number of chemicals has been produced so far and the cyanobacterial biocatalysts are currently hampered by low yields which prevents their application on a large scale. Overcoming these limitations requires a holistic strategy, which includes increasing the photosynthetic and CO_2 fixation efficiencies; optimizing photobioreactors; but also a better understanding of the cyanobacterial physiology, coupled with the use of systems and synthetic biology approaches. Here, we briefly review recent advances in cyanobacterial biotechnology. We elaborate on some of the open problems in the field and call attention to the need for systems-level metabolic optimization endeavors in order to further develop these promising biocatalysts.

Cyanobacteria as cell-factories

Cyanobacteria are able to synthetize an array of value added compounds and they have been used as a source of drugs, toxins and fine chemicals for decades.[4] The dawn of the genomic age, recent developments in genetic tools and the need to find alternatives to oil-based products have resulted in a significant increase in biotechnological studies of cyanobacteria in the last five years. These recent efforts have been targeted at the production of: i) alcohols and related biofuel compounds, ii) lipid based biofuels, iii) sugars, iv) biomaterials and v) high value compounds (Table S1).

Alcohols and related biofuel compounds

Many cyanobacteria produce small quantities of ethanol, however they lack the repertory of NADH-dependent fermentative pathways required to synthetize ethanol and other biofuel-like compounds at high titers. Compounding the problem is a NAD:NADP ratio, ranging from 1:10 to 1:5, depending on the growth conditions. The NAD(P)H/NAD(P) reduction states are likely to be an important factor as well but more studies are needed since the results reported so far are conflicting. [16-18] These properties have significant implications for biofuel production in cyanobacteria due to the preference for NADH of most the biofuel biosynthetic pathways. In the last decade, several non-native pathways have been introduced in cyanobacteria, resulting in the successful production of a number of alcohol based biofuels. Deng and Coleman first demonstrated the feasibility of this engineered “photofermentative” metabolism. By introducing pyruvate decarboxylase and alcohol dehydrogenase II from *Zymomonas mobilis* in *Synechococcus* sp PCC 7942 (PCC7942), it was possible to produce up to 5 mg/ml of ethanol. [19] Although the yields were extremely low compared to heterotrophic organisms, significantly higher yields were obtained in later attempts[20, 21] (Table S1). For instance, Gao et al recently reported the production of up to 5500 mg/L of ethanol in *Synechocystis* sp PCC 6803 (PCC6803).[21] The high productivity was achieved by overexpressing the pyruvate decarboxylase from *Z. mobilis as well as* the native NADPH-dependent alcohol dehydrogenase Slr1192, thereby overcoming the low NAD:NADP ratio. Atsumi and colleagues had previously employed a similar approach to produce isobutyraldehyde and isobutanol from pyruvate in PCC7942 by expressing a synthetic pathway including an acetoacetate synthase (AlsS), an acetohydroxy acid isomeroreductase (IlvC), a dihydroxy-acid dehydratase (IlvD), a ketoacid decarboxylase (Kivd) and an NADPH-dependent alcohol dehydrogenase (YqhD).[22] They obtained increased yields of isobutyraldehyde by keeping the concentration low,

reducing the toxicity to the cells, by *in situ* product removal. In a later study, the Liao group increased the yield of isobutanol by a factor of 2.5 by blocking the glycogen biosynthetic pathway which is a major sink of carbon in the autotrophic metabolism. [23]

Many non-native pathways exhibit high sensitivity to oxygen, reducing the possibility of heterologous expression in oxygenic photosynthetic organisms. Lan and Liao transferred a modified CoA-dependent 1-butanol pathway into PCC7942, obtaining 14.5 mg/L of 1-butanol in the dark under anoxic conditions.[24] No 1-butanol production was observed under photosynthetic conditions which was attributed to a lack of a thermodynamic driving force and oxygen sensitivity of the enzymes in the pathway.[24] In a follow-up study Lan and Liao engineered an ATP driving force by expressing an acetoacetyl-CoA synthase [25] resulting in 6.5 mg/L of 1-butanol under photosynthetic conditions. By utilizing a NADPH-dependent acetoacetyl-CoA synthase and 3-ketobutyryl-CoA reductase in the pathway, the final titer reached 30 mg/L. [25] In their most recent work, Lan and Lio addressed the oxygen sensitivity of the pathway by replacing the CoA-acylating butyraldehyde dehydrogenase with a CoA-acylating propionaldehyde dehydrogenase, resulting in 400 mg/L of 1-butanol.[26] By combining an oxygen-insensitive pathway, cofactor optimization and the introduction of irreversible enzymatic steps, 2.38 g/L of 2,3-butanediol were obtained from CO_2.[27] The above examples serve to illustrate the importance of taking the specific properties of the phototrophic metabolism into account in the metabolic engineering of cyanobacteria.

Lipid-based biofuels

Oleaginous algae have largely dominated the production of lipid based biofuels until now because of their ability to produce high amounts of triacylglycerol.[28] However, cells containing triacylglycerol require complex and expensive downstream processing.[3] Cyanobacteria have recently been engineered to produce and secrete free fatty acids (FFA) and long chain alkenes into the culture medium and they may turn out to be a viable alternative to algae. Liu and colleges employed six successive rounds of genetic modifications of PCC6803 and achieved 200 mg/L of extracellularly secreted fatty acids[29] (Table S1). The modifications included overproduction of acetyl-CoA carboxylase to funnel more carbon flux to fatty acids biosynthesis, the heterologous expression of engineered thioesterases, knockouts of genes encoding for competing pathways, e.g. cyanophycin biosynthesis, and weakening of the cell wall layer.

Non-native fatty alcohol biosynthetic pathways, including heterologous fatty acyl-CoA reductases (FARs) from different sources, were used to produce

hexadecanol, octadecanol and other fatty alcohols in PCC6803. [30] The most promising producer strain included a FAR from jojoba, resulting in up to 10 μg.$OD^{-1}L^{-1}$ (≈ 26.2 μg/gDW) of fatty alcohols and scale-up experiments revealed that the production of fatty alcohols was effectively doubled under high light conditions. A follow-up study by the same group reported significantly improved yields, 761 μg/gDW. This was achieved by knocking out competing pathways (e.g., glycogen biosynthesis), promoter engineering and overexpressing multiple FARs. [31]

Long chain alka(e)nes are produced naturally by several cyanobacterial species as part of their lipidic membranes (up to 0.1% of the cell dry weight). They are ideal biofuels for several reasons. They have minimal downstream processing requirements, good combustion properties and the infrastructure for storage and distribution is already in place. An alkane biosynthesis pathway in cyanobacteria involving an acyl-acyl carrier protein reductase (AAR) and an aldehyde decarbonylase (AD) was recently discovered. [32] The overexpression of these enzymes together with an acetyl-CoA carboxylase in PCC6803 resulted in 26 mg/L of alka(e)nes. [33] Furthermore, the overexpression of a class-3 aldehyde-dehydrogenase in conjunction with AAR and AD in PCC7942 [34] shifted the production from alkanes towards fatty acids, with the fatty acids being secreted from the cell. It has been shown that the production of fatty acids in cyanobacteria induces oxidative stress, which could limit both the efficiency and the lifetime of the biocatalyst. Comparative transcriptomics analysis identified up to 15 genes involved in fatty acid-induced stress defense in PCC7942.[35] Interestingly, targeted mutagenesis and/or overexpression of some of the genes reduced fatty acid toxicity and subsequently led to an increase in fatty acid production.

