Multicomponent engineering of cyanobacteria towards a photoproduction platform

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Maximilian Dietsch
from Wuppertal

Münster, December 2022

from the institute for Synthetic Microbiology
at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Ilka Maria Axmann
Co-supervisor: Prof. Dr. Michael Feldbrügge

Date of the oral examination:

“It’s the job that’s never started as takes longest to finish.”

J.R.R. Tolkien, The Lord of the Rings

Zusammenfassung

Cyanobakterien sind vielversprechende Kandidaten für die biotechnologische Umwandlung von CO_2 in Plattformchemikalien sowie hochwertige Produkte. Neben ihrer photosynthetischen Lebensweise, die maßgeblich zur Fixierung von atmosphärischem Kohlenstoff beiträgt, sind sie morphologisch vielfältig, bevölkern viele Nischen auf der Erde und haben ein enormes Potenzial als alternative mikrobielle *Chassis*-Organismen. Ihr umfangreiches Membransystem und die Verwandtschaft zu pflanzlichen Chloroplasten macht sie besonders geeignet für die Produktion von Terpenoiden, einer großen Klasse sekundärer Metabolite. Die Biosynthese von Terpenoiden hat in letzter Zeit aufgrund ihres attraktiven Potenzials als Inhaltsstoff in Lebensmitteln, Medikamenten und Kraftstoffen mehr Aufmerksamkeit auf sich gezogen. Allerdings sind die Ausbeuten in heterologen mikrobiellen Produktionssystemen oft gering. Um die Produktivität des mikrobiellen *Chassis* zu steigern, gibt es viele *metabolic engineering* Strategien, die angewendet werden können. Dazu gehören *in silico*-Ansätze wie *flux balance* Analysen, Überexpression von relevanten Genen oder gezieltes Entfernen unerwünschter Nebenreaktionen, aber auch Strategien, die auf eine globalere Veränderung des Organismus abzielen. In meiner Dissertation beschreibe verschiedene Ansätze, den einzelligen Modellorganismus *Synechocystis* sp. PCC 6803 zur Herstellung verschiedener Terpenoide zu modifizieren. Im ersten Ansatz wurde der metabolische Fluss durch CRISPR-vermitteltes Gen-*Silencing* vom Carotinoid-Biosyntheseweg zum Vorläufer FPP umverteilt. Das neu verfügbare FPP wurde dann durch Überexpression der Valencen-Synthase aus einer Kiefer, *Callitropsis nootkatensis*, in das heterologe Sesquiterpen Valencen umgewandelt. In einem zweiten Ansatz wurde die Akkumulation von Squalen, der universelle Vorläufer für Triterpene, optimiert. Dies erfolgte durch Auswertung der metabolischen Flussverteilung durch *flux balance* Analysen. Die identifizierten Ziele wurden systematisch überexprimiert, und ihre Wirksamkeit wurde durch Messung des Produktoutputs bewertet. Im letzten Teil wurde ein globalerer Ansatz gewählt. Die Schlüsselenzyme, Topoisomerasen, die auf die DNA-*Supercoiling*-Homöostase von Bakterien einwirken, wurden manipuliert, um das globale negative *Supercoiling* zu reduzieren und dadurch die *Supercoiling*-Homöostase zu verändern. Die Auswirkungen auf die Verteilung des ATP/ADP-Verhältnisses sowie des Energiespeichers Glykogen wurden ausführlich charakterisiert. Die gezielte Manipulation des DNA-*Supercoiling* ermöglichte es, Veränderungen in der Physiologie, den Expressionsmustern und der Morphologie herbeizuführen. Insgesamt stellt diese Arbeit einen wichtigen Baustein für die Entwicklung einer cyanobakteriellen Plattform für die Photoproduktion von Terpenoiden dar.

Abstract

Cyanobacteria are highly promising candidates for biotechnological conversion of CO_2 to platform chemicals, as well as high-value products. Next to their photosynthetic lifestyle, which contributes significantly to the fixation of atmospheric carbon, they are morphologically diverse, inhabit many niches on Earth and have a huge potential as alternative microbial chassis organisms. Their extensive membrane system and ancestral relation to plant chloroplasts makes them especially suitable for the production of terpenoids, a large class of secondary metabolites. The biosynthesis of terpenoids has drawn more attention recently due to its appealing potential as an ingredient in food, medicine, and fuel. However, yields are often low in heterologous microbial production systems. To increase the productivity of microbial chassis, there are many metabolic engineering strategies that can be applied. This includes *in silico* approaches such as flux distribution analyses, overexpression of genes of interest or targeted removal of undesired side reactions, but also strategies aiming towards a more global change in the organism. In my thesis, I describe my efforts to modify the unicellular model cyanobacterium *Synechocystis* sp. PCC 6803 to produce various terpenoids. In the first approach, metabolic flux was redistributed from the carotenoid biosynthesis pathway towards the precursor FPP by CRISPR-mediated gene silencing. The newly available FPP was then converted to the heterologous sesquiterpene valencene by overexpressing valencene synthase from a pine tree, *Callitropsis nootkatensis*. In a second approach, accumulation of squalene, the universal precursor for triterpenes, has been optimized. This was done by evaluating metabolic flux distribution through flux balance analysis. The identified targets were systematically overexpressed, and their effectiveness was evaluated by measuring product output. In the final part, a more global approach was chosen. Targeting the DNA supercoiling homeostasis of bacteria, the key enzymes, topoisomerases, were manipulated to reduce global negative supercoiling, thereby altering supercoiling homeostasis. The effects on the distribution of the ATP/ADP ratio, as well as the energy storage compound glycogen, were characterized extensively. The targeted manipulation of DNA supercoiling made it possible to induce changes in physiology, expression patterns, and morphology. Overall this work represents an important building block to the development of a cyanobacterial platform for photoproduction of terpenoids.

Table of Contents

1 Chapter 1 Introduction

1.1 Cyanobacterial potential for biotechnological application

Cyanobacteria are photosynthetic prokaryotes that use sunlight to transfer electrons from water to carbon dioxide. These reduced carbon molecules provide the basis for the formation of all organic molecules in the cell[1]. Morphologically, they comprise a very diverse but at the same time highly specialized group. They inhabit a vast number of different natural environments and the types of appearance range from unicellular, filamentous to colonial strains[2]. Since some of them can fix atmospheric nitrogen, they are often used as model organisms for studying nitrogen fixation[3]. At the same time, the focus of cyanobacterial research has shifted more and more toward the development of industrially usable platform organisms in the last decade. The photosynthetic way of life and the simultaneous morphological and physiological diversity make cyanobacteria a rich source of various naturally occurring molecules such as pigments and vitamins[4]. Since new synthetic biology tools for cyanobacteria have been developed over the past ten years, they are not only suitable as a source for a natural variety of potentially bioactive molecules but are also suitable for applied biotechnology[5]. Through targeted modification of the genome of cyanobacteria, the captured chemical energy of photosynthesis can be redirected, for example, into the production of low-structural molecules such as alcohols, sugars, and fatty acids, but also for the synthesis of highly complex structures such as plant secondary metabolites, e.g. terpenoids[6(p2)]. Terpenoids are derived from the five-carbon compound isoprene, and are the largest class of secondary plant metabolites, many of which have important biological activities[7]. Extracting them from plant material is challenging in that the secondary metabolites are only present in small amounts - therefore this method is inefficient. Using cyanobacteria instead of plants to produce valuable compounds, or even phytochemicals has several advantages[8]. The energy efficiency to generate solar energy in biomass is much lower in land plants compared to cyanobacteria. The efficiency of cyanobacteria is 3-9% whereas that of land plants is only <0.5% and at the same time no energy is used for the production of stems and roots[1]. This saves energy and waste in downstream processing. They are easier to genetically modify and can grow all year. Another advantage is that the cultivation of cyanobacteria does not compete with the production of food because they do not require arable land. Areas that are completely unsuitable for agriculture as a growing area could also potentially be suitable. Desert regions with high solar intensity or coastal regions in which some species can be cultivated with seawater are two examples[9]. In addition to the advantages of cultivation and efficiency of energy use, cyanobacteria are highly suitable for the synthesis of heterologous terpenoids due to their physiological properties stemming from the phototropic way of life. Specifically, the large intracellular membrane system required for photosynthesis is excellent for synthesizing these hydrophobic terpenes[10,11]. Many of the terpene synthases required for synthesis of e.g. triterpenes form multiple products from a single substrate because of multiple potential rearrangements of the carbon backbone. In addition, further modifications can be carried out by heme-containing cytochrome P450 monooxygenases (CP450), thus resulting in a further variety of terpenes[12]. These CP450 are directly dependent on the reduction of NADPH by a NADPH-P450 oxidoreductase and NADPH can be provided directly by photosynthesis in cyanobacteria[13]. At the same time, the intrinsic

isoprene metabolism responsible for native carotenoid biosynthesis can effectively provide precursor molecules for the synthesis of heterologousterpenoids[14]

1.2 Isoprenoid metabolism

There are two ways to produce the terpenoid precursor molecules isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). On the one hand there is the mevalonate (MVA) pathway used by eukaryotes, archaea and some bacteria[15]. On the other hand, there is the methylerythritol-4-phosphate (MEP) pathway used by most bacteria, including cyanobacteria, as well as plant plastids, who share a common ancestor with cyanobacteria. In higher plants, both the MEP and MVA pathways generate precursors for the biosynthesis of isoprenoids[16].

1.2.1 Terpenoid precursor synthesis through the mevalonate pathway

In the first step of the MVA biosynthetic pathway, two acetyl-CoA molecules are condensed by a thiolase to form acetoacetyl-CoA[17]. Thereafter, acetoacetyl-CoA and a third molecule of acetyl-CoA are converted to HMG-CoA by a β-hydroxy-β-methylglutaryl-CoA synthase (HMGS). HMG-CoA is then reduced to mevalonate by HMG-CoA reductase (HMGR)[18]. After two phosphorylation steps, mevalonate is converted to IPP by decarboxylation and dehydration. Isopentenyl pyrophosphate isomerase (IPI) then establishes the equilibrium between IPP and DMAPP since both are required for the later reaction steps[19]. Nevertheless, the MEP route has a higher utilization of carbon with an almost 50% higher stoichiometric carbon yield compared to the MVA route[20].

1.2.2 Terpenoid biosynthesis in *Synechocystis* 6803

The cyanobacterium *Synechocystis* sp. PCC 6803 (*Synechocystis* hereafter) has become the focus of many areas of research. This freshwater cyanobacterium was the first phototrophic organism to be completely sequenced. It is easy to genetically modify due to its ability to take up native exogenous DNA and incorporate it into the genome via homologous recombination. In addition, certain broad-host-range plasmids can be used in this organism[21]. In the MEP pathway of *Synechocystis* (Fig. 1), DMAPP and IPP are formed from pyruvate and glyceraldehyde 3-phosphate (G3P) via seven enzymatic steps[22]. In the first step, the rate-limiting enzyme 1-deoxy-D-xylulose-5-phosphate synthase (Dxs) catalyzes G3P and pyruvate to 1-deoxyxylulose-5-phosphate (DXP)[23]. DXP is converted into 2-methylerythritol-4-phosphate (MEP) by DXP-reductase (Dxr) under consumption of NADPH. By coupling the co-factor cytidine triphosphate (CTP) to MEP, 4-pyrophosphocytidyl-2-C-methyl-D-erythritol (CDP-ME) is formed , which is phosphorylated to 4-pyrophosphocytidyl-2-C-methyl-D-erythritol (CDP-MEP), then cyclized to 2-C-methyl-D-erythritol 2,4-cyclopyrophosphate (MEcPP) via the 2-C-methyl-D-erythritol 2,4-cyclopyrophosphate synthase (IspF) and reductively dehydrated to form 4-hydroxyl-3-methylbut-2-enylpyrophosphate (HMB-PP) via

the 4-hydroxy-3-methylbut-2-enyl pyrophosphate synthase (IspG). HMB-PP is further converted to IPP and DMAPP by HMB-PP reductase (IspH)[24]. The bacterial homolog of the enzyme IPI previously mentioned in the MVA pathway, which shares the same name, also maintains the stoichiometry between IPP and DMAPP, preventing over-accumulation or depletion of either. The 5-carbon (C5) molecules IPP and DMAPP serve as the universal precursor for all terpenoids. The first C10 geranyl pyrophosphate (GPP) is formed by combination of the two monomers by the enzyme geranylgeranyl pyrophosphate synthase (CrtE). CrtE is responsible for the further responsible for the enzymatic polymerization by chain elongation[25]. In this process, C15 farnesyl pyrophosphate (FPP) is formed from GPP by the addition of another IPP molecule. By addition of one more IPP molecule to FPP, C20 geranylgeranyl pyrophosphate (GGPP) is created. While there is a significant carbon flux towards this pathway in *Synechocystis*, the majority is directed towards growth and biomass formation and therefore is shuttled towards molecules relevant for molecules involved in photosynthesis and membrane biosynthesis[26]. FPP serves as a central precursor for hopanoid synthesis. First, squalene synthase (Sqs) catalyzes the linear precursor squalene from two FPP molecules via the intermediate presqualene pyrophosphate in an NADPH-dependent step[27]. Squaleneis then cyclized by squalene-hopene cyclases (shc) to form the pentacyclic hopanoid[28]. In *Synechocystis*, hopanoids modify plasma membrane properties, functionally analogous to steroids found in eukaryotes. GGPP, on the other hand, is the main branching point of carotenoids, the phytol side chain of chlorophyll, phylloquinone and the start of tocopherol synthesis. At the beginning of the biosynthesis pathway towards carotenoids, the colorless tetraterpene phytoene, which is formed from two units of GGPP, is condensed by the phytoene synthase (crtB). This is followed by sequential desaturation and isomerization to lycopene. Lycopene represents a branching point in the carotenoid biosynthesis pathway. By gradually introducing hydroxyl and keto groups, a large number of different carotenoids can be produced. In *Synechocystis*, β-carotene, myxoxanthophyll, zeaxanthin and echinenone are among the most common carotenoids[11]. Four tocopherols (α, β, δ, γ) are produced in *Synechocystis*. Although the precise function of each tocopherol in cells is not fully understood, it is known that they are involved in protection against lipid peroxidation, cold tolerance, and may optimize photosynthetic activity[29,30]. All tocopherols are synthesized from the precursor 6-methyl-6-phytyl-1,4-benzoquinol using phytyl pyrophosphate and homogentisate. Phytyl pyrophosphate, in turn, is synthesized by chlP via GGPP[31,32]. Phytyl pyrophosphate also serves as a precursor for the synthesis of the phytol tail of chlorophyll and phytyl pyrophosphate, together with chorismate, form the electron acceptor phylloquinone, both of which are integral to photosynthesis[33,34].

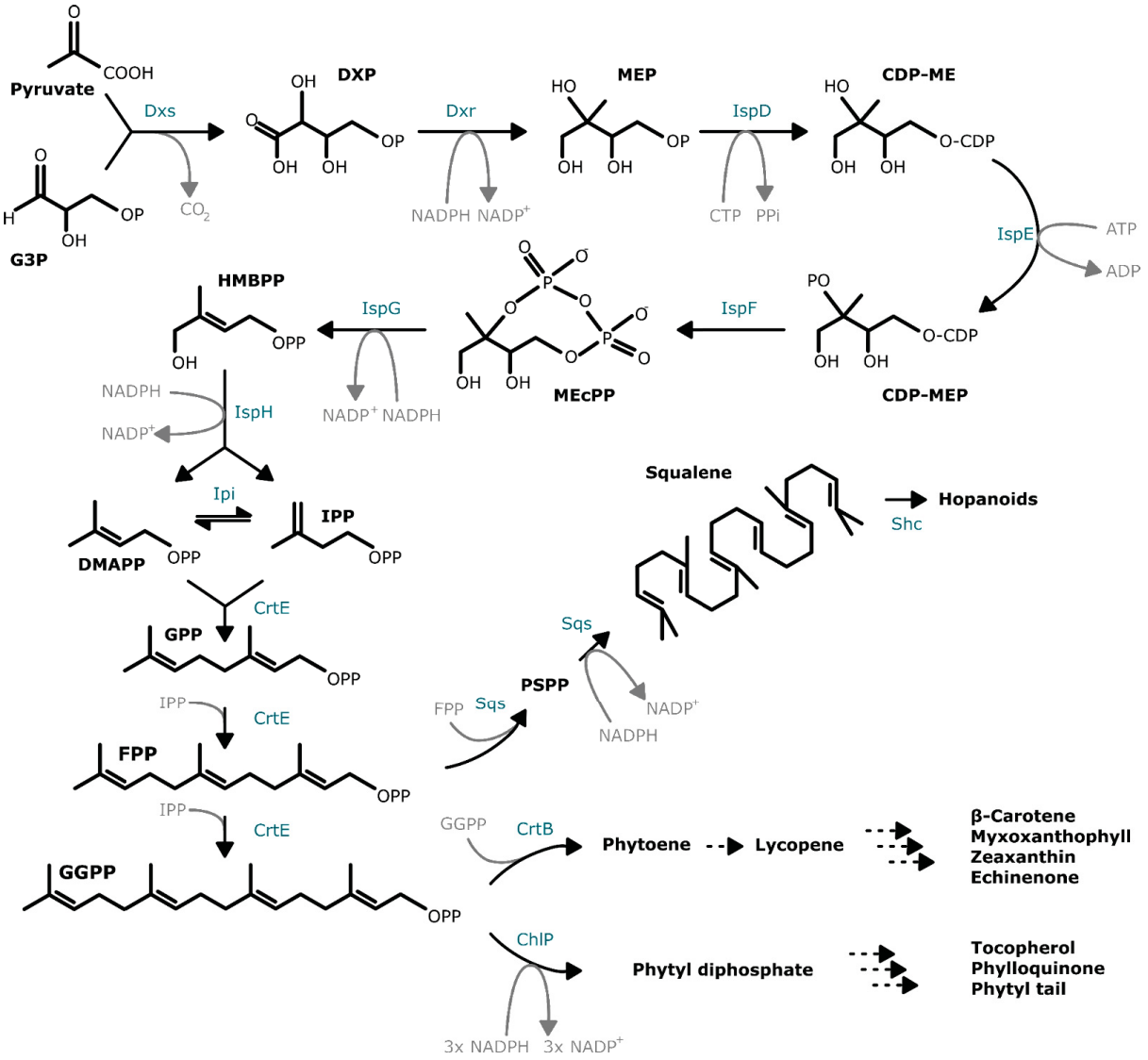


Figure 1: Bacterial terpenoid metabolism. Black arrows indicate enzymatic reactions, grey arrows indicate conversions involving co-factors. Protein names are shown in blue, intermediates are indicated in bold. Dotted arrows indicate multiple intermediate reaction steps, while multiple dotted arrows additionally indicate multiple branching steps. Metabolite abbreviations: G3P: glyceraldehyde 3-phosphate , DXP: 1-deoxyxylulose-5-phosphate , MEP: 2-methylerythritol-4-phosphate, CDP-ME: 4-pyrophosphocytidyl-2-C-methyl-D-erythritol, CDP-MEP: methylerythritol-4-phosphate, MEcPP: 2-C-methyl-D-erythritol 2,4-cyclopyrophosphate, HMBPP: (E)-4-Hydroxy-3-methyl-but-2-enyl pyrophosphate , DMAPP: dimethylallyl pyrophosphate, IPP: isopentenyl pyrophosphate, GPP: geranyl pyrophosphate, FPP: farnesyl pyrophosphate, GGPP: geranylgeranyl pyrophosphate, Enzyme abbreviations: Dxs: 1-deoxy-D-xylulose-5-phosphate synthase, Dxr: DXP-reductase, IspD: 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase , IspE: 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, IspF: 2-C-methyl-D-erythritol 2,4-cyclopyrophosphate synthase, IspG: 4-hydroxy-3-methylbut-2-enyl pyrophosphate synthase, IspH: HMB-PP reductase, Ipi: Isopentenyl pyrophosphate isomerase, CrtE: geranylgeranyl pyrophosphate synthase, Sqs: squalene synthase, Shc: squalene-hopene cyclases, CrtB: phytoene synthase, ChlP geranylgeranyl reductase.

1.3 Plant terpenoids and their industrial application

Terpenes, a chemically diverse group of metabolites, are utilized in plants for mutual and antagonistic interactions between plants and the environment and have an impact on their developmental physiology[35]. They also play a role in protection from phototoxicity in photosynthesis, regulation of membrane fluidity, and electron transfer in the respiratory chain

and in photosynthesis[36]. Plants form one of the richest sources of various terpenes, and together with aromatic compounds, form the essential oils. The highest concentrations in plant tissues are found in the specialized storage cavities of the leaves[37]. The synthesis of terpenoids is catalyzed by specific terpene synthases, which convert prenyl pyrophosphates into terpenoids of the various terpene classes[38]. So far, over 55,000 different structures are known and the terpene classes can be divided into hemiterpenes, monoterpenes, sesquiterpenes, diterpenes, triterpenes and tetraterpenes based on the number of carbon atoms[24,39](Fig. 2). Hemiterpenes, which are made up of a single terpenoid unit containing five carbon atoms, are the smallest terpenoids. The enzyme isoprene synthase, which catalyzes the conversion of DMAPP to isoprene, has been discovered to be present in a variety of plant species[40,41]. When plants are under heat stress, isoprene production is induced, and plants release it into the environment[42]. The largest class of plant secondary metabolites is called monoterpenes[43]. They are produced by condensation of IPP and DMAPP or by condensation of two DMAPP monomers and have a carbon content of ten. These substances have a vast variety and are found in significant amounts in plant extracts and essential oils[44]. Sesquiterpenes, which have 15 carbons, have been discovered to support a variety of biological processes[45]. They are produced by condensation of one more IPP monomer with the C10 monoterpene GPP to produce FPP, the precursor of all sesquiterpenes. Numerous species, including representatives of the cyanobacterial phylum, have been used for the production of the sesquiterpene-derived alcohol known as geosmin, known for its typical earthy or musty odor present after rain[46]. The C20 hydrocarbons known as diterpenoids are created when an IPP and an FPP molecule are combined to form GGPP. Diterpene synthases and CP450 are two sets of enzymes that, in combination, cause a high structural variability[47]. Through the condensation of two molecules of FPP, the 30-carbon hydrocarbon squalene is formed. In turn, squalene acts as a precursor for the synthesis of triterpenoids, such as eukaryotic sterols which are vital to cell membrane structure, and functions as a precursor for fat-soluble vitamins[48]. Phytoene, which is produced by two C20 GGPP, is the source of tetraterpenes with 40 carbons. The family of substances known as carotenoids is one group of tetraterpenes with roles in light absorption, antioxidative activity, synthesis of plant hormones, and as structural elements in membranes[49].

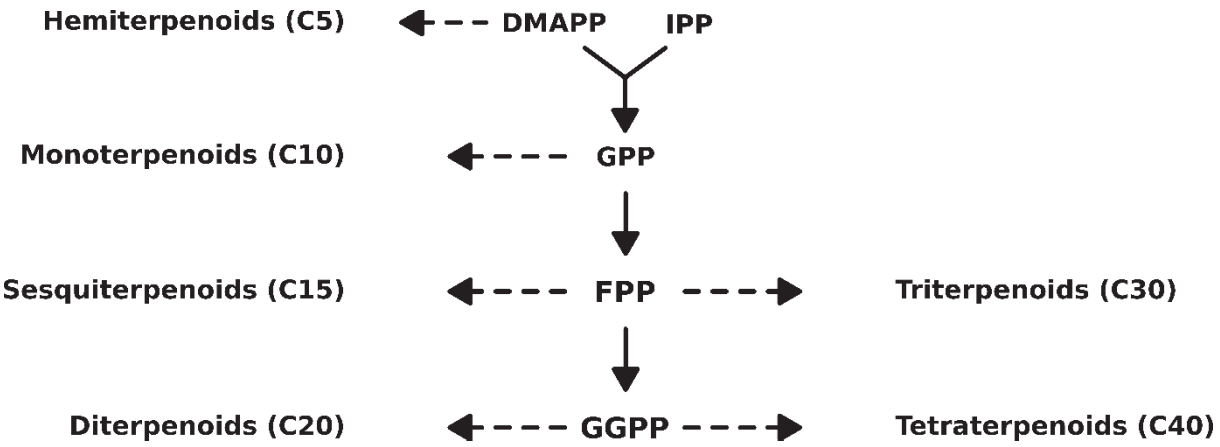


Figure 2: Lower isoprene precursor pathway. Dotted arrows indicate the terpenoid families branching off of the central precursors. DMAPP: dimethylallyl pyrophosphate, IPP: isopentenyl pyrophosphate, GPP: geranyl pyrophosphate, FPP: farnesyl pyrophosphate, GGPP: geranylgeranyl pyrophosphate

Many of these terpenes are industrially extracted and commercially used as agrochemicals, fragrances, nutraceuticals, and pharmaceuticals[18,50]. Terpenoids such as lutein, lycopene and astaxanthin are used as nutraceuticals and a variety of monoterpenes and sesquiterpenes such as menthol, d-limonene, α-farnesene, valencene and nootkatone are used as fragrances and flavors[51]. Also, terpenes such as limonene, myrcene and farnesene can be used in their hydrogenated forms as alternative fuels[52,53]. Artemisinin is particularly noteworthy. This sesquiterpene is naturally produced by the *Artemisia annua* plant, and its derivative acts as a potential antimalarial agent[54]. Using a heterologous mevalonate pathway derived from yeast, high levels of amorpha-4,11-diene, an artemisinin precursor, were produced in *Escherichia coli* (*E. coli*)[55]. These plant-derived terpenes represent a sustainable alternative to those chemically produced from petrochemicals, as no additional CO2 is released during production[56]. Chemical synthesis of these compounds is often non-trivial due to their complex structural and stereochemical properties[57]. The availability of natural terpenoids is typically constrained due to the time-consuming extraction processes required to separate them from their original plant sources[58]. Furthermore, naturally derived terpenes are often not feasible from a commercial standpoint due to the low yield of the specific terpenes. For example, approximately 400,000 kg of grapefruit are required to extract 1 kg of nootkatone[59].Although in recent years there have been significant breakthroughs in new technologies in plant research, such as the improvement of genome modification tools like CRISPR/Cas, the heterologous production of these plant terpenes in microorganisms represents a commercially cheaper and more productive alternative[60]. The reason for this is, for example, that constant availability, as a result, a stable price cannot always be realized due to the varying availability of agricultural raw material. At the same time, volatile substances such as isoprene or monoterpenes are difficult to capture from plants. Microorganisms that are grown in a bioreactor have a clear advantage due to the closed system[61,62]. Production in heterologous microbial hosts is an attractive option as they are easy to handle, grow, and generally have a simpler metabolism compared to plant hosts. Engineering genetically tractable microbial systems has led to significant advances in the reconstruction of metabolic pathways in microorganisms[63]

"The discovery of artemisinin was an example of successful collective efforts"
Tu Youyou (Winner of the Nobel Prize in Physiology or Medicine 2015)

1.4 Heterologous photoproduction of terpenoids

Heterotrophic production chassis such as *E. coli* and *Saccharomyces cerevisiae (S. cerevisiae)* offer many advantages for sustainable terpene production[64]. Especially *S. cerevisiae* is particularly interesting for the synthesis of terpenoids. Due to the endoplasmic reticulum in eukaryotes, yeast is particularly suitable for the heterologous expression of membrane-localized cytochrome P450s. At the same time, however, the availability of NADPH in most heterotrophic bacteria and yeast cells is a limiting factor required for enzyme activity of, for example, CP450s[65]. The unique advantage of phototrophic microorganisms like cyanobacteria is that they obtain energy from light and NADPH can be provided directly through

photosynthesis. Since terpenoid synthesis requires relatively more NADPH than ATP, lowering the ATP/NADPH ratio could provide a balanced cofactor pool. One example to increase NADPH production is to overexpress specific enzymes in the malic enzyme pathway and simultaneously knocking out the alternative pyruvate kinase pathway to generate a net increase NADPH and net decrease in ATP[66]. In addition, the use of atmospheric CO2 as a carbon source offers a particularly sustainable production chassis, since a reduced carbon source must be made available for heterotrophic organisms. As already described in the previous chapter, cyanobacteria have relevant building blocks for heterologous terpenoid synthesis through the intrinsic photopigment metabolism and, through the thylakoid membrane, offer storage for hydrophobic compounds and enzymes embedded in membranes[67,68].

At the same time, the photosynthetic way of life is the biggest bottleneck for commercial production. In phototrophic cultivation, self-shadowing effects and CO2 availability may lead to limited cell growth and reduced production. In addition to the expansion of new strategies in metabolic engineering, new approaches to lighting and cultivation conditions are essential. This optimization of lighting conditions can help to improve the biosynthesis of terpenoids in cyanobacteria. Approaches include adjustment of the light intensity, the light emission spectrum or to run different light-dark cycles. Thus, by cultivating *Synechocystis* under high light conditions, the yield of patchoulol production could be doubled compared to cultivation under low light conditions[69]. Also, cultivation in a simulated outdoor light-dark cycle could increase α-bisabolene productivity compared to a continuous light cycle in *Synechocystis*[70].

Another challenge for heterologous terpenoid production lies in the carbon distribution during photosynthesis, since in *Synechocystis* only about 5% of the photosynthetic carbon is available for terpene biosynthesis. By far the largest portion goes into the sugar biosynthetic pathways, with more than 80% of the fixed carbon being used for the accumulation of biomass[26]. A common carbon flux redirection strategy is to remove competing pathways for the desired product. By removing the glycogen biosynthetic pathway, preventing the formation of the energy storage compound glycogen, an important carbon sink, lactic acid production in *Synechocystis* could be doubled[71,72]. At the same time, the production of terpenes competes with the essential photosynthetic pigments such as chlorophyll and carotenoids, since the same precursor metabolites are used[25]. It could be shown that the production of terpenes reduced the photosynthetic pigments and reduced cell growth as a result[73]. This also affects the production of terpenoids, so in addition to increasing terpene precursors, there is also a need to balance carbon flow between product and pigment synthesis[74].

1.5 Strategies in metabolic engineering of bacteria towards terpenoid production

The field of Metabolic Engineering is primarily focused on enhancing the biological production of value-added chemicals and optimizing metabolic pathways, rerouting intracellular fluxes and improving cell properties for industrial applications[75]. To increase the production of terpenoids, it is crucial to increase the amount of the precursors IPP and DMAPP[76,77]. In cyanobacteria, most studies focus on optimizing the native MEP pathway to increase terpene precursors[51]. By overexpressing key enzymes such as Dxs and Ipi, the production of, for example, isoprene, α-farnesene, lycopene and β-carotene in *E. coli* could be increased. Similar results could be

achieved by overexpression of the native *dxs* gene in *Synechococcus elongatus* PCC 7942. There, the isoprene yield could be increased significantly[78]. Although Dxs and Idi were the main genes of interest, the overexpression targets for the MEP pathway varied depending on the products[79–81]. When metabolic engineering improves a reaction rate, the remaining reaction rates share the pathway control, except when there is a single rate-limiting step controlling the pathway. A combinatorial and systematic optimization of the metabolic pathway is required to avoid this problem. The DXP synthase in the MEP pathway, the entry enzyme that acts as a pathway-controlling step, is tightly regulated by IPP and DMAPP via feedback inhibition[82]. In addition to the MEP pathway, access to these precursors could also be provided via the MVA pathway, which can be expressed heterologously in bacteria[83]. In the MVA pathway, the rate-limiting steps are tightly regulated by free CoA and HMG in the conversion of intermediates by HMGR and HMGS[84] (see 1.2.1). In order to reduce the inhibition of their expression, these key enzymes are the focus of metabolic engineering and are investigated by overexpression or the introduction of mutations[85]. Moreover, the terpene production can be influenced by the carbon source and the energy balance. In the MEP-pathway, generation of 1 IPP or DMAPP requires NADPH in three reaction steps, while only one step requires NADPH in the MVA pathway. However, the MEP route is more efficient in producing one mole of IPP: For production, 1.5 mol glucose is used in the MVA pathway, whereas the MEP pathway only consumes 1.25 mol glucose[51].

Another approach in metabolic engineering is varying expression levels of each individual metabolic gene, taking into account individual speed of each reaction and thereby optimizing metabolic flux through the entire pathway. Using titratable genetic systems, lycopene production could be optimized iteratively in *E. coli*. Different metabolic genes were each put under the control of a different inducible system and each was subjected to different inducer concentrations and productivity was assessed. After each iteration, the optimal combination of inducer concentrations was identified and refined to a higher resolution. After four iterations, lycopene productivity was increased by 90 mg/L[86]. While this approach is more of a bottom-up approach making use of rationally designed systems, the top-down approach takes into account metabolic network models that make use of organism-wide information. Metabolic network models and metabolic fluxes are essential concepts in metabolic engineering. Both computational and analytical experimental approaches can help explore the capabilities of biological systems through the analysis of metabolic network models. Furthermore, stoichiometric data are sufficient for certain genome-scale metabolic models, and have the advantage that they are computationally inexpensive[87,88]. For example, constraint-based flux balance analysis (FBA) has been successfully used as a tool to understand and analyze metabolic pathways and, moreover, to identify bottlenecks and narrow down the options of potential overexpression and gene deletions targets to increase terpenoid production[75,89,90]. Along with the rapid advances in microbial engineering, the identification of novel terpene synthases can lead to the discovery of new terpenes and the resulting new applications[64]. To screen potential terpene synthase genes, high-throughput prediction approaches based on amino acid sequence similarity using metagenomic sequencing data have emerged[91]. However, this may exclude completely novel enzymes. In a functional metagenomic approach, screening of a library made it possible to isolate a novel β-farnesene synthase from a metagenomic library that has no amino acid sequence similarity to known β-farnesene

synthases[92]. At the same time, it has proven difficult to engineer known terpene synthases towards novel functions, since a clear relationship between phylogenetic organization and catalytic specificities has yet to be deciphered[93]. Directed evolution has proven to be a suitable method for optimizing the performance of enzymes[94]. For example, the thermostability of a terpene synthase and the reaction conditions of another enzyme could be optimized by screening for cyclization activity by colorimetric read-out in cell lysates[95].

In summary, the selection of a suitable chassis organism depends strongly on the nature of the product to be synthesized[64] and finding new microbial hosts to produce terpenoids is therefore intriguing. Cyanobacteria offer a good host for the production of plant-derived terpenes, given that the required enzymes' natural environment, the chloroplast, is closely related to cyanobacteria[96].

1.5.1 Trade-off between growth and production

Since the production of a desired metabolite is in resource allocation conflict with biomass production, it can be generalized that strains with high production grow slowly and strains with low product yield grow faster[97]. Both cell growth and product formation consume the same resources and require the same precursor metabolites. The strategy to avoid this dilemma is to establish a two-stage bioprocess. The aim is to decouple growth from production. In the first step, the cells with a fast growth rate are grown to the ideal biomass density[98]. In the next step, growth is stopped so that the resources that have become available can be used for product build-up. In *Corynebacterium acetoacidophilum*, in which the alternation of growth and production is part of its natural life cycle, succinate yield could be increased by changing from an aerobic growth stage to an anaerobic production stage[99]. Other strategies are used for other bacteria that do not have this natural regulation. In *E. coli*, product synthesis could be sustained while growth was inhibited by nutrient limitation[100,101]. Using a growth switch based on a serine recombinase from the bacteriophage phiC31, the origin of replication (oriC) could be permanently removed from the chromosome of *E. coli*. As a result, growth was stopped by the absence of replication and the metabolism was kept active. The change from the growth phase to growth arrest resulted in a reporter protein content that was five times higher than in the non-switched cells[102]. Other approaches are the use of inhibitors or even the deletion of growth-relevant genes. Because some of these genes are essential, the current approach is to take advantage of conditional lethality. The bacteria will only survive with or without certain substances[103]. A single gene deletion study performed in *E. coli* found that even the deletion of non-essential genes produced a loss in overall fitness[104]. An elegant method for temporarily silencing genes is enabled by CRISPR interference technology (CRISPRi). The catalytically inactive Cas9 blocks transcription using a specifically coding single-guide RNA on the desired gene[105]. A genome-wide CRISPRi library was used to systematically identify targets in *E. coli* that exhibited inducible reduced growth. Simultaneously, the production capacity of these strains was measured using GFP, identifying strains with growth-production decoupling capacity[106]. One identified candidate was *gyrA*, which as an essential gene cannot be deleted, but leads to inhibited growth when inactivated. In cyanobacteria, the two main topoisomerases are encoded by *gyrA/B* (gyrase) and *topA* (topoisomerase I). While type I topoisomerases are involved in the relaxation of negatively supercoiled DNA, type II topoisomerases, also known

as DNA gyrases, actively introduce negative supercoils by hydrolyzing ATP. These two major players involved in the global and local regulation of DNA supercoiling are themselves intrinsically regulated by the topological state in the genomic region where they are encoded; While the gyrase genes are preferentially expressed from relaxed DNA along with GC-rich genes for growth and anabolic processes, topoisomerase I genes and AT-rich genes involved in catabolism are preferentially expressed from negatively supercoiled DNA, resulting in DNA topology homeostasis[107]. In cyanobacteria, a connection between the diurnal expression pattern and oscillations in DNA supercoiling could be shown[108]. A trend towards an increasing ATP/ADP ratio at lower gyrase expression levels in *E. coli* could be shown in Jensen et. al 1999. It was argued that the gyrase uses ATP as a substrate and the resulting reduced ATP requirement has an impact on the ATP/ADP ratio. However, at the same time, it could be shown that the growth rate was reduced at lower gyrase expression levels and that there was possibly an additional reduced ATP requirement[109]. Based on these observations, studying the interaction of supercoiling, growth rate and the ATP/ADP ratio may offer an interesting perspective for future production.

1.6 Aim

The aim of my work was to establish a biotechnologically usable expression host, exemplified by the photoautotrophic model organism *Synechocystis* sp. PCC 6803.

- To examine how competing pathways release precursor molecules by reducing the transcripts via CRISPRi, and whether the production of the desired metabolites can be increased, the carotenoid biosynthetic pathway was specifically downregulated as proof-of-concept by reducing the transcript of the key enzyme in the terpenoid metabolism of *Synechocystis*. At the same time, the gap to the precursor molecules was to be closed by compensatory genes, which are heterologously overexpressed. The success of each modification, as well as combinations thereof, was assessed by measuring valencene production.
- Based on an *in silico* flux-balance analysis potential overexpression targets had been predicted. These predicted targets needed to be experimentally validated through the integration of the native genes in the genome of *Synechocystis*, resulting in overexpression. The target molecule squalene was used as a readout to verify the *in silico* predictions.
- The question whether reducing the growth rate of *Synechocystis can* free up resources for the production of metabolites remains unanswered. For this purpose, the regulatory mechanism of DNA supercoiling was decoupled from the energetic state of the cell by targeted overexpression or removal of two opposing topoisomerases. The resulting strains were characterized in detail with regards to their physiology and transcriptional response.

The final goal was designing an organism by using metabolic engineering technology to simultaneously up- and downregulate relevant synthesis pathways that have a positive effect on production rates and that release further resources through controlled growth.

2 Chapter 2 Key results

2.1 Photoproduction of the sesquiterpene valencene

Using the cyanobacterium *Synechocystis* sp. In PCC 6803, we effectively present a multicomponent engineering strategy for the photoproduction of valencene. First, we used markerless genomic deletions of *shc* and *sqs* to increase flux towards valencene. Second, we used inducible CRISPRi instead of gene deletion on *crtE* to inhibit the production of carotenoids, which are necessary for cell survival. To ensure continuous production of the precursor FPP for valencene production, crtE was replaced by the heterologous ispA, which catalyzes the same reaction as crtE, but only up to FPP. Subsequently, by designing a fusion protein, we aimed to increase the spatial proximity of the two enzymes involved in the production of valencene, ispA and CnVS. *Synechocystis* was able to produce 19 mg/g DCW valencene by combining the most effective techniques.

2.2 A systematic overexpression approach to increase squalene

Flux balance analysis revealed possible genetic targets that were predicted to have a beneficial effect on squalene synthesis upon overexpression. Selected target genes were inserted into the *Synechocystis* Δ*shc* genome, which is incapable of producing hopanoids, thereby rendering the cell incapable of further metabolizing squalene. As a proxy for increased metabolic flux towards terpene synthesis, squalene production was assessed. The genes, ispH, ispE, and ipi, which are genes of the MEP pathway, resulted in improvements in squalene production. The biggest gain in production was achieved by overexpression of the native squalene synthase gene (*sqs*) in *Synechocystis* Δ*shc,* reaching a production titer of 13.72 mg L^{-1} Squalene.

2.3 Topoisomerase Expression Inhibits Cell Division but not Growth

We modified DNA supercoiling in *Synechocystis* sp. PCC 6803 by either overexpression of topoisomerase I (topAOX) or targeted transcriptional downregulation of the two gyrase subunits using CRISPRi. All strains showed a characteristic halt on cell division, although cell growth persisted, resulting in a 4-fold increase in cell volume. The increased cell volume was further supported by flow cytometry and microscopy in all strains, which also showed a rise in the proportion of 8-shaped cells, further supporting a stop in cell division but not growth. These findings are in line with the fundamental assumptions of both the homeostasis and twin-domain models of bacterial supercoiling. Initially, topAOX resulted in the global upregulation of A+T-rich genes and the downregulation of G+C-rich genes. Loss of DNA supercoiling had an effect on the amount of pigment, and the amount of ATP. Pigment levels dropped, while the ATP+ADP content increased. The strain topAOX, which showed the strongest effects, had significantly elevated levels of glycogen, comparable to those of cells that had been deprived of nitrogen. By manipulating DNA supercoiling, we were able to successfully reroute cellular resources, creating a viable platform for photoproduction.

3 Chapter 3
Manuscript I

3.1 Author's contributions

M. Dietsch*, A. Behle*, P. Westhoff, and I. M. Axmann, ''Metabolic engineering of *Synechocystis* sp. PCC 6803 for the photoproduction of the sesquiterpene valencene,'' Metab. Eng. Commun., vol. 13, p. e00178, Dec. 2021, doi: 10.1016/j.mec. 2021.e00178

* Shared first authorship.

Maximilian Dietsch: Conceptualization, Investigation, Methodology, Writing - Original Draft, Writing - Review & Editing, Visualization, Data Curation

M.D. and A.B. designed the overall study with input from I.M.A. M.D. and A.B. constructed all plasmids and strains used in this publication. M.D. performed cultivation experiments, as well as biomass measurements and pigment quantification. M.D. performed all RNA extractions; he performed qRT-PCR experiments and data analysis thereof in cooperation with A.B. Literature research, and manuscript preparation was performed by A.B and M.D with input and support of I.M.A.

The attribution categories described for each journal have been summarized as following:

Study design: 75%
Experimental contribution: 60%
Data analysis: 30%
Manuscript preparation: 30%

3.2 Metabolic engineering of *Synechocystis* sp. PCC 6803 for the photoproduction of the sesquiterpene valencene

Metabolic Engineering Communications 13 (2021) e00178

Contents lists available at ScienceDirect

Metabolic Engineering Communications

journal homepage: www.elsevier.com/locate/mec

Metabolic engineering of *Synechocystis* sp. PCC 6803 for the photoproduction of the sesquiterpene valencene

Maximilian Dietsch [a,1], Anna Behle [a,1], Philipp Westhoff [b], Ilka M. Axmann [a,*]

[a] Institute for Synthetic Microbiology, Department of Biology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[b] Plant Metabolism and Metabolomics Laboratory, Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University Düsseldorf, D-40001, Düsseldorf, Germany

ARTICLE INFO

Keywords:
Metabolic engineering
Cyanobacteria
Synechocystis
Valencene
Sesquiterpene

ABSTRACT

Cyanobacteria are extremely adaptable, fast-growing, solar-powered cell factories that, like plants, are able to convert carbon dioxide into sugar and oxygen and thereby produce a large number of important compounds. Due to their unique phototrophy-associated physiological properties, i.e. naturally occurring isoprenoid metabolic pathway, they represent a highly promising platform for terpenoid biosynthesis. Here, we implemented a carefully devised engineering strategy to boost the biosynthesis of commercially attractive plant sequiterpenes, in particular valencene. Sesquiterpenes are a diverse group of bioactive metabolites, mainly produced in higher plants, but with often low concentrations and expensive downstream extraction. In this work we successfully demonstrate a multi-component engineering approach towards the photosynthetic production of valencene in the cyanobacterium *Synechocystis* sp. PCC 6803. First, we improved the flux towards valencene by markerless genomic deletions of *shc* and *sqs*. Secondly, we downregulated the formation of carotenoids, which are essential for viability of the cell, using CRISPRi on *crtE*. Finally, we intended to increase the spatial proximity of the two enzymes, *ispA* and *CnVS*, involved in valencene formation by creating an operon construct, as well as a fusion protein. Combining the most successful strategies resulted in a valencene production of 19 mg/g DCW in *Synechocystis*. In this work, we have devised a useful platform for future engineering steps.

1. Introduction

Cyanobacteria are known for their unique ability of oxygenic photosynthesis among bacteria. Thus, they are becoming increasingly important in biotechnological applications and for generating sustainable energy. Unlike plants, cyanobacteria can be cultivated in huge salt water basins, even in desert regions, solely with sunlight and CO_2 from the air or from connected power plants and, thus, do not compete with agricultural land and food production Furthermore, extraction of plant secondary metabolites has proven to be inefficient, as it has yielded only small amounts of the desired products thus far. Here, cyanobacteria represent excellent candidates for the expression of plant biosynthetic genes and gene clusters due to their ancestral relationship to plant chloroplasts. In recent years, continuous efforts have been put into developing industrially viable strains of cyanobacteria for the sustainable production of various fine chemicals, secondary metabolites, and other compounds (Jodlbauer et al., 2021; Liu et al., 2021). Advances in synthetic microbiology and increasing availability of new genetic tools for this important group of organisms enable even more innovative solutions.

In terms of structural diversity, terpenoids comprise an extremely versatile class of compounds. Naturally, the terpenoid backbones in cyanobacteria are generated via the methyl-erythritol-phosphate (MEP-) pathway, which produces the central terpenoid precursors IPP and DMAPP. By subsequent addition of another precursor, GPP (C10), the precursor for monoterpenes, FPP (C15) the precursor for sesquiterpenes and triterpenes, such as hopanoids, and GGPP (C20), the precursor for di- and tetraterpenes, to which the carotenoids belong, are generated. One prominent example for natural sesquiterpene production is geosmin found in several Cyanobacteria species, which is responsible for the characteristic earthy smell in water bodies (Lee et al., 2017). Sesquiterpenes are especially convenient for the heterologous production in microorganisms because they are often volatile, eliminating the necessity for costly extraction methods and downstream processing.

* Corresponding author.
E-mail address: ilka.axmann@hhu.de (I.M. Axmann).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.mec.2021.e00178
Received 2 March 2021; Received in revised form 5 July 2021; Accepted 22 July 2021
Available online 13 August 2021

Naturally, they are often found in plants, where they may function as defensive agents against predators. In industry, sesquiterpenes are used as flavor and fragrance additives and have been successfully produced in numerous microbial hosts, with very different yields.

The first metabolic engineering efforts for the production of sesquiterpenoids were made in *Escherichia coli (E. coli)*, where amorphadiene, the precursor of the antimalarial drug artemisinin, was produced via heterologous expression of the complete mevalonate pathway from the yeast *Saccharomyces cerevisiae (S. cerevisiae)* (Newman et al., 2006). Using a combination of metabolic engineering and a two-phase cultivation system, a total yield of ~0.5 g/L product was achieved. This product titer was even further improved by the introduction of metabolically more active enzymes, as well as an improved growth media composition (Tsuruta et al., 2009).

S. cerevisiae, as well as other fungal species, has also successfully been applied for the production of sesquiterpenes such as valencene. Similar to previous efforts, enhancing the flux through the native isoprenoid biosynthesis pathway by overexpressing each gene had an advantageous effect on valencene product yield. In addition, by repressing essential genes that normally diverted some of the FPP precursor away from the desired product, valencene yield increased even more. The highest product yield in yeast, ~540 mg/L valencene, was again achieved by a combination of genetic engineering and optimization of media composition and cultivation (Chen et al., 2019). Recently, the corn smut fungus *Ustilago maydis* was explored as a microbial production host for sesquiterpenes, due to previous successes in biotechnological applications using this host (Lee et al., 2020).

Another strategy to increase the utilization of precursors towards desired products instead of native off-target pathways is to increase the spatial proximity of two sequential enzymes. In yeast, this was achieved by fusing FPP-synthase, encoded by *erg20*, with the heterologous sesquiterpene synthase, producing germacrene A in one case, and patchoulol in another (Albertsen et al., 2011; Chen et al., 2019).

This method of creating a chimeric enzyme showed success, with all fusion variants leading to an overall increase in product yield. In another study, heterologous production of carotenoids was achieved through bi- and tridomain fusion proteins, further demonstrating the importance of spatial proximity in metabolic pathways (Rabeharindranto et al., 2019).

Photoautotrophic bacteria, cyanobacteria in particular, have shown promising results in terms of production (Angermayr et al., 2015). Since they are natural terpene producers, they are excellent candidate chassis for metabolic engineering. In terms of product yield, they were able to compete with heterologous hosts under standard laboratory conditions. For example, the two sesquiterpenoids bisabolene and patchoulol were produced under high density conditions, yielding ~179.4 mg/L and 17.3 mg/L, respectively (Dienst et al., 2020). Another study showed successful production of various triterpenes from one key precursor in *Rhodobacter capsulatus* and *Synechocystis* sp. PCC 6803 (*Synechocystis* hereafter), indicating efficient exploitation of the native terpene pathway of photosynthetic organisms through genetic engineering (Loeschcke et al., 2017). Here, we present a multi-component approach towards the photosynthetic production of valencene. First, we applied metabolic engineering to generate a strain with a more favorable flux towards the precursor FPP by markerless genomic deletions. Secondly, we used CRISPRi to downregulate the formation of carotenoids, which are essential for viability of the cell. Finally, we applied two strategies to increase the spatial proximity of the two enzymes involved in valencene formation by creating an operon construct, as well as a fusion protein to increase the flux from FPP to the final precursor, valencene. This work successfully demonstrates heterologous production of the sesquiterpene valencene in the cyanobacterium *Synechocystis* using different engineering approaches.

2. Material & methods

2.1. Plasmid and strain construction

A detailed list of all relevant genetic modules and information regarding their origin, is provided in the Supporting Information (Table S2).

The previously published pSHDY-Prha-mVenus_rhaS (Behle et al., 2020) (Addgene #137662) was slightly modified by excising the spectinomycin resistance cassette and replacing it with a nourseothricin resistance cassette, thereby creating an alternative plasmid we termed pSNDY.

Synthetic, codon-optimized genes were synthesized by IDT. Relevant genetic components were amplified and fused using overlap extension PCR when necessary, (dx.doi.org/10.17504/protocols.io.psndnde).

and integrated into the pSNDY backbone, either via Gibson assembly (dx. doi.org/10.17504/protocols.io.n9xdh7n), or using restriction/ligation cloning.

Plasmids were transferred to *Synechocystis* sp. PCC 6803 wild-type using triparental mating (dx.doi.org/10.17504/protocols.io.psndnde).

pMD19T-psba1-Ppsba2-dCas9-SpR was a gift from Paul Hudson (Addgene plasmid # 73220; http://n2t.net/addgene:73220; RRID: Addgene_73220).

2.2. Culture conditions

For pre-culturing and growth experiments, *Synechocystis* was cultivated in BG11 medium (Stanier et al., 1979). Standard cultivation was performed at 30 °C with 150 rpm shaking and continuous illumination of ~80 μE m–2 s–1. Aeration was ensured by continuous shaking and CO_2 enriched air (0.5%). Whenever necessary, appropriate antibiotics were added to the different strains. Pre-culturing was performed in 100 ml baffle-free Erlenmeyer shaking flasks with 20 ml cell suspension for three days. After adjusting all different strains on the OD growth experiments were performed after one additional day of pre-culturing. For this, 4 ml cultures were incubated in 6-well plates for 48 h with a start OD_{750} of 0.5 in biological triplicates. To avoid loss of the volatile product valencene, cultures were overlaid with 20 % dodecane.

2.3. Biomass measurements (DCW, OD, spectra)

Optical density and whole cell spectra measurements were performed in the SpEcoRd 200 plus and diluted if necessary. To determine the cell dry weight (CDW) 2–3.5 ml cell culture was pelleted for 3 min at maximum speed. After washing the pellet with PBS buffer, the pellet was resuspended in ~50 μl water and transferred to a pre-weighed PCR tube, where it was dried at 60° overnight prior to weighing.

2.4. Microscopy

Cells were analyzed phenotypically using the bright field setting of a Zeiss AxioScope.A1, under 400-fold magnification.

2.5. Pigment quantification

0.2–0.5 ml of each culture was sampled after 48 h at the end of the growth experiment. The sample was centrifuged for 5 min at 14,000 g and 4 °C. The supernatant was discarded and the pellet resuspended in 100 μl water. The samples were frozen at −20 °C until further processing. 900 μl of 100% methanol was added and the sample was mixed by vortexing. After incubation with gentle shaking for 30 min at 4 °C, the sample was centrifuged at 14,000 g for 5 min. The supernatant was transferred to a cuvette and the absorbance spectrum was measured from 400 nm to 750 nm. The absorbance spectra were divided by the OD_{750} or CDW and the amount of chlorophyll *a* in the sample was quantified by the absorbance maximum of chlorophyll *a* at 665 nm

(A_{665nm}) using following equation (Lichtenthaler and Buschmann, 2001):

2.6. Chlorophyll content [μg/ml] = 12.66 μg/ml * $A_{665\ nm}$

The amount of carotenoids in the sample was quantified by the absorbance maximum of the sum of carotenoids at 470 nm (A_{470nm}) and a correction term considering absorbance of chlorophyll *a* at 470 nm (c (Chl a): concentration of chlorophyll *a* in the sample) using the following equation:

Carotenoid content [mg/ml] = (1000 μg/ml * $A_{470\ nm}$ −1.91 * c (Chl))/225.

2.7. RNA extraction & qRT-PCR

RNA extraction was performed according to (Pinto et al., 2009). Briefly, 0.2–1 ml cell culture was collected and pelleted for 3 min at maximum speed at 4 °C. After discarding the supernatant, the pellet was resuspended with 0.5 ml PGTX and incubated at 95 °C for 5 min. After cooling on ice, 350 μl chloroform/isoamyl alcohol were added and the mixture was incubated shaking gently at room temperature for 10 min. To separate the aqueous from organic phases the mixture was centrifuged for 10 min at maximal speed at 4 °C. The upper phase was transferred to a fresh tube and 1 vol chloroform/isoamyl alcohol added. After repeating the centrifugation step the upper phase was again transferred and precipitated with 3 vol of 100 % ethanol sodium acetate at −20 °C overnight. The RNA was pelleted for 30 min at maximum speed and 4 °C, washed twice with 70% ethanol and resuspended in RNase-free water.

RNA was DNaseI-digested using commercial DNaseI from ThermoFisher (EN0525), according to the manufacturer's specifications. DNaseI-digested RNA was phenol/chloroform extracted again to remove the DNaseI.

For cDNA synthesis, the commercial RevertAid RT from ThermoFisher (K1621) was used according to the manufacturer's specifications.

qRT-PCR was performed using the DyNAmo ColorFlash SYBR™ Green qPCR-Kit (ThermoFisher, F416L), according to the manufacturer's specifications.

2.8. GC-MS for the quantification of volatile sesquiterpenoids

100 μL dodecane overlay fractions were collected in micro inserts inside 1.5 mL clear glass GC vials. 2 μL of the sample were diluted 1:50 in HPLC grade hexane (Th. Geyer GmbH, Germany) prior to injection. 1 μl of the diluted was injected with an MPS autosampler with automatic liner exchange system in conjunction with a cold injection system (Gerstel) in splitless mode (ramping from 50 °C to 250 °C at 12 °C s^{-1}) into the GC with a helium flow of 1 ml min^{-1}. Chromatography was performed using a 7890B GC system (Agilent Technologies) with a HP-5MS column with (5%-phenyl)-methylpolysiloxane film (Agilent, 19091S-433, 30 m length, 0.25 mm internal diameter, 0.25 μM film). The oven temperature was held constant at 70 °C for 2 min and then ramped at 12.5 °C min^{-1} to 320 °C at which it was held constant for 5 min; resulting in a total run time of 27 min. Metabolites were ionized with an electron impact source at −70 eV and 200 °C source temperature and recorded in a mass range of *m/z* 60 to *m/z* 800 at 20 scans per second with a 7200 GC-QTOF (Agilent Technologies) after a solvent delay time of 8 min. Compound identification was conducted via MassHunter Qualitative (v b08.00, Agilent Technologies) by comparison of mass spectra to the NIST14 Mass Spectral Library (https://www.nist.gov/srd/nist-standard-reference-database-1a-v14) and validated by retention time comparison with chemical reference substances (Sigma-Aldrich, #06808). Peaks were integrated using MassHunter Quantitative (v b08.00, Agilent Technologies). The concentration was determined via external calibration. The calibration curve was generated with 8 points from 0.1 μM to 20 μM with a quadratic curve fit and 1/x curve fit weight. After direct measurement of valencene in the dodecane layer, molar concentrations were calculated to mg valencene/L *Synechocystis* culture. To determine whether valencene was lost over time via evaporation or degradation, dodecane with 225 μM valence was measured directly and compared to a dodecane layered cell culture with 225 μM valence and cultivated for 48 h. Technical triplicates were cultured and measured, and there was no significant difference detected between the samples (Fig. S6).

2.9. Total protein isolation and Western Blot analysis

Total protein was extracted from cultured and induced *Synechocystis* cultures as described (dx.doi.org/10.17504/protocols.io.ps6dnhe). Protein concentration was determined according to Lowry et al. using a BSA standard. 20 μg total protein was loaded on an SDS gel, transferred to a PVDF membrane, UV-crosslinked, and the presence of the IspA:CnVS fusion protein, as well as IspA only from the operon construct, was detected using a monoclonal anti-FLAG-M2-alkaline-phosphatase antibody (Sigma, A9469) as primary, and an anti-mouse antibody as the secondary antibody.

3. Results & discussion

The central terpenoid pathway in *Synechocystis* starts with IPP and DMAPP, which are derived from the MEP-pathway. A single gene, *crtE*, is responsible for the elongation of terpene precursors towards GPP, FPP, and GGPP (Fig. 1). Next to GGPP, another downstream metabolic product of FPP is squalene, which is converted to hopanoids.

In the following, we demonstrate various strategies to divert metabolic flux towards heterologous sesquiterpenes (Fig. 1, black components), effectively eliminating undesired side products (Fig. 1, gray components).

3.1. Modulating the internal precursor pool by genomic gene deletion of squalene synthase and squalene hopane cyclase

In order to divert metabolic flux away from undesired side products and towards farnesyl pyrophosphate (FPP), which is the central precursor for sesquiterpenes (Fig. 1), we applied two strategies. First, we performed markerless deletions of two genes, squalene synthase (*sll0513*, *sqs*), which is responsible for the conversion of FPP to the triterpene squalene, and the gene directly downstream, squalene hopane cyclase (*slr2089*, *shc*), which further converts squalene to hopanoids. A *shc* knock-out was previously performed in order to accumulate squalene, and a 70-fold increase was demonstrated using this deletion mutant (Englund et al., 2014). We hypothesize that an additional *sqs* deletion might lead to an accumulation in FPP in a similar manner. We further improved the initial strain design by performing markerless gene deletions, which are of special interest because resistance cassettes can be recycled, instead of being occupied indefinitely within the genome. Thereby, multiple different alterations in one strain are possible. Each markerless deletion was performed in two sequential steps as previously described (Viola et al., 2014); first, a CmR-sacB cassette, flanked by the neighboring genomic regions and including a partially overlapping fragment thereof, was introduced into wild type *Synechocystis*. By gradually selecting on higher chloramphenicol concentrations, complete genome segregation was achieved. In a second step, counter-selection of segregated clones on solid media containing sucrose, but no chloramphenicol, was carried out, thereby eliminating cells still carrying the CmR-sacB cassette and selecting for a second double-crossover event between the partial overlaps. The schematic genotype of the double mutant is shown in Fig. 2A.

To ensure complete deletion of both genes, the lack of transcripts was verified via qRT-PCR (Table S1) correctly segregated mutants were verified via colony PCR (Fig. S1A).

The fully segregated strain was then compared to the wild type strain

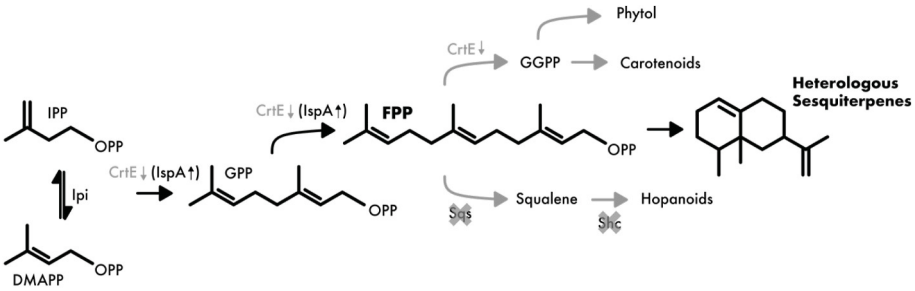


Fig. 1. Isoprene pathway from *Synechocystis* with optimizations done in this work. Abbreviations used: IPP = isopentenyl diphosphate; DMAPP = dimethylallyl diphosphate; GPP = geranyl diphosphate; FPP = farnesyl diphosphate; Ipi = isopentenyl diphosphate delta isomerase; CrtE = geranylgeranyl pyrophosphate synthase; Sqs = squalene synthase; Shc = squalene hopene cyclase; IspA = farnesyl diphosphate synthase. Crossed out target = gene deletion. Down arrow = repression target. Upward arrow = overexpression/rescue.

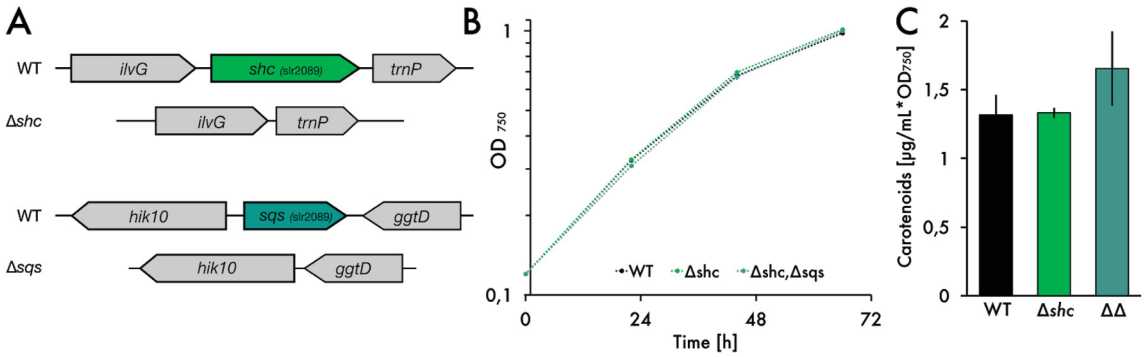


Fig. 2. Knockout strategy and growth/pigment analysis of mutants. A: Schematic overview of markerless mutant genotypes. B: Comparison of growth between wild type and mutants. C: Carotenoid content of the three different strains. Results represent the mean of three biological replicates. WT = wild type, Δshc = squalene-hopene-cyclase deletion, ΔΔ = squalene-hopene-cyclase and squalene synthase deletion.

in terms of growth and pigment composition. While there was no discernible difference in growth (Fig. 2B), the double mutant showed a visible shift in carotenoids (Fig. S1 C, D), suggesting an increase in metabolic flux towards the central carotenoid precursor, GGPP, which is derived from FPP (Fig. 1). Upon further investigation via pigment extraction, this observation was confirmed; the double mutant showed an increase in carotenoid content (Fig. 2C). Due to the previously described toxicity of FPP in *E. coli* (Dahl et al., 2013), an enrichment of this intermediate seems implausible and a further conversion to harmless carotenoids seems to be likely.

To test whether this increased carotenoid content could be translated to an increased FPP availability for sesquiterpene production, the wild type and mutant expressing valencene synthase (CnVS) from *Callitropsis nootkatensis* (Beekwilder et al., 2014) under the rhamnose-inducible promoter (Behle et al., 2020) (Fig. 3A) were cultured alongside in 6-well plates for two days. To avoid loss of the volatile product valencene, cultures were overlaid with 20 % dodecane. The dodecane layer was then sampled and quantified directly using GC-MS.

The identification of valencene in all strains was performed by comparing retention time and mass spectra with those of a commercial standard (Fig. S4). In Fig. 3B, the WT expressing *CnVS* as well as a negative control is shown in the extracted ion chromatogram (m/z 161.12) as an example.

Remarkably, the double mutant showed a ~40% increase in valencene production compared to the wild type (Fig. 3C). In contrast to before (Fig. 2C), the mutant now expressing *CnVS* did not show an increase in carotenoid content (Fig. 3D). In all likelihood, the excess precursor pool that was diverted towards carotenoid production before was now successfully used by CnVS. Due to the promising production increase in the mutant, we exclusively used the double knockout mutant background for the following experiments.

3.2. Enhancing the FPP precursor pool by conditional repression of the essential gene crtE

To further exploit the available carotenoid pool, we aimed at reducing the conversion of FPP to GGPP. In *Synechocystis*, a single gene, *crtE*, is responsible for the consecutive condensation of IPP and DMAPP to GPP, FPP and finally to GGPP (Lin et al., 2017), the precursor for diterpenoids, including the chlorophyll phytol tails, and tetraterpenoids, such as carotenoids (Fig. 1A). In contrast, genes from heterotrophic species, such as *ispA* from *E. coli*, only perform these conversions up until FPP (Reiling et al., 2004). Since GGPP-derived pigments are essential for cyanobacterial viability, we decided to reduce *crtE* expression via inducible, dCas9-based CRISPRi, and then introduce a heterologous FPP-synthase to increase the relative amount of FPP compared to GGPP.

We chose an sgRNA from a previously published work to target *crtE* (Yao et al., 2020), as well as the aTc-inducible dCas9 system from Yao et al. (2016) (Fig. 4A).

Interestingly, the qRT-PCR with *crtE*-specific oligonucleotides shows a repression down to <10 % of the wild type level at much lower inducer concentrations of 10 ng/mL aTc, despite a reported >90 % repression at concentrations as low as 100 ng/mL aTc. Notably, the uninduced *crtE*-repression strain already shows a 40 % reduction of gene expression compared to the wild type (Fig. 4B). Consistent with published results, induction with 100 ng/mL aTc shows almost complete repression of *crtE*. While the pigment composition of the uninduced strain resembles the wild type, an aTc-dependent effect on both chlorophyll and carotenoids can be observed (Fig. 2C). General pigmentation is severely affected at 100 ng/mL aTc, whereas only carotenoids are affected at 10 ng/mL aTc. This was further confirmed via pigment extraction (Fig. S2 A). In addition, a severe photoprotective phenotype, where the cells form aggregates, was observed at 100 ng/mL (Fig. 4D). This also occurred at 10 ng/mL aTc, but much less frequent and with smaller clumps (Fig. 4C). Interestingly, when culturing the strains in 6-well plates, OD_{750} was almost not affected at all (Fig. S2 B).

It is possible that the slight phenotype observed at 10 ng/mL aTc

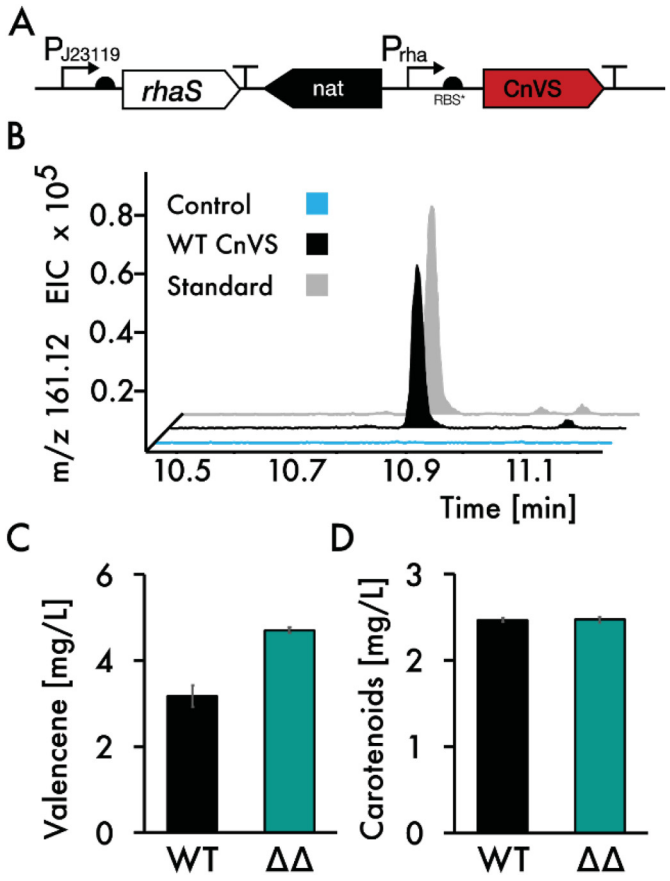


Fig. 3. Comparison of productivity between wild type and mutant. A: Construct overview. B: Detection of valencene by GC-MS analysis. Dodecane layer of the engineered Synechocystis strain (WT CnvS) after 48 h cultivation with 5 μM rhamnose induction, compared with a standard (225 μM) and the dodecane layer of the cultivated wild type strain (Extracted ion chromatogram, m/z 161.12). C: Valencene production in wild type (WT) and the Δshc/Δsqs mutant strain (ΔΔ). D: Carotenoid content in wild type (WT) and the Δshc/Δsqs mutant strain (ΔΔ). Results represent the mean of three biological replicates.

might become more severe over longer periods of time, since the expression levels appear to be saturated at the lowest concentration used. Likely, this is due to phenotypical changes taking longer than changes in expression. Another possibility is faster repression by higher aTc concentrations, resulting in a higher initial amount of CrtE protein at 10 ng/mL aTc compared with 100 ng/mL, which would take longer to be diluted out via cell division. On the other hand, since aTc is light-sensitive, the effect is likely transient and cells may recover from both their phenotype and their carotenoid deficit. Since industrial applications rely on robust strains, ideally without the necessity of adding costly inducer compounds, further fine-tuning might be of interest to achieve a constitutively downregulated *crtE* gene, while still maintaining cell viability and productivity. Since the uninduced control already shows a noticeable decrease in expression, using a stronger promoter for the control of the CRISPRi system might already be sufficient.