Sugars

Cyanobacteria accumulate high levels of sucrose as an osmoprotectant under salt stress.[36] This property, combined with the overexpression of key enzymes involved in sucrose biosynthesis has been used in PCC6803 to obtain a significant increase in intracellular sucrose levels.[37] Niederholtmeyer and colleagues obtained secretion of glucose and fructose in PCC7942 by overexpressing both invertase InvA, which produces glucose and fructose from sucrose, and the GLF sugar transport. [38] Ducat et al. expressed a CscB-dependent sucrose export system from *E. coli*, knocked out invertase InvA and blocked the glycogen biosynthesis pathway and obtained up to 10 mM (approx. 3.5 g/L) of sucrose under osmotic stress. [39] The scope of sugar production continues to grow. For instance, a synthetic pathway for mannitol biosynthesis was recently expressed in *Synechoccocuus sp*. PCC 7002 (PCC7002),

yielding up to 10% of cell dry weight.[40] By blocking glycogen biosynthesis, the yield increased to 32%.

The potential of cyanobacteria to secrete sugars has been suggested as a feasible strategy to genetically engineer multispecies microbial cell factories, where the cyanobacteria provide oxygen and organic substrates and the heterotroph partner acts as an efficient biocatalyst.[41] Although Niederholtmeyer et al. found that the secreted sugars supported *E. coli* growth in a co-culture with PCC7942, [38] theoretical estimates indicate that the development of highly efficient multispecies biocatalysts will be very challenging. [42, 43]

Biomaterials and chemical building blocks

Many cyanobacteria accumulate polyhydroxybutyrate at a high rate under nitrogen and/or phosphorus starvation.[44] However, the addition of external carbon sources such as acetate was frequently required to achieve yields comparable to those found in heterotrophs. Recent efforts combining inverse metabolic engineering with high-throughput screening,[45] as well as systems biology approaches, including transcriptomic and carbon flux rerouting, have been successfully applied, resulting in significant increase in yields.[46] In addition, the production and extracellular secretion of the polyhydroxybutyrate intermediate (S),(R)-3-hydroxybutyrate has been achieved in PCC6803 under autotrophic conditions.[47]

Isoprene is a versatile building block derived from crude oil which is mainly used for synthetic rubber but also in flavorings and perfumes. Small amounts of isoprene in PCC6803 were obtained by the heterologous expression of a codon optimized version of the *ispS* gene from *kudzu*, encoding for the isoprene synthase.[48] Ethylene, another major building block in the chemical industry has received considerable interest. For instance, ethylene has been produced in PCC7942 harboring the Ethylene-Forming Enzyme (EFE) from *Pseudomonas syringae*.[49] However, the production of ethylene was not stable over time, a frequent problem in the metabolic engineering of cyanobacteria. This was due to recurrent mutations of the encoding gene, even under inducible expression.[50] In two recent studies, stable and continuous production of ethylene was achieved through rigorous codon-use and promoter optimization.[51, 52] While these results are promising, the best yields achieved so far are 171 mg/L/day[52] and further efforts are clearly needed.

L-lactate has been produced in titers up to 290 mg/L in PCC6803, by expressing the *ldh* gene from *Bacillus subtilis* and a soluble transhydrogenase. [53] In a

later study Angermayr and Hellingwerf used metabolic control analysis to show that L-lactic production was linearly dependent on the lactate dehydrogenase enzymatic capacity. Significantly higher yields were achieved by expressing the *ldh* gene of *Lactococcus lactis* under the control of the promoter *trc* [54] (Table S1). Finally, D-lactate has been obtained by expressing a mutated glycerol dehydrogenase with D-lactate dehydrogenase activity in PCC6803. A titer of 1140 mg/L was achieved by increasing the NADH pool through the expression of a soluble transhydrogenase and codon optimization.[55]

High value compounds

Given the low yields of metabolites obtained in cyanobacteria so far, a reasonable alternative is to focus biotechnological efforts on the production of high-value compounds instead of high-volume, low-value compounds such as biofuels. [43] In a pioneering study, Yu et al, obtained 2.24 mg/L of eicosapentaenoic acid (EPA), a polyunsatured fatty acid of clinical importance, by expressing the EPA biosynthetic pathway from *Shewanella* sp. SCRC2738 in *Synechococus* sp. NKBG15041c.[56] Reinsvold and colleagues engineered a recombinant PCC6803 strain harboring the β-caryophyllene synthase gene (*QHS1*) from *Artemisia annua* resulting in the production of the non-native secondary metabolite β-caryophyllene which is used in the cosmetic industry. [57] Squalene, a 30-carbon natural isoprenoid, used in cosmetics and vaccines, was successfully produced in PCC6803 by disrupting the hopanoids biosynthetic pathway.[58]

Metabolic features specific to cyanobacterial networks

It is evident from the above discussion that there is a growing interest in the use of cyanobacteria as biocatalysts. Many foundational problems have been addressed (BOX 1), resulting in an increased variety of target chemicals, as well as in improved titers (Table S1). However, there is a long way to go before cyanobacteria can be widely applied as biocatalysts in industrial settings. Concerns stem from the many problems that are still unsolved, including low yields and the difficulties in designing efficient photobioreactors. [43, 59] It is not our aim here to discuss in detail all of the challenges facing cyanobacterial biotechnology since excellent reviews have already been published. [43,60-63] We focus instead on the metabolic features specific to

cyanobacterial networks and how they affect the potential use of cyanobacteria as efficient biocatalysts.

Knowledge gaps

Although cyanobacteria have been studied in considerable detail, PCC6803 in particular, knowledge about their metabolism is lacking in many aspects, even the central metabolism. For instance, the TCA cycle was considered "incomplete" for almost 50 years because alpha-ketoglutarate dehydrogenase (AKGDH) is missing, and only very recently has an alternative pathway been identified.[64] For the most widely used organisms in biotechnology, such as *E. coli*, there exists a large number of knowledgebases, including metabolic,[65] gene-expression[66] and transcriptional regulation [67] databases, together with extensive gene knock out[68] and gene knock in[69] libraries. In comparison, the availability of cyanobacterial resources is fairly limited [70, 71] and cyanobacteria-specific genetic tools have only recently become available.[72, 73] From successful engineering projects involving *E. coli* and yeast it is clear that improving productivity requires deep physiological, genetic and metabolic characterization of the host strain, as well as strain-specific genetic tools.[60, 74] Therefore, it is reasonable to assume that the metabolic engineering efforts undertaken in cyanobacteria until now have been hampered by significant knowledge gaps. This may explain to some extent the relatively small number of target chemicals explored so far and the low titers obtained in most of the studies.

Chemical space

An analysis of the chemical space covered by metabolically engineered cyanobacteria illustrates that many targets of industrial importance such as dicarboxylic acids, organic acids and amino acids in particular, remain to be explored (Table S1, Fig. 1a). While a significant portion of the chemical space relevant to industry has already been covered in metabolic engineering workhorses such as *E. coli*, the space covered by cyanobacteria is mostly restricted to alcohols, a few organic acids, sugars and terpenes (Fig. 1a). It is likely that the limited coverage is mainly due to lack of attempts so far and the recent example of p-coumaric acid production from tyrosine in PCC6803 [75] demonstrates how the scope can be extended to a new family of chemicals such as aromatic acids.

Theoretical yields

The low yields frequently reported in metabolic engineering studies involving cyanobacteria raise the question of whether they are caused by some inherent limitations of the metabolic network. To investigate this possibility, we computed the theoretical yields for selected chemicals. A genome-scale model of PCC6803, iJN678, was used to compute the maximum yields of under auto- and heterotrophic conditions and the yields were then compared to *E. coli.* The yields were defined as the number of carbon atoms converted to target product versus the number of carbon atoms consumed. The metabolic burden of the biosynthetic pathways was taken into account by requiring a certain amount of biomass to be produced at the same time as previously reported. [99] Accordingly, here the fraction was set to 20% of the maximum. Qualitatively similar results were obtained with other values of the biomass fraction. Under heterotrophic conditions the yields resembled those of *E. coli*, ranging from 0.53 (ethylene) to 0.86 (lactate). With the exception of lactate, the loss of CO_2 results in a significant decrease in yields in *E. coli*. Interestingly, the yields were very stable under autotrophic conditions and considerably higher in PCC6803 than in *E. coli* (Fig. 1b). The reason is that the light-driven metabolism avoids the loss of carbon in the form of CO_2. This is an exclusive trait of CO_2 fixing organisms [76] and it represents an important advantage over heterotrophs. The conservation of carbon increases the productivity of the target compound since carbon fixation is a major bottleneck in biotechnology. [63] The analysis also showed that the chemical production required fewer photons than biomass production, a finding previously reported by Maarleveld et al. [77] This indicates that increased product yields will lead to better light usage which can explain, to some extent, the high CO_2 fixation ratios found in several overproducer strains. [39] In summary, the high theoretical yields obtained under autotrophic conditions suggest that the topology of the photosynthetic metabolic networks, *per se*, is not directly responsible for the low yields obtained so far in cyanobacteria.

Carbon flux

An important but often overlooked issue in metabolic engineering efforts is the unique nature of cyanobacterial metabolic networks. For instance, the carbon flux and carbon partitioning in the central metabolism under autotrophic conditions have only recently been determined. [78] It was shown that under autotrophic conditions, the total CO_2 fixed in the form of 3-phosphoglycerate (3PG), is split between phosphoglycerate mutase and phosphoglycerate kinase in a ratio of 1:10. A similar ratio was later

predicted using a computational approach[12]. This carbon flux distribution differs considerably from heterotrophic metabolism and as a consequence only one out of ten CO_2 molecules that are fixed are funneled to pyruvate. In addition, studies carried out by the Melis group suggest that up to 80% of the carbon is funneled to sugar biosynthesis while only 5% and 10% is allocated to biotechnologically relevant pathways such as terpenoid and fatty acid biosynthesis, respectively (unpublished results).[48, 79] The flux distribution present in cyanobacteria appears to be a consequence of the high carbon flux across the Calvin cycle which is required to optimize CO_2 fixation. Furthermore, since respiration is not needed for energy production under autotrophic conditions, the carbon flux towards important biotechnological precursor metabolites, such as pyruvate and tricarboxylic acids is reduced. [43] This intrinsic difference with respect to heterotrophic networks is likely to be one of the reasons why attempts to transfer existing engineering strategies from heterotrophic organisms to cyanobacteria have had limited success so far. It has been suggested that in order to maximize the carbon partitioning towards the target product, a heterologous pathway should obtain the carbon as close as possible to the fixation pathway.[61] Interestingly, genetic engineering approaches which have focused on increasing the driving-force of the synthetic pathway, have increased the carbon flux to the key precursor, pyruvate, (which is only three steps away from 3PG) resulting in significant yields for ethanol (≈70%)[21] and 2,3-butanediol (>60%).[27] Although the issue of which precursors are best to "tap" from is unresolved, the above discussion indicates that achieving comparable productivity for other industrially relevant chemicals whose synthesis starts far away from 3PG, may be a considerable challenge.

Light metabolism

A unique property of the photoautotrophic metabolism is its dependence on light for energy supply. In addition to the photosynthetic linear electron flow pathway, phototrophs are equipped with a large number of alternate electron flow pathways which assist in balancing the ATP:NADPH ratio as a function of metabolic demand.[12, 80] This flexibility, or robustness of the photosynthetic system, allows for fine-tuning of light to energy conversion by photosystems I and II and the metabolic reactions and provides an ATP:NADPH ratio close to 1.5 which is required for optimal carbon fixation. In addition, systems biology studies employing gene essentiality prediction have suggested that cyanobacteria possess significantly reduced metabolic robustness compared to heterotrophic organisms[81] [12] and that there is a tradeoff between high photosynthetic robustness and low metabolic robustness.[12] If these predictions are

confirmed, the above properties would have a significant impact on biotechnology. First, the existence of multiple electron flow pathways involved in ATP and redox balancing limits the effects of the extreme ATP:NADPH ratios required to produce non-native chemicals.[61] However it is important to keep in mind that the photosynthetic robustness could also limit the use of biotechnologically relevant pathways as non-native electron sinks, a strategy that otherwise could be used to increase yields. Second, although reduced metabolic robustness could be an advantage for carbon flux rerouting since it implies fewer redundant and/or competing pathways, it also limits the possibilities to remove non-desirable, potentially toxic, byproducts from engineered pathways. The byproducts may contribute to the genetic instability found often in many of the heterologous pathways explored in cyanobacteria. The systems properties described above have so far mostly been unexplored in metabolic engineering efforts. Systems biology opens up the possibility to engineer these properties which could in turn increase the applicability of cyanobacteria as cell-factories.

In summary, it seems reasonable to assume that in addition to well-known optimization targets, such as photosynthetic efficiency, carbon fixation and bioreactor design, a deeper knowledge of the metabolism and physiology at the systems level will lead to further advances in cyanobacterial biotechnology. Systems metabolic engineering approaches, have successfully been applied to heterotrophic hosts to modify carbon partitioning and to increased yields,[74, 82] Such approaches have yet to be applied to cyanobacteria, suggesting that a large part of the metabolic "production space" is yet to be explored (Fig. 2). Systems approaches are likely to play an important role in future efforts, since the resulting gains are mostly independent of improvements of the classical optimization targets (Fig. 2).

Taking advantage of systems biology in cyanobacterial biotechnology

The metabolic optimization of cyanobacteria as biocatalysts requires a systems level understanding of their metabolic and physiological processes as well as cyanobacterial-specific metabolic engineering designs. Genome-scale metabolic network reconstructions (BOX 2) may turn out to be essential in achieving these goals. These reconstruction have been extremely useful as platforms for biological knowledge discovery, contextualization of omics data,[83] and they are increasingly being used in metabolic engineering[84]. Metabolic reconstructions now exist for several cyanobacterial

species, PCC6803,[12, 81, 85-89] PCC7002,[90] *Cyanothece* sp. ATCC51142 (ATCC51142) [13, 88] and *Spirulina platensis* C1.[14] The earliest models involved mainly the central metabolism, but later reconstructions included detailed modeling of photosynthetic processes, the diurnal cycle, synthesis of lipids and photosynthetic pigments. [10, 12, 13] In addition, 35 cyanobacterial metabolic networks have been reconstructed from their respective genome sequences using an automatic algorithm.[91] It is reasonable to expect that the arrival of cyanobacteria reconstructions and the emergence of synthetic biology will play a key role in resolving the multiple problems that currently hamper the biotechnological potential of cyanobacteria. Until now, the cyanobacteria models have been used for three main proposes: i) as tools for increasing biological knowledge, ii) as platforms for omics data integration and contextualization and iii) as a test bed for biotechnological applications.

Models as tools for increasing biological knowledge

Metabolic reconstructions are increasingly being used to gain insights into cyanobacterial metabolism and the TCA cycle has been studied in some detail. Nogales et al. proposed the GABA shunt as an alternative to close the TCA cycle in PCC6803 and computational analysis suggested that the GABA shunt provided an advantage over AKGDH under photoautotrophic conditions.[12] Knoop et al. evaluated several alternatives proposed to close the TCA cycle, and provided computational and experimental evidence for the absence of a functional glyoxlyate shunt.[10] These computational studies have recently been validated and the important role of GABA shunt closing the TCA cycle in PCC6803 has been demonstrated, despite the presence of the alternative TCA shortcut as well.[64 92, 93] By studying the photosynthetic processes in PCC6803 under varying light and carbon conditions and analyzing network robustness under genetic perturbations, it was concluded that high photosynthetic robustness, including multiple alternate electron flow pathways, is required for optimal photosynthetic performance and that this comes at the expense of reduced metabolic robustness.[12] Building on these initial systems-level analyses of photosynthetic networks, future studies employing metabolic models are likely to increase our current understanding of cyanobacterial metabolism and facilitate metabolic engineering efforts (Fig. 3).