Nonetheless, we were able to demonstrate that tuned down-regulation of *crtE* leads to a reduction of carotenoids, while maintaining almost wild type levels of chlorophyll, as well as a wild type-like performance in terms of cell growth, and that by applying this strategy, we likely were able to enhance precursor availability for heterologous biosynthetic pathways upon introduction of alternative prenyltransferases.

3.3. Exploiting the carotenoid pool for the production of valencene

Since the newly engineered *crtE* knock-down strain lacks GGPP, but also the desired FPP precursor, we introduced the heterologous *ispA* gene from *E. coli*, which is functionally homologous to *crtE*, but unable to produce GGPP (Reiling et al., 2004).

To favor conversion of IPP and DMAPP towards valencene, we applied two strategies. First, we generated an ispA-CnVS protein fusion construct with a GGGGS linker in between to strongly increase proximity between the two enzymes the linker was chosen because it showed the most promising results in previous works (Hu et al., 2017). Second, we cloned the same genes in an operon (Fig. 5A). In the operon case, the enzymes were able to retain their full functions, while still being translated from the same mRNA, thereby optimizing spatial and temporal proximity to each other without potential compromise of function. The constructs were designated IspA:CnVS-fus and IspA:CnVS-op, respectively. In all variants, heterologous genes were controlled by the strong inducible promoter, P_{rha}. Fig. 5A outlines the construct design. To verify the production of soluble protein, we included an N-terminal FLAG-tag upstream of *ispA*. Western Blot analysis confirmed the presence of both the chimeric protein in ispA:CnVS-fus, as well as *ispA* in ispA:CnVS-op (Fig. S3 A). We also included a control with only CnVS to quantify the performance of the enzyme on its own in each background strain.

Cultures were grown as described earlier, and dodecane fractions were sampled after 48 h, before quantifying OD_{750} and density-adjusted spectra, as well as sampling for pigment extraction, quantification of dry cell weight (DCW).

As hypothesized, *crtE*-knockdown strains expressing only *CnVS* performed poorly in terms of valencene production. Induction of *crtE* repression via aTc led to a decrease in both valencene yield and carotenoids. This was expected, since all intermediates within the terpenoid pathway should be affected by a repression of *crtE*.

Coexpression of *ispA* and *CnVS*, both as an operon and a fusion protein, resulted in an increased amount of valencene. Especially in IspA:CnVS-op, production increased by about 3.5-fold compared to the strain expressing only *CnVS*. The increase in IspA:CnVS-fus was less apparent with a 1.7-fold change in valencene.

It is unclear at this point why the protein fusion construct had a smaller effect than the operon construct. Transcript analysis of *ispA* and *CnVS* in the two strains showed similar expression levels (Fig. S3 B); *ispA* was expressed slightly higher in the operon construct. It is therefore unlikely that different transcript levels play a role in metabolic output, although this might be a hint that it could be beneficial to find the correct balance of expression between all enzymes involved - higher levels of *ispA* lead to higher conversion of IPP and DMAPP toward FPP. The most likely reason for the poorer performance of the fusion protein is therefore a loss in efficiency due to impeded enzyme function or misfolding of the protein. Since other studies showed great promise in this area of research (Daletos and Stephanopoulos, 2020; Wang et al., 2021), it might therefore be interesting to further investigate different protein fusion constructs, for example by switching the order of the enzymes, as well as exploring different protein linkers. While the use of *ispA* in combination with *CnVS* was briefly described earlier (Matsudaira et al., 2020), we show that this combination of genetic components is even more productive in combination with metabolic engineering of the native pathways in *Synechocystis*, yielding improved levels of valencene. Strikingly, additional *crtE* repression of *ispA*-expressing strains with aTc further increased valencene titer up to 17.6 mg/L and 12.5 mg/L valencene. In contrast, the strains producing more valencene also show a noticeable reduction in carotenoid content, indicating that the pool of the precursors IPP and DMAPP, which are normally diverted towards carotenoid production are now available and successfully used as a substrate by CnVS. This is also consistent with earlier works, in which a common carotenoid precursor was diverted towards production of manoyl oxide (Englund et al., 2015). Surprisingly, despite the reduced

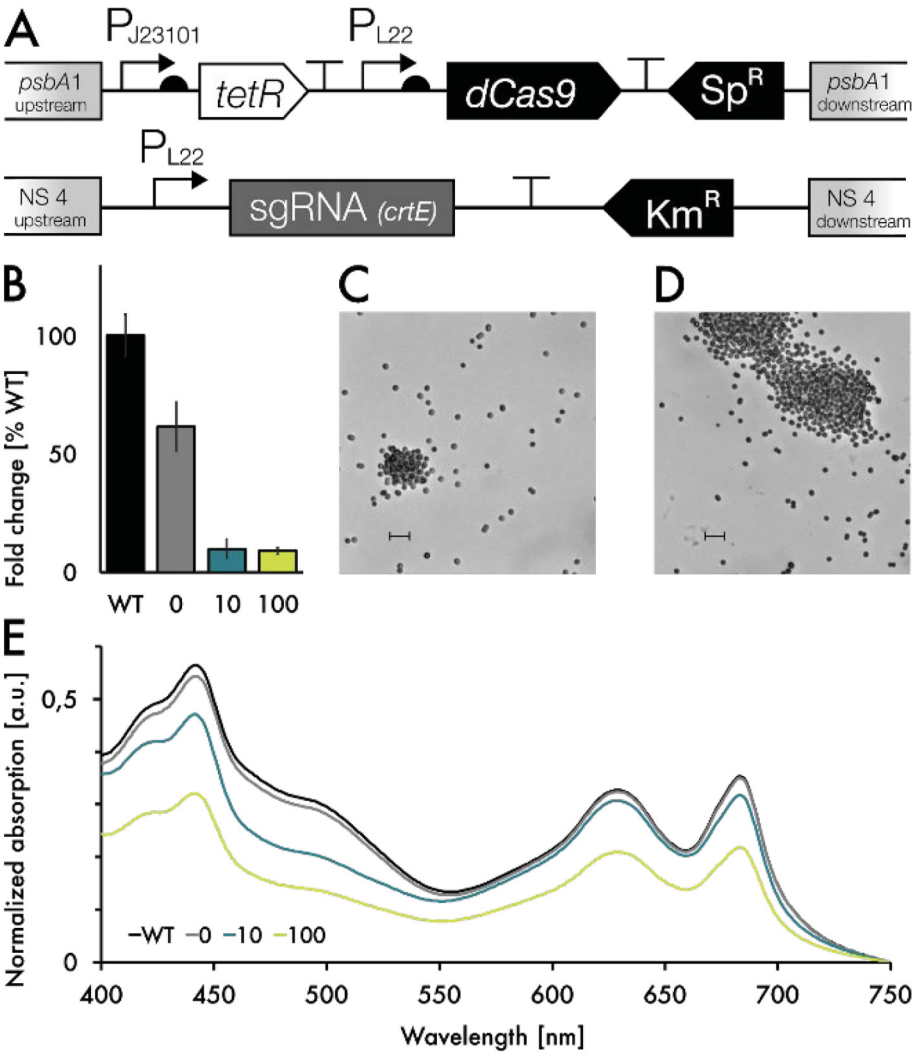


Fig. 4. CrtE gene repression in Synechocystis. A: Construct overview. B: CRISPRi knockdown of Geranylgeranyl pyrophosphate synthase (CrtE) using the PL22 promoter with 0, 10 and 100 ng/ml anhydrotetracycline (aTc). Transcripts measured by RT-qPCR after 24h of cultivation compared to the induced (100 ng/ml) control strain denoted as WT (containing only dCas9, but no sgRNA). Results represent the mean and standard deviation of three biological replicates and three technical replicates each. C/D: Bright field microscopy picture after 24 h cultivation of the strain with 10 ng/ml (C) or 100 ng/ml (D) aTc induction. Magnification ×400, scale bar 10 μm. E: Whole cell absorption spectra analysis. Cultures were adjusted for OD750 prior for measurement and values were baseline corrected. CrtE reduction leads to a blueish culture color. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

pigment content coupled with the metabolic burden of valencene production, the aTc-induced cells grew remarkably well, reaching an OD_{750} of ~2.5 compared to uninduced cells, which reached an OD_{750} of ~3 after 48 h. It is possible that aTc-mediated *crtE*-repression is, in fact, transient due to the light-sensitive properties of aTc, and that after an initial rerouting of the precursor pool towards valencene, the cell returns back to its initial balanced state. While *crtE* was expected to be an essential gene due to carotenoids being an essential part of light harvesting and photoprotection, it remains unclear at this point whether the effect is transient. Nevertheless, the decrease in carotenoid levels clearly shows the expected metabolic effect. It is therefore likely that the introduced genetic alterations function as hypothesized and that a majority of the terpenoid precursor pool is in fact diverted towards valencene production. However, the molar increase in valencene corresponds to roughly three times the amount of FPP that would be made available by the reduction of carotenoids alone. Since the phytol tail of chlorophyll is also derived from GGPP, this is likely partially responsible for the discrepancy. Furthermore, metabolic feedback regulation probably plays an important role, both within the isoprenoid biosynthetic pathway and the carotenoid pathway (Cazzonelli and Pogson, 2010). Since carotenoids are heavily involved in the response to light stress (Llewellyn et al., 2020; Steiger et al., 1999), reduced carotenoid content could lead to the accumulation of ROS, thereby possibly triggering increased flux towards GGPP.

It would be highly interesting to investigate valencene production over time in order to assess whether the generated strain produces stable metabolic output over a longer amount of time, or whether the cell returns to its pigmented state. We therefore observed the behavior of the best-performing strain, ΔΔ *crtE*↓ IspA:CnVS-op + aTc, over five days. Three replicates were precultured in 30 mL BG11 in non-baffled flasks, induced with 5 mM L-rhamnose and 10 ng/mL aTc, overlaid with 3 mL dodecane, and observed over five days. Fig. S5 shows the volumetric daily production rates of the strain, as well as total valencene accumulation and OD_{750}. While the cell density reaches a plateau after four days, valencene is continuously produced. There is a strong depletion of pigments in the production strain (Fig. S5B), both in carotenoid and chlorophyll content. Despite this strong phenotype, the cells appear to retain some level of productivity. However, the pigmentation, as well as the growth halt further indicates that the strain can be further optimized to regain some productivity likely lost due to the loss of photosynthetic efficiency.

The individual yields of each strain in terms of culture volume, dry cell weight (DCW), and cell density are summarized in Table 1.

4. Conclusion & outlook

For the redirection of metabolic flux towards the heterologous production of terpenoids, in this case the sesquiterpene valencene, we identified the native carotenoid pool of *Synechocystis* as a major target. We were able to demonstrate the capability of *Synechocystis* to divert terpene precursors by I. Deletion of native metabolic pathways not essential to the central metabolism, markerless Δ*shc* and Δ*sqs*, II.

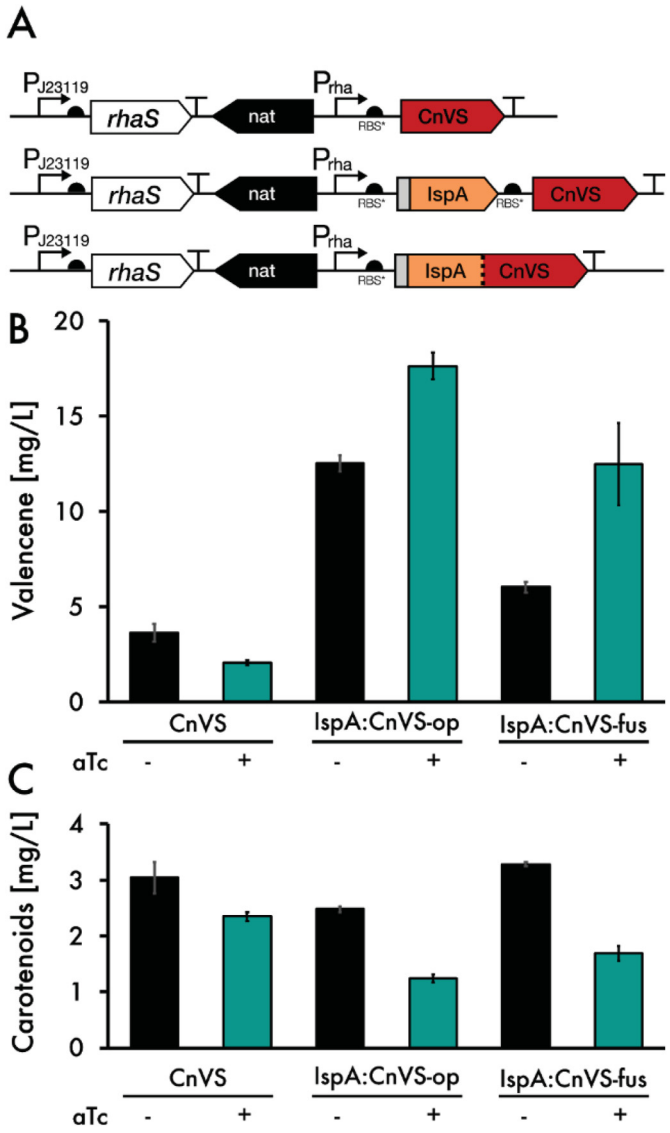


Fig. 5. Performance of the crtE repressed valence production. A: Construct overview. B: Valencene production in the CnVS, the CnVS-ispA-operon and CnVS-ispA-fusion strain. All strains are in the ΔΔ and dCas9/CrtE sgRNA background. + indicates the induction with 10 ng/ml aTc. Additionally, all strains were induced with 5 mM rhamnose. C: Carotenoid level of strains described. Results represent the mean and standard deviation of three biological replicates.

Conditional gene repression of a major component in the terpenoid pathway, *crtE*, and III. Introduction of heterologous enzymes, *ispA* and CnVS, with functions tailored to the specific production of our target molecule. With these strategies, we were able to successfully overcome some of the native pathway bottlenecks in cyanobacteria, while simultaneously exploiting their native ability of producing terpene compounds. Observing the best-producing strain over time also showed that there is even more potential towards optimization towards a more robust production strain. We believe that this delicate balance between cell viability in terms of conversion of light to energy, but also protection from light stress on the one hand, and improved productivity is an important step towards utilizing these photosynthetic organisms in a more continuous, industrial-scale application. Future studies of long-term effects of metabolic engineering of strains will certainly help improve engineering strategies towards industrially relevant utilization of cyanobacterial chassis.

Furthermore, valencene also serves as an intermediate, which can be converted to nootkatone via cytochrome P450 enzymes (CYPs). CYPs that function as monooxygenases are found in many plant species, often membrane-bound and dependent on the availability of oxygen and NADPH. In terms of functionality, cyanobacteria might be especially suitable for the application of engineered CYPs: They already contain endogenous cytochromes P450, for which oxygen and NADPH are readily available via photosynthetic activity, while this can be a limiting factor in heterotrophs.

Table 1
Individual valencene production performance of strains investigated in this work. Downward arrow represents CRISPRi-mediated repression. ΔΔ represents the Δshc, Δsqs double mutant. All values shown represent the mean ± the standard deviation of three biological replicates.

Strain	Genotype	Genes expressed from plasmid	Yield [mg/L]	Yield [mg/gDCW]	Yield [mg/OD_{750}]
WT	Non-motile wild type Synechocystis sp. PCC 6803	–	n.d.	n.d.	n.d.
Δ*shc*, Δ*sqs*	Δ*shc*, Δ*sqs*	–	n.d.	n.d.	n.d.
ΔΔ *crtE*↓	Δ*shc*, Δ*sqs*, Δ psbA1 :: *crtE*↓	–	n.d.	n.d.	n.d.
WT CnVS	–	CnVS	3.2 ± 0.25	4.5 ± 0.43	1.7 ± 0.13
ΔΔ CnVS	Δ*shc*, Δ*sqs*	CnVS	4.7 ± 0.06	6.4 ± 0.52	2.4 ± 0.06
ΔΔ *crtE*↓CnVS-aTc	Δ*shc*, Δ*sqs*, *crtE*↓	CnVS	3.6 ± 0.47	3.7 ± 0.42	1.5 ± 0.18
ΔΔ *crtE*↓CnVS + aTc	Δ*shc*, Δ*sqs*, *crtE*↓	CnVS	2.0 ± 0.12	2.3 ± 0.18	0.9 ± 0.06
ΔΔ *crtE*↓CnVS-op-aTc	Δ*shc*, Δ*sqs*, *crtE*↓	ispA, CnVS (operon)	12.5 ± 0.44	9.8 ± 0.54	3.9 ± 0.12
ΔΔ *crtE*↓CnVS-op + aTc	Δ*shc*, Δ*sqs*, *crtE*↓	ispA, CnVS (operon)	17.6 ± 0.71	19.0 ± 0.62	7.1 ± 0.14
ΔΔ *crtE*↓CnVS-fus-aTc	Δ*shc*, Δ*sqs*, *crtE*↓	ispA, CnVS (fusion)	6.0 ± 0.27	4.9 ± 0.25	1.9 ± 0.08
ΔΔ *crtE*↓CnVS-fus + aTc	Δ*shc*, Δ*sqs*, *crtE*↓	ispA, CnVS (fusion)	12.5 ± 2.15	12.6 ± 2.18	5.0 ± 0.84

Funding

This project was financially supported by the CLIB-Competence Centre Biotechnology (CKB) funded by the European Regional Development Fund ERDF (34.EFRE-0300096).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author statement

Maximilian Dietsch: Conceptualization, Investigation, Methodology, Writing - Original Draft, Writing - Review & Editing, Visualization, Data Curation. Anna Behle: Conceptualization, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization, Data Curation. Philipp Westhoff: Methodology, Data Curation. Ilka M. Axmann: Supervision, Writing- Reviewing and Editing.

Acknowledgements

Metabolite analyses were supported by the CEPLAS Plant

Metabolism and Metabolomics laboratory, which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2048/1 – project ID 390686111. We acknowledge the excellent technical assistance by E. Klemp for GC-MS measurements.

We appreciate the help of Thomas Drepper, Oliver Klaus and Arsenij Kokorin with discussion and technical advice.

We thank Paul Hudson for providing us with the dCas9 plasmid, Nicolas Schmelling for critical comments on the manuscript, Dennis Dienst for helpful advice on cultivation issues and Jungho Lee for providing us with the Nourseothricin resistance gene.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mec.2021.e00178.

References

Albertsen, L., Chen, Y., Bach, L.S., Rattleff, S., Maury, J., Brix, S., Nielsen, J., Mortensen, U.H., 2011. Diversion of flux toward sesquiterpene production in *Saccharomyces cerevisiae* by fusion of host and heterologous enzymes. Appl. Environ. Microbiol. 77, 1033–1040.

Angermayr, S.A., Gorchs Rovira, A., Hellingwerf, K.J., 2015. Metabolic engineering of cyanobacteria for the synthesis of commodity products. Trends Biotechnol. 33, 352–361.

Beekwilder, J., van Houwelingen, A., Cankar, K., van Dijk, A.D.J., de Jong, R.M., Stoopen, G., Bouwmeester, H., Achkar, J., Sonke, T., Bosch, D., 2014. Valencene synthase from the heartwood of Nootka cypress (*Callitropsis nootkatensis*) for biotechnological production of valencene. Plant Biotechnol. J. 12, 174–182.

Behle, A., Saake, P., Germann, A.T., Dienst, D., Axmann, I.M., 2020. Comparative dose–response analysis of inducible promoters in cyanobacteria. ACS Synth. Biol. 9, 843–855.

Cazzonelli, C.I., Pogson, B.J., 2010. Source to sink: regulation of carotenoid biosynthesis in plants. Trends Plant Sci. 15, 266–274.

Chen, H., Zhu, C., Zhu, M., Xiong, J., Ma, H., Zhuo, M., Li, S., 2019. High production of valencene in *Saccharomyces cerevisiae* through metabolic engineering. Microb. Cell Factories 18, 195.

Dahl, R.H., Zhang, F., Alonso-Gutierrez, J., Baidoo, E., Batth, T.S., Redding-Johanson, A. M., Petzold, C.J., Mukhopadhyay, A., Lee, T.S., Adams, P.D., et al., 2013. Engineering dynamic pathway regulation using stress-response promoters. Nat. Biotechnol. 31, 1039–1046.

Daletos, G., Stephanopoulos, G., 2020. Protein engineering strategies for microbial production of isoprenoids. Metab. Eng. Commun. 11, e00129.

Dienst, D., Wichmann, J., Mantovani, O., Rodrigues, J.S., Lindberg, P., 2020. High density cultivation for efficient sesquiterpenoid biosynthesis in *Synechocystis* sp. PCC 6803. Sci. Rep. 10, 5932.

Englund, E., Pattanaik, B., Ubhayasekera, S.J.K., Stensjö, K., Bergquist, J., Lindberg, P., 2014. Production of squalene in *Synechocystis* sp. PCC 6803. PloS One 9, e90270.

Englund, E., Andersen-Ranberg, J., Miao, R., Hamberger, B., Lindberg, P., 2015. Metabolic engineering of *Synechocystis* sp. PCC 6803 for production of the plant diterpenoid manoyl oxide. ACS Synth. Biol. 4, 1270–1278.

Hu, Y., Zhou, Y.J., Bao, J., Huang, L., Nielsen, J., Krivoruchko, A., 2017. Metabolic engineering of *Saccharomyces cerevisiae* for production of germacrene A, a precursor of beta-elemene. J. Ind. Microbiol. Biotechnol. 44, 1065–1072.

Jodlbauer, J., Rohr, T., Spadiut, O., Mihovilovic, M.D., Rudroff, F., 2021. Biocatalysis in green and blue: cyanobacteria. Trends Biotechnol., S0167779920303346

Lee, J., Rai, P.K., Jeon, Y.J., Kim, K.-H., Kwon, E.E., 2017. The role of algae and cyanobacteria in the production and release of odorants in water. Environ. Pollut. 227, 252–262.

Lee, J., Hilgers, F., Loeschke, A., Jaeger, K.-E., Feldbrügge, M., 2020. *Ustilago maydis* serves as a novel production host for the synthesis of plant and fungal sesquiterpenoids. Front. Microbiol. 11, 1655.

Lichtenthaler, H.K., Buschmann, C., 2001. Chlorophylls and carotenoids: measurement and characterization by UV-VIS spectroscopy. Curr. Protoc. Food Anal. Chem. 1, F4.3.1–F4.3.8.

Lin, P.-C., Saha, R., Zhang, F., Pakrasi, H.B., 2017. Metabolic engineering of the pentose phosphate pathway for enhanced limonene production in the cyanobacterium *Synechocysti* s sp. PCC 6803. Sci. Rep. 7.

Liu, Y., Cruz-Morales, P., Zargar, A., Belcher, M.S., Pang, B., Englund, E., Dan, Q., Yin, K., Keasling, J.D., 2021. Biofuels for a sustainable future. Cell, S0092867421000957.

Llewellyn, C.A., Airs, R.L., Farnham, G., Greig, C., 2020. Synthesis, regulation and degradation of carotenoids under low level UV-B radiation in the filamentous cyanobacterium *Chlorogloeopsis fritschii* PCC 6912. Front. Microbiol. 11, 163.

Loeschcke, A., Dienst, D., Wewer, V., Hage-Hülsmann, J., Dietsch, M., Kranz-Finger, S., Hüren, V., Metzger, S., Urlacher, V.B., Gigolashvili, T., et al., 2017. The photosynthetic bacteria *Rhodobacter capsulatus* and *Synechocystis* sp. PCC 6803 as new hosts for cyclic plant triterpene biosynthesis. PloS One 12, e0189816.

Matsudaira, A., Hoshino, Y., Uesaka, K., Takatani, N., Omata, T., Usuda, Y., 2020. Production of glutamate and stereospecific flavors, (S)-linalool and (+)-valencene, by *Synechocystis* sp. PCC6803. J. Biosci. Bioeng. 130, 464–470.

Newman, J.D., Marshall, J., Chang, M., Nowroozi, F., Paradise, E., Pitera, D., Newman, K. L., Keasling, J.D., 2006. High-level production of amorpha-4,11-diene in a two-phase partitioning bioreactor of metabolically engineered *Escherichia coli*. Biotechnol. Bioeng. 95, 684–691.

Pinto, F., Thapper, A., Sontheim, W., Lindblad, P., 2009. Analysis of current and alternative phenol based RNA extraction methodologies for cyanobacteria. BMC Mol. Biol. 10, 79.

Rabeharindranto, H., Castaño-Cerezo, S., Lautier, T., Garcia-Alles, L.F., Treitz, C., Tholey, A., Truan, G., 2019. Enzyme-fusion strategies for redirecting and improving carotenoid synthesis in S. cerevisiae. Metab. Eng. Commun. 8, e00086.

Reiling, K.K., Yoshikuni, Y., Martin, V.J.J., Newman, J., Bohlmann, J., Keasling, J.D., 2004. Mono and diterpene production in *Escherichia coli*. Biotechnol. Bioeng. 87, 200–212.

Stanier, R.Y., Deruelles, J., Rippka, R., Herdman, M., Waterbury, J.B., 1979. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. Microbiology 111, 1–61.

Steiger, S., Schäfer, L., Sandmann, G., 1999. High-light-dependent upregulation of carotenoids and their antioxidative properties in the cyanobacterium *Synechocystis* PCC 6803. J. Photochem. Photobiol., B 52, 14–18.

Tsuruta, H., Paddon, C.J., Eng, D., Lenihan, J.R., Horning, T., Anthony, L.C., Regentin, R., Keasling, J.D., Renninger, N.S., Newman, J.D., 2009. High-level production of amorpha-4,11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. PloS One 4, e4489.

Viola, S., Rühle, T., Leister, D., 2014. A single vector-based strategy for marker-less gene replacement in *Synechocystis* sp. PCC 6803. Microb. Cell Factories 13, 4.

Wang, X., Pereira, J.H., Tsutakawa, S., Fang, X., Adams, P.D., Mukhopadhyay, A., Lee, T. S., 2021. Efficient production of oxidized terpenoids via engineering fusion proteins of terpene synthase and cytochrome P450. Metab. Eng. 64, 41–51.

Yao, L., Cengic, I., Anfelt, J., Hudson, E.P., 2016. Multiple gene repression in cyanobacteria using CRISPRi. ACS Synth. Biol. 5, 207–212.

Yao, L., Shabestary, K., Björk, S.M., Asplund-Samuelsson, J., Joensson, H.N., Jahn, M., Hudson, E.P., 2020. Pooled CRISPRi screening of the cyanobacterium Synechocystis sp PCC 6803 for enhanced industrial phenotypes. Nat. Commun. 11, 1666.

Metabolic engineering of *Synechocystis* sp. PCC 6803 for the photoproduction of the sesquiterpene valencene

Maximilian Dietsch[1*], Anna Behle[1*], Philipp Westhoff[2], Ilka M. Axmann[1]

[1]Institute for Synthetic Microbiology, Department of Biology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.
[2]Plant Metabolism and Metabolomics Laboratory, Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University Düsseldorf, D-40001 Düsseldorf, Germany
*These authors contributed equally to this work

Supporting Information

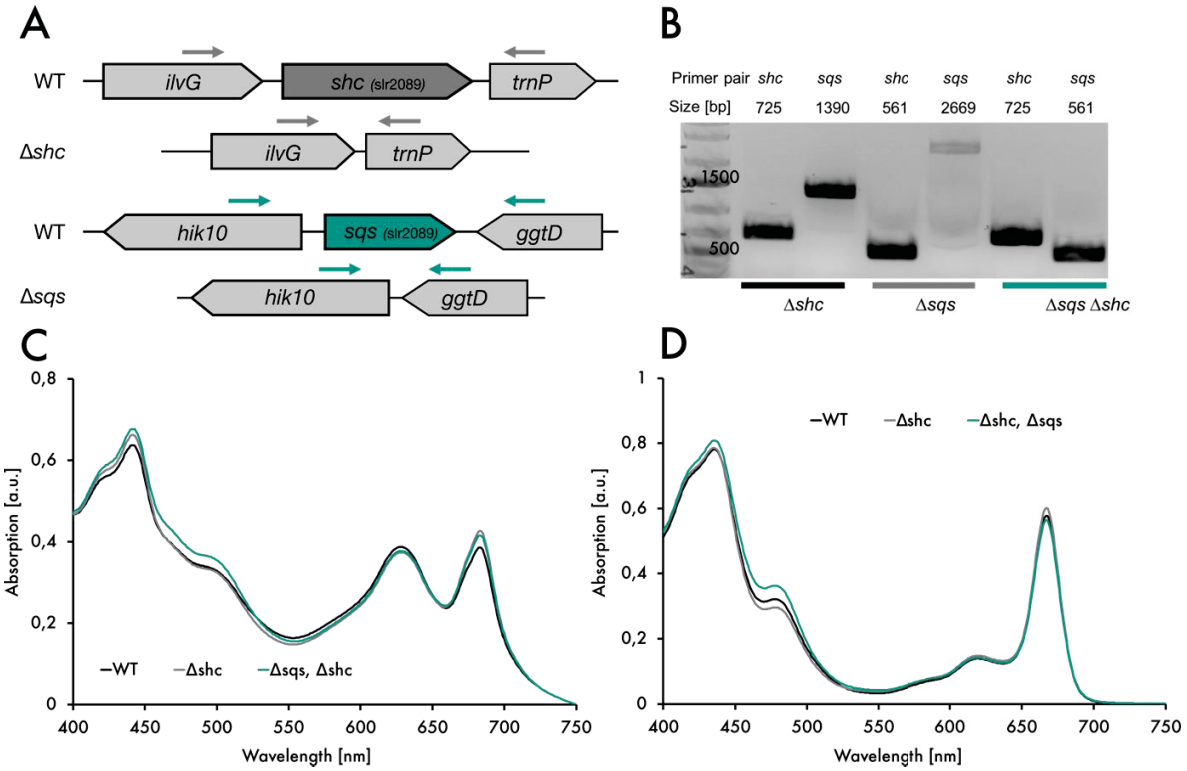


Fig. S1: Supplementary information on the markerless mutants Δ*shc* and Δ*shc*, Δ*sqs*.
A: Schematic overview of markerless mutant genotypes. Arrows denote primers used for colony PCR.
B: PCR analysis of single and double mutant strains using oligonucleotides that bind outside of the affected area. Primer pair and expected sizes are shown above. Thermo 1kb+ ladder was used as size standard. C: Whole cell spectra of WT, Δ*shc* and double mutant. Spectra were baseline-corrected by subtracting the absorption at 750 nm. D: Spectra of methanol-extracted cells from WT, Δ*shc* and double mutant strains.

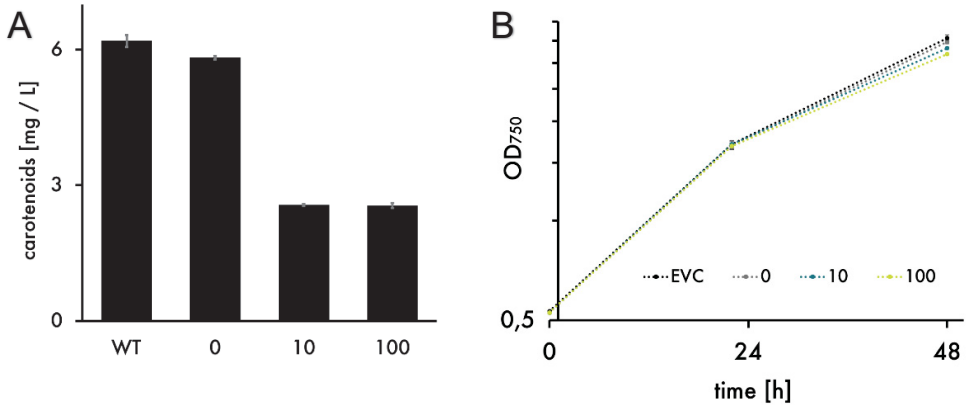


Fig. S2: Pigment quantification and growth behavior of the *crtE* knock-down strain.
A: Carotenoid content of Δ*shc*, Δ*sqs* mutant expressing dCas9 only (WT) compared to Δ*shc*, Δ*sqs* mutant expressing both dCas9 and the *crtE* sgRNA, induced with 0, 10, and 100 ng/mL aTc. Carotenoids were quantified as described in Material & Methods, section 2.5. B: Growth behavior of aforementioned strains.

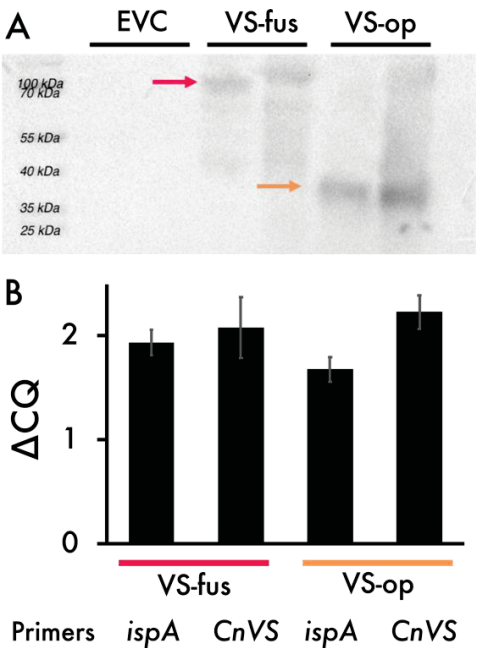


Fig. S3: Western Blot and qRT-PCR analysis of IspA:CnVS fusion vs. operon strains. A: Western Blot analysis of IspA:CnVS protein fusion (VS-fus) and IspA:CnVS operon (VS-op). The fusion protein N-FLAG-IspA-CnVS corresponds to a size of ~105 kDa, while N-FLAG-IspA in the operon construct corresponds to about ~35 kDa. B: ΔCQ values of qRT-PCR performed on both strains using either *ispA* or *CnVS* primers, as denoted below. ΔCQ values were calculated by subtracting the CQ value of the housekeeping gene *rnpB* from each CQ value. A higher ΔCQ value corresponds to a lower transcript amount.