Models as platforms for omics data integration and contextualization

In recent years the biological sciences have been hit by a data avalanche in the form of “omics” data sets. Transcriptomics proteomics, fluxomics and metabolomics

data are now routinely collected during biological experiments. Metabolic reconstructions provide a very useful context for omics data sets since they enable mechanistic interpretation of the data.[83, 94] The resulting condition-specific models can be used to prioritize hypothesis for experimental validation, they may lend support to observations that are otherwise difficult to validate experimentally, lead to biological discovery and more accurate metabolic engineering designs.[83]

Several condition-specific models already exist for cyanobacteria, and they have been used to increase our understanding of cyanobacterial physiology.[10, 13, 88, 90] Protein expression data over light and dark phases was used to model the diurnal rhythm of ATCC51142 by Saha et al.[88] Vu et al. applied mRNA and protein expression datasets to construct light and ammonium limited models of ATCC51142. The condition-specific models not only reduced the prediction uncertainty, but they also guided the discovery of proline as an alternative nitrogen source. [13] By combining metabolic and transcriptomic networks, Montagud et al. identified the first steps of pyrimidine synthesis and oxidative phosphorylation as regulatory hubs for transcriptional changes in light availability in PCC6803.[95] Knoop et al. modeled the diurnal cycle using measurements of transcriptional expression in response to light variability. This approach allowed the dynamic simulation and estimation of carbon flux in PCC6803 as a function of light availability during the day.[10] The potential of these condition specific models in biotechnology is largely untapped. A large number of transcriptome data sets collected for PCC6803 has recently been compiled [96] and awaits further study in the context of metabolic networks. The same holds true for a recently published web-based database for interactive exploration and visualization of transcriptomic data in PCC6803. [70]

Models as a test bed in biotechnological applications

Several algorithms for metabolic engineering based on network reconstructions are available, including algorithms that couple the secretion of the target compound to growth by gene knockouts.[97, 98] Growth-coupling is a highly desirable trait since it alleviates the problems of selection and genetic instability and enables the use of adaptive evolution to further increase the production rate.[99] Algorithms such as OptForce[100] can be used to identify which fluxes must increase or decrease to achieve a pre-specified overproduction target. In addition, algorithms for designing synthetic pathways can be used to optimize the production of native and non-native compounds and to devise strategies for the removal of toxic byproducts.[101] Such algorithms have

been instrumental in the success of several recent metabolic engineering projects including the overproduction of 1,4-butanediol[74] and L-valine[82] in *E. coli*.

In-silico studies of the capabilities of cyanobacteria for chemical production have mostly focused on the overproduction of hydrogen[81, 88] and ethanol.[86-88]. Other computational studies have demonstrated some of the significant challenges involved in the use of cyanobacteria as biocatalysts and have identified key bottlenecks. A search for growth-coupled knockout strains in PCC6803 showed that carbon flux re-routing was considerably more difficult under autotrophic conditions than under mixotrophic and heterotrophic conditions.[102] Furthermore, for growth-coupling under mixotrophic conditions it was necessary to reduce the photosynthetic robustness by blocking the light-driven metabolism but in this case net CO_2 fixation was absent for most of the strains and the mixotrophic metabolism resembled that of heterotrophs. [102] Vu et al. studied the capabilities of PCC7002 for producing several native and non-native compounds, including succinate, alanine, isoprene, butanol and ethanol.[90] Computational experiments showed that single deletions in the central metabolism were predicted to improve the production of the target chemicals but the production was not coupled to growth. The computational search for growth coupled mutants found that a large number of knockouts were needed under autotrophic conditions in both the strains, most strategies required 9 to 10 deletions.[12, 90] High-quality models of cyanobacteria have only recently become available[103] and their uses in metabolic engineering appear to be limited to computational analysis. To the best of our knowledge, no model-driven experimental attempts to overproduce chemicals have been undertaken so far, however the implementation of these computational approaches is now possible.

Outlook

Significant advances have been made in cyanobacterial biotechnology in the last few years. The field has matured rapidly and it is now possible to use cyanobacteria for sustainable production of biofuels and fine chemicals. However, the current approaches are still far from being economically feasible and are unlikely to replace crude oil-based processes in the near future. Many challenges remain and multiple steps need to be optimized before phototroph-based biotechnology becomes competitive, such as photosynthetic efficiency and photobioreactor design. [43, 59, 104] The optimization of the cyanobacterial metabolism has barely begun. Increasing our

knowledge about the metabolism in cyanobacteria is likely to lead to significant improvements in strain design in the same way as happened with *E. coli* and yeast. Existing metabolic reconstructions are extremely useful tools for these purposes. To obtain further insights into biotechnologically relevant processes, the next generation of models need to include additional modules, such as reactive oxygen species and a more comprehensive diurnal cycle description. The effects of varying light wavelengths can be modeled in a fairly straightforward manner[105] and light-quality could then be included as an additional environmental factor in future model-driven biotechnology efforts. These systems biology efforts, combined with modern synthetic biology approaches may lead to the long awaited economic feasibility of cyanobacterial cell factories (BOX 2).

Acknowledgments

This work was supported by the SYNPOL (FP7-KBBE 311815; http://www.synpol.org/) EU project. The authors would like to thank the anonymous reviewers for their critical comments and valuable suggestions which improved the quality of the paper considerably.

BOX 1. Foundational problems in cyanobacterial biotechnology, the approaches that are currently used to address them and potential approaches based on systems biology.

Problem	Current approach	Systems biology approach
Lack of industrially relevant pathways	Heterologous expression from heterotrophic organisms[22, 104]	Model-driven design of pathways optimized for cyanobacteria, e.g. using OptStrain[106]
Low NADH/NADPH ratio	Co-factors optimization using engineered biosynthetic pathways[25, 107]	Model based optimization of co-factor swapping[108]
Toxicity of target metabolite	Continuous removal of the metabolite[22]	Systematically increase tolerance, e.g. by, overexpression of toxicity induced genes, heterologous expression of solvent pumps.
Toxicity of byproducts	Not addressed	Model-based use of pathway design algorithms to obtain strategies to convert byproducts into non-toxic intermediates[100, 106]
Oxygen sensitivity of target biosynthetic pathways	Local replacement of oxygen sensitive steps. [26]	Design of new-to-nature oxygen tolerant pathway guided by computational algorithms such as BNICE [109] and PathPred.[110]
Low carbon flux through target biosynthetic pathways	Removal of competing biosynthetic pathways[23, 40]	Systems metabolic engineering of pathways to increase flux.[100, 106] Systematic removal of competing pathways.[97, 98]
Lack of physiological and metabolic knowledge	Employ molecular biology approaches to increase biological knowledge.[32]	Systems understanding of cyanobacterial physiology and metabolism through model driven analysis and integration of omics data[12, 13]
Low genetic stability	Optimization of codon use and expression[52]	Model-driven growth-coupled overproducer strains. [97, 98] Adaptive laboratory evolution efforts [111]

BOX 2. Genome-scale metabolic reconstructions.