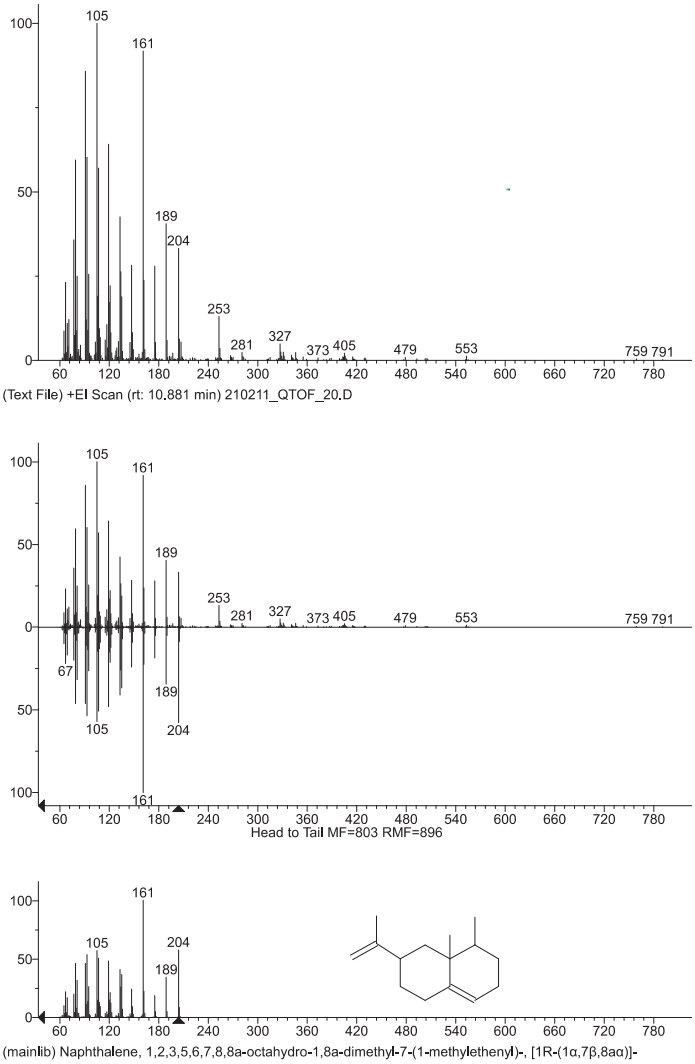


Fig. S4: Mass spectra comparison of samples with reference.
Top: Mass spectra of wild type expressing CnVS. Bottom: Reference mass spectra of (+)-valencene. Middle: Direct comparison of top and bottom spectra.

Table S1: Cq values for *shc* and *sqs* in WT and knock-out strains.
Values represent the mean and standard deviation of three biological replicates. Cq values were obtained via qRT-PCR. Sample values above 30 were defined as not containing any template.

Strain	Target gene	Cq value
Wild type	*sqs*	21.2 ± 0.3
	shc	21.8 ± 0.2
Δ*shc*	*sqs*	21.1 ± 0.2
	shc	33.4 ± 0.6
Δ*shc*; Δ*sqs*	*sqs*	31.5 ± 0.8
	shc	35.5 ± 1.2
No template control	*sqs*	34.4
	shc	n. def.

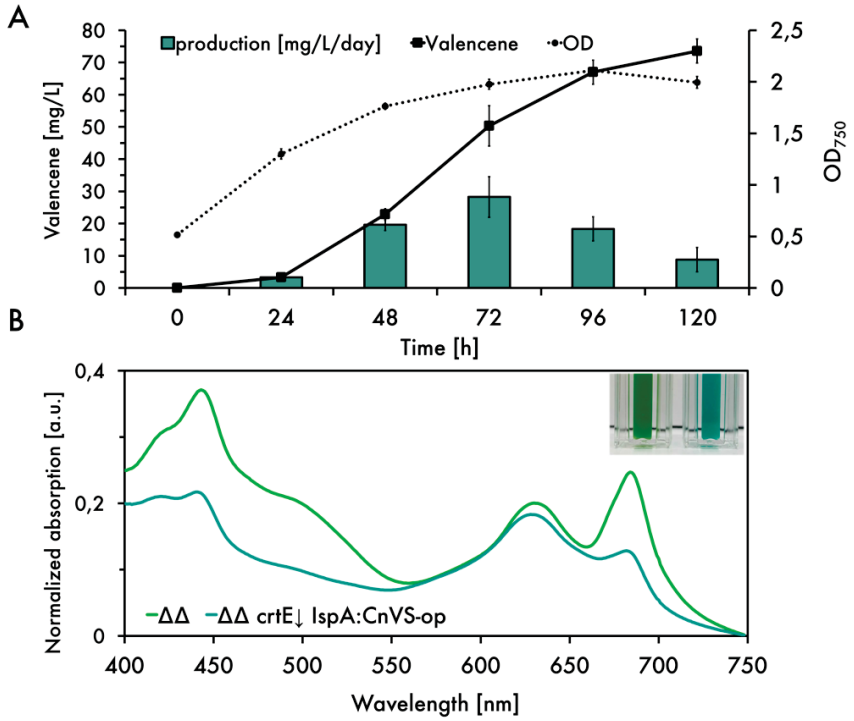


Fig. S1: Physiological changes and valencene production in ΔΔ *crtE*↓ IspA:CnVS-op +aTc. The strain was cultured in biological triplicates over five days in shake flasks overlaid with 10% dodecane. The dodecane layer, as well as the culture, were sampled daily for valencene quantification and cell density (OD_{750}), respectively. A: Volumetric accumulation of valencene (square symbols, continuous line) and cell density (round symbols, dotted line). The volumetric production per day is shown as blue bars, corrected for the sample removed each day. B: Whole cell spectra of double mutant (green) and ΔΔ *crtE*↓ IspA:CnVS-op +aTc (blue) after 120 h cultivation. Spectra were baseline-corrected by subtracting the absorption at 750 nm. An image of the cuvettes is embedded for better visualization of the color difference (left: double mutant, right: ΔΔ *crtE*↓ IspA:CnVS-op +aTc).

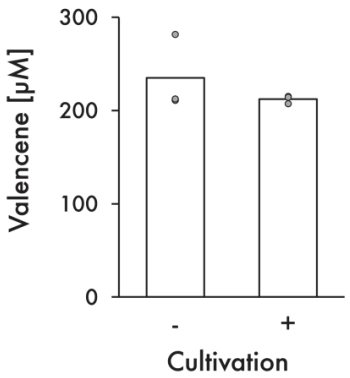


Fig. S2: Quantification of possible valencene loss via evaporation or degradation. A dodecane sample containing 225 µM valencene was used to overlay a *Synechocystis* wild type culture, which was grown for 48 h in technical triplicates. Both the initial sample (-) and the sample recovered from the culture (+) was measured and compared. The bars represent the mean of the three technical replicates, each of which is shown in grey. No significant difference was observed (t-test, P=0.39).

Table S2: Detailed descriptions and sequences of all relevant genetic modules used in this work.

Name	Part type	Sequence	Origin	Notes
P_{rha}	Promoter	gccacaattcagcaaattgtgaacatcatcacgttcatctttccctggttgccaatggcccattttcctgtcagtaacgagaaggtcgcgaattcaggcgctttttagactggtcgtaatgaa	(Behle *et al.*, 2020)	
RBS*	RBS	tactagagtagtggaggttactag	(Behle *et al.*, 2020)	
P_{J23119}	Promoter	ttgacagctagctcagtcctaggtataatgctagc	(Behle *et al.*, 2020)	Bba_J23119 in iGEM registry
Bba_0034	RBS	aaagaggagaaatactag	(Behle *et al.*, 2020)	Bba_B0034 in iGEM registry
rhaS	CDS	atgaccgtattacatagtgtggatttttttccgtctggtaacgcgtccgtggcgatagaaccccggctcccgcaggcggatttttcctgaacatcatcatgattttcatgaaattgtgattgtcgaacatggc acgggtattcatgtgtttaatgggcagccctataccatcaccggtggcacggtctgtttcgtacgcgatcatgatcggcatctgtatgaacataccgataatctgtgtctgaccaatgtgctgtatcgctc gccggatcgatttcagtttctcgccgggctgaatcagttgctgccacaagagctggatgggcagtatccgtctcactggcgcgttaaccacagcgtattgcagcaggtgcgacagctggttgcaca gatggaacagcaggaaggggaaaatgatttaccctcgaccgccagtcgcgagatcttgtttatgcaattactgctcttgctgcgtaaaagcagtttgcaggagaacctggaaaacagcgcatcac gtctcaacttgcttctggcctggctggaggaccattttgccgatgaggtgaattgggatgccgtggcggatcaattttctctttcactgcgtacgctacatcggcagcttaagcagcaaacgggactga cgcctcagcgatacctgaaccgcctgcgactgatgaaagcccgacatctgctacgccacagcgaggccagcgttactgacatcgcctatcgctgtggattcagcgacagtaaccacttttcgacg ctttttcgccgagagtttaactggtcaccgcgtgatattcgccagggacgggatggctttctgcaataa	*E. coli*	Arg214Leu mutation
natR	CDS	ctaggggcagggcatgctcatgtagagcgcctgccgctcgccgtccgaggcggtgccgtcgtacagggcggtgtccaggccgcagagggtgaaccccatccgccggtacgcgtggatcgcc ggtgcgttgacgttggtgacctccagccagaggtgcccggcgccccgctcgccggcgaactccgtcgcgagccccatcaacgcgcgcccgaccccgtgcccccggtgctccgggggcgacctc gatgtcctcgacggtcagccggcggttccacgccgagtacgagatgaccacgaagcccgccaggtcgccgtcgtccccgtacgcgacgaacgtccgggagtccgggtcgccgtcctccccgt cgtccgattcgtcgtccgattcgtcgtcggggaacaccttggtcaggggcgggtccaccggcacctcccgcagggtgaagccgtccccggtggcggtgacgcggaagacggtgtcggtggtga aggacccatccagtgcctcgatggcctcggcgtccccgggacactggtgcggtagcggtaagccgtgtcgtcaagagtggtcat	*Streptomyces*	
N-FLAG	N-terminal FLAG tag	ATGGCTAGCgattataaagatcatgatggcgattataaagatcatgatattgattataaagatgatgatgataaaGGT	(Wiegard *et al.*, 2013)	Contains an NheI restriction site after the start codon; and a single glycin linker at the end (ggt)
ispA	CDS	atggacttcccgcagcagctagaggcctgtgtgaagcaagcaaatcaagcgctcagccgttttatcgccccactacccttccaaaaacactcctgtcgtggaaactatgcaatacggcgcgttactg gaggagcaaacgattgcggccgtttttggtctatgccaccggccacatgttcggagtgagcaccaacacactcgacgctccagctgcggcagttgagtgtattcacgcatattctttgatccatgatga tctacccgcaatggatgacgatgacctgcgtcgaggcttacctacttgtcacgtcaagttcggcgaagcgaatgcaattctagcgggcgacgcgttgcaaacttttggctttctccatttttatccgatgc cgatatgccggaagtctctgaccgcgaccgaatttccatgatctctgaactcgccagcgcgagcggaatcgccggaatgtgcggtggtcaagccctggacttggatgcggaaggaaaacacgtt cctctcgatgcgttagaacgaatccatcgtcataagactggtgctctcatccgtgccgcagttcgactcggggcgttgagcgctggtgacaagggacgacgcgcttttacctgtgctcgacaaatac gctgagtctattggtctcgcattccaagttcaggacgacattctcgacgttgtgggagacacggccacgctggggaaacggcagggtgcggaccagcagctcgggaaaagcacgtacccggct ttactcggtctcgaacaagcacgcaagaaggcgcgggacttgattgatgatgctcgccagtccctgaagcagttagcagaacagtctctcgacacttccgctttggaggctctcgcggacatatatta ttcagcggaataaataa	*Callitropsis nootkatensis*	Gene was codon optimized for *Synechocystis* sp. PCC 6803
CnVS	CDS	atggcggaaatgtttaatgggaacagcagtaatgacgggtcctcttgtatgcctgtgaaggacgctctccgtcggactggaaaccatcatccaaacttatggacggatgactttatccaaagcctga atagtccatattctgactctagttatcacaaacatcgagaaattctaatcgacgaaatccgcgacatgtttagcaatggtgagggtgatgaatttggcgtgttggaaaacatctggtttgtggatgtagt acagcggctcgggattgaccgtcatttccaggaagagatcaaaacagcactggattacatctataaattttggaaccacgattctatttttggggatctgaacatggtcgctctcggtttttcgtattttgcg gctaaatcgttacgtagcgagtagcgatgtgttcaaaaaattcaagggtgaggagggccaattttctggcttcgagtccagtgaccaagacgccaagttggaaatgatgttaaatttgtataaggca tccgaattggatttcccagacgaggatattctaaaggaggcccgcgcgttcgcaagtatgtacctcaagcacgttatcaaggaatatgggacatccaagaaagcaagaacccactcctaatgg aaattgaatacacttttaagtatccgtggcgctgtcgcttaccacgtttagaggcgtggaacttcatccatattatgcgccaacaggactgtaatattagcctggcgaacaacctatacaagattccaa aaatttacatgaaaaaattctggagttggccattctggatttcaacattctacagagtcagcatcaacacgagatgaagctcatctctacatggtggaaaaactcttccgctattcaactcgatttttttc gccatcgacatatcgagagctacttttggtgggcctccccctatttgaacctgagttctccacctgtcgcattaactgcacgaagctgagcactaagatgtttctattggacgacatctatgatacgtat ggtacggtcgaggaattaaagccttttacgacaactctaactcgatgggatgtatccacggttgataatcatccggattatatgaaaattgcgttcaatttttcctacgaaatctataaagagattgcatc cgaagccgagcgtaagcacgggccttttgtttacaaatatctgcaaagttgttggaagagttacattgaagcatatatgcaggaagcggagtggattgcttccaaccacattcctggattcgatgaat acctcatgaacggtgtcaagtcctccggtatgcgtattttgatgattcatgctctcatcctcatggatacgccattgagcgatgaaatcttagaacagctggacattccaagttctaaaagtcaggcgct cttgagcctaatcacacggctagtcgatgacgtcaaggatttcgaagatgagcaggcccacggagagatggcaagtagtatcgagtgttatatgaaagataaccacggttccacacgggagga tgcgttaaactacctccaagatccggatcgagagctgtgttcaggaactaaataaagaactgctagaaccaagcaacatgcacggcagtttccggaacttgtatttgaacgtaggaatgcgtgtcat cttcttcatgctcaacgacggtgatttgtttacgcattccaaccgtaaagaaatccaagacgcaatcactaaatttttgtcgaacccattatcccgaagctctaa		
GGGGS	protein linker	ggcggtggcggatcc	(Hu *et al.*, 2017)	
sgRNA *crtE*	sgRNA	AAGTCGGTTCGTGTTTGTT		
	total sgRNA part	tccctatcagtgatagagattgacatccctatcagtgatagagatactgggagctaaagtcggttcgtgtttgttgttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaa aaagtggcaccgagtcggtgctttttt	(Yao *et al.*, 2020)	P_{L22}, sgRNA, dCas9-binding
T7-term	Terminator	ggctcaccttcgggtgggcctttctgcg		

Supplementary References

Behle, Anna, Pia Saake, Anna T. Germann, Dennis Dienst, und Ilka M. Axmann. „Comparative Dose–Response Analysis of Inducible Promoters in Cyanobacteria“. *ACS Synthetic Biology* 9, Nr. 4 (17. April 2020): 843–55. https://doi.org/10.1021/acssynbio.9b00505.

Hu, Yating, Yongjin J. Zhou, Jichen Bao, Luqi Huang, Jens Nielsen, und Anastasia Krivoruchko. „Metabolic Engineering of *Saccharomyces Cerevisiae* for Production of Germacrene A, a Precursor of Beta-Elemene“. *Journal of Industrial Microbiology & Biotechnology* 44, Nr. 7 (Juli 2017): 1065–72. https://doi.org/10.1007/s10295-017-1934-z.

Wiegard, Anika, Anja K. Dörrich, Hans-Tobias Deinzer, Christian Beck, Annegret Wilde, Julia Holtzendorff, und Ilka M. Axmann. „Biochemical Analysis of Three Putative KaiC Clock Proteins from *Synechocystis* Sp. PCC 6803 Suggests Their Functional Divergence“. *Microbiology* 159, Nr. Pt_5 (1. Mai 2013): 948–58. https://doi.org/10.1099/mic.0.065425-0.

Yao, Lun, Kiyan Shabestary, Sara M. Björk, Johannes Asplund-Samuelsson, Haakan N. Joensson, Michael Jahn, und Elton P. Hudson. „Pooled CRISPRi Screening of the Cyanobacterium *Synechocystis* Sp PCC 6803 for Enhanced Industrial Phenotypes“. *Nature Communications* 11, Nr. 1 (Dezember 2020): 1666. https://doi.org/10.1038/s41467-020-15491-7.

4 Chapter 4 Manuscript II

4.1 Author's contributions

Anna T. Germann1, Andreas Nakielski1, Maximilian Dietsch1, Tim Petzel1, Daniel Moser2, Sebastian Triesch3, Philipp Westhoff4, Ilka M. Axmann1 ''A systematic overexpression approach reveals native targets to increase squalene production in *Synechocystis*,'' Submitted to frontiers, in revision

Maximilian Dietsch: Conceptualization, Methodology, Experimental work, Writing – review & editing

M.D. designed the experimental part of the paper with input from A.N. and A.T.G.
M.D constructed all plasmids and strains used in this publication. (with input from A.N., D.M. and S.T.). In cooperation with A.N., he performed the cultivation and subsequent qRT-PCR experiments. M.D. was involved in establishing the method for squalene quantification in cooperation with P.W. Literature research for preparation of the manuscript was supported by M.D.

The attribution categories described for each journal have been summarized as following:

Study design: 40%
Experimental contribution: 30%
Data analysis: 0%
Manuscript preparation: 10%

4.2 A systematic overexpression approach reveals native targets to increase squalene production in *Synechocystis*

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 30 May 2023
DOI 10.3389/fpls.2023.1024981

OPEN ACCESS

EDITED BY
Talia Karasov,
The University of Utah, United States

REVIEWED BY
Saowarath Jantaro,
Chulalongkorn University, Thailand
Pia Lindberg,
Uppsala University, Sweden

*CORRESPONDENCE
Ilka M. Axmann
ilka.axmann@hhu.de

†These authors have contributed equally to this work and share first authorship

RECEIVED 22 August 2022
ACCEPTED 28 April 2023
PUBLISHED 30 May 2023

CITATION
Germann AT, Nakielski A, Dietsch M, Petzel T, Moser D, Triesch S, Westhoff P and Axmann IM (2023) A systematic overexpression approach reveals native targets to increase squalene production in *Synechocystis* sp. PCC 6803. *Front. Plant Sci.* 14:1024981. doi: 10.3389/fpls.2023.1024981

A systematic overexpression approach reveals native targets to increase squalene production in *Synechocystis* sp. PCC 6803

Anna T. Germann[1†], Andreas Nakielski[1†], Maximilian Dietsch[1], Tim Petzel[1], Daniel Moser[2], Sebastian Triesch[3], Philipp Westhoff[4] and Ilka M. Axmann[1*]

[1]Institute for Synthetic Microbiology, Department of Biology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, [2]Institute for Plant Sciences and Cluster of Excellence on Plant Sciences (CEPLAS), University of Cologne, Cologne, Germany, [3]Institute of Plant Biochemistry, Cluster of Excellence on Plant Science (CEPLAS), Heinrich Heine University, Düsseldorf, Germany, [4]Plant Metabolism and Metabolomics Laboratory, Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Cyanobacteria are a promising platform for the production of the triterpene squalene (C30), a precursor for all plant and animal sterols, and a highly attractive intermediate towards triterpenoids, a large group of secondary plant metabolites. *Synechocystis* sp. PCC 6803 natively produces squalene from CO_2 through the MEP pathway. Based on the predictions of a constraint-based metabolic model, we took a systematic overexpression approach to quantify native *Synechocystis* gene's impact on squalene production in a squalene-hopene cyclase gene knock-out strain (Δ*shc*). Our *in silico* analysis revealed an increased flux through the Calvin-Benson-Bassham cycle in the Δ*shc* mutant compared to the wildtype, including the pentose phosphate pathway, as well as lower glycolysis, while the tricarboxylic acid cycle predicted to be downregulated. Further, all enzymes of the MEP pathway and terpenoid synthesis, as well as enzymes from the central carbon metabolism, Gap2, Tpi and PyrK, were predicted to positively contribute to squalene production upon their overexpression. Each identified target gene was integrated into the genome of *Synechocystis* Δ*shc* under the control of the rhamnose-inducible promoter P_{rha}. Squalene production was increased in an inducer concentration dependent manner through the overexpression of most predicted genes, which are genes of the MEP pathway, *ispH*, *ispE*, and *idi*, leading to the greatest improvements. Moreover, we were able to overexpress the native squalene synthase gene (*sqs)* in *Synechocystis* Δ*shc*, which reached the highest production titer of 13.72 mg l^{-1} reported for squalene in *Synechocystis* sp. PCC 6803 so far, thereby providing a promising and sustainable platform for triterpene production.

KEYWORDS

Synechocystis, squalene, MEP pathway, FBA, metabolic engineering, metabolic modeling

10.3389/fpls.2023.1024981

Introduction

Cyanobacteria are the only known prokaryotes capable of oxygenic photosynthesis (Mulkidjanian et al., 2006; Lau et al., 2015). The gram-negative bacteria exhibit a large ecological variety as well as a broad morphological diversity (Bennett and Bogorad, 1973; Rippka et al., 1979; Schirrmeister et al., 2013). Their physiological diversity makes them promising biological chassis for the synthesis of a variety of natural products, including bioactive metabolites like cytotoxins and potential pharmaceutical lead compounds, food supplements, animal feed, pigments, as well as biofuels (Pulz and Gross, 2004; Hays and Ducat, 2015; Jain et al., 2017; Hudson et al., 2021; Barone et al., 2023). Their ability to convert sunlight and atmospheric CO_2 directly into valuable organic compounds could make the chemical and pharmaceutical industry more sustainable and therefore mitigate climate change if high production yields are achieved (Choi et al., 2020; Posten and Schaub, 2009; Oliver and Atsumi, 2015; Oliver et al., 2016).

Metabolic engineering tools for the production of desired compounds are particularly well established for laboratory model strains like *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*) (Knoop and Steuer, 2015; Ramey et al., 2015). The unicellular organism was the first entirely sequenced cyanobacterium (Kaneko et al., 1996) and is one of the best characterized model organisms regarding cyanobacterial biosynthesis (Pils and Schmetterer, 2001; Leplat et al., 2013). *Synechocystis* is easy and inexpensive to cultivate, and is genetically modifiable with high success rates and predictability (Rippka et al., 1979; Berla et al., 2013).

One promising class of compounds to be produced in cyanobacteria are terpenoids, a large heterogeneous group of naturally occurring organic carbon compounds with over 80,000 known structures (Karunanithi and Zerbe, 2019), having applications in nutrition, medicine and chemistry but also as potential biofuels. Triterpenoids are a group of secondary metabolites, which are composed of 6 isoprene units (Moss et al., 1995), existing in a huge variety of structures with nearly 200 distinct triterpene skeletons, all deriving from the precursor squalene (Xu et al., 2004; Connolly and Hill, 2010). Squalene is typically extracted from shark liver oil, but this method poses serious ecological risks and is not sufficient to sustainably meet increasing demands (Gohil et al., 2019; O'Hagan et al., 2021; Mendes et al., 2022). The diverse applications of squalene include its use as an ingredient in cosmetic products (Gohil et al., 2019), as an antioxidant (Kohno et al., 1995) and an emulsion adjuvant in vaccines (O'Hagan et al., 2021). It has recently been used in several COVID-19 vaccines (Liang et al., 2020; Ho et al., 2021), introducing a surge in demand for this terpenoid. Other reported properties of squalene include tumor-suppressing (Smith et al., 1998; Yang et al., 2014), immunity improving (Ronco and De Stéfani, 2013), cholesterol-lowering (He et al., 2003), as well as antibacterial and antifungal effects (Katabami et al., 2015). Squalene has attracted attention as a feasible source of biofuels (Hellier et al., 2013) as well, if it could be produced sustainably and in large quantities.

In most plants, algae and prokaryotes, terpenoids can be synthesized via the methyl-erythritol-4-phosphate (MEP) pathway, also called the non-mevalonate pathway (Okada and Hase, 2005; Sawai and Saito, 2011). While the mevalonate pathway is present in most eukaryotic cells, the MEP pathway was acquired through endosymbiotic or horizontal gene transfer in plastid-bearing organisms. As the progenitors of plastids, cyanobacteria can serve as both a model organism for chloroplastic terpenoid synthesis through the MEP pathway and are promising production hosts for plant terpenoids (Hemmerlin et al., 2006; Loeschcke et al., 2017). The MEP pathway produces isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which are the universal precursors for terpenoid synthesis. In *Synechocystis*, a single gene, *crtE*, is responsible for the elongation of terpene precursors towards geranyl pyrophosphate (GPP), farnesyl pyrophosphate (FPP) and geranylgeranyl pyrophosphate (GGPP) in consecutive condensation reactions. The enzyme squalene synthase (Sqs) catalyzes the condensation of two molecules of FPP to presqualene diphosphate (PSPP), which is then converted into squalene via reduction by NADPH. In *Synechocystis*, squalene is then cyclized by squalene hopene cyclase (Shc) to hopene (Englund et al., 2014). Englund and colleagues (Englund et al., 2014) used a modified strain of *Synechocystis* to produce squalene by inactivating the *shc* gene. By preventing the generation of hopene, squalene is accumulated in this mutant. We previously constructed a markerless deletion of *shc* in *Synechocystis* to minimize the number of antibiotic resistances carried by each strain and enable further engineering (Dietsch et al., 2021). The availability of inducible promoter systems, such as the rhamnose, anhydrotetracycline or copper inducible promoters in *Synechocystis* allows fine-tuning of gene expression levels, making improvement of metabolic pathways possible by identifying optimal expression levels for each involved gene (Behle et al., 2020).

A straightforward approach to achieve higher yields of desired products is the overexpression of certain genes to increase the flux towards these metabolites, with many strategies already reported for increasing heterologous terpenoid production (Klaus et al., 2022). Despite some strategies proving successful, the regulation and bottlenecks of the MEP pathway are still not entirely understood (Klaus et al., 2022). Most initial pathway modification approaches result in relatively low product yield, and optimization is often dependent on heuristic techniques (Choi et al., 2017; Steuer et al., 2012). While the identification of genes to be modified is an essential step in metabolic engineering for strain improvement toward the enhanced production of desired bioproducts, it is still difficult to decide which genes to insert or modify, due to the vast number of possibilities of potential targets and the consideration of complex regulation of metabolic networks. In order to rationally identify and overcome bottlenecks for the improvement of strain designs, the use of *in silico* models has gained increasing significance over the past years (King et al., 2015; Broddrick et al., 2016; Lin et al., 2017; Hendry et al., 2020). Genome-scale metabolic modeling requires only stoichiometric information and no kinetic parameters, which are usually not available even for small reaction networks. In contrast, stoichiometric information is readily and reliably available for a large number of annotated genes (Raman and Chandra, 2009). Additionally, the estimation of stoichiometric yield is computationally not expensive and feasible even for large models involving hundreds of reactions (Knoop and Steuer, 2015).

Constraint-based reconstruction and analysis (COBRA) utilizes these models to calculate flux distributions under certain environmental or internal conditions (Schellenberger et al., 2011). COBRA methods are mainly used for the prediction of maximum theoretical yield for native and non-native pathways (Shen and Liao, 2013; King et al., 2015), the effect of gene deletions on biomass or other target compounds (e.g. OptKnock) (Burgard et al., 2003) as well as the identification of bottlenecks and potential targets for up- and downregulation in order to increase product yield (e.g. flux scanning based on enforced objective flux (FSEOF) or minimization of metabolic adjustment (MOMA)) (Segrè et al., 2002; Choi et al., 2010). Previous studies (Choi et al., 2010; Englund et al., 2018; Park et al., 2018), aiming to increase terpenoid production, have shown that constraint-based flux balance analysis (FBA) can be a helpful tool to not only understand and analyze metabolic pathways, but to identify bottlenecks and narrow down the options of potential amplification or knock-down targets for increased product yield, without having a negative influence on the growth rate. Englund et al. successfully employed genome-scale metabolic flux analysis to identify amplification targets that increased isoprene production in *Synechocystis* (Englund et al., 2018).

In our work, we used an algorithm called FSEOF to screen for potential overexpression targets increasing the terpenoid concentration in *Synechocystis* (Choi et al., 2010). *In silico* analysis predicted an increased flux through reactions of the MEP pathway, terpenoid synthesis, the light-dependent reactions of photosynthesis as well as the central carbon metabolism. Additionally, metabolic modeling proposed elevated requirements for the efficient supply, balance and regeneration of cofactors. Twelve of the identified amplification targets were tested *in vivo* to assess the biological relevance of the model, all of which positively impacted squalene production. The overexpression of *sqs* led to the strongest increase, with a yield of 13.72 mg l^{-1}, the highest production titer reported for squalene in *Synechocystis* to date.

Materials and methods

Plasmid and strain construction

A detailed list of all relevant genetic modules and information regarding their origin, is provided in the Supporting Information (Table S1 (SI)).

To investigate the computationally identified genes' effect on squalene production, the pEERM4 plasmid was used to integrate each gene into the neutral site 2 (NS2) under control of the rhamnose promoter P_{rha} (Englund et al., 2015; Behle et al., 2020). The plasmid pEERM4 Cm was a gift from Pia Lindberg (Addgene plasmid # 64026; http://n2t.net/addgene:64026; RRID : Addgene_64026) (Englund et al., 2015). This plasmid was used to clone each gene of interest under the control of P_{rha}. It contains 500 bp DNA homologous to the upstream and downstream region of NS2, between which a chloramphenicol resistance and the gene of interest are located, flanked by the rhamnose promoter and the T7 terminator. Each gene of interest was cloned into the plasmid using the restriction enzymes *Nhe*I and *Pst*I, with the *Nhe*I cutting site located after the start codon. The genes of interest were amplified from the *Synechocystis* genome, using Q5-Polymerase (NEB # M0491) according to manufacturer's instructions with oligonucleotides shown in Table S2 (SI). In two cases, an NheI restriction site was removed from the native gene sequence without changing the amino acid sequence (*gap2, sqs*). The *sqs* gene is annotated as starting with GTG as a start codon in the published Kazusa genome and this codon was changed to ATG for the purposes of this study.

To enable induction of the P_{rha} promoter, the rhamnose activator rhaS was constitutively expressed by the J23119 promoter from the replicative plasmid pSHDY (AddGene Plasmid #137661, (Behle et al., 2020)), which was transferred to *Synechcoystis* via triparental mating (Behle et al., 2020). This plasmid was constructed using the restriction sites of the BioBrick and NeoBrick standards and carries a spectinomycin resistance.

Synechocystis was transformed with the pEERM4 plasmids (Table S1 (SI)) using a protocol based on its natural competence (dx.doi.org/10.17504/protocols.io.mdrc256). Successful integration of the plasmid into the genome through heterologous recombination into the neutral site 2 (NS2) (Satoh et al., 2001) was verified by colony PCR (Figure S1 (SI)). The plasmid pSHDY carrying the rhamnose activator *rhaS* was then transferred to *Synechocystis* using triparental mating (dx.doi.org/10.17504/protocols.io.psndnde).

Culture conditions

The *Synechocystis* strains were inoculated in 30 ml BG11 liquid cultures containing 20 µg/ml spectinomycin and for overexpression strains 10 µg/ml chloramphenicol in 100 ml Erlenmeyer flasks from agar plates. The cultures were diluted twice to an OD_{750} of 0.2 to equalize their cell densities and growth phases. Two days before the start of the experiment, cultures were again diluted to an OD_{750} of 0.2 after which they were transferred to 6-well plates with 5 ml per well. L-Rhamnose was then added to the cultures and they were grown for 72 h at 30°C with 150 rpm shaking, 0.5% CO_2 and 80 µE m^{-2} s^{-1} of continuous light. After 72 h, cell samples were taken and stored for further processing.

For measurements of squalene production over time, 30 ml of BG11 were inoculated from a pre-culture to OD_{750} = 0.4, supplemented with 5 mM of rhamnose and incubated over 14 days. Samples were taken daily for the first four days, every second day for the following six days and after 14 days. The lost culture volume from sampling was replaced with fresh BG11 containing appropriate antibiotics and 5 mM of rhamnose.

Biomass measurements (DCW, OD, spectra)

Cell dry weight measurements were carried out by transferring the cell pellet of 2 ml of cyanobacterial culture to a pre-weighed

PCR tube, which was incubated at 60°C for 20 h. The tube was weighed and the difference noted as the cell dry weight, with measurements carried out in triplicates.

Absorption spectra and OD measurements were carried out in 1 ml polystyrene cuvettes in a SPECORD 200 Plus Spectrophotometer (Analytik Jena) with BG11 as a blank and as a reference sample. Samples were diluted with BG11 to be within an absorption range of 0.1 to 1.0 to ensure accurate measurements. Cell densities for *Synechocystis* were measured at 750 nm.

Pigment quantification

Each culture (300 µL) was sampled after 72 hours at the end of the growth experiment. The sample was centrifuged at 14,000 g for 5 minutes and 4°C. The supernatant was discarded and the pellet was resuspended in 100 µl water. The samples were frozen at -80°C until further processing. 900 µl of 100% methanol were added to the sample and the sample was mixed by vortexing. After incubation in the dark under gentle agitation for 1 h at 4°C the sample was centrifuged at 14,000 g for 5 minutes. The supernatant was transferred into a cuvette and an absorbance spectrum was measured from 400 nm to 750 nm. The absorbance spectra were divided by the OD_{750} or CDW and the amount of chlorophyll *a* in the sample was quantified by the absorbance maximum of chlorophyll *a* at 665 nm (A_{665nm}) using following equation (Lichtenthaler and Buschmann 2001):

$$\text{Chlorophyll content}[\mu g/ml] = 12.66\ \mu g/ml * A_{665\ nm}$$

The amount of carotenoids in the sample was quantified by the absorbance maximum of the sum of carotenoids at 470 nm (A_{470nm}) and a correction term considering absorbance of chlorophyll *a* at 470 nm (c(Chl *a*): concentration of chlorophyll *a* in the sample) using following Equation (Lichtenthaler and Buschmann 2001):

$$\text{Carotenoid content}[mg/ml] = (1000\ \mu g/ml * A_{470\ nm} - 1.91 * c(Chl))/225$$

GC-MS measurements for the quantification of squalene

Each culture (1.5 ml) was sampled after 72 hours at the end of the growth experiment. The sample was centrifuged at 14,000 g, for five minutes and 4°C. The supernatant was discarded and the pellet was frozen at -80°C until further processing. The pellet was extracted with 500 µL acetone, containing 25 µM β-sitosterol as internal standard, under agitation at 1000 rpm and 50°C for 10 min. 500 µL of 1 M NaCl was added and mixed by vortexing. After adding 250 µL hexane, the sample was vigorously mixed for 1 min and centrifuged for phase separation (1 min at 1,780 g and 4°C). The upper hexane phase was transferred into GC-MS vials and stored at -20°C until the analysis.

GC-MS analysis was carried out using a Gerstel automatic liner exchange system with multipurpose sample MPS2 dual rail and two derivatization stations, used in conjunction with a Gerstel CIS cold injection system (Gerstel, Muehlheim, Germany). For every 10-12 samples, a fresh multibaffled liner was inserted. Chromatography was performed using the Agilent 7890B GC. Metabolites were separated on an Agilent HP-5MS column (30ml x 0.25mm), the oven temperature was ramped with 12.5 °C/min from 70 °C (initial temp for 2 min) to 320 °C (final temp hold 5 min). Metabolites were ionized and fragmented in an EI source (70V, 200 °C source temp) and detected using 7200 accurate mass Q-TOF GC-MS from Agilent Technologies. Data analysis was performed using Agilent MassHunter Quantitative Analysis B.09.00. Peaks were identified using already available EI-MS fragmentation data. Peaks were identified using characteristic fragment ions (Bhatia et al., 2013) and retention times of standards (Squalene: mass/charge (m/z) = 81.07, retention time (RT) = 9.5 min; β-sitosterol: m/z = 107.09, RT = 13.6 min). Squalene concentrations in the measured samples were calculated using a calibration curve with a squalene standard (Figure S2 (SI)).

Quantitative real-time PCR (qRT-PCR)

Cultures were sampled (0.5 ml) after 72 hours at the end of the growth experiment. The pellet was processed for RNA extraction using the PGTX method (dx.doi.org/10.17504/protocols.io.jm3ck8n, Pinto et al., 2009). The remaining DNA in the extracted RNA was removed by DNase digestion using the TURBO DNA-free™ (ThermoFischer) kit according to the manufacturer's instructions. Extracted RNAs (250 ng) were used in a reverse transcriptase reaction using the RevertAid First Strand cDNA Synthesis Kit (ThermoFischer) according to the manufacturer's instructions. The resulting cDNA was diluted 1:20. For performing qPCR, the DyNAmo ColorFlash SYBR Green qPCR Kit was used according to the manufacturer's instructions. Primers for *sqs*, *dxs* and the housekeeping gene *rpoA* are shown in Table S2 (SI). Primer efficiencies were tested before performing qRT-PCR and were deemed sufficient to yield quantitative information (Figure S3; Table S3 (SI)). Changes in gene expression as fold changes compared to the control were determined using the $2^{-\Delta\Delta CT}$ method, using *rpoA* as a housekeeping gene and the Δ*shc* strain subjected to the same rhamnose concentration as a control.

Metabolic modeling for the identification of amplification targets

All simulations are based on a genome-scale stoichiometric network model of *Synechocystis* published by Knoop and colleagues (Knoop and Steuer, 2015). A modified, extended version was used, kindly provided by Ralf Steuer. All flux distributions have been calculated with constraint-based flux analysis using COBRApy (v.0.25.0) (Ebrahim et al., 2013). To simulate phototrophic growth, different constraints were applied to the model of *Synechocystis* (see Table S5 (SI)).

FSEOF (Choi et al., 2010) was used to find amplification targets by simulating the transition from a wildtype to a production phenotype. All isoreactions were excluded for the transition experiments (Knoop and Steuer, 2015). The initial fluxes of all

reactions were calculated by using the objective function to maximize the growth rate. Then, the theoretical maximum squalene production rate was calculated by setting the objective function as maximizing squalene flux. Subsequently, under constant light flux, the product formation flux rate was stepwise increased from 0% to 67% of the maximum achievable rate, while the growth rate was maximized. Only targets for which the overall mean flux rate from maximum biomass synthesis to maximum product synthesis increases were chosen. Additionally, only reactions that did not change flux direction during transition were considered. To confirm the results, flux variability analysis was performed for the selected targets, by stepwise increasing squalene flux from 0% to 67% of the maximum rate and subsequently maximizing biomass synthesis. For each simulation step, the variability of all selected targets was determined. To visualize the flux distributions a simplified network was implemented with d3flux (v.0.2.7) (St. John, 2016), a d3.js based visualization tool for COBRApy models.

Results

Metabolic modeling predicts overexpression targets in MEP pathway and central carbon metabolism

In this study our goal was to enhance squalene production by identifying potential amplification targets and systematically overexpressing selected genes of interest. Due to the extensive number of possible genes to modify, we chose to first screen for suitable targets *in silico*. We used a genome-scale metabolic network of *Synechocystis* to find the ideal flux distribution for the optimal production of squalene while maintaining at least 1/3 of the growth rate. Reactions that increase in flux, when more squalene is produced, were chosen as potential amplification targets.

To predict these targets, we applied an algorithm called FSEOF, developed by Choi and colleagues (Choi et al., 2010). The transition from a wildtype phenotype to a production phenotype was simulated by stepwise increasing squalene flux while the growth rate was maximized. This way, all resources beyond the growth rate are directed towards the forced product synthesis. We only considered targets for which the overall mean flux rate increases and that did not change flux direction during transition (Knoop and Steuer, 2015). An overview of reactions meeting these criteria can be found in Figure 1 and in Table S6 (SI). Excluding transport, export and spontaneous reactions, 39 potential overexpression targets were identified. Flux variability analysis was performed for these targets to validate the results (Table S7 (SI)). Upon enforced objective flux different flux patterns could be observed. 22 fluxes showed an increasing pattern without any variability. Among the remaining fluxes two increased within a narrow range and seven increased with broad variability. Three showed a pattern, where the minimal flux increased and the maximum possible flux decreased. One reaction showed a pattern indicating a change in flux direction and four were unbound.

The results of the metabolic modeling suggest an increase of flux through the MEP pathway, from the decarboxylative carboligation of glyceraldehyde 3-phosphate (G3P) and pyruvate to deoxyxylulose 5-phosphate (DXP) by deoxyxylulose 5-phosphate synthase (Dxs), to the formation of isopentenyl diphosphate (IPP) by IspH as well its conversion to DMAPP by isomerization (Idi). The following reactions towards terpenoid synthesis, catalyzed by CrtE and Sqs, are also showing a flux increase.

Additionally, the model suggests an increased flux through cytidine monophosphate kinase (CMPk) and cytidine diphosphate kinase (CDPk) as well as an increased activity of inorganic diphosphatase (Ppa), converting diphosphate to monophosphate.

The reactions of lower glycolysis are proposed to be upregulated as well. 3-Phospho-D-glycerate (PG3), obtained by the RuBisCO reaction, is converted to 2-phospho-D-glycerate (PG2) by phosphoglycerate mutase (Pgm), to phosphoenolpyruvate (PEP) by enolase (Eno) and finally to pyruvate by pyruvate kinase (PyrK).

To provide the carbon needed for enhanced terpenoid synthesis, the total flux through the Calvin-Benson-Bassham (CBB) cycle is expected to increase. Especially the flux through RuBisCO, fixing CO_2 as PG3 shows a strong increase. The same applies to the phosphorylation of PG3 to 1,3-bisphosphoglycerate (13DPG) by phosphoglycerate kinase (PgK), its reduction to the MEP pathway precursor G3P by glyceraldehyde 3-phosphate dehydrogenase (Gap2), the conversion to dihydroxyacetone-phosphate (DHAP) by triosephosphate isomerase (Tpi) as well as the subsequent regeneration of D-ribulose 1,5-bisphosphate (RuDP) through the reductive pentose phosphate pathway.

The model predicts an increased flux through the light-dependent reactions of oxygenic photosynthesis in the thylakoid membrane. However, photosynthetic pigments like chlorophyll and carotenoids display a reduced flux. Whereas NADPH production via ferredoxin-$NADP^+$ reductase (FNR) is suggested to be increased, ATPase activity is supposed to be decreased. This goes along with a reduced TCA cycle activity.

As enhanced flux through the light reactions and CBB cycle are predicted, photorespiratory metabolism increases as well. However, the model suggests utilizing the glycerate photorespiratory bypass via tartronate semialdehyde. Phosphoglycolate as an inevitable by-product of the photorespiratory chain is converted to first glycolate, then glyoxylate, tartronate semialdehyde and afterwards to glycerate, which can be used to synthesize pyruvate or be recycled into the CBB cycle.

In summary, the model proposes an increased flux through the MEP pathway and terpenoid synthesis as well as the CBB cycle, lower glycolysis and the light-dependent reactions of photosynthesis to enhance squalene production. Additionally, an increased regeneration of the cofactor CTP and a decreased ATP/NADPH ratio are predicted to positively correlate with squalene synthesis. Subsequently, we systematically overexpressed 11 selected targets.

Systematic overexpression study confirms predictions of FSEOF through altered squalene and pigment content

To test the predictions made by the FBA modeling, 11 genes of interest were inducibly overexpressed in *Synechocystis* Δ*shc* markerless deletion mutant (Dietsch et al., 2021) using the previously characterized rhamnose-inducible promoter system

(Behle et al., 2020) including all genes of the MEP-pathway except for *dxr*, which was previously reported to negatively impact terpenoid production upon overexpression (Choi et al., 2016). In addition to genes directly involved in terpenoid synthesis, *gap2*, *pyrK* and *tpi* were overexpressed to increase availability of pyruvate and G3P. All strains, including the control and wild type possessed the pSHDY *rhaS* replicative plasmid. *Synechocystis* Δ*shc* pSHDY *rhaS* serves as the control due to its ability to accumulate squalene. The functionality of the overexpression system was confirmed by performing qRT-PCR on two representative genes, showing increased transcript levels upon induction (Figure S4 (SI)).

The experimental work showed that squalene production was increased, albeit in some cases only slightly compared to the Δ*shc* strain for all overexpressed genes identified by FSEOF, as shown in Figure 2. The amounts of squalene produced in these strains vary widely, depending on the overexpressed gene (Table S4 (SI)).

Overexpression of genes strongly altered the pigmentation in several strains, with the relative changes compared to the *Synechocystis* Δ*shc* control strain shown in Figure 3. Growth was slightly reduced in the overexpression strains, as shown in Figure S5 (SI).

Most overexpression strains showed a reduction in chlorophyll and carotenoid content, as shown in Figure 3 likely due to disruption of optimal pigment synthesis caused by genetic modification, leading to accumulation of metabolites which may inhibit enzymes upstream in their respective pathway. The reduction in pigments may also be caused by the presence of an additional antibiotic resistance cassette in the overexpression strains. However, the strong pigment variations between overexpression mutants suggest only a weak effect of the antibiotic compared to the genetic change. The overexpression of *ispE*, *ispH* and *gap2* are notable exceptions to this observation, showing little changes in pigmentation upon overexpression. Absorbance spectra of all strains can be seen in Figure S6 (SI). The overexpression of *sqs* led to a blue-colored phenotype, which is characterized by reduced chlorophyll and strongly reduced carotenoid concentrations, as shown in Figure 4.

Overexpression of *sqs* is the most effective way to improve squalene production

In the MEP pathway, overexpression of *dxs* did not have a strong effect on squalene production, shown in Figure 2, while the downstream reactions catalyzed by IspD, IspE, IspF and IspH

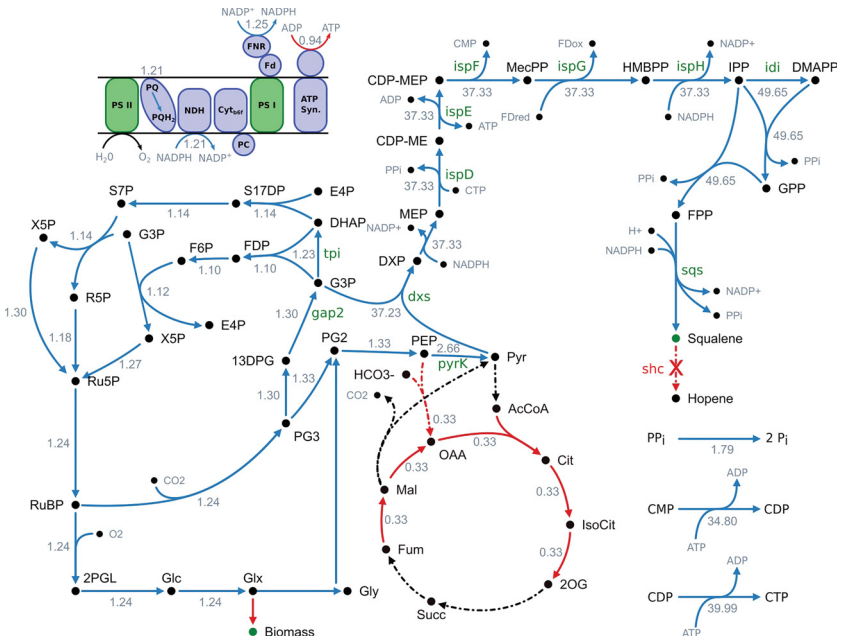


FIGURE 1
Overview of fluxes predicted to change upon increased squalene production. Blue arrows indicate an increased flux and red arrows a decreased flux, respectively. Black arrows indicate no change. Reactions with no flux have a dotted line. The numbers indicate the maximum fold change of the corresponding flux. It is stated that this is not a minimal network but a part of the genome-scale model and not all active reactions are shown. 13DPG, 1;3-bisphosphoglycerate; 2OG, 2-oxoglutarate; 2PGL, 2-phosphoglycolate; AcCoA, acetyl-CoA; ATP synth., ATP synthase; CDP-ME, 4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol; CDP-MEP, 2-phospho-4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol; Cit, citrate; Cyt_{b6f}, cytochrome b_6f complex; DHAP, dihydroxyacetone phosphate; DMAPP, dimethylallyl diphosphate; DXP, 1-deoxy-D-xylulose 5-phosphate; E4P, erythrose 4-phosphate; F6P, fructose 6-phosphate; Fd_{ox}, ferredoxin (oxidized); Fd_{red}, ferredoxin (reduced); FDP, fructose 1;6-biphosphate; FNR, ferredoxin-$NADP^+$ reductase; FPP, farnesyl pyrophosphate; Fum, fumarate; G3P, glyceraldehyde 3-phosphate; Glc, D-glycerate; Glx, glyoxylate; Gly, glycolate; GPP, geranyl pyrophosphate; HMBPP, 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate; IsoCit, isocitrate; IPP, isopentenyl diphosphate; Mal, malonate; MEcPP, 2-C-methyl-D-erythritol 2;4-cyclodiphosphate; MEP, 2-C-methyl-D-erythritol 4-phosphate; NDH, NADPH dehydrogenase; OAA, oxaloacetate; PC, plastocyanin; PEP, phosphoenolpyruvate; PG2, 2-phosphoglycerate; PG3, 3-phosphoglycerate; Pi, orthophosphate; PPi, diphosphate; PQ, plastoquinone; PQH_2, plastohydroquinone; PSI, photosystem I; PSII, photosystem II; Pyr, pyruvate; R5P, ribose 5-phosphate; Ru5P, ribulose 5-phosphate; RuBP, ribulose 1;5-biphosphate; S17DP, sedoheptulose 1;7-bisphosphate; S7P, sedoheptulose 7-phosphate; Succ, succinate; X5P, xylulose 5-phosphate.

10.3389/fpls.2023.1024981

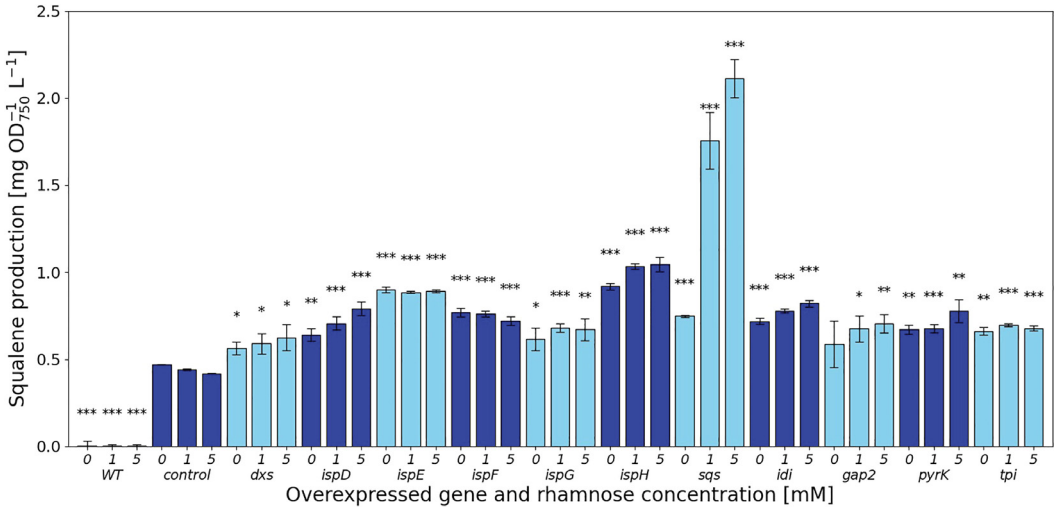


FIGURE 2
Squalene concentrations [mg l^{-1} OD_{750}^{-1}] in response to gene overexpressions. Values are represented as the means of three biological replicates, standard deviations are shown. WT represents the *Synechocystis* sp. PCC 6803 wild type, while the control strain is *Synechocystis* sp. PCC 6803 Δ*shc* pSHDY *rhaS*, from which the overexpression strains were constructed by inserting an additional copy of the specified gene under the control of the rhamnose-inducible promoter P_{rha} into its genome. Asterisks (*) represent the p-value of the two-sided t-test between the respective strain and the control strain at the same rhamnose concentration (* denotes a value of p<0.05, ** denotes p<0.01 and *** denotes p<0.001). Samples were measured after three days of incubation with the specified concentration of rhamnose as an inducer.

showed more positive effects. The overexpression of both *ispE* and *ispH* led to increased squalene production, even without any inducer present, suggesting that the leakiness of the rhamnose promoter provided sufficient additional enzyme to relieve a bottleneck from the MEP pathway. Overexpressing *ispD* on the other hand increased squalene concentrations in a more concentration-dependent manner. Overexpression of *ispG* did not increase squalene content by a large amount, also showing the strongest negative impact on pigment concentrations of the overexpression of genes in the MEP pathway.

The overexpression of *idi* increased squalene concentrations by 1.96-fold, reducing pigment concentrations in the process, likely by shifting the IPP/DMAPP ratio.

The overexpression of *sqs* led to a 5-fold increased squalene concentration, yielding 2.11 mg OD_{750}^{-1} l^{-1}. *Synechocystis'* Sqs may be a particularly active enzyme, as cloning in *E. coli* proved challenging, possibly due to toxicity of squalene to *E. coli*, limiting cell densities in liquid culture and favoring mutated or truncated versions of the gene on agar plates. *E. coli* DH5α cells transformed with a non-mutated version of *sqs* typically required around 36 h of incubation at 37°C to form colonies of normal size. The overexpression of *sqs* led to a decrease in chlorophyll and a strong decrease in carotenoid content in *Synechocystis*, as shown by its cell absorbance spectrum in Figure 4.

Of the genes not involved in terpenoid synthesis, overexpressing *gap2*, a central enzyme involved in the CBB cycle only led to a small increase in squalene concentrations, but did so while keeping pigment concentrations approximately at the same level as in the Δ*shc* base strain. Upon induction, squalene and pigment content increased with the inducer concentration. Overexpression of *pyrK*, a glycolysis gene, increased squalene concentrations after induction, but reduced pigment concentrations. Expression of *tpi* did not have a strong positive effect on squalene concentrations and also reduced pigmentation.

The strain overexpressing *sqs* was additionally investigated regarding its squalene production in 30 ml Erlenmeyer flasks, with its growth and squalene production over time shown in Figure 5. The yield of the culture after 14 days was 2.78 mg l^{-1} OD_{750}^{-1}/13.72 mg l^{-1}.

In summary, all overexpression strains showed increased squalene production, with the overexpression of *sqs* leading to a 5-fold improvement compared to the base strain, with longer term cultivation leading to a yield of 13.72 mg l^{-1}. The most efficient overexpressions regarding squalene yield were all in the MEP-pathway and terpenoid synthesis, but overexpression of *gap2* was also notable, as it showed a positive impact on both pigment and squalene concentrations with increasing induction.

Discussion

In the pursuit of identifying potential amplification targets, improving cyanobacterial squalene production upon upregulation, we used a genome scale constraint-based metabolic model. Twelve of the identified overexpression targets were experimentally confirmed.

The analysis suggested an up-regulation of fluxes through the MEP pathway and terpenoid synthesis. Since terpenoids are exclusively synthesized via the MEP pathway in *Synechocystis*, this is very intuitive and shows the robustness of our *in silico* analysis pipeline. Flux through the light-dependent reactions of photosynthesis as well as the CBB cycle were proposed to increase, to meet the demand of fixed carbon and produce the cofactor NADPH. In fact, the FNR reaction showed an increased flux, while the ATPase reaction is proposed to be downregulated,

10.3389/fpls.2023.1024981

leading to an overall decreased ATP/NADPH ratio. Previous *in silico* studies suggested lower ATP/NADPH ratio requirements for many biofuels as well (Erdrich et al., 2014; Shabestary and Hudson, 2016; Englund et al., 2018). Artificially creating imbalances in this ratio could increase product formation. This can be achieved via different approaches, e.g., by blocking cyclic and other alternative electron flows. As a result, ATP and NADPH are forced to be synthesized exclusively via linear electron flow, whose ATP/NADPH ratio is below the ratio required for biomass synthesis. Knocking out ATP producing reactions or introducing ATP futile cycles or other waste reactions could force the organism to use terpenoid production as a sink for excess reduction equivalents (Erdrich et al., 2014). The carboxylation and oxygenation reaction of RuBisCO were coupled in our model, so the flux through the photorespiratory reactions was forced to increase as well. The model suggested using the glycerate photorespiratory bypass. Nonetheless, since it is an undesired reaction to occur, we did not consider it as an amplification target. On the contrary, Zhou et al. could show that impairing photorespiration leads to a redirection of excess energy and isoprene production could be doubled (Zhou et al., 2021). Additionally, cytidine triphosphate (CTP), a cofactor in the MEP pathway, has to be regenerated by multiple phosphorylations of cytidinmonophosphat (CMP) and the model proposed it might become rate limiting upon high squalene flux. Furthermore, the conversion of diphosphate to monophosphate was predicted to be upregulated because of the increased levels of diphosphate released by IspD, CrtE and Sqs. These results are in accordance with previous *in silico* studies aiming to increase terpenoid production in cyanobacteria, although different models as well as different algorithms were utilized (Lin et al., 2017; Englund et al., 2018). The TCA cycle was indicated to decrease in flux, since it is a sink for carbon, which also was found by previous studies in cyanobacteria (Englund et al., 2018). This finding is in contrast to studies in *E. coli* (Choi et al., 2010), where the TCA cycle is called to be upregulated. This difference can be explained by the fact that heterotrophs like *E. coli* have to regenerate their cofactors via the TCA cycle and autotrophs, in contrast, are able to generate cofactors via the light reactions of photosynthesis. The metabolic model suggested a decrease in photosynthetic pigments as well, since increased squalene production diverts carbon flux away from pigments and they compete for the precursor GPP (Lin et al., 2017).

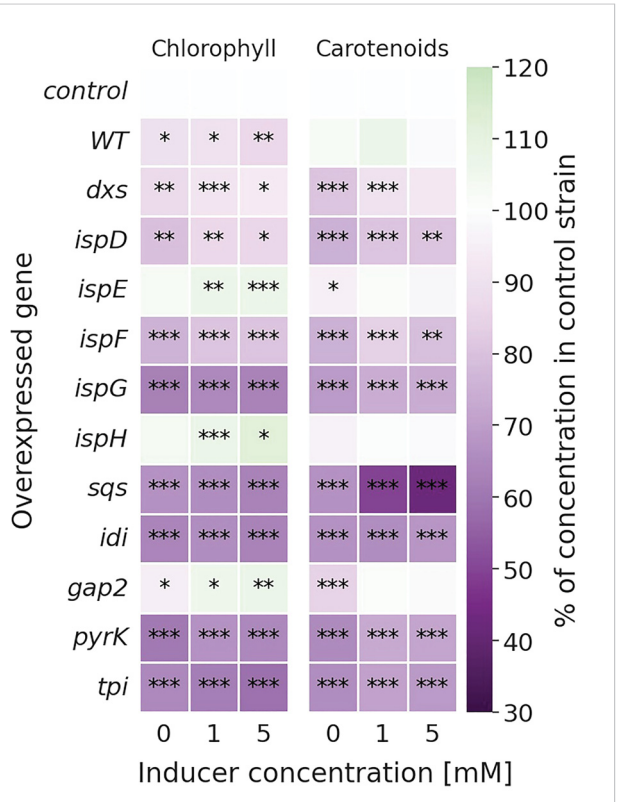


FIGURE 3
Relative change in chlorophyll (*left*) and carotenoid (*right*) concentrations [mg l^{-1} OD_{750}^{-1}] of the overexpression strains compared to the Δ*shc* control strain. Values are represented as the means of three biological replicates. WT represents the *Synechocystis* sp. PCC 6803 wild type, while the control strain is *Synechocystis* sp. PCC 6803 Δ*shc* pSHDY *rhaS*, from which the overexpression strains were constructed by inserting an additional copy of the specified gene under the control of the rhamnose-inducible promoter P_{rha} into its genome. Asterisks (*) represent the p-value of the two-sided t-test between the respective strain and the control strain at the same rhamnose concentration (* denotes a value of p<0.05, ** denotes p<0.01 and *** denotes p<0.001). Samples were measured after three days of incubation with the specified concentration of rhamnose as an inducer.

Flux variability analysis found 31 fluxes to be consistent with the results of the FSEOF analysis and increase upon enforced squalene flux. Among the remaining eight fluxes, three showed a pattern where the maximum flux decreased, while the minimum flux increased, and were thus not considered as amplification targets (Choi et al., 2010). The reaction catalyzed by Idi was the only one to display a sign change, where the maximum flux is positive and the minimum flux is negative, implicating the direction of the metabolic flux is unknown and cannot be determined by flux variability analysis (Flowers et al., 2018). Incorporating kinetic parameters as constraints would help to increase the accuracy of the prediction (Moulin et al., 2021). Since in either direction, the flux through Idi positively correlates with squalene production, we chose to verify the result experimentally. Four reactions were completely unbound, thus providing no further information. Previous studies (Choi et al., 2010) showed that the overexpression of unbound targets also had a positive impact on product synthesis.

To confirm the predictions made by the FSEOF analysis, 11 genes were experimentally overexpressed in *Synechocystis* Δ*shc*. All overexpression strains showed elevated squalene levels, but the degrees to which squalene concentrations were increased varied widely. Overall, growth of the overexpression strains was slightly reduced compared to the control (Figure S5 (SI)), possibly as a result of the additional antibiotic present in the growth medium. The differences between the overexpression of genes in the MEP pathway are of particular interest, as regulations and feedback mechanisms in the pathway are not entirely known. Since protein tags can affect both activity and stability of enzymes, we chose to use the native protein sequences for overexpression. Since no antibodies are reported for these native enzymes, we were not able to quantify their protein concentrations. However, we confirmed the functionality of our expression system via qRT-PCR for the *dxs* and *sqs* overexpression strains which both showed increased

10.3389/fpls.2023.1024981

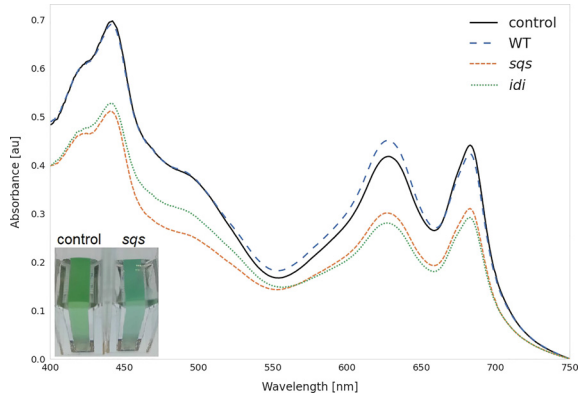


FIGURE 4
Whole cell spectra of selected strains. Spectra were measured after three days incubation with 5 mM rhamnose in 1 cm cuvettes after equalizing the OD_{750} across all strains and subtracting the OD_{750} as a baseline correction. WT represents the *Synechocystis* sp. PCC 6803 wild type, while the control strain is *Synechocystis* sp. PCC 6803 Δ*shc* pSHDY *rhaS*, from which the overexpression strains were constructed. The visually blue phenotype of the *sqs* overexpression strain in the cuvette is pictured in the bottom left.

transcript levels upon induction with rhamnose (Figure S4 (SI)). Quantitative differences in RNA abundance between genes are unavoidable due to differing mRNA lengths and stabilities, but it can be safely assumed that induction of expression through addition of rhamnose yields higher gene expression levels than the control in all cases. Since no protein amounts could be quantified in this study, no conclusions can be drawn regarding the metabolic efficiency of each overexpression relative to the expression strength. Instead, changes in the final metabolite concentrations are caused by the combined effect of the strength of the overexpression and the catalytic activity of the protein.

Dxs is reported to be the rate limiting enzyme of the MEP pathway in many previous studies, but did not prove to be a particularly beneficial overexpression target in this study. This may be attributed to the native Sqs activity not providing a strong enough carbon sink downstream of the MEP pathway, leading to the accumulation of intermediates, such as IPP and DMAPP, which are reported to act as inhibitors to the Dxs enzyme (Banerjee et al., 2013; Álvarez-Vasquez et al., 2021; Di et al., 2022). Kudoh et al., 2017 showed that overexpression of *dxs* led to aggregation of the inactivated protein via allosteric inhibition by IPP and DMAPP, ultimately diminishing the impact of the overexpression on the protein level (Kudoh et al., 2017). Protein inactivation may also explain the small effect of *dxs* overexpression on metabolites in this study.

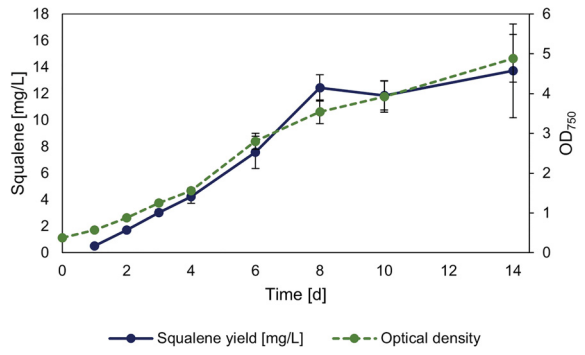


FIGURE 5
Timeseries of squalene production in *sqs* overexpression strain. Squalene production and OD_{750} of *Synechocystis* Δ*shc* pEERM P_{rha} *sqs* pSHDY *rhaS* in a 30 ml flask culture in mg l^{-1} over a period of two weeks after induction with 5 mM rhamnose to trigger overexpression of the squalene synthase (*sqs*). Means and standard deviations of three biological replicates are shown.

Other studies reported MEcPP to accumulate upon overexpression of *dxs*, so a dual overexpression with *ispG* may show a more positive effect (Gao et al., 2016; Volke et al., 2019). The overexpression of *ispG* alone did not lead to increased production however, the dependence of IspG on reduced ferredoxin units may be a limiting factor to the conversion of MEcPP to HMBPP (Wang and Oldfield, 2014). If the concentrations of MEcPP were too low for IspG to be limiting, the binding of ferredoxin to the additional enzyme may be the cause of the reduction in photosynthetic pigments, with the overexpression of *ispG* leading to the lowest chlorophyll concentrations among the genes of the MEP pathway.

In contrast to most overexpressions, *ispE* and *ispH* showed a positive effect on squalene production even without inducer present, but not strongly increasing upon induction. This relationship suggests that these enzymatic steps represent a metabolite bottleneck, which can be relieved through a slight overexpression and then yields diminishing returns upon stronger overexpression.

The overexpression of *idi* led to an increase in squalene while decreasing pigments, which is in accordance with the predictions made by the constraint-based model. IspH favors production of IPP from HMBPP over DMAPP, while Idi favors the isomerization of IPP towards DMAPP, leading to an IPP : DMAPP ratio of 3:1 *in vivo*, the optimal ratio for production of GGPP (Chaves et al., 2016; Volke et al., 2019). The overexpression of *idi* likely shifted the IPP : DMAPP ratio in favor of DMAPP.

Overexpressing *sqs* led to the strongest increase in squalene and reduction in carotenoids. Sqs competes with the native GGPP synthase *crtE* for the intermediate FPP, with the overexpression leading to a shift in favor of squalene, away from GGPP, the precursor for carotenoids and phytol. Since there was only little change in growth over three days (Figure S5 (SI)), the reduced pigmentation did not seem to have a significant negative effect on the cell, but shifting the balance further towards squalene might lead to decreased photosynthetic performance. Overexpression of *ispH* and *ispE* on the other hand increased squalene concentrations as well as the total amount of measured terpenoids, defined as the sum of squalene, carotenoid and the phytol chain of chlorophyll by up to 18%. Combining the overexpressions of *sqs* with *ispH* and *ispE*, may be a promising strategy moving forward, as increased total flux through the MEP pathway can both compensate for the reduced pigmentation caused by *sqs* overexpression and increase squalene titers.

In conclusion, FSEOF allowed us to choose biologically relevant amplification targets computationally, all of which had a positive effect on squalene synthesis upon experimental validation.