Systems biology attempts to obtain a detailed understanding of biological processes by a bottom-up approach where biochemical information about sub-cellular processes are integrated to form computational models of cellular activity at higher levels, culminating in comprehensive models of single cell activity or even groups of cells. The computational nature of the models enables quantitative predictions of cellular behavior under different conditions. Models of metabolism in many prokaryotic and eukaryotic organisms have been constructed[112] and models of transcription, regulation and signaling networks have also been developed, although to a lesser extent.[113] Multi-scale models combining different types of networks, e.g. metabolic and transcription/translation networks[114] are starting to become available and are expected to increase the predictive accuracy even further.

Genome-scale metabolic models are constructed from genetic, genomic and biochemical data obtained from online databases and primary literature, following a standardized protocol.[115] The models enable quantitative predictions in terms of *fluxes* through individual reactions. The predictions are frequently made using *flux balance analysis*, a computational algorithm which calculates flux values in all the reactions corresponding to a particular cellular objective such as the maximization of biomass, production of ATP or synthesis of the target compound.[99] The effects of heterologous gene insertion are easily simulated by adding the corresponding reaction(s) to the model. An example is the production of lactate, a non-native compound in *Synechocystis*. In this case, reactions describing the synthesis of lactate from pyruvate (via lactate dehydrogenase) and the transport of lactate out of the cell would be added. A gene knockout is simulated by simply removing the corresponding reaction(s) from the model.

Figure 1. Maximum theoretical yields and the chemical space covered by cyanobacteria. a. The chemical space covered by cyanobacteria and *E. coli*. The figure represents chemical similarities between compounds in such a way that two points that are close together have a similar chemical structure while compounds that are separated by a large distance are dissimilar. The cyanobacterial metabolites (gray points) and metabolic engineering targets in cyanobacteria (blue and green points) and *E. coli* (blue and orange points) reveal that the productive potential of cyanobacteria is largely unexplored compared to *E. coli*. The cyanobacteria targets fall mostly into the top left quadrant whereas the *E. coli* targets cover a considerably larger area. The axis units are arbitrary. *Figure details*: The metabolites for the cyanobacteria were derived from metabolic reconstructions of *Synechocystis* sp. PCC6803,[12] *Cyanothece* sp. ATCC 51142[13] and *Spirulina* platensis C1[14] after removing unstable intermediates. They correspond to carbon containing compounds with KEGG IDs and matching IUPAC International Chemical Identifiers. Chemical similarities between the compounds are represented by Tanimoto coefficients derived from FP2 path-based fingerprints of the compounds obtained with the obabel program.[116] The resulting 364 by 364 similarity matrix was visualized using the t-SNE algorithm.[117] **b**. Maximum theoretical yields of selected compounds in *Synechocystis* under autotrophic conditions (green), heterotrophic conditions (red) and *E. coli* growing on glucose (orange). The yields are defined as the ratio of the number of carbon atoms converted to the target product versus the number of carbon atoms consumed. The low yields of ethylene are due to the generation of guanidine, a byproduct that does not appear to be metabolized. *Figure details*: The reconstructed networks of *Synechocystis,* iJN678[12] and *E. coli*, iJO1366[118] were used to calculate the theoretical yields after adding the necessary pathways to the models and fixing the biomass to 20% of the maximum. Light uptake under autotrophic conditions corresponded to the amount required for maximal growth in order to avoid unrealistic energy production due to extra light uptake. *Abbreviations*: 1,2-propanediol (12ppd), p-coumaric acid (4ca) (R)-3-hydroxybutyrate (3hba), isobutanol (iBuOH), isobutyraldehyde (iBAL).

Figure 2. The unexplored potential of cyanobacteria as biocatalysts. Living organisms have evolved by prioritizing growth. Under nutrient-rich conditions this leads to maximum biomass production (**a**). Increasing the production of the target compound under these conditions is only possible by a corresponding decrease in growth since the two are conflicting biological objectives. Biotechnological efforts focus on movement towards maximum metabolite production (**b**) from low yielding strains (**c**), preferably along the Pareto frontier (**red line**). Typical metabolic optimized *E. coli* production strains combine limited growth with high levels of the target compound (**d**). Theoretical estimates suggest that a large fraction of the “production space” remains unexplored in cyanobacteria and is amenable to optimization (**gray area**) with the best designs lying on the Pareto frontier. Optimization of bioreactor design and photosynthetic efficiency expands the production space outwards (**green area**). Combined with metabolic optimization the production can be improved significantly (**e**) in the absence of metabolic optimization only a limited increase is possible (**f**).

Figure 3. Workflow for systems metabolic engineering. The **metabolic network reconstruction** process (BOX 2) frequently leads to increased biological knowledge of the target host and the identification of knowledge gaps, some of which may require

experimental studies to resolve. The network is then used to **analyze** the metabolic capabilities of the target organism from its genotype. This includes exploration of the feasible metabolic space under various environmental and genetic conditions, emergent properties of the network and metabolic bottlenecks. Once the metabolic capabilities of the host have been well defined, the model is used for **systems metabolic engineering** which involves the use of computational algorithms to design synthetic pathways, identify enzyme targets for up- and down-regulation and to block competing pathways. This step also involves the creation of context-specific models using available omics data. Finally a sophisticated **synthetic biology approach** is needed to implement *in vivo* the model-driven designs, as well as the expression of transporter, and synthetic regulatory networks.