Considering squalene is synthesized via the linear MEP pathway, most of the identified targets are rather intuitive. Since we were able to confirm all selected targets, we suggest a validation of the non-intuitive targets outside the MEP pathway, such as Ppa and FNR for further studies. These two targets have previously been predicted and tested in *Synechocystis* (Englund et al., 2018). Our findings propose that constraint-based metabolic models could aid in the selection of targets improving the production of desired metabolites. This could be of particular interest for the prediction of combinatorial interventions. However, classic FBA does not account for regulatory mechanisms like feedback inhibition or potential metabolite toxicity (Knoop and Steuer, 2015). It could be helpful for future studies to test FSEOF in combination with dynamic extensions of FBA (Mahadevan et al., 2002) and hybrid kinetic and constraint-based models (Shameer et al., 2022). The incorporation of metabolite concentrations, flux rates or kinetic parameters could drastically improve the precision and reliability of the results (Mahadevan et al., 2002; Moulin et al., 2021; Shameer et al., 2022). For future studies we suggest the combinatorial expression of the identified amplification targets, especially with the native *sqs*, since it seems to be the rate limiting enzyme in the present study, as well as the combination of amplification and knock-down targets. Our experiments are in accordance with the central idea of synthetic biology, where experimental designs are determined by metabolic modeling and experimental results can feed back data into models to increase their accuracy, leading to deterministic, steady improvement. The overexpression of the native *sqs* gene of *Synechocystis* proved to be the most successful strategy for squalene production to date, with the strain reaching a higher production titer than heterologous *sqs* expression in *Synechocystis* (Pattanaik et al., 2020).

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

AG: Conceptualization, Investigation, Methodology, Computational work, Writing – Original draft, review & editing, Data Visualization. AN: Conceptualization, Investigation, Experimental work, Writing – Original draft, review & editing, Data Visualization. MD: Conceptualization, Methodology, Experimental work, Writing – review & editing. TP: Computational work, Writing -Data visualization, review. DM: Experimental work, Writing – Review & Editing. ST: Experimental work, Writing – Review & Editing. PW: Methodology, Data Curation. IA: Supervision, Writing – review & editing. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2048/1 - project ID 390686111, and further, by project ID 391465903/GRK 2466.

Acknowledgments

We thank Ralf Steuer for the kind provision of the metabolic model. We also thank Anna Behle for the pSHDY plasmid and for assistance with the design of genetic constructs, as well as Pia Lindberg for providing the pEERM4 plasmid.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1024981/full#supplementary-material

References

Álvarez-Vasquez, F., González-Alcón, C., Gallego-Jara, J., de Diego, T., Cánovas, M., and Torres, N. V. (2021). Model based optimization of the terpenoids biosynthesis in *E. coli*. *bioRxiv*. doi: 10.1101/2021.08.11.455928

Banerjee, A., Wu, Y., Banerjee, R., Li, Y., Yan, H., and Sharkey, T. D. (2013). Feedback inhibition of deoxy-D-xylulose-5-phosphate synthase regulates the methylerythritol 4-phosphate pathway. *J. Biol. Chem.* 288, 16926–16936. doi: 10.1074/jbc.M113.464636

Barone, G. D., Cernava, T., Ullmann, J., Liu, J., Lio, E., Germann, A. T., et al. (2023). Recent developments in the production and utilization of photosynthetic microorganisms for food applications. *Heliyon* 9 (4). doi: 10.1016/j.heliyon.2023.e14708

Bhatia, A., Bharti, S. K., Tewari, S. K., Sidhu, O. P., and Roy, R. (2013). Metabolic profiling for studying chemotype variations in Withania somnifera (L.) Dunal fruits using GC-MS and NMR spectroscopy. *Phytochemistry* 93:105–115. doi: 10.1016/j.phytochem.2013.03.013

Behle, A., Saake, P., Germann, A. T., Dienst, D., and Axmann, I. M. (2020). Comparative dose-response analysis of inducible promoters in cyanobacteria. *ACS Synth. Biol.* 9, 843–855. doi: 10.1021/acssynbio.9b00505

Bennett, A., and Bogorad, L. (1973). Complementary chromatic adaptation in a filamentous blue-green alga. *J. Cell Biol.* 58, 419–435. doi: 10.1083/jcb.58.2.419

Berla, B. M., Saha, R., Immethun, C. M., Maranas, C. D., Moon, T. S., and Pakrasi, H. B. (2013). Synthetic biology of cyanobacteria: unique challenges and opportunities. *Front. Microbiol.* 4. doi: 10.3389/fmicb.2013.00246

Broddrick, J. T., Rubin, B. E., Welkie, D. G., Du, N., Mih, N., Diamond, S., et al. (2016). Unique attributes of cyanobacterial metabolism revealed by improved genome-scale metabolic modeling and essential gene analysis. *Proc. Natl. Acad. Sci. U. S. A.* 113, E8344–E8353. doi: 10.1073/pnas.1613446113

Burgard, A. P., Pharkya, P., and Maranas, C. D. (2003). Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* 84, 647–657. doi: 10.1002/bit.10803

Chaves, J. E., Romero, P. R., Kirst, H., and Melis, A. (2016). Role of isopentenyl-diphosphate isomerase in heterologous cyanobacterial (*Synechocystis*) isoprene production. *Photosynth. Res.* 130, 517–527. doi: 10.1007/s11120-016-0293-3

Choi, S. Y., Lee, H. J., Choi, J., Kim, J., Sim, S. J., Um, Y., et al. (2016). Photosynthetic conversion of CO_2 to farnesyl diphosphate-derived phytochemicals (amorpha-4,11-diene and squalene) by engineered cyanobacteria. *Biotechnol. Biofuels* 9, 202. doi: 10.1186/s13068-016-0617-8

Choi, H. S., Lee, S. Y., Kim, T. Y., and Woo, H. M. (2010). *In silico* identification of gene amplification targets for improvement of lycopene production. *Appl. Environ. Microbiol.* 76, 3097–3105. doi: 10.1128/AEM.00115-10

Choi, S. Y., Sim, S. J., Ko, S. C., Son, J., Lee, J. S., Lee, H. J., et al. (2020). Scalable cultivation of engineered cyanobacteria for squalene production from industrial flue gas in a closed photobioreactor. *J. Agric. Food Chem.* 68, 10050–10055. doi: 10.1021/acs.jafc.0c03133

Choi, S. Y., Wang, J. Y., Kwak, H. S., Lee, S. M., Um, Y., Kim, Y., et al. (2017). Improvement of squalene production from CO_2 in *Synechococcus elongatus* PCC 7942 by metabolic engineering and scalable production in a photobioreactor. *ACS Synth. Biol.* 6, 1289–1295. doi: 10.1021/acssynbio.7b00083

Connolly, J. D., and Hill, R. A. (2010). Triterpenoids. *Nat. Prod. Rep.* 27, 79–132. doi: 10.1039/b808530g

Di, X., Ortega-Alarcon, D., Kakumanu, R., Iglesias-Fernandez, J., Diaz, L., Baidoo, E. E. K., et al. (2022). MEP pathway products allosterically promote monomerization of deoxy-D-xylulose-5-phosphate synthase to feedback-regulate their supply. *Plant Commun.*, 100512. doi: 10.1016/j.xplc.2022.100512

Dietsch, M., Behle, A., Westhoff, P., and Axmann, I. M. (2021). Metabolic engineering of *Synechocystis* sp. PCC 6803 for the photoproduction of the sesquiterpene valencene. *Metab. Eng. Commun.* 13, e00178. doi: 10.1016/j.mec.2021.e00178

Ebrahim, A., Lerman, J. A., Palsson, B. O., and Hyduke, D. R. (2013). COBRApy: COnstraints-based reconstruction and analysis for Python. *BMC Syst. Biol.* 7, 74. doi: 10.1186/1752-0509-7-74

Englund, E., Andersen-Ranberg, J., Miao, R., Hamberger, B., and Lindberg, P. (2015). Metabolic engineering of *Synechocystis* sp. PCC 6803 for production of the plant diterpenoid manoyl oxide. *ACS Synth. Biol.* 4, 1270–1278. doi: 10.1021/acssynbio.5b00070

Englund, E., Pattanaik, B., Ubhayasekera, S. J. K., Stensjö, K., Bergquist, J., and Lindberg, P. (2014). Production of squalene in *Synechocystis* sp. PCC 6803. *PLoS One* 9 (3), e90270. doi: 10.1371/journal.pone.0090270

Englund, E., Shabestary, K., Hudson, E. P., and Lindberg, P. (2018). Systematic overexpression study to find target enzymes enhancing production of terpenes in *Synechocystis* PCC 6803, using isoprene as a model compound. *Metab. Eng.* 49, 164–177. doi: 10.1016/j.ymben.2018.07.004

Erdrich, P., Knoop, H., Steuer, R., and Klamt, S. (2014). Cyanobacterial biofuels: new insights and strain design strategies revealed by computational modeling. *Microb. Cell Fact.* 13, 128. doi: 10.1186/s12934-014-0128-x

Flowers, J. J., Richards, M. A., Baliga, N., Meyer, B., and Stahl, D. A. (2018). Constraint-based modelling captures the metabolic versatility of *Desulfovibrio vulgaris*. *Environ. Microbiol. Rep.* 10, 190–201. doi: 10.1111/1758-2229.12619

Gao, X., Gao, F., Liu, D., Zhang, H., Nie, X., and Yang, C. (2016). Engineering the methylerythritol phosphate pathway in cyanobacteria for photosynthetic isoprene production from CO_2. *Energy Environ. Sci.* 9, 1400–1411. doi: 10.1039/c5ee03102h

Gohil, N., Bhattacharjee, G., Khambhati, K., Braddick, D., and Singh, V. (2019). Engineering strategies in microorganisms for the enhanced production of squalene: advances, challenges and opportunities. *Front. Bioeng. Biotechnol.* 7. doi: 10.3389/fbioe.2019.00050

Hays, S. G., and Ducat, D. C. (2015). Engineering cyanobacteria as photosynthetic feedstock factories. *Photosynth. Res.* 123, 285–295. doi: 10.1007/s11120-014-9980-0

He, H.-P., Corke, H., and Cai, J.-G. (2003). Supercritical carbon dioxide extraction of oil and squalene from amaranthus grain. *J. Agric. Food Chem.* 51, 7921–7925. doi: 10.1021/jf030488y

Hellier, P., Al-Haj, L., Talibi, M., Purton, S., and Ladommatos, N. (2013). Combustion and emissions characterization of terpenes with a view to their biological production in cyanobacteria. *Fuel* 111, 670–688. doi: 10.1016/j.fuel.2013.04.042

Hemmerlin, A., Tritsch, D., Hartmann, M., Pacaud, K., Hoeffler, J. F., van Dorsselaer, A., et al. (2006). A cytosolic *Arabidopsis* d-xylulose kinase catalyzes the phosphorylation of 1-deoxy-D-xylulose into a precursor of the plastidial isoprenoid pathway. *Plant Physiol.* 142 (2), 441–457. doi: 10.1104/pp.106.086652

Hendry, J. I., Bandyopadhyay, A., Srinivasan, S., Pakrasi, H. B., and Maranas, C. D. (2020). Metabolic model guided strain design of cyanobacteria. *Curr. Opin. Biotechnol.* 64, 17–23. doi: 10.1016/j.copbio.2019.08.011

Ho, H.-M., Huang, C.-Y., Cheng, Y.-J., Shen, K.-Y., Tzeng, T.-T., Liu, S.-J., et al. (2021). Assessment of adjuvantation strategy of lipid squalene nanoparticles for enhancing the immunogenicity of a SARS-CoV-2 spike subunit protein against COVID-19. *Int. J. Pharm.* 607, 121024. doi: 10.1016/j.ijpharm.2021.121024

Hudson, P., Lee, S. Y., Nielsen, J., and Stephanopoulos, G. (2021). *Cyanobacteria biotechnology* (Weinheim, Germany: Wiley-VCH Verlag).

Jain, S., Prajapat, G., Abrar, M., Ledwani, L., Singh, A., and Agrawal, A. (2017). Cyanobacteria as efficient producers of mycosporine-like amino acids. *J. Basic Microbiol.* 57, 715–727. doi: 10.1002/jobm.201700044

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., et al. (1996). Sequence analysis of the genome of the unicellular cyanobacterium synechocystis sp. strain PCC6803. II. sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* 3, 109–136. doi: 10.1093/dnares/3.3.109

Karunanithi, P. S., and Zerbe, P. (2019). Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.01166

Katabami, A., Li, L., Iwasaki, M., Furubayashi, M., Saito, K., and Umeno, D. (2015). Production of squalene by squalene synthases and their truncated mutants in *Escherichia coli*. *J. Biosci. Bioeng.* 119, 165–171. doi: 10.1016/j.jbiosc.2014.07.013

King, Z. A., Lloyd, C. J., Feist, A. M., and Palsson, B. O. (2015). Next-generation genome-scale models for metabolic engineering. *Curr. Opin. Biotechnol.* 35, 23–29. doi: 10.1016/j.copbio.2014.12.016

Klaus, O., Hilgers, F., Nakielski, A., Hasenklever, D., Jaeger, K.-E., Axmann, I. M., et al. (2022). Engineering phototrophic bacteria for the production of terpenoids. *Curr. Opin. Biotechnol.* 77, 102764–102022. doi: 10.1016/j.copbio.2022.102764

Knoop, H., and Steuer, R. (2015). A computational analysis of stoichiometric constraints and trade-offs in cyanobacterial biofuel production. *Front. Bioeng. Biotechnol.* 3. doi: 10.3389/fbioe.2015.00047

Kohno, Y., Egawa, Y., Itoh, S., Nagaoka, S.-I., Takahashi, M., and Mukai, K. (1995). Kinetic study of quenching reaction of singlet oxygen and scavenging reaction of free radical by squalene in n-butanol. *Biochim. Biophys. Acta (BBA) - Lipids Lipid Metab.* 1256, 52–56. doi: 10.1016/0005-2760(95)00005-w

Kudoh, K., Hotta, S., Sekine, M., Fujii, R., Uchida, A., Kubota, G., et al. (2017). Overexpression of endogenous 1-Deoxy-D-Xylulose 5-phosphate synthase (DXS) in cyanobacterium synechocystis sp. PCC6803 accelerates protein aggregation. *J. Biosci. Bioeng.* 123 (5), 590–596. doi: 10.1016/j.jbiosc.2017.01.001

Lau, N.-S., Matsui, M., and Abdullah, A. A.-A. (2015). Cyanobacteria: photoautotrophic microbial factories for the sustainable synthesis of industrial products. *Biomed. Res. Int.* 2015, 754934. doi: 10.1155/2015/754934

Leplat, C., Champeimont, R., Saenkham, P., Cassier-Chauvat, C., Jean-Christophe, A., and Chauvat, F. (2013). Genome-wide transcriptome analysis of hydrogen production in the cyanobacterium *Synechocystis*: towards the identification of new players. *Int. J. Hydrogen Energy* 38, 1866–1872. doi: 10.1016/j.ijhydene.2012.11.118

Liang, Z., Zhu, H., Wang, X., Jing, B., Li, Z., Xia, X., et al. (2020). Adjuvants for coronavirus vaccines. *Front. Immunol.* 11. doi: 10.3389/fimmu.2020.589833

Lin, P.-C., Saha, R., Zhang, F., and Pakrasi, H. B. (2017). Metabolic engineering of the pentose phosphate pathway for enhanced limonene production in the cyanobacterium *Synechocystis* sp. PCC 6803. *Sci. Rep.* 7, 17503. doi: 10.1038/s41598-017-17831-y

Lin, P.-C., Saha, R., Zhang, F., and Pakrasi, H. B. (2017). Metabolic engineering of the pentose phosphate pathway for enhanced limonene production in the cyanobacterium *Synechocystis* sp. PCC 6803. *Sci. Rep.* 7, 17503. doi: 10.1038/s41598-017-17831-y

Lichtenthaler, H. K., and Buschmann, C. (2001). Chlorophylls and carotenoids: Measurement and characterization by UV-VIS spectroscopy. *Curr. Protoc. Food Anal. Chem.* 1, F4.3.1–F4.3.8. doi: 10.1002/0471142913.faf0403s01

Mahadevan, R., Edwards, J. S., and Doyle, F. J.3rd (2002). Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* 83, 1331–1340. doi: 10.1016/S0006-3495(02)73903-9

Mendes, A., Azevedo-Silva, J., and Fernandes, J. C. (2022). From sharks to yeasts: squalene in the development of vaccine adjuvants. *Pharmaceuticals* 15 (3), 265. doi: 10.3390/ph15030265

Moss, G. P., Smith, P. A. S., and Tavernier, D. (1995). Glossary of class names of organic compounds and reactivity intermediates based on structure (IUPAC Recommendations 1995). *J. Macromol. Sci. Part A Pure Appl. Chem.* 67, 1307–1375. doi: 10.1351/pac199567081307

Moulin, C., Tournier, L., and Peres, S. (2021). Combining kinetic and constraint-based modelling to better understand metabolism dynamics. *Processes* 9, 1701. doi: 10.3390/pr9101701

Mulkidjanian, A. Y., Koonin, E. V., Makarova, K. S., Mekhedov, S. L., Sorokin, A., Wolf, Y. I., et al. (2006). The cyanobacterial genome core and the origin of photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13126–13131. doi: 10.1073/pnas.0605709103

O'Hagan, D. T., van der Most, R., Lodaya, R. N., Coccia, M., and Lofano, G. (2021). "World in motion" - emulsion adjuvants rising to meet the pandemic challenges. *NPJ Vaccines* 6, 158. doi: 10.1038/s41541-021-00418-0

Okada, K., and Hase, T. (2005). Cyanobacterial non-mevalonate pathway: (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase interacts with ferredoxin in *Thermosynechococcus elongatus* BP-1. *J. Biol. Chem.* 280, 20672–20679. doi: 10.1074/jbc.M500865200

Oliver, J. W. K., and Atsumi, S. (2015). A carbon sink pathway increases carbon productivity in cyanobacteria. *Metab. Eng.* 29, 106–112. doi: 10.1016/j.ymben.2015.03.006

Oliver, N. J., Rabinovitch-Deere, C. A., Carroll, A. L., Nozzi, N. E., Case, A. E., and Atsumi, S. (2016). Cyanobacterial metabolic engineering for biofuel and chemical production. *Curr. Opin. Chem. Biol.* 35, 43–50. doi: 10.1016/j.cbpa.2016.08.023

Park, S. Y., Binkley, R. M., Kim, W. J., Lee, M. H., and Lee, S. Y. (2018). Metabolic engineering of *Escherichia coli* for high-level astaxanthin production w ith high productivity. *Metab. Eng.* 49, 105–115. doi: 10.1016/j.ymben.2018.08.002

Pattanaik, B., Englund, E., Nolte, N., and Lindberg, P. (2020). Introduction of a green algal squalene synthase enhances squalene accumulation in a strain of *Synechocystis* sp. PCC 6803. *Metab. Eng. Commun.* 10, e00125. doi: 10.1016/j.mec.2020.e00125

Pils, D., and Schmetterer, G. (2001). Characterization of three bioenergetically active respiratory terminal oxidases in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *FEMS Microbiol. Lett.* 203, 217–222. doi: 10.1111/j.1574-6968.2001.tb10844.x

Pinto, F. L., Thapper, A., Sontheim, W., and Lindblad, P. (2009). Analysis of current and alternative phenol based RNA extraction methodologies for cyanobacteria. *BMC Molecular. Biol.* 10, 79. doi: 10.1186/1471-2199-10-79

Posten, C., and Schaub, G. (2009). Microalgae and terrestrial biomass as source for fuels–a process view. *J. Biotechnol.* 142, 64–69. doi: 10.1016/j.jbiotec.2009.03.015

Pulz, O., and Gross, W. (2004). Valuable products from biotechnology of microalgae. *Appl. Microbiol. Biotechnol.* 65, 635–648. doi: 10.1007/s00253-004-1647-x

Raman, K., and Chandra, N. (2009). Flux balance analysis of biological systems: applications and challenges. *Brief. Bioinform.* 10, 435–449. doi: 10.1093/bib/bbp011

Ramey, C. J., Barón-Sola, Á., Aucoin, H. R., and Boyle, N. R. (2015). Genome engineering in cyanobacteria: where we are and where we need to go. *ACS Synth. Biol.* 4, 1186–1196. doi: 10.1021/acssynbio.5b00043

Rippka, R., Stanier, R. Y., Deruelles, J., Herdman, M., and Waterbury, J. B. (1979). Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology* 111, 1–61. doi: 10.1099/00221287-111-1-1

Ronco, A. L., and De Stéfani, E. (2013). Squalene: a multi-task link in the crossroads of cancer and aging. *Funct. Foods Health Dis.* 3, 462. doi: 10.31989/ffhd.v3i12.30

Satoh, S., Ikeuchi, M., Mimuro, M., and Tanaka, A. (2001). Chlorophyll *b* expressed in cyanobacteria functions as a light-harvesting antenna in photosystem I through flexibility of the proteins. *J. Biol. Chem.* 276, 4293–4297. doi: 10.1074/jbc.M008238200

Sawai, S., and Saito, K. (2011). Triterpenoid biosynthesis and engineering in plants. *Front. Plant Sci.* 2. doi: 10.3389/fpls.2011.00025

Schellenberger, J., Que, R., Fleming, R. M. T., Thiele, I., Orth, J. D., Feist, A. M., et al. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox v2.0. *Nat. Protoc.* 6, 1290–1307. doi: 10.1038/nprot.2011.308

Schirrmeister, B. E., de Vos, J. M., Antonelli, A., and Bagheri, H. C. (2013). Evolution of multicellularity coincided with increased diversification of cyanobacteria and the great oxidation event. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1791–1796. doi: 10.1073/pnas.1209927110

Segrè, D., Vitkup, D., and Church, G. M. (2002). Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15112–15117. doi: 10.1073/pnas.232349399

Shabestary, K., and Hudson, E. P. (2016). Computational metabolic engineering strategies for growth-coupled biofuel production by *Synechocystis*. *Metab. Eng. Commun.* 3, 216–226. doi: 10.1016/j.meteno.2016.07.003

Shameer, S., Wang, Y., Bota, P., Ratcliffe, R. G., Long, S. P., and Sweetlove, L. J. (2022). A hybrid kinetic and constraint-based model of leaf metabolism allows predictions of metabolic fluxes in different environments. *Plant J.* 109, 295–313. doi: 10.1111/tpj.15551

Shen, C. R., and Liao, J. C. (2013). Synergy as design principle for metabolic engineering of 1-propanol production in *Escherichia coli*. *Metab. Eng.* 17, 12–22. doi: 10.1016/j.ymben.2013.01.008

Smith, T. J., Yang, G. Y., Seril, D. N., Liao, J., and Kim, S. (1998). Inhibition of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone-induced lung tumorigenesis by dietary olive oil and squalene. *Carcinogenesis* 19, 703–706. doi: 10.1093/carcin/19.4.703

Steuer, R., Knoop, H., and Machné, R. (2012). Modelling cyanobacteria: from metabolism to integrative models of phototrophic growth. *J. Exp. Bot.* 63, 2259–2274. doi: 10.1093/jxb/ers018

St. John, P. (2016). "d3flux," in *Openbase*. Available at: https://openbase.com/python/d3flux.

Volke, D. C., Rohwer, J., Fischer, R., and Jennewein, S. (2019). Investigation of the methylerythritol 4-phosphate pathway for microbial terpenoid production through metabolic control analysis. *Microb. Cell Fact.* 18, 1–15. doi: 10.1186/s12934-019-1235-5

Wang, W., and Oldfield, E. (2014). Bioorganometallic chemistry with IspG and IspH: structure, function, and inhibition of the [Fe(4)S(4)] proteins involved in isoprenoid biosynthesis. *Angew. Chem. Int. Ed. Engl.* 53, 4294–4310. doi: 10.1002/anie.201306712

Xu, R., Fazio, G. C., and Matsuda, S. P. T. (2004). On the origins of triterpenoid skeletal diversity. *Phytochemistry* 65, 261–291. doi: 10.1016/j.phytochem.2003.11.014

Yang, Y.-F., Jan, Y.-H., Liu, Y.-P., Yang, C.-J., Su, C.-Y., Chang, Y.-C., et al. (2014). Squalene synthase induces tumor necrosis factor receptor 1 enrichment in lipid rafts to promote lung cancer metastasis. *Am. J. Respir. Crit. Care Med.* 190, 675–687. doi: 10.1164/rccm.201404-0714OC

Zhou, J., Yang, F., Zhang, F., Meng, H., Zhang, Y., and Li, Y. (2021). Impairing photorespiration increases photosynthetic conversion of CO_2 to isoprene in engineered cyanobacteria. *Bioresources Bioprocessing* 8, 1–13. doi: 10.1186/s40643-021-00398-y

frontiers

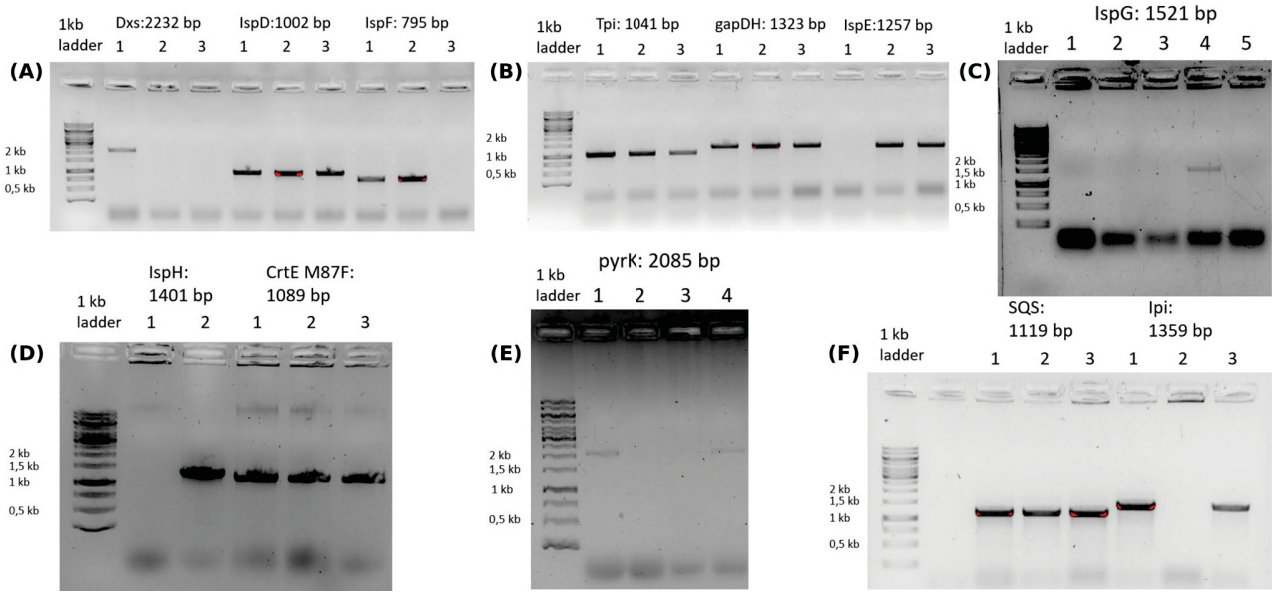
Supplementary Material

Supplementary Table 1. Plasmids used in this study and information regarding their origin.

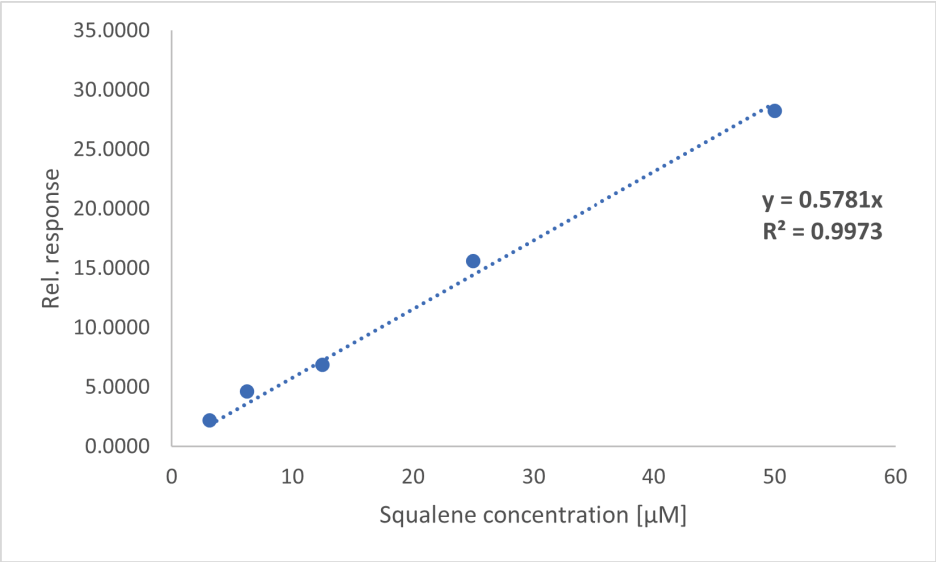
Plasmid name	**Source**
pSHDY rhaS	(Behle *et al.* 2020)
pEERM4	(Englund *et al.* 2015)
pEERM4 Prha dxs	This study
pEERM4 Prha ispD	This study
pEERM4 Prha ispE	This study
pEERM4 Prha ispF	This study
pEERM4 Prha ispG	This study
pEERM4 Prha ispH	This study
pEERM4 Prha idi	This study
pEERM4 Prha sqs	This study
pEERM4 Prha crtE	This study
pEERM4 Prha gap2	This study
pEERM4 Prha pyrK	This study
pEERM4 Prha tpi	This study

Supplementary Table 2. DNA sequences of primers used in this study and their modifications for cloning purposes.

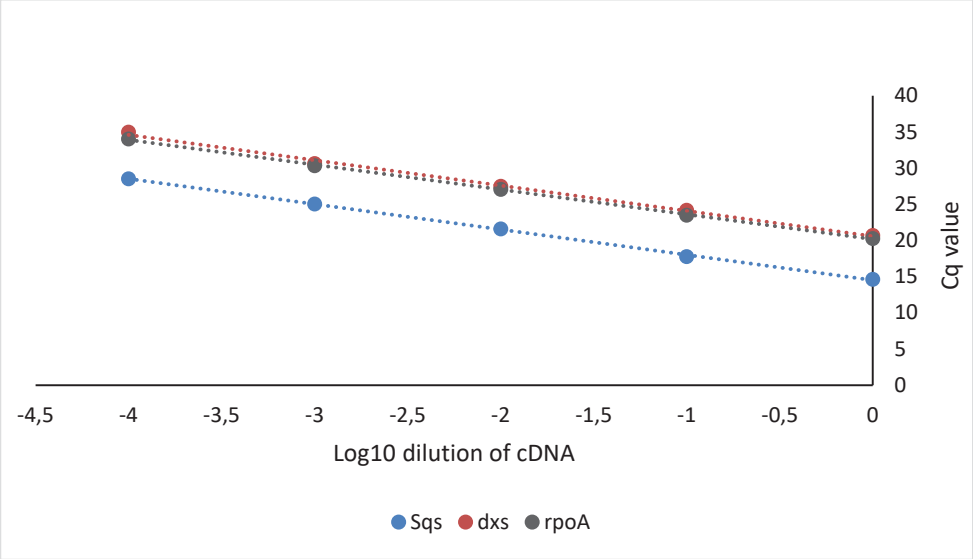
Gene	P fwd (5'-3')	P rev (5'-3')	Other modifications
idi	TGACATGGCTAGCGATA GCACCCCCCACCGTAA	AGCCTGCAGTTAAGGTT TAGTTAACCTTT	
dxs	TGACATGGCTAGCCACATC AGCGAACTGACCCACCCCAA TGAG	GCTACTGCAGCTAACTAACTC CAGGAGCGACAACTG	
sqs	TGACATGGCTAGCTCAG GAGTTGATCGCATGAGC	AGCTACTGCAGCTAACTGG CAATAACCCGATTAA	silent mutation in 110L to remove NheI
ispD	TGACATGGCTAGCCATTT ACTAATTCCAGCGGC	GCTACTGCAGTCAGGCGGA TTTTGCCGACC	
ispE	TGACATGGCTAGCCATT CCTACACCCTCCATGCCCCG	GCTACTGCAGTCAATTATTC ATAATTTGGATGCCG	
ispF	TGACGCTAGCACTGCTC TACGCATCGGCAACGG	GCTACTGCAGTTACCCTTCT TTGATTAACAAAGCCACG	
ispG	TGACATGGCTAGCGT AACCGCTTCCCTGCCGACC	GCTACTGCAGTTAAGGGTCA ACCCAACGGC	
ispH	TGACATGGCTAGCGATACCA AAGCTTTTAAACGGTCTCTGC	GCTACTGCAGCTATCCCGCA ATTTCTAGGACG	
gap2	TGACATGGCTAGCACTA GAGTAGCAATTAACGG	GCTACTGCAGCTATTTCCAGTT TTTAGCCAC	silent mutation in 192A to remove NheI
pyrK	TGACATGGCTAGCCAAA CGTCTCCCCTTCCCCGTCG	GCTACTGCAGCTATCCTTTGG ACACCGGGGGTAATGC	
tpi	TGACATGGCTAGCGTGC GAAAAATCATTATTGC	GCTACTGCAGTCAGGGCTGA AAATTAACAA	
***dxs* qPCR**	CCCATACCAGACTAATGGTG ATT	TGCTGAGGCGGACTTTATTT	
***sqs* qPCR**	GCGATCGATGAAGTGGAAGA	CGTCGCACTCTGGAGATTAAG	
***rpoA* qPCR**	CCATGAGTTCGCCACTATTCT	GGCTGATCGGTGTAGCTTT	
Colony PCR	ATGCGAATTCGCGGCCGCTTC TAGAG	CTGCAGCGGCCGCTACTAGT ATATAAACGCAGAAAGGCC CACCCGAAGG	Colony PCR primers for insert in pEERM4



Supplementary Figure 1. Agarose gel electrophoresis of colony PCR products to prove the integration of the respective gene into the genome through heterologous recombination into the neutral site 2 (NS2). Number denote tested colonies, sizes of the expected PCR bands are shown. PCR was carried out with the colony PCR primers shown in Supplementary Table 2 (A) Dxs: 2232 bp, IspD: 1002 bp, IspF: 795 bp (B) Tpi: 2041 bp, gapDH: 1323 bp, IspE: 1257 bp (C) IspG: 1521 bp (D) IspH: 1401 bp, CrtE M87F: 1089 bp (E) PyrK:2085 bp (F) Sqs: 1119 bp, Ipi: 1359 bp



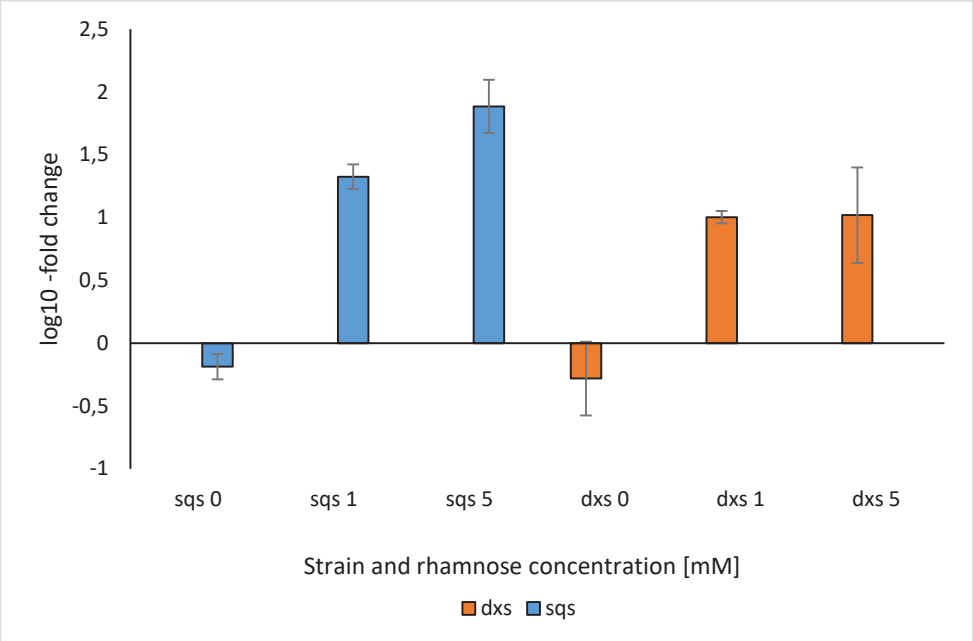
Supplementary Figure 2. GC-MS calibration curve for squalene after extraction of 50, 25, 12.5, 6.25 and 3.125 µM of squalene using the method for squalene extraction from *Synechocystis* cells. Relative response is in relation to the 25 µM β-sitosterol standard, which was solved in the acetone used for extraction.



Supplementary Figure 3: Cq values of qRT-PCR primer pairs used with dilution series of cDNA. Primers for *sqs, dxs* and *rpoA* were tested with cDNA extracted after 3 days from *Synechocystis* Δ*shc* pEERM P_{rha} *sqs* pSHDY rhaS, *Synechocystis* Δ*shc* pEERM P_{rha} *dxs* pSHDY rhaS and *Synechocystis* Δ*shc* pSHDY rhaS, induced with 5 mM rhamnose respectively. Primer sequences are shown in Suppl. Table 2.

Supplementary Table 3: Primer efficiencies of qRT-PCR primers used with dilution series of cDNA. Primers for *sqs, dxs* and *rpoA* were tested with cDNA extracted after 3 days from *Synechocystis* Δ*shc* pEERM P_{rha} *sqs* pSHDY rhaS, *Synechocystis* Δ*shc* pEERM P_{rha} *dxs* pSHDY rhaS and *Synechocystis* Δ*shc* pSHDY rhaS, induced with 5 mM rhamnose respectively. Primer sequences are shown in Suppl. Table 2.

Primer target gene	**Efficiency**
Sqs	**92.96103**
Dxs	**92.99725**
RpoA	**95.64115**

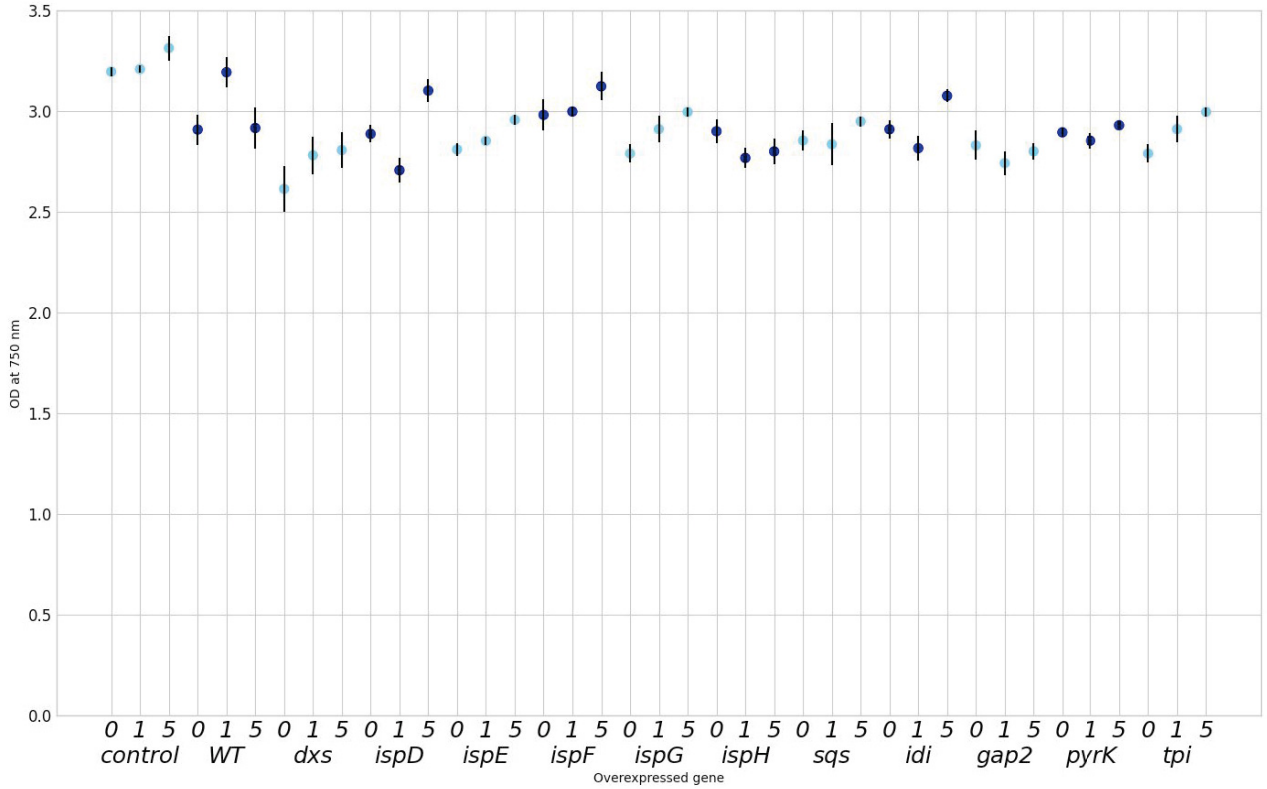


Supplementary Figure 4: Results of qRT-PCR for genes *sqs* and *dxs* as log10-fold changes compared to the control strain in the strains *Synechocystis* Δ*shc* pEERM P_{rha} *sqs* pSHDY rhaS, *Synechocystis* Δ*shc* pEERM P_{rha} *dxs* pSHDY rhaS, respectively. Values were calculated *via* the $2^{-\Delta\Delta C_T}$ method, using *rpoA* as a housekeeping gene and cDNA extracted from *Synechocystis* Δ*shc* pSHDY rhaS, treated with the same rhamnose concentration as a control strain. Primer sequences are shown in Suppl. Table 2. The mean and standard deviation of two biological replicates is shown, which were measured in technical triplicates.

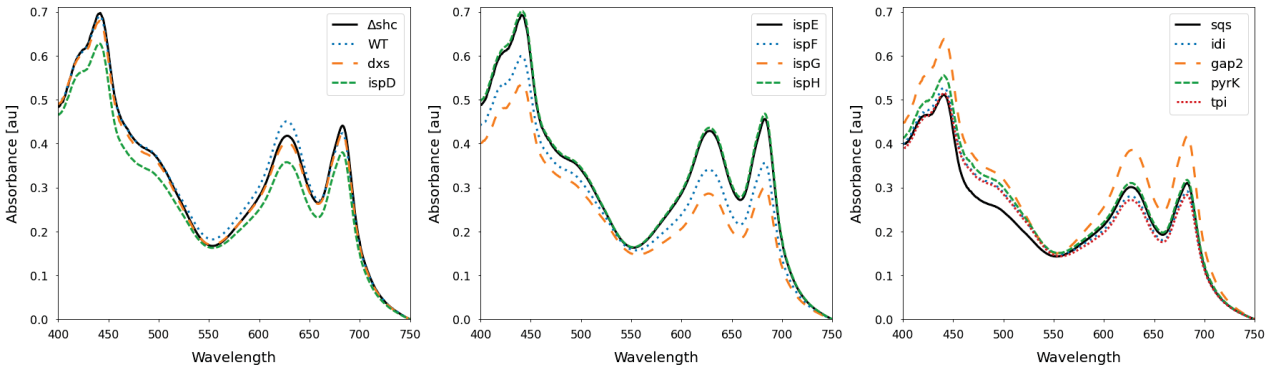
Supplementary Table 4. Squalene yield of all controls and overexpression strains under three different inducer concentrations. WT = Wild type *Synechocystis* sp. PCC 6803, Δshc is the control strain, in which overexpression of the specified genes took place. Mean values and standard deviations of three biological replicates are shown.

Strain	**Rhamnose concentration [mM]**	**Yield [mg L-1]**	**Yield [mg OD750 -1 L-1]**	**[mg gCDW-1]**
WT	**0**	0.003 ± 0.101	0.001 ± 0.023	0.002 ± 0.058
	1	0.01 ± 0.03	0.003 ± 0.008	0.006 ± 0.011
	5	0.004 ± 0.06	0.001 ± 0.017	0.003 ± 0.021
Δshc	**0**	1.5 ± 0.1	0.47 ± 0.03	0.94 ± 0.06
	1	1.41 ± 0.02	0.44 ± 0.01	0.89 ± 0.02
	5	1.39 ± 0.06	0.42 ± 0.01	0.84 ± 0.02
dxs	**0**	1.47 ± 0.04	0.56 ± 0.04	1.13 ± 0.07
	1	1.64 ± 0.22	0.59 ± 0.06	1.18 ± 0.12
	5	1.76 ± 0.27	0.62 ± 0.08	1.26 ± 0.15
ispD	**0**	1.85 ± 0.12	0.64 ± 0.04	1.29 ± 0.07
	1	1.91 ± 0.15	0.71 ± 0.04	1.42 ± 0.08
	5	2.45 ± 0.14	0.79 ± 0.04	1.59 ± 0.08
	0	2.53 ± 0.07	0.9 ± 0.02	1.81 ± 0.03

ispE	**1**	2.52 ± 0.03	0.89 ± 0	1.78 ± 0.01
	5	2.64 ± 0.03	0.89 ± 0.01	1.79 ± 0.02
ispF	**0**	2.29 ± 0.02	0.77 ± 0.03	1.55 ± 0.05
	1	2.28 ± 0.06	0.76 ± 0.01	1.53 ± 0.03
	5	2.25 ± 0.08	0.72 ± 0.02	1.45 ± 0.05
ispG	**0**	1.72 ± 0.19	0.62 ± 0.06	1.24 ± 0.13
	1	1.98 ± 0.03	0.68 ± 0.03	1.37 ± 0.05
	5	2.01 ± 0.2	0.67 ± 0.06	1.35 ± 0.12
ispH	**0**	2.66 ± 0.05	0.92 ± 0.02	1.85 ± 0.03
	1	2.86 ± 0.04	1.03 ± 0.02	2.08 ± 0.03
	5	2.93 ± 0.17	1.05 ± 0.04	2.1 ± 0.08
sqs	**0**	2.13 ± 0.03	0.75 ± 0	1.5 ± 0.01
	1	4.96 ± 0.29	1.76 ± 0.16	3.53 ± 0.33
	5	6.23 ± 0.31	2.11 ± 0.11	4.25 ± 0.22
Idi	**0**	2.09 ± 0.09	0.72 ± 0.02	1.44 ± 0.04
	1	2.19 ± 0.07	0.78 ± 0.01	1.56 ± 0.02
	5	2.52 ± 0.04	0.82 ± 0.02	1.65 ± 0.04
gap2	**0**	1.66 ± 0.36	0.59 ± 0.13	1.18 ± 0.27
	1	1.85 ± 0.22	0.67 ± 0.07	1.36 ± 0.15
	5	1.98 ± 0.14	0.71 ± 0.05	1.42 ± 0.11
pyrK	**0**	1.94 ± 0.06	0.67 ± 0.03	1.35 ± 0.05
	1	1.93 ± 0.05	0.68 ± 0.02	1.36 ± 0.05
	5	2.28 ± 0.19	0.78 ± 0.06	1.57 ± 0.13
tpi	**0**	1.84 ± 0.03	0.66 ± 0.02	1.33 ± 0.05
	1	2.03 ± 0.06	0.7 ± 0.01	1.4 ± 0.02
	5	2.03 ± 0.03	0.68 ± 0.01	1.36 ± 0.03



Supplementary Figure 5. Effect of overexpressions on growth of the different strains after 3 days of growth with the indicated rhamnose concentration. Control denotes the Δ*shc* strain in which the overexpression strains were constructed, WT denotes the *Synechocystis* sp. PCC 6803 wild type. Average values from three biological replicates, error bars represent the standard deviation.



Supplementary Figure 6: Spectra of *Synechocystis* cells after 3 days' incubation with 5 mM rhamnose, measured in 1 cm cuvettes. OD_{750} values were equalized across all measurements in the cuvettes, then the spectra were baseline corrected by subtracting the OD_{750} value.

5 Chapter 5
Manuscript III

5.1 Author's contributions

Anna Behle*, Maximilian Dietsch*, Louis Goldschmidt, Wandana Murugathas, Lutz C Berwanger, Jonas Burmester, Lun Yao, David Brandt, Tobias Busche, Jörn Kalinowski, Elton P Hudson, Oliver Ebenhöh, Ilka M Axmann, Rainer Machné, Manipulation of topoisomerase expression inhibits cell division but not growth and reveals a distinctive promoter structure in *Synechocystis*, Nucleic Acids Research, 2022; gkac1132, https://doi.org/10.1093/nar/gkac1132

* Shared first authorship.

Maximilian Dietsch: Conceptualization, Investigation, Methodology, Writing – Original Draft, Writing– Review & Editing, Data Curation

M.D. designed the study in cooperation with R.M. and A.B. Strains and plasmids were designed and constructed by M.D. and A.B. Endpoint cultivation experiments were performed by M.D. with support by W.M. L.G. set up the continuous cultivation with support by M.D., R.M. and A.B. Biomass measurements were performed by M.D. in cooperation with A.B. M.D. performed experiments for the measurement of single cells and their size distribution, as well as glycogen measurements and quantification and extraction of total DNA and plasmids. M.D. performed chloroquine agarose gel electrophoresis of plasmids in cooperation with A.B, with input by R.M. RNA extraction and processing was performed by M.D. Literature research, and manuscript preparation was performed by R.M. with input and support of M.D an A.B.

The attribution categories described for each journal have been summarized as following:

Study design: 30%
Experimental contribution: 60%
Data analysis: 20%
Manuscript preparation: 10%

5.2 Manipulation of topoisomerase expression inhibits cell division but not growth and reveals a distinctive promoter structure in *Synechocystis*

12790–12808 *Nucleic Acids Research, 2022, Vol. 50, No. 22* *Published online 19 December 2022*
https://doi.org/10.1093/nar/gkac1132

Manipulation of topoisomerase expression inhibits cell division but not growth and reveals a distinctive promoter structure in *Synechocystis*

Anna Behle [1,2,†], Maximilian Dietsch [1,†], Louis Goldschmidt [3], Wandana Murugathas [1], Lutz C. Berwanger [1], Jonas Burmester [1], Lun Yao [4], David Brandt [5], Tobias Busche [5], Jörn Kalinowski [5], Elton P. Hudson [4], Oliver Ebenhöh [3,6], Ilka M. Axmann [1] and Rainer Machné [1,3,*]

[1]Institut f. Synthetische Mikrobiologie, Heinrich-Heine Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany, [2]Photanol B.V, Science Park 406, 1098 XH Amsterdam, The Netherlands, [3]Institut f. Quantitative u. Theoretische Biologie, Heinrich-Heine Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany, [4]School of Engineering Sciences in Chemistry, Biotechnology and Health, Science for Life Laboratory, KTH – Royal Institute of Technology, Stockholm, Sweden, [5]Centrum für Biotechnologie (CeBiTec), Universität Bielefeld, Universitätsstrasse 27, 33615 Bielefeld, Germany and [6]Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

Received September 10, 2021; Revised November 03, 2022; Editorial Decision November 04, 2022; Accepted November 10, 2022

ABSTRACT

In cyanobacteria DNA supercoiling varies over the diurnal cycle and is integrated with temporal programs of transcription and replication. We manipulated DNA supercoiling in *Synechocystis* sp. PCC 6803 by CRISPRi-based knockdown of gyrase subunits and overexpression of topoisomerase I (TopoI). Cell division was blocked but cell growth continued in all strains. The small endogenous plasmids were only transiently relaxed, then became strongly supercoiled in the TopoI overexpression strain. Transcript abundances showed a pronounced 5'/3' gradient along transcription units, incl. the rRNA genes, in the gyrase knockdown strains. These observations are consistent with the basic tenets of the homeostasis and twin-domain models of supercoiling in bacteria. TopoI induction initially led to downregulation of G+C-rich and upregulation of A+T-rich genes. The transcriptional response quickly bifurcated into six groups which overlap with diurnally co-expressed gene groups. Each group shows distinct deviations from a common core promoter structure, where helically phased A-tracts are in phase with the transcription start site. Together, our data show that major co-expression groups (regulons) in *Synechocystis* all respond differentially to DNA supercoiling, and suggest to re-evaluate the long-standing question of the role of A-tracts in bacterial promoters.

INTRODUCTION

In vivo, the DNA double helix exists in a torsionally strained and underwound state, often denoted as 'negative DNA supercoiling'. In bacteria, a homeostatic feedback system of DNA supercoiling is coupled to differential expression of large gene groups. Supercoiling is high during times of high metabolic flux, such as during exponential growth, and is required to express rRNA and G+C-rich growth-related genes and for DNA replication (1). Supercoiling arises as a consequence of DNA transcription and replication and is regulated by two enzymes: gyrase, a heterotetramer of *gyrA* and *gyrB* gene products, can remove positive supercoiling and introduce negative supercoiling, using energy from ATP hydrolysis; and topoisomerase I (TopoI, gene: *topA*) can remove negative supercoiling without any co-factors. The transcription of both enzymes is itself regulated by supercoiling-sensitive promoters in a negative feedback, leading to a homeostatic control of supercoiling (2–6). The ATP dependence of gyrase (7–9) and the control over the expression of growth-related (rRNA, ribosomal proteins, biosynthesis) and G+C-rich genes, and catabolism-related and A+T-rich genes (10–16) extends this homeostatic system to metabolism (Figure 1A).

However, the relation between DNA transcription and replication to supercoiling is mutual and complex (1).

*To whom correspondence should be addressed. Tel: +49 211 81 12923; Email: machne@hhu.de
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

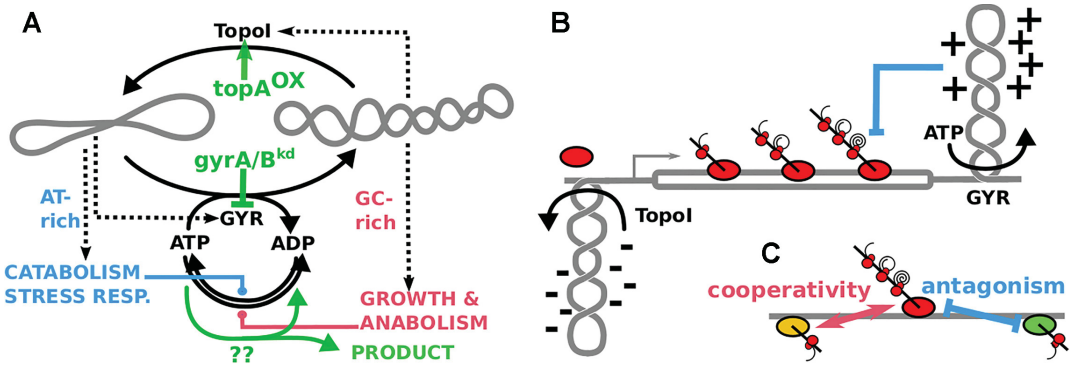


Figure 1. Homeostasis and *Twin-Domain* Models of DNA Supercoiling. (**A**) Global homeostasis of supercoiling by direct feedback on the expression of topoisomerases (GYR: Gyrase holoenzyme; TopoI: topoisomerase I) and G+C-rich anabolic/growth genes and A+T-rich catabolic and stress-response genes. The gray coils reflect relaxed (left) or supercoiled DNA (right). Dashed arrows indicate transcription and solid arrows catalytic conversions. Green arrows indicate the manipulations of this core regulatory hub studied in this work and the underlying hypothesis that these could be used to redirect metabolic energy towards desired products. (**B**) Transcription-dependent DNA supercoiling accumulates downstream (positive) and upstream (negative) of the RNA polymerase, widely known as the *twin-domain* model. If unresolved by TopoI and gyrase, this would lead to RNA polymerase stalling (blue arrow) and R-loop formation. (**C**) The torsional stress exerted by transcription can lead to long-distance cooperative and antagonistic effects, where negative supercoiling upstream facilitates and positive supercoiling downstream blocks transcription from adjacent loci.

According to the twin-domain model (17) of transcription-dependent supercoiling (Figure 1B), negative supercoiling accumulates upstream and positive supercoiling downstream of RNA polymerases, leading to cooperative and antagonistic long-range effects between transcription loci (18) (Figure 1C). Strong transcriptional activity requires downstream activity of gyrase to set the elongation rate and avoid RNA polymerase stalling (19–21) and upstream activity of TopoI to avoid R-loop formation and genome instability (22,23). Such cooperative long-range effects can underpin temporal expression programs; locally in the *leu* operon (24,25) and globally as a spatio-temporal gradient along the origin-terminus axis of the *Escherichia coli* genome (26). The DNA sequence properties of a short region (discriminator) just upstream of the transcription start site are suspected to underlie the common response of many promoters to both supercoiling and to guanosine tetraphosphate (ppGpp) (27–35). Additionally, bacterial and bacteriophage promoters often show a pattern of short repeats of A and T nucleotides (A-tracts) upstream of the core promoter, repeated at distances that match the pitch of the DNA helix (helically phased) (5,36–46), e.g. in light-responsive genes of cyanobacteria (43). On a genome-wide scale (47–54) the helically phased enrichment of A-tract-related dinucleotide motifs is especially pronounced in genomes of polyploid cyanobacteria, including *Synechocystis* sp. PCC6803 (hereafter: *Synechocystis*), where it is found in both intergenic and protein-coding regions (53).

In cyanobacteria and chloroplasts (descendants of cyanobacteria) supercoiling fluctuates with the light/dark (LD) cycle (55,56), and supercoiling homeostasis is integrated with the transcriptional output of the cyanobacterial circadian clock in *Synechococcus elongatus* PCC 7942 (*S. elongatus*) (12,57,58). Recently, the focus has shifted towards the role of ppGpp (59), in dark-phase transcription shutdown (60) and light-phase modulation of the diurnal transcription program (61) of *S. elongatus*, but ppGpp and supercoiling affect the same type of promoters in the same direction in many species (32). In *Synechocystis*, a cold-shock induced increase in supercoiling was found to affect fatty acid synthesis (62), and the transcriptome response to the gyrase inhibitor novobiocin (NB) largely overlapped with the response to various stress conditions (13).

Its central position in metabolism- and growth-related transcription suggests supercoiling as a biotechnological target, where an artificial de-construction and re-construction of this homeostatic feedback may allow to control cellular resource allocation and channel metabolic energy into desired products. Here, we tested the current models of supercoiling in *Synechocystis*, also with respect to this biotechnological potential (Figure 1A). Overexpression of *topA* (63) and CRISPRi-based knockdown of gyrase subunits (64) induced a pleiotropic phenotype, where cell division was blocked but cell volume growth continued. Glycogen and ATP+ADP content increased only upon *topA* overexpression. The transcriptome changed quickly and globally upon induction, compatible with the global homeostatic model (Figure 1A), then remained locked in a state reflecting the dark/light transition at dawn. A graded response at rRNA loci and growth-related transcription units in gyrase knockdown strains is compatible with the twin-domain model (Figure 1B, C), where gyrase activity downstream of RNA polymerase facilitates strong transcription. Coexpressed groups of transcription units show significant deviations from a common core promoter structure.

MATERIALS AND METHODS

Strains and plasmids

The *Synechocystis* parental strain used for all genetic alterations is described by Yao *et al.* (64) and contains an CRISPRi-based gene knockdown system where both the dCas9 protein and the sgRNA expression are inducible by anhydrotetracycline (aTc). All strains further contained our pSNDY plasmid for rhamnose-inducible overexpression (63). Supplementary Table S1 provides details on strain construction and plasmid design. The sgRNA protospacer sequences (Supplementary Table S2) were designed with

CHOPCHOP (65), and potential off-targets were predicted after Cui *et al.* (66).

Batch culture conditions

For pre-culturing, growth and endpoint experiments, strains were cultivated in BG11 medium (67) in 100 ml Erlenmeyer flasks in an Infors HT multitron chamber, at 30°C with 150 rpm shaking, continuous illumination of ≈80 $\mu mol\, m^{-2}\, m^{-1}$ and CO_2 enriched air (0.5%). Pre-culturing was performed in 20 ml cell suspension for three days. For experiments, cultures were adjusted to $OD_{750} \approx 0.4$ in 30 ml, grown for one additional day, then density was again adjusted to a start OD_{750} ≈0.25 and all inducers were added (100 ng/ml aTc, 1 mM L-rhamnose) to each strain at time $t = 0$ h. Antibiotics were added to liquid and solid media as required to maintain genetic constructs, i.e., 25 $\mu g\, ml^{-1}$ (liquid) and 50 $\mu g\, ml^{-1}$ (solid) nourseothricin, 20 $\mu g\, ml^{-1}$ spectinomycin and 25 $\mu g\, ml^{-1}$ kanamycin.

Continuous culture, online measurements and calculations

The continuous culture was performed in a Lambda Photobioreactor (PBR) (Appendix A, Figure A1) in BG11 medium, supplemented with the required antibiotics, at culture volume $V_\ell = 1$ l, aeration with 1 l min^{-1} of CO_2-enriched (0.5%) air, agitation by the Lambda fish-tail mixing system at 5 Hz, temperature control at 30°C, and pH 8, with 0.5 M NaOH and 0.5 M H_2SO_4 as pH correction solutions. After equilibration to these conditions the reactor was inoculated to a start OD_{750} ≈0.5, from 100 ml preculture. White light from the Lambda LUMO module was calibrated to $\mu mol\, m^{-2}\, m^{-1}$ (Figure A2E and F) and intensity adjusted to achieve ≈90 $\mu mol\, m^{-2}\, m^{-1}$ per OD_{750} over the experiment (Figure A2F). For evaporation control and continuous culture mode, the total weight of the reactor setup was kept constant using the built-in Lambda reactor mass control module and automatic addition of fresh culture medium through the feed pump. Continuous culture was performed by setting the waste pump to a fixed speed. The PBR was equipped with additional monitoring of optical density by a DASGIP OD4 module, calibrated to offline OD_{750} (Figure A2A and B), and monitoring of offgas O_2 and CO_2 concentrations and the weights of feed and pH control bottles by Arduino-based custom-built data loggers (Figure A1). Culture evaporation and dilution rates and biomass growth rates were calculated from the slopes of measured data (Figure A3) using piecewise linear segmentation with our R package `dpseg` (https://cran.r-project.org/package=dpseg, version 0.1.2 at https://gitlab.com/raim/dpseg/). Cell volume growth rate was calculated as the rate of change of the peaks of the CASY cell volume distributions.

Biomass measurements: OD, spectra and cell dry weight

The optical density (OD_{750}) and absorbance spectra were measured on a Specord200 Plus (Jena Bioscience) dual path spectrometer, using BG11 as blank and reference. Samples were appropriately diluted with BG11 before measuring. All $topA^{OX}$ time series samples were diluted 1:4 before recording OD_{750}. For absorbance spectra the OD_{750} was adjusted to 0.5 before measurement. The spectra were all divided by the absorbance at 750 nm.

To determine the cell dry weight (CDW) 5 ml cell culture was filtered through a pre-dried and pre-weighed cellulose acetate membrane (pore size 0.45 μm) using a filtering flask. After that the membrane was dried at 50°C for 24 h and weighed after cooling. 5 ml of filtered and dried growth medium served as a blank. For normalization of glycogen measurements by biomass and for estimation of the biomass density of cells (g_{DCW}/ml_{cell}), Figure 5A) the OD_λ signal was calibrated to CDW (Figure A2C and D).

Cell count and size distributions

To determine cell counts and size distributions, 10 μl cyanobacteria culture, pre-diluted for OD_{750} measurement, were dispensed in 10 ml CASYton and measured with a Schaerfe CASY Cell Counter (Modell TTC) using a diameter 45 μm capillary. Cell size was recorded in the diameter range 0–10 μm. Each sample was measured with 400 μl in triplicate runs. Analysis of the raw data was performed in R. Counted events in the CASY are a mix of live cells, dead cells, cell debris and background signals. Only counts with diameter $d > 1.5$ μm and $d < 5$ μm were considered for the time series experiment (Figure 5) while a lower cut-off $d > 1.25$ μm was used for the endpoint measurements (Figure 2B) to avoid cutting the distribution of the slightly smaller $topA^{kd}$ cells. Since *Synechocystis* cells are spherical, the cell volumes were calculated from the reported cell diameters d as $V_{cell} = (\frac{d}{2})^3 \pi \frac{4}{3}$.

Glycogen measurement

0.5 ml of cell culture was harvested into reaction vessels that had been pre-cooled on ice, samples were centrifuged at maximum speed (5 min, 4°C). The pellets were flash-frozen in liquid nitrogen and stored at –80°C. The pellets were resuspended in 400 μl KOH (30% w/v) and incubated (2 h, 95°C). For precipitation, 1200 μl ice cold ethanol was added and the mix incubated (over night, –20°C). After centrifugation (10 min, 4°C, 10 000 g), the pellet was washed once with 70% ethanol and again with pure ethanol. The pellets were dried in a Concentrator Plus (Eppendorf) speed-vac (20 min, 60°C). To degrade glycogen to glucose units, pellets were resuspended in 1 ml 100 mM sodium acetate (pH 4.5) supplemented with amyloglucosidase powder (Sigma-Aldrich, 10115) at a final concentration of 35 U/ml and incubated (2 h, 60°C). The sucrose/D-glucose assay kit from Megazyme (K-SUCGL) was applied according to the manufacturer's specifications to measure the total glucose content, but omitting the fructosidase reaction step and scaling down the total reaction volume to 850 μl. Absorbance at 510 nm was measured using a BMG Clariostar photospectrometer.

ATP and ADP measurement

2 ml tubes were preloaded with 250 μl of buffer BI (3 M $HClO_2$, 77 mM EDTA). 1 ml culture sample was added, vortexed and incubated (lysis, 15 min on ice). 600 μl of BII

(1 M KOH, 0.5 M KCl, 0.5 M Tris) were added (neutralization). Samples vortexed and incubated (10 min, on ice), centrifuged (10 min, 0°C, 12 000 g), flash-frozen in liquid nitrogen and stored at –80°C. Extracts were thawed on ice and centrifuged (10 min, 0°C, 12 000 g). 200 µl samples were added either to 320 µl of BIII/PEP (100 mM HEPES, 50 mM $MgSO_4 \cdot 7H_2O$, adjusted to pH 7.4 with NaOH, and 1.6 mM phosphoenolpyruvate (Sigma-Aldrich, 860077)) for ATP quantification or BIII/PEP + PK (BIII/PEP with 2 U/µl pyruvate kinase, (Sigma-Aldrich, P1506)) for ATP + ADP quantification, incubated (30 min, 37°C), and heat-inactivated (10 min, 90°C). ATP concentrations were determined using the Invitrogen ATP determination kit was used (ThermoFisher: A22066). 10 µl of each PEP or PEP + PK-treated sample was loaded in a white 96-well plate with solid bottom and kept on ice until the reaction was started. The luciferase master mix was scaled down in volume, and 90 µl of master mix was added to each well. Luminescence was recorded using a BMG Clariostar. ATP concentrations were calculated from a standard curve on the same plate.

Microscopy

500 µl cell culture was sampled four days after induction and mixed with glutaraldehyde to 0.25%. After incubating for 15 min at room temperature (RT) samples were flash-frozen in liquid nitrogen and stored at –80°C. Cells were thawed on ice for 2 h and additionally 30 min at RT. Then washed twice with 1 ml 1× PBS (phosphate buffered saline) and stained with HOECHST 33342 (1 µg ml^{-1}, ThermoFisher: 62249) and propidium iodide (30 µm, ThermoFisher: L13152). After 15 min incubation cells were washed with 1 ml 1× PBS. Coverslips (18 × 18 mM, IDL: 19 00 02460) were covered with poly-L-lysine solution for 5 min. Poly-L-lysine solution was removed with a pipette. Coverslips were placed in six-well plates and covered with 1 ml 1× PBS, 10 µl cell suspension was added, and the well plates centrifuged at 1500 g for 15 min. Coverslips were placed on slides and images were captured with the Olympus FluoView FV3000 confocal microscope. HOECHST fluorescence was excited with a 405 nm laser and emission was captured from 430 to 470 nm. Chlorophyll was excited with a 640 nm laser and emission was captured from 650 to 750 nm. Images were analyzed using Fiji (ImageJ, version: 2.1.0/1.53f51). To automatically detect cells and measure cell dimensions the plugin ObjectJ (version: 1.04z) and its Coli-Inspector macro were used (68), with minimum and maximum widths of 0.5 and 3.5 µm, applied to the chlorohpyll fluorescence images. Some objects were manually edited: undetected cells were added, 8-shaped cells recognized as two single cells were merged, and artifacts marked as objects were deleted. To determine the ratios of single cells and 8-shaped cells, cells were counted manually. Images for publication were prepared following the QUAREP-LiMi guidelines (69).

Flow cytometry and analysis

Samples were fixed in 4% para-formaldehyde in 1× PBS, washed three times in 1× PBS, and stained with the SYTO9 green fluorescent nucleic acid stain from the LIVE/DEAD BacLight kit (ThermoFisher, L13152) according to manufacturer's instructions. The flow cytometric measurements were taken at the FACS Facility at the Heinrich-Heine University (Dipl.-Biol. Klaus L. Meyer) using a BD FACSAria III. Forward scatter (FSC) and side-scatter (SSC) were recorded. Syto9 was measured with a 530/30 nm filter, and chlorophyll fluorescence was measured with 695/40 nm filter. For each sample 10 000 events (cells, debris and background) were recorded. Data was exported in `.fcs` format, parsed and analyzed using the `flowCore` R package (70), and plotted using our in-house `segmenTools` R package.