References

1. J. M. Melillo, *et al*., *Science*, 2009, **326**, 1397-1399.
2. M. A. Delucchi, *Ann N Y Acad Sci,* 2010, **1195**, 28-45.
3. E. Molina Grima, E. H. Belarbi, F. G. Acién Fernández, A. Robles Medina and Y. Chisti, *Biotechnol Adv*, 2003, **20**, 491-515.
4. N. W. Lem and B. R. Glick, *Biotechnol Adv*, 1985, **3**, 195-208.
5. R. M. M. Abed, S. Dobretsov and K. Sudesh, *J Appl Microbiol*, 2009, **106**, 1-12.
6. R. Y. Stanier and G. C. Bazine, *Annu Rev Microbiol*, 1977, **31**, 225-274.
7. N. Quintana, F. Kooy, M. Rhee, G. Voshol and R. Verpoorte, *Appl Microbiol Biotechnol*, 2011, **91**, 471-490.
8. G. C. Dismukes, D. Carrieri, N. Bennette, G. M. Ananyev and M. C. Posewitz, *Current Opinion in Biotechnology*, 2008, **19**, 235-240.
9. X.-G. Zhu, S. P. Long and D. R. Ort, *Curr Opin Biotechnol*, 2008, **19**, 153-159.
10. H. Knoop, , *et al*., *PLoS Comput Biol*, 2013, **9**, e1003081.
11. T. Mueller, B. Berla, H. Pakrasi and C. Maranas, *BMC Syst Biol*, 2013, **7**, 142.
12. J. Nogales, S. Gudmundsson, E. M. Knight, B. O. Palsson and I. Thiele, *Proc Natl Acad Sci U S A*, 2012, **109**, 2678-2683.
13. T. T. Vu, *et al*., *PLoS Comput Biol*, 2012, **8**, e1002460.
14. A. Klanchui, C. Khannapho, A. Phodee, S. Cheevadhanarak and A. Meechai, *BMC Syst Biol*, 2012, **6**, 71.
15. T. Heidorn, *et al*., in *Methods Enzymol*, ed. V. Chris, Academic Press, 2011, vol. Volume 497, pp. 539-579.
16. M. Tamoi, T. Miyazaki, T. Fukamizo and S. Shigeoka, *Plant J*, 2005, **42**, 504-513.
17. H. Takahashi, H. Uchimiya and Y. Hihara, *J Exp Bot*, 2008, **59**, 3009-3018.
18. J. W. Cooley and W. F. J. Vermaas, *J Bacteriol*, 2001, **183**, 4251-4258.
19. M.-D. Deng and J. R. Coleman, *Appl Environ Microbiol*, 1999, **65**, 523-528.
20. J. Dexter and P. Fu, *Energy Environ Sci*, 2009, **2**, 857-864.
21. Z. Gao, H. Zhao, Z. Li, X. Tan and X. Lu, *Energy Environ Sci*, 2012, **5**, 9857-9865.
22. S. Atsumi, W. Higashide and J. C. Liao, *Nat Biotech*, 2009, **27**, 1177-1180.
23. X. Li, C. Shen and J. Liao, *Photosynth Res*, 2014, 1-10.
24. E. I. Lan and J. C. Liao, *Metab Eng*, 2011, **13**, 353-363.
25. E. I. Lan and J. C. Liao, *Proc Natl Acad Sci U S A*, 2012, **109**, 6018-6023.
26. E. I. Lan, S. Y. Ro and J. C. Liao, *Energy Environ Sci*, 2013, **6**, 2672-2681.
27. J. W. K. Oliver, I. M. P. Machado, H. Yoneda and S. Atsumi, *Proc Natl Acad Sci U S A*, 2013, **110**, 1249-54
28. Q. Hu, *et al*., *Plant J*, 2008, **54**, 621-639.
29. X. Liu, J. Sheng and R. Curtiss III, *Proc Natl Acad Sci U S A*, 2011, **108**,6899-904.
30. X. Tan, *et al*., *Metab Eng*, 2011, **13**, 169-176.
31. F. Qi, L. Yao, X. Tan and X. Lu, *Biotechnol Lett*, 2013, **35**, 1655-1661.
32. A. Schirmer, M. A. Rude, X. Li, E. Popova and S. B. del Cardayre, *Science*, 2010, **329**, 559-562.
33. W. Wang, X. Liu and X. Lu, *Biotechnol Biofuels*, 2013, **6**, 69.
34. B. K. Kaiser, *et al*., *PLoS ONE*, 2013, **8**, e58307.
35. A. Ruffing, *Biotechnol Biofuels*, 2013, **6**, 113.
36. M. Hagemann and K. Marin, *J Plant Physiol*, 1999, **155**, 424-430.
37. W. Du, F. Liang, Y. Duan, X. Tan and X. Lu, *Metab Eng*, 2013, **19**, 17-25.
38. H. Niederholtmeyer, B. T. Wolfstädter, D. F. Savage, P. A. Silver and J. C. Way, *Appl Environ Microbiol,* 2010, **76**, 3462-3466.

39. D. C. Ducat, J. A. Avelar-Rivas, J. C. Way and P. A. Silver, *Appl Environ Microbiol*, 2012, **78**, 2660-2668.
40. J. H. Jacobsen and N.-U. Frigaard, *Metab Eng*, 2014, **21**, 60-70.
41. J. C. F. Ortiz-Marquez, M. Do Nascimento, J. P. Zehr and L. Curatti, *Trends Biotechnol*, 2013, **31**, 521-529.
42. C. M. Agapakis, *et al*., *PLoS ONE*, 2011, **6**, e18877.
43. D. C. Ducat, J. C. Way and P. A. Silver, *Trends Biotechnol*, 2011, **29**, 95-103.
44. Y. Asada, M. Miyake, J. Miyake, R. Kurane and Y. Tokiwa, *Int J Biol Macromol,* 1999, **25**, 37-42.
45. K. E. J. Tyo, Y.-S. Jin, F. A. Espinoza and G. Stephanopoulos, *Biotechnol Prog*, 2009, **25**, 1236-1243.
46. T. Osanai, *et al*., *DNA Research*, 2013, **20**, 525-535.
47. B. Wang, S. Pugh, D. R. Nielsen, W. Zhang and D. R. Meldrum, *Metab Eng*, 2013, **16**, 68-77.
48. P. Lindberg, S. Park and A. Melis, *Metab Eng*, 2010, **12**, 70-79.
49. M. Sakai, T. Ogawa, M. Matsuoka and H. Fukuda, *J Ferment Bioeng*, 1997, **84**, 434-443.
50. K. Takahama, M. Matsuoka, K. Nagahama and T. Ogawa, *J Biosci Bioeng*, 2003, **95**, 302-305.
51. F. Guerrero, V. Carbonell, M. Cossu, D. Correddu and P. R. Jones, *PLoS ONE*, 2012, **7**, e50470.
52. J. Ungerer, L. Tao, M. Davis, M. Ghirardi, P.-C. Maness and J. Yu, *Energy Environ Sci*, 2012, **5**, 8998-9006.
53. S. A. Angermayr, M. Paszota and K. J. Hellingwerf, *Appl Environ Microbiol*, 2012, **78**, 7098-7106.
54. S. A. Angermayr and K. J. Hellingwerf, *J Phys Chem B*, 2013, **117**, 11169-11175.
55. A. Varman, Y. Yu, L. You and Y. Tang, *Microb Cell Fact*, 2013, **12**, 1-8.
56. R. Yu, *et al*., *Lipids*, 2000, **35**, 1061-1064.
57. R. E. Reinsvold, R. E. Jinkerson, R. Radakovits, M. C. Posewitz and C. Basu, *J Plant Physiol*, 2011, **168**, 848-852.
58. E. Englund, *et al*., *PLoS ONE*, 2014, **9**, e90270.
59. V. H. Work, S. D'Adamo, R. Radakovits, R. E. Jinkerson and M. C. Posewitz, *Curr Opin Biotechnol,* 2012, **23**, 290-297.
60. N. E. Nozzi, J. W. K. Oliver and S. Atsumi, *Front Bioeng Biotechnol*, 2013, **1**.
61. J. K. Oliver and S. Atsumi, *Photosynth Res*, 2014, **120**, 249-261.
62. K. J. Hellingwerf and M. J. Teixeira de Mattos, *J Biotechnol*, 2009, **142**, 87-90.
63. L. Rosgaard, A. J. de Porcellinis, J. H. Jacobsen, N.-U. Frigaard and Y. Sakuragi, *J Biotechnol*, 2012, **162**, 134-147.
64. S. Zhang and D. A. Bryant, *Science*, 2011, **334**, 1551-1553.
65. I. M. Keseler, *et al*., *Nucleic Acids Res*, 2013, **41**, D605-D612.
66. *E. coli* Gene Expression Database (GenExpDB). <http://genexpdb.ou.edu/main/>.
67. H. Salgado, *et al*., *Nucleic Acids Res*, 2013, **41**, D203-D213.
68. T. Baba, *et al*., *Mol Syst Biol. 2006, **2**:2006.0008*.
69. M. Kitagawa, *et al*., *DNA Research*, 2006, **12**, 291-299.
70. M. A. Hernandez-Prieto and M. E. Futschik, *Bioinformation*, 2012, **8**, 634-638.
71. M. A. Hernández-Prieto, T. A. Semeniuk and M. E. Futschik, *Front Genet*, 2014, **5**.
72. A. Taton, *et al*., *Nucleic Acids Res*, 2014.
73. B. M. Berla, R. Saha, C. M. Immethun, C. D. Maranas, T. S. Moon and H. Pakrasi, *Front Microbiol*, 2013, **4**.
74. H. Yim, *et al*., *Nature chemical biology*, 2011, **7**, 445-452.
75. Y. Xue, Y. Zhang, D. Cheng, S. Daddy and Q. He, *Proc Natl Acad Sci U S A*, 2014, **111**, 9449-9454.

76. I. W. Bogorad, T.-S. Lin and J. C. Liao, *Nature*, 2013, **502**, 693-697.
77. T. R. Maarleveld, J. Boele, F. J. Bruggeman and B. Teusink, *Plant Physiol*, 2014, **164**, 1111-1121.
78. J. D. Young, A. A. Shastri, G. Stephanopoulos and J. A. Morgan, *Metab Eng*, 2011, **13**, 656-665.
79. A. Melis, *Current Opinion in Chemical Biology*, 2013, **17**, 453-456.
80. D. M. Kramer and J. R. Evans, *Plant Physiol*, 2011, **155**, 70-78.
81. A. Montagud, E. Navarro, P. Fernandez de Cordoba, J. F. Urchueguia and K. R. Patil, *BMC Syst Biol*, 2010, **4**, 156.
82. J. H. Park, K. H. Lee, T. Y. Kim and S. Y. Lee, *Proc Natl Acad Sci U S A*, 2007, **104**, 7797-7802.
83. D. R. Hyduke, N. E. Lewis and B. O. Palsson, *Mol Biosyst*, 2013, **9**, 167-174.
84. D. McCloskey, B. O. Palsson and A. M. Feist, *Mol Syst Biol*, 2013, **9**.
85. A. A. Shastri and J. A. Morgan, *Biotechnology Prog*, 2005, **21**, 1617-1626.
86. K. Yoshikawa, *et al*., *Appl Microbiol Biotechnol*, 2011, **92**, 347-358.
87. P. Fu, *J Chem Technol Biotechnol*, 2009, **84**, 473-483.
88. R. Saha, *et al*., *PLoS ONE*, 2012, **7**, e48285.
89. H. Knoop, Y. Zilliges, W. Lockau and R. Steuer, *Plant Physiol*, 2010, **154**, 410-422.
90. T. T. Vu, *et al*., *Biotechnol J*, 2013, **8**, 619-630.
91. E. Vitkin and T. Shlomi, *Genome Biol*, 2012, **13**, R111.
92. L. You, B. Berla, L. He, H. B. Pakrasi and Y. J. Tang, *Biotechnol J*, 2014, **9**, 684-692.
93. W. Xiong, D. Brune and W. F. J. Vermaas, *Mol Microbiol*, 2014, **93**, 786-796.
94. R. Saha, A. Chowdhury and C. D. Maranas, *Curr Opin Biotechnol*, 2014, **29C**, 39-45.
95. A. Montagud, *et al*., 2011, **6**, 330-342.
96. A. K. Singh, *et al*., *BMC Syst Biol*, 2010, **4**, 105.
97. A. P. Burgard, P. Pharkya and C. D. Maranas, *Biotechnol Bioeng*, 2003, **84**, 647-657.
98. C. Cotten and J. L. Reed, *Biotechnol J*, 2013, **8**, 595-604.
99. A. M. Feist, *et al*., *Metab Eng*, 2010, **12**, 173-186.
100. S. Ranganathan, P. F. Suthers and C. D. Maranas, *PLoS Comput Biol*, 2010, **6**, e1000744.
101. M. H. Medema, R. van Raaphorst, E. Takano and R. Breitling, *Nat Rev Micro*, 2012, **10**, 191-202.
102. J. Nogales, S. Gudmundsson and I. Thiele, *Bioengineered*, 2013, **4**, 158-163.
103. R. Steuer, H. Knoop and R. Machne, *J Exp Bot*, 2012, **63**, 2259-2274.
104. R. H. Wijffels, O. Kruse and K. J. Hellingwerf, *Curr Opin Biotechnol,* 2013, **24**, 405-413.
105. R. L. Chang, *et al*., *Mol Syst Biol*, 2011, **7**, 518.
106. P. Pharkya, A. P. Burgard and C. D. Maranas, *Genome res,* 2004, **14**, 2367-2376.
107. P. E. Savakis, S. A. Angermayr and K. J. Hellingwerf, *Metab Eng*, 2013, **20**, 121-130.
108. Z. A. King and A. M. Feist, *Metab Eng*, 2014.
109. S. D. Finley, L. J. Broadbelt and V. Hatzimanikatis, *Biotechnol Bioeng*, 2009, **104**, 1086-1097.
110. Y. Moriya, *et al*., *Nucleic Acids Res*, 2010, **38**, W138-W143.
111. T. M. L. Conrad, Nathan E. Palsson, Bernhard Ø, *Mol Syst Biol*, 2011, **7**, 509.
112. J. Monk, J. Nogales and B. O. Palsson, *Nat Biotech*, 2014, **32**, 447-452.
113. E. Goncalves, *et al*., *Mol Biosyst*, 2013, **9**, 1576-1583.
114. J. A. Lerman, *et al*., *Nat Commun*, 2012, **3**, 929.
115. I. Thiele and B. O. Palsson, *Nature protocols*, 2010, **5**, 93-121.
116. N. M. O'Boyle, *et al*., *J Cheminform*, 2011, **3**, 33.
117. L. J. P. v. d. Maaten and G. E. Hinton, *J Mach Learn Res*, 2008, **9**, 2579-2605.
118. J. D. Orth, *et al*., *Mol Syst Biol*, 2011, **7**, 535.

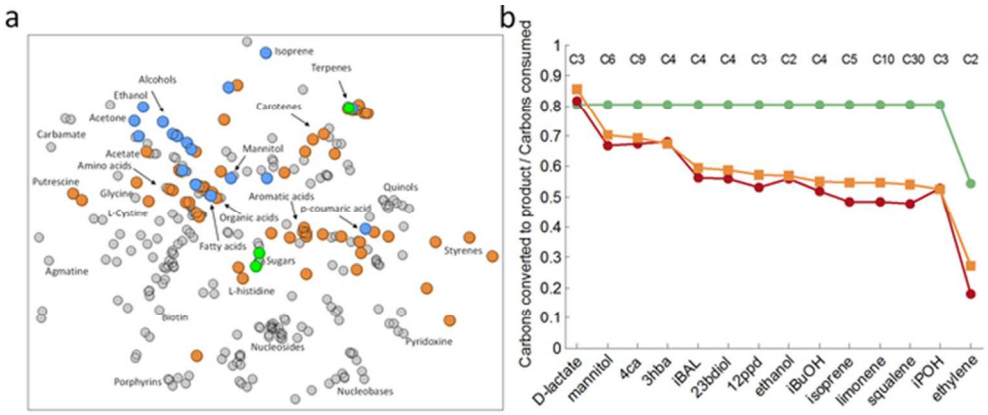


Figure 1. Maximum theoretical yields and the chemical space covered by cyanobacteria. a. The chemical space covered by cyanobacteria and E. coli. The figure represents chemical similarities between compounds in such a way that two points that are close together have a similar chemical structure while compounds that are separated by a large distance are dissimilar. The cyanobacterial metabolites (gray points) and metabolic engineering targets in cyanobacteria (blue and green points) and E. coli (blue and orange points) reveal that the productive potential of cyanobacteria is largely unexplored compared to E. coli. The cyanobacteria targets fall mostly into the top left quadrant whereas the E. coli targets cover a considerably larger area. The axis units are arbitrary. Figure details: The metabolites for the cyanobacteria were derived from metabolic reconstructions of Synechocystis sp. PCC6803,12 Cyanothece sp. ATCC 5114213 and Spirulina platensis C114 after removing unstable intermediates. They correspond to carbon containing compounds with KEGG IDs and matching IUPAC International Chemical Identifiers. Chemical similarities between the compounds are represented by Tanimoto coefficients derived from FP2 path-based fingerprints of the compounds obtained with the obabel program.116 The resulting 364 by 364 similarity matrix was visualized using the t-SNE algorithm.117 b. Maximum theoretical yields of selected compounds in Synechocystis under autotrophic conditions (green), heterotrophic conditions (red) and E. coli growing on glucose (orange). The yields are defined as the ratio of the number of carbon atoms converted to the target product versus the number of carbon atoms consumed. The low yields of ethylene are due to the generation of guanidine, a byproduct that does not appear to be metabolized. Figure details: The reconstructed networks of Synechocystis, iJN67812 and E. coli, iJO1366118 were used to calculate the theoretical yields after adding the necessary pathways to the models and fixing the biomass to 20% of the maximum. Light uptake under autotrophic conditions corresponded to the amount required for maximal growth in order to avoid unrealistic energy production due to extra light uptake. Abbreviations: 1,2-propanediol (12ppd), p-coumaric acid (4ca) (R)-3-hydroxybutyrate (3hba), isobutanol (iBuOH), isobutyraldehyde (iBAL).
57x23mm (300 x 300 DPI)

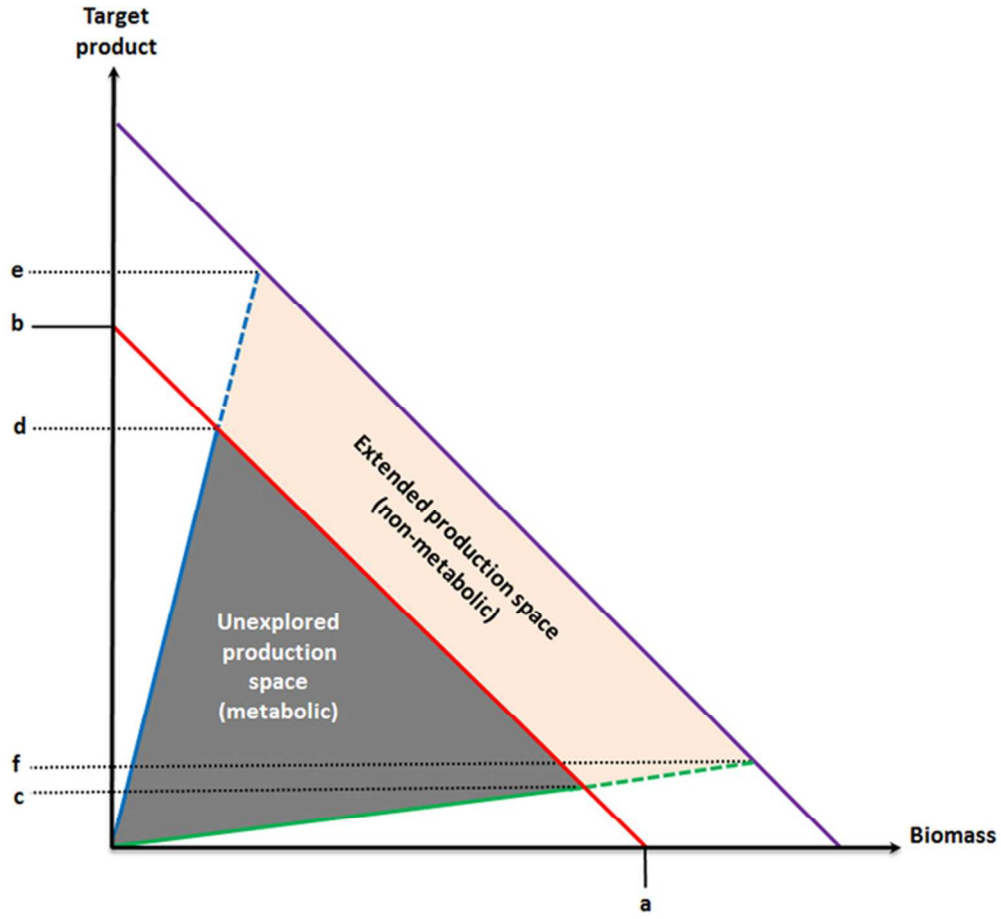


Figure 2. The unexplored potential of cyanobacteria as biocatalysts. Living organisms have evolved by prioritizing growth. Under nutrient-rich conditions this leads to maximum biomass production (a). Increasing the production of the target compound under these conditions is only possible by a corresponding decrease in growth since the two are conflicting biological objectives. Biotechnological efforts focus on movement towards maximum metabolite production (b) from low yielding strains (c), preferably along the Pareto frontier (red line). Typical metabolic optimized E. coli production strains combine limited growth with high levels of the target compound (d). Theoretical estimates suggest that a large fraction of the "production space" remains unexplored in cyanobacteria and is amenable to optimization (gray area) with the best designs lying on the Pareto frontier. Optimization of bioreactor design and photosynthetic efficiency expands the production space outwards (green area). Combined with metabolic optimization the production can be improved significantly (e) in the absence of metabolic optimization only a limited increase is possible (f).

59x55mm (300 x 300 DPI)

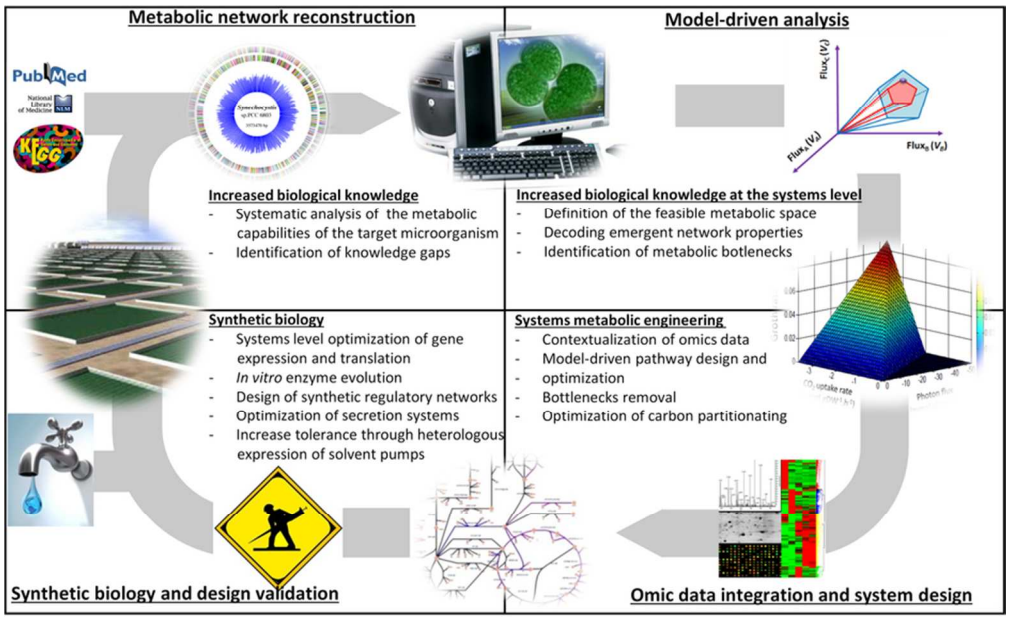


Figure 3. Workflow for systems metabolic engineering. The metabolic network reconstruction process (BOX 2) frequently leads to increased biological knowledge of the target host and the identification of knowledge gaps, some of which may require experimental studies to resolve. The network is then used to analyze the metabolic capabilities of the target organism from its genotype. This includes exploration of the feasible metabolic space under various environmental and genetic conditions, emergent properties of the network and metabolic bottlenecks. Once the metabolic capabilities of the host have been well defined, the model is used for systems metabolic engineering which involves the use of computational algorithms to design synthetic pathways, identify enzyme targets for up- and down-regulation and to block competing pathways. This step also involves the creation of context-specific models using available omics data. Finally a sophisticated synthetic biology approach is needed to implement in vivo the model-driven designs, as well as the expression of transporter, and synthetic regulatory networks.

79x48mm (300 x 300 DPI)