Total DNA and plasmid extractions

To isolate total DNA, 1 ml culture was centrifuged at maximum speed (10 min, 4°C), flash-frozen in liquid nitrogen and stored at −80°C. Thawed samples were resuspended in 1 ml 1× TE buffer, and incubated (1 h, 37°C) with 100 µl lysozyme (50 mg/ml stock solution). 10 µl Proteinase K (20 mg/ml) and 100 µl 20% SDS were added and samples incubated (20 h, 37°C). DNA was extracted in Phasemaker Tubes (ThermoFisher: A33248) with one volume of phenol/chloroform/isoamyl alcohol, centrifugation at maximal speed (10 min, 4°C). The upper phase was transferred, mixed with 100 ng/µl RNAse A and incubated (15 min, 37°C). After addition of 1 volume of chloroform/isoamyl alcohol, the centrifugation step was repeated. DNA was precipitated from the upper phase with 1 volume 2-propanol (over night, −20°C), and pelleted by centrifugation at maximal speed (10 min, 4°C). The pellet was washed twice with 500 µl ice-cold 70% EtOH and centrifuged at maximal speed (10 min, at 4°C), dried at room temperature, and resuspended in 30 µl MilliQ water.

To isolate the small endogenous plasmids, 20 ml of cell culture were mixed with 20 ml of undenatured 99.5% ethanol, pre-cooled to −80°C, in 50 ml centrifuge tubes and stored at −80°C until processing. Samples were thawed on ice, centrifuged (10 min, 4°C, 4000 g). The QIAprep Spin Miniprep kit was modified to extract plasmids from the pellet. The cell pellet was resuspended in 250 µl Qiagen P1 solution and transferred to 1.5 ml reaction tubes, 50 µl lysozyme solution (50 mg ml^{-1}) was added, and the mix incubated (1 h, 37°C). Then 55 µl of 20% SDS and 3 µl of proteinase K (20 mg ml^{-1}) were added and the mix incubated (16 h, 37°C). Further extraction proceeded with alkaline lysis (Qiagen P2) as per manufacturer's instruction but with volumes adjusted. To enrich covalently closed circular DNA, the samples where digested with the T5 exonuclease (NEB: M0363, 30 min, 37°C), and purified with the QIAprep Spin Miniprep kit.

Chloroquine agarose gel electrophoresis of plasmids

Agarose gels (1.2%) with 20 µg ml^{-1} chloroquine diphosphate (CQ, Sigma: C6628-50G, CAS: 50-63-5 in 0.5× TBE buffer) were performed as detailed at protocols.io (https://dx.doi.org/10.17504/protocols.io.rbcd2iw) and in a `bioRxiv` preprint (71), Briefly, gels were run at 1.8 V cm^{-1}, protected from light and for 18 h–22 h (as indicated, Supplementary Figure S2), stained with SYBR Gold (ThermoFisher: S11494) and imaged on a BioRad Imaging System (ChemiDoc MP). Electropherograms of each lane were

extracted in ImageJ and processed in R, with smoothing and peak detection functions from the `msProcess` R package (version 1.0.7) (https://cran.r-project.org/web/packages/msProcess/). A baseline was determined in two steps using the `msSmoothLoess` function. The first step used the full signal and served to determine the coarse positions of peaks. The final baseline was then calculated from the signal after removal of peak values and subtracted from the total signal and subtracted from all electropherograms.

RNA extraction and processing

1 ml culture was added to 250 μl pure ethanol supplemented with 5% phenol, flash-frozen in liquid nitrogen and stored at −80°C. RNA was extracted after (72) with some modifications. Frozen samples were centrifuged (10 min, 4°C, maximum speed), and the pellet resuspended in 1 ml PGTX (per 1 l: 39.6 g phenol, 6.9 ml glycerol, 0.1 g 8-hydroxyquinoline, 0.58 g EDTA , 0.8 g sodium acetate, 9.5 g guanidine thiocyanate, 4.6 g guanidine hydrochloride and 2 ml Triton X-100) and incubated (5 min, 95°C). After cooling on ice for 2 min, 700 μl chloroform:isoamyl alcohol (24:1) was added and the mixture incubated (10 min, room temperature) while shaking gently. The mixture was centrifuged (10 min, 4°C, maximal speed). The upper phase was transferred to a fresh tube and 1 volume chloroform:isoamyl alcohol was added. After repeating the centrifugation step, the upper phase was transferred and precipitated with 3 volumes of 99.5% ethanol and 1/2 volume 7.5 M ammonium acetate and (time series only) 1 μl RNA-grade glycogen at −20°C over night. The RNA was pelleted by centrifugation (30 min, 4°C, maximum speed), washed twice with 70% ethanol and resuspended in 30 μl RNase-free water. Volumes were adjusted to contain 2 μg of nucleic acid (Nanodrop), and DNA was removed by DNaseI (ThermoFisher: EN0525) according to the manufacturer's specifications but at 2× reaction buffer concentration. RNA was extracted as above but using 1/10 volume of 3 M sodium acetate (pH 5.3) instead of ammonium acetate.

Quantitative RT-PCR

100 ng DNaseI-digested RNA samples were reverse-transcribed to cDNA using the RevertAid RT (ThermoFisher: K1621) according to the manufacturer's specifications in a reaction volume of 20 μl, and RT-qPCR performed with the DyNAmo ColorFlash SYBR Green qPCR-Kit (ThermoFisher: F416L). Briefly, 60 μl RNase-free water was added to the cDNA reaction mix. 2 μl (2.5 ng) were transferred into qPCR 96-well microplates and 8 μl Master Mix added. Primer efficiencies (Supplementary Table S3) were assessed from calibration curves. Primers were added at a final concentration of 0.5 mM. The thermal cycling conditions were: 7 min at 95°C, followed by 40 cycles of 5 s at 95°C and 30 s at 60°C. Melting curves were recorded for each sample to ensure sample purity. RT-negative controls and no-template-controls (distilled water) were included for each run. Each sample was loaded in technical triplicates. Gene expression changes at indicated time points were then quantified by the $\Delta\Delta Ct$ method (73), using *rpoA* as a reference gene (74), and a time point before induction of genetic construct (time series) or the empty vector control (EVC) strain (batch culture endpoint experiments) as the reference expression state. $\Delta\Delta Ct$ is then the $\log_2$ fold-change with respect to this reference state. MIQE guidelines were followed where applicable.

RNAseq: total RNA analysis, library generation and sequencing

DNaseI-digested RNA samples (25 μl) were sent for sequencing analysis. RNA quality was evaluated spectrometrically by Trinean Xpose (Gentbrugge, Belgium) and by size separation by capillary gel electrophoresis on an Agilent 2100 Bioanalyzer with the RNA Nano 6000 kit (Agilent Technologies, Böblingen, Germany). For total RNA analysis, electropherograms were parsed from exported XML files using the R package `bioanalyzeR` (v 0.9.1, obtained from https://github.com/jwfoley/bioanalyzeR) (75), and each lane was divided by the total RNA content as reported by the Agilent 2100 Bioanalyzer software. The Illumina Ribo-Zero Plus rRNA Depletion Kit was then used to remove the ribosomal RNA, and removal confirmed by capillary gel electrophoresis as above. Preparation of cDNA libraries was performed according to the manufacturer's instructions for the TruSeq stranded mRNA kit (Illumina, San Diego, CA, United States). Subsequently, each cDNA library was sequenced on an Illumina NextSeq 500 system (2 × 75 nt, PE high output v2.5).

RNAseq: read mapping

The resulting sequence reads were quality trimmed with Trimmomatic v0.33 (76) using standard setting. The quality trimmed reads were subsequently mapped to coding genes of the Synechocystis sp. PCC 6803 reference genome, its seven endogenous plasmids and our pSNDY construct (Supplementary Table S4) using Bowtie 2 (77). For the endpoint measurements from batch cultures the $\log_2$ fold changes with respect to the control (EVC) were calculated with the DESeq2 algorithm (78) *via* the ReadXplorer software version 2.0 (79), based on three replicate measurements for each strain ('M-value'), and these values are denoted $\log_2$(<strain>/EVC) in figures, where <strain> is the tested strain and EVC is the control strain. For the analysis of the expression gradient within transcription units, the difference of these values between the first and the last transcribed gene of each TU was taken. This difference equals the $\log_2$ ratio of the fold changes. For the time series read count data were normalized by library sizes to the transcripts per kilobase million (TPM) unit. Missing values at individual time points were interpreted as 0 TPM. For plots, the $\log_2$ fold change of each time point to the mean of the two pre-induction time points was calculated, denoted as $\log_2(x_i/\overline{x_{1,2}})$ in figures.

Clustering analyses

For clustering the time series into co-expressed groups, a previously established pipeline was used (80,81). The input time series were RNAseq samples 2 to 16 (from –0.5 to 72 h around the time of induction at 0 h), i.e., without the first pre-induction time-point and ignoring the two long-term

response samples (Supplementary Figure S12C). Briefly, the time-series of TPM values was `arcsinh`-transformed, the Discrete Fourier Transform (DFT) X_k was calculated, each DFT component $k > 0$ normalized (X'_k) to the mean of amplitudes at all other components $k > 1$. The real and imaginary parts of selected components $X'_{k=1,\ldots,6}$ were then clustered with the `flowClust` algorithm (82) over cluster numbers $K = 2, \ldots, 10$. The clustering with the maximal Bayesian Information Criterion, as reported by `flowClust` (Supplementary Figure S12A), was selected for further analyses. Data transformation and clustering were performed by the `processTimeseries` and the `clusterTimeseries2` functions of `segmenTier` and `segmenTools` R packages (81), respectively. The resulting clusters were sorted and colored based on the comparison with diurnal co-expression cohorts (Figure 6 and Supplementary Figure S17) for informative plots of the subsequent analyses. To map this clustering from genes to transcription units (TU) (83), the mean expression of all coding genes in each TU was calculated. The resulting TU time-series were then clustered by `k-means`, using the cluster centers from the gene-based clustering as input (Supplementary Figure S18). To estimate the immediate transcriptional response to *topA* overexpression the log2 ratio of the means of the two post-induction time points (5 min, 20 min, or as indicated) to the means the two pre-induction time points (−1 day, −35 min) were calculated ($\log_2(\overline{x_{3,4}}/\overline{x_{1,2}})$). Transcripts with negative values ($<-\theta$) were labeled as 'down', with positive values ($>\theta$) as 'up', and all others as 'nc' (for 'no change'). A low threshold $\theta = 0.01$ was used for the gene-level analysis (Figure 6D, E), since here a comprehensive picture of directionality was desired, and a stricter $\theta = 0.15$ for TU-level analysis (Supplementary Figure S18C) since the extremes were of interest for promoter structural analysis. Diurnal expression data (84) were obtained from GEO (GSE79714) and genes summarized as the mean over all associated probes. These expression values were clustered (Supplementary Figure S17) as described for the RNAseq data, but using the `flowclusterTimeseries` function.

Cluster enrichment profiles

Categorical enrichments, e.g., coding gene co-expression cohorts vs. gene annotations, were analyzed by cumulative hypergeometric distribution tests (R's `phyper`) using `segmenTools`'s `clusterCluster` function and the `clusterAnnotation` wrapper for GO and and protein complex analysis, which compares overlaps of each pair of two distinct classifications into multiple classes, and stores overlap counts and *P*-values (enrichment tables). To analyze $\log_2$ fold-changes by clusters two-sided `t-test` were performed (R base function `t-test`, incl. Welch approximation for different sample sizes), comparing the distribution of values of the cluster with all other values (function `clusterProfile`).

For intuitively informative plots the enrichment table rows were sorted along the other dimension (table columns) such that all categories enriched above a certain threshold p_{sort} in the first column cluster are moved to the top, and, within, sorted by increasing *p*-values. Next, the same sorting was applied to all remaining row clusters for the second column cluster, and so on until the last column cluster. Remaining row clusters are either plotted unsorted below a red line or removed. This is especially useful to visualize enrichment of functional categories along the temporal program of co-expression cohorts, e.g., Figure 6B. This sorting is implemented in `segmenTools`' function `sortOverlaps`.

Sorted enrichment tables were visualized as colored table plots (Enrichment Profiles) (e.g. Figure 6B, C), using `segmenTools`' function `plotOverlaps`. For the categorical overlap tests, the total counts of overlapping pairs are plotted as text, and for t-test profiles the rounded *t* statistic. The text color is black or white based on a p-value cutoff p_{txt} (as indicated). The field background color intensities scale with $\log_2(p)$ of the reported p-values, where the full color corresponds to a minimal p-value $p_{\min}$ cutoff (as indicated) and white reflects $p = 1$. For categorical enrichment tests the full color is black. For numerical tests, the sign of the *t* statistic is used to determine a color to indicate the direction of change: red for negative values ($t < 0$, downregulated) and blue for positive values ($t > 0$, upregulated).

Promoter nucleotide frequency profiles

Only transcription units from the main chromosome were considered for promoter structure analysis. The genome sequence was converted into a vector of 0 and 1, where 1 indicates occurrence of the motif under consideration. Motif occurrence vectors upstream and downstream of transcription start sites were extracted from the genome vector and aligned into a matrix (columns: positions around the alignment anchor, rows: all genomic sites under consideration). The occurrence of a motif in all sequences of a cluster were counted at each position in 66 or 5 bp windows surrounding the position. Cumulative hypergeometric distribution tests (R's `phyper`) were performed to analyze statistical enrichment or deprivation within the window of all TU in a cluster vs. the same window in all TU. The mean position-wise motif occurrence (frequency, in %) was plotted on the y-axis and the size of the plotted data point was scaled by the enrichment and deprivation p-values to emphasize regions of significant difference. The maximal size was determined by the minimum p-value in each test series, as indicated in the Figure legends. The point style (closed or open circles) indicates the directionality of the test (enriched or deprived). These significance points are shown at every third or tenth position to avoid overlaps.

Other data sources

Genome sequences and annotation were downloaded from NCBI (Supplementary Table S4). The gene 'categories' annotation was downloaded on 2017-09-23 from `CyanoBase`, file category.txt (85). Gene Ontology annotation was downloaded from the `UniProt` database (2021-03-20, `organism:1111708`) (86). Datasets from other publications were all obtained from the supplemental materials of the indicated publications.

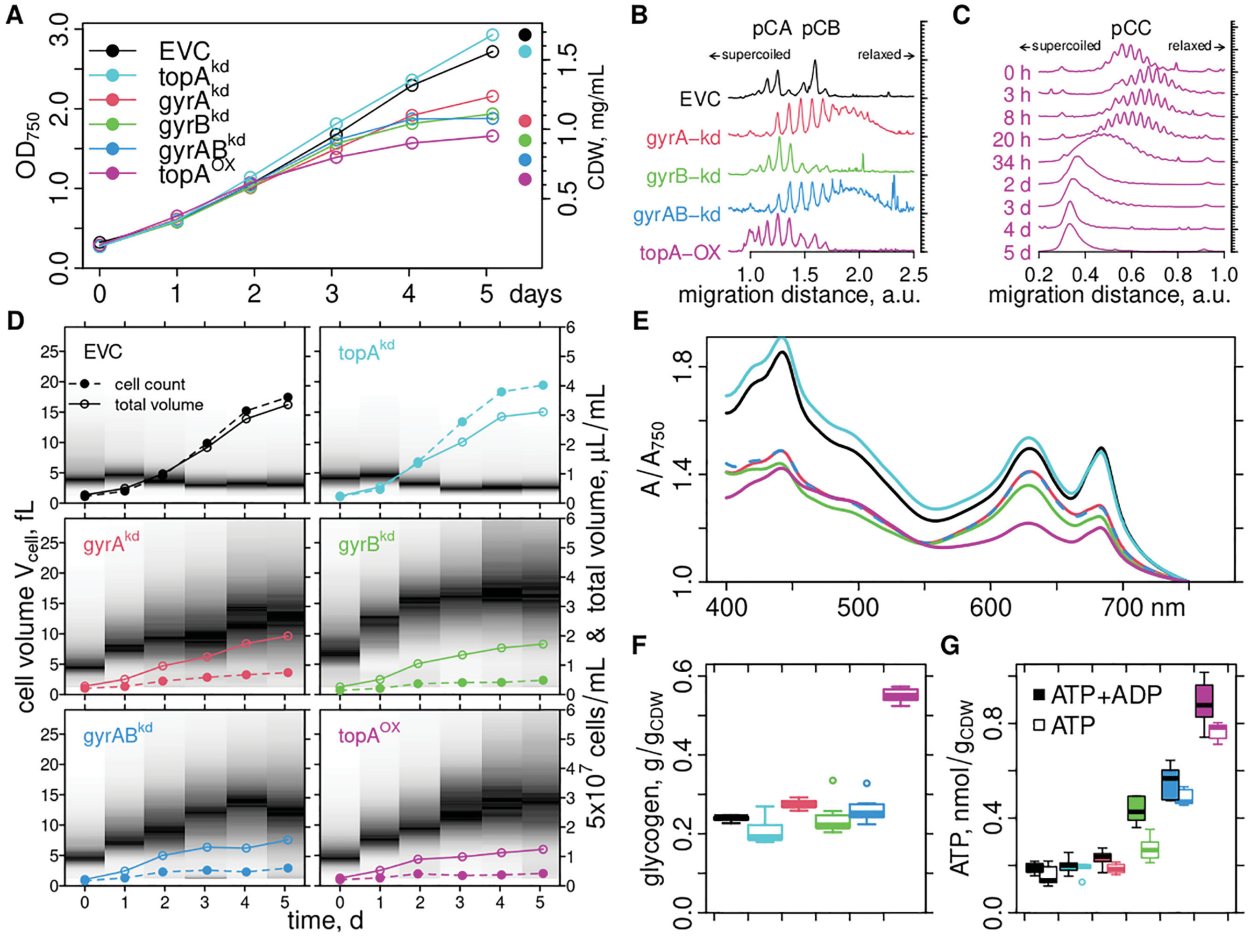


Figure 2. Batch culture endpoint measurements. Overexpression and knockdown strains of this study where grown for 5 days in BG11 medium supplemented with all required antibiotics, and all inducers for the plasmid constructs in each experiment (100 ng/ml aTc, 1 mM L-rhamnose). (**A**) The optical density at 750 nm (OD_{750}) was measured daily and cell dry weight (CDW) determined directly after the last measurement on day 5. (**B**, **C**) Electropherograms of chloroquine-supplemented agarose gels (1.2% agarose, 20 $\mu g\ ml^{-1}$ chloroquine) of plasmids extracted at harvest time (B) of the cultures in (A), or as a time series (growth curve, Supplementary Figure S2D) of the $topA^{OX}$ strain (C). The migration direction of more supercoiled and more relaxed topoisomers is indicated. See Supplementary Figure S2 for the original gel images. (**D**) Cell counts and size distributions were measured daily in the CASY cell counter and plotted as a gray-scale gradient (black: more cells at this volume). (**E**) Absorption spectra after the harvest on day 5. See Supplementary Figure S1B for spectra at inoculation time. All spectra were divided by the absorption at 750 nm. (**F**) Glycogen content at harvest time was determined by a colorimetric assay after harvest, and boxplots of 18 technical replicates (three samples, each measured 3× in two assays) are shown. (**G**) ATP and ATP+ADP contents at harvest time were determined by a luciferase-based assay, and boxplots of six technical replicates (three samples and two measurements) are shown.

RESULTS

Cell division block and redirection of cellular resources

Manipulation of gyrase and Topoisomerase I expression. Based on the current models of the role of DNA supercoiling homeostasis in bacteria (Figure 1A), we hypothesized that artificial genome relaxation should inhibit growth and redirect metabolic flux. To test this idea, we constructed three strains (Supplementary Table S1) to inducibly repress (knockdown, kd) gyrase subunits with the dCas9-mediated CRISPR-interference system (64), and one strain to overexpress TopoI: strains $gyrA^{kd}$ (target: *slr0417*), $gyrB^{kd}$ (*sll2005*) and $gyrAB^{kd}$ (both subunits), all inducible by anhydrotetracycline (aTc); and strain $topA^{OX}$ with *slr2058* with a rhamnose-inducible promoter on the pSNDY plasmid (63). As controls, we included a TopoI knockdown strain ($topA^{kd}$), and an empty vector control (EVC) strain, bearing all plasmids but without the sgRNA or the *topA* gene. All six strains were induced with aTc and rhamnose and cultured in continuous light for 5 days (Figure 2A), then harvested for quantification of plasmid supercoiling, cell dry weight, ATP + ADP, and glycogen. Reverse transcription quantitative PCR (RT-qPCR) verified the functionality of our inducible genetic constructs, but two reference genes gave disparate results (Supplementary Figure S1A). This points to global changes of the transcriptome and precludes quantification in terms of fold changes by RT-qPCR, which we resolve below by RNAseq analysis.

Hypernegative plasmid supercoiling in the $topA^{OX}$ strain. To analyze the specificity of our manipulations, we first investigated the effects on plasmid supercoiling by agarose gel electrophoresis in the presence of an intercalator. Samples taken from the harvest time of the batch growth ex-

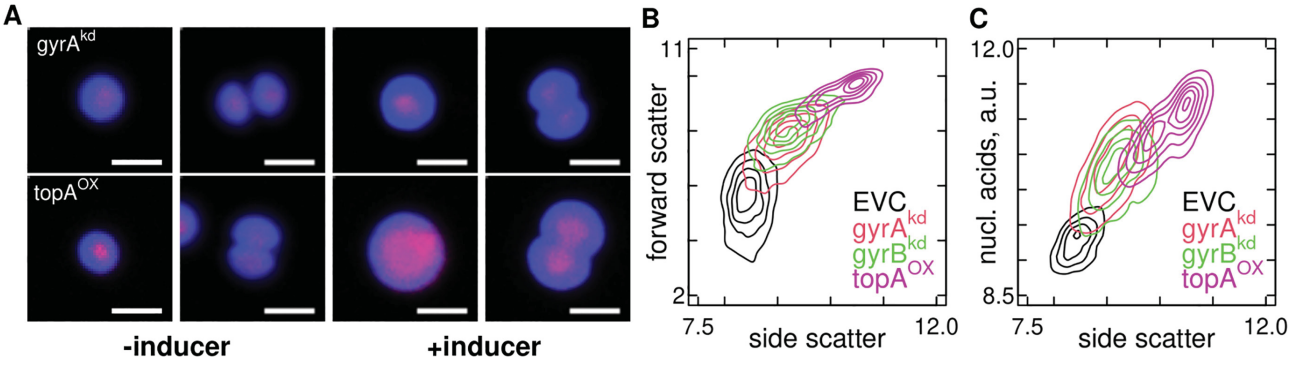


Figure 3. Microscopy and Flow Cytometry Confirm the Volume Growth Phenotype. (**A**) Fluorescence microscopy images of typical round and dividing cells, after 4 days of growth with or without the inducers (Supplementary Figure S3). The bar indicates 2 μm. Chlorophyll-specific fluorescence is shown in blue and DNA-specific (HOECHST 33342) fluorescence in red. Bright-field and single channel images are provided in Supplementary Figure S4. (**B**, **C**) Flow cytometry after 6 days of growth in the presence of the inducers (Supplementary Figure S6). The natural logarithms of forward scatter, side scatter (B) and nucleic acid stain Syto9 (C) were calculated and 2D distributions plotted as contour plots (flow cytometry raw data: Supplementary Figure S7).

periments showed three sets of topoisomer bands (Supplementary Figure S2A), consistent with the presence of three annotated small plasmids of *Synechocystis*, pCA2.4_M, pCB2.4_M and pCC5.2_M. Electropherograms of the two smaller plasmids indicate that only strains $gyrA^{kd}$ and $gyrAB^{kd}$ showed plasmid relaxation (Figure 2B). We could not extract plasmids from the $topA^{kd}$ strain (Supplementary Figure S2A). Unexpectedly, plasmids in the $topA^{OX}$ and $gyrB^{kd}$ strains had a higher level of supercoiling. To investigate this effect, we measured plasmid supercoiling as a time series of the $topA^{OX}$ strain after inoculation in fresh medium with and without the inducer (Supplementary Figure S2B–F). The gel run time was increased to better separate topoisomers of pCC5.2_M. All three plasmids were more relaxed after induction for 3 h (Figure 2C). Already after 8 h the trend had reversed, and at 20–34 h plasmids were more supercoiled than at time 0 h and in the uninduced control time series (Supplementary Figure S2C). Then plasmids became further supercoiled to an extent where topoisomers were not separable anymore. In summary, the effects on plasmids verify the functionality of our constructs on protein activity level, and indicate quick compensatory reactions.

Cell volume growth, and increased adenosine and glycogen content. Next, we investigated the phenotypes to test the hypothesis that genome relaxing manipulations could set free cellular energy for potential use in bioproduction. Initially, all cultures showed comparable growth. After three days all strains except $topA^{kd}$ grew slower than the EVC; and $topA^{OX}$ showed the strongest growth defect. The cell dry weight (CDW) at harvest time correlated with the final OD_{750} of the cultures (Figure 2A), but was relatively higher for the EVC and $topA^{kd}$ strains. Cell volume distributions of the EVC and $topA^{kd}$ strains showed a transient small increase (≈ 10%) on the first day of cultivation and were stable thereafter (Figure 2D). In contrast, cell volumes of the gyr^{kd} and $topA^{OX}$ strains increased over time, from 4 –5 fL to 12–15 fL after four days of cultivation. Total cell numbers increased only slightly. Thus, strains where gyrase subunits were knocked down or TopoI was overexpressed showed inhibition of cell division but not of cell growth. Absorption spectra (Figure 2E, Supplementary Figure S1B) showed an overall decrease of all pigments in $topA^{OX}$. The gyr^{kd} strains showed a stronger decrease at chlorophyll-specific wavelengths than at phycocyanin-specific wavelengths. All knockdown strains showed glycogen levels similar to the EVC, with 25 % of the total CDW (Figure 2F). In contrast, $topA^{OX}$ contained more than twice as much glycogen, 55% of the CDW, and more than four times as much ATP+ADP as the EVC (Figure 2G). $gyrB^{kd}$ and $gyrAB^{kd}$ accumulated about twice as much ATP+ADP as the EVC; $topA^{kd}$ and $gyrA^{kd}$ showed no difference to the EVC control. While the strains show clear differences in their metabolic phenotype, the volume growth phenotype is consistent for all manipulations that should (in principle) decrease supercoiling and not seen in the two controls; a further verification of the functionality of our constructs.

Confirmation by microscopy and flow cytometry. The conductivity-based cell sizes provided by the CASY system does not distinguish cell shape. We thus confirmed the volume growth phenotype by fluorescence microscopy and flow cytometry, each with nucleic acid staining. Cell volumes were increased in the $gyrA^{kd}$ and $topA^{OX}$ strains only in the presence of the inducers (Figure 3A, Supplementary Figures S3 and S4). The cell size distributions, measured from microscopy images with the `Coli-Inspector` (68) (Supplementary Figure S3D), agreed well with the CASY data (Supplementary Figure S5). Manual counting of cells in division (8-shaped) or estimation from the distribution of cell widths and lengths showed an increase from < 10% to ≈ 20% after four days of growth in the presence of the inducers (Supplementary Figure S5D, E). The phenotype was further confirmed by flow cytometry (Figure 3B, Supplementary Figure S7): forward scatter (FSC), which reflects cell size, was increased in all strains, and most in $topA^{OX}$. Total nucleic acid content (RNA+DNA) also increased with cell size (Figure 3C).

Reduction of rRNA, global remodeling of mRNA & homeostatic regulation of supercoiling enzymes. To investigate the effects on transcription, the same cultures that were used for flow cytometry (Supplementary Figure S6A) were sub-

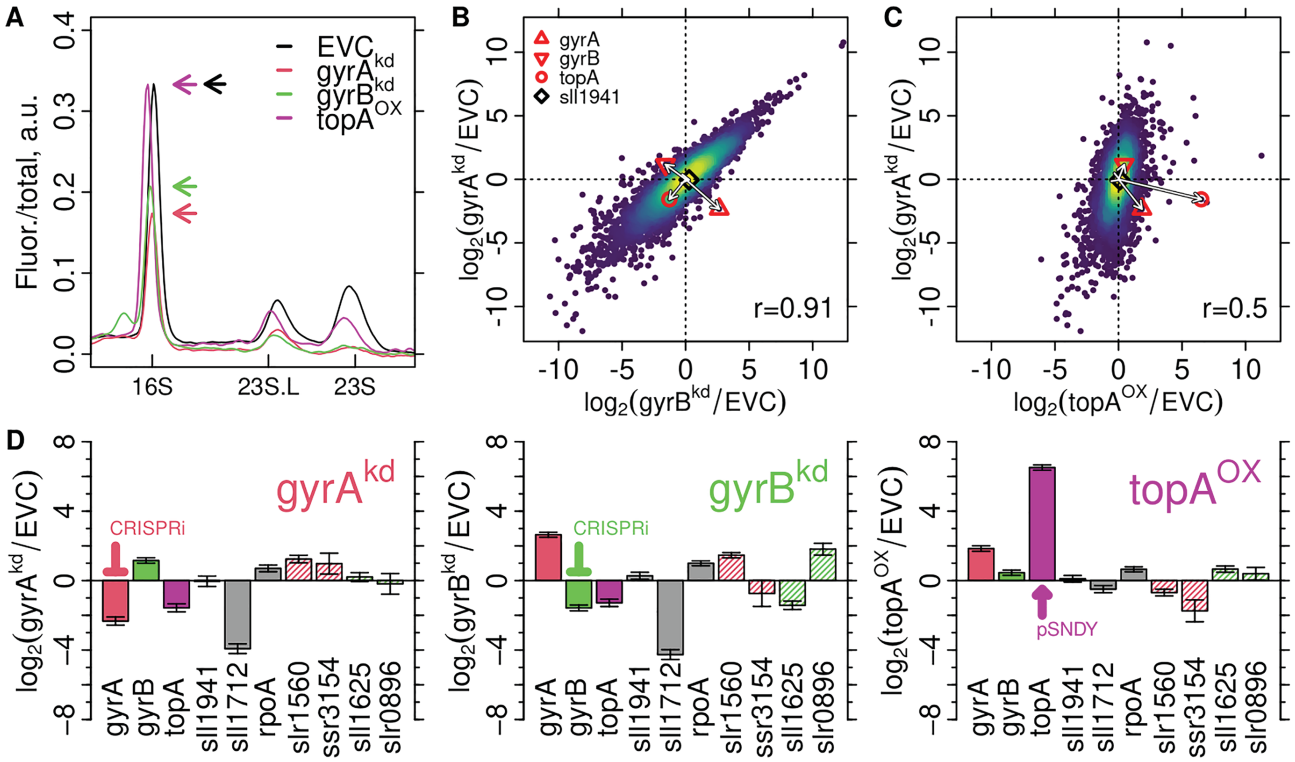


Figure 4. Global transcriptome changes and homeostatic regulation of topoisomerase genes. (**A**) Electropherograms of the capillary gel electrophoresis analysis of extracted RNA used for RNAseq. The fluorescence signal of each lane was normalized by the total RNA content as reported by the Bioanalyzer 2100 software (Supplementary Figure S6C). Lines are the means of three replicates (Supplementary Figures S8 and S9). Locations of the 16S, the 23S and the large fragment of the 23S rRNA (23S.L) are indicated on the x-axis. Arrows indicate the maxima of the 16S rRNA peaks. (**B**, **C**) Expression changes of coding genes in induced strains relative to the control strain (EVC) were derived as the $\log_2$ ratio of RPKM normalized read counts and then compared between the three different strains by 2D histograms (yellow: highest and purple: lowest local density of genes). The Pearson correlations (r) are indicated in the bottom right corner. (**B**) $gyrA^{kd}$ (y-axis) versus $gyrB^{kd}$ (x-axis) strains. (C) $gyrA^{kd}$ (y-axis) versus $topA^{OX}$ (x-axis). The induction/repression and the homeostatic responses of *gyrA*, *gyrB* and *topA* are highlighted by arrows from the origin to indicate the direction of change. (**D**) Expression changes of the targeted topoisomerase genes, the *gyrA/parC* homolog *sll1941*, the HU protein (*sll1712*), the qPCR reference *rpoA*, and the predicted CRISPRi off-targets (indicated by colored stripes). Error bars are standard errors reported by DESeq2.

jected to transcriptome analysis. Total RNA composition and the relative abundances of rRNA and mRNA were analyzed by capillary gel electrophoresis (Supplementary Figures S6B, C, S8). Ribosomal RNA species were strongly reduced in the $gyrA^{kd}$ and $gyrB^{kd}$ strains and less reduced in $topA^{OX}$ (Figure 4A, Supplementary Figure S9). Interestingly, the reduction was stronger for the 23S than for the 16S subunit, even though they are synthesized as one transcript, with 16S upstream and 23S downstream, and processed into subunits co-transcriptionally (87). The same RNA samples were further processed (rRNA species depleted) and sequenced on the Illumina platform, and transcript abundances relative to the EVC control strain (fold change) evaluated with DESeq2 (78). All strains showed overall similar expression changes, but the extent was lower in $topA^{OX}$ (Figure 4B, C). However, this difference could also just reflect normalization effects (88) by the decreased rRNA content in the gyrase knockdowns. In all strains, the targeted manipulation was still observable at harvest time (Figure 4B–D), i.e., *gyrA* transcripts were reduced in $gyrA^{kd}$, *gyrB* transcripts in $gyrB^{kd}$ and *topA* transcripts were increased in $topA^{OX}$. The non-manipulated genes showed the compensatory response expected from homeostatic regulation, i.e., *topA* was repressed in both gyr^{kd} strains, and all non-manipulated gyrase subunits were upregulated in all experiments. Transcription of the DNA binding protein HU (*sll1712*) was strongly downregulated in the gyr^{kd} strains but only weakly in the $topA^{OX}$ strain. In contrast, the *sll1941* gene, annotated either as a second gyrase A subunit or as the topoisomerase IV ParC subunit (89,90), showed no response in either experiment. The upregulation of the qPCR reference gene *rpoA* in all strains explains the disparate results of RT-qPCR (Supplementary Figure S10B–D). Both CRISPRi constructs have potential off-targets (Supplementary Table S2). Indeed, the succinate dehydrogenase gene *sll1625* (91,92) a predicted off-target of the *gyrB*-specific sgRNA was downregulated in the $gyrB^{kd}$ strain (Figure 4D), while other off-targets were not systematically affected.

Dynamic response and adaptation to topoisomerase I overexpression

Compensatory regulation of the non-manipulated topoisomerase genes in each strain was observed even five days after induction. Ribosomal RNA synthesis was strongly impaired. These observations are compatible with the established models of the role of supercoiling in bacterial tran-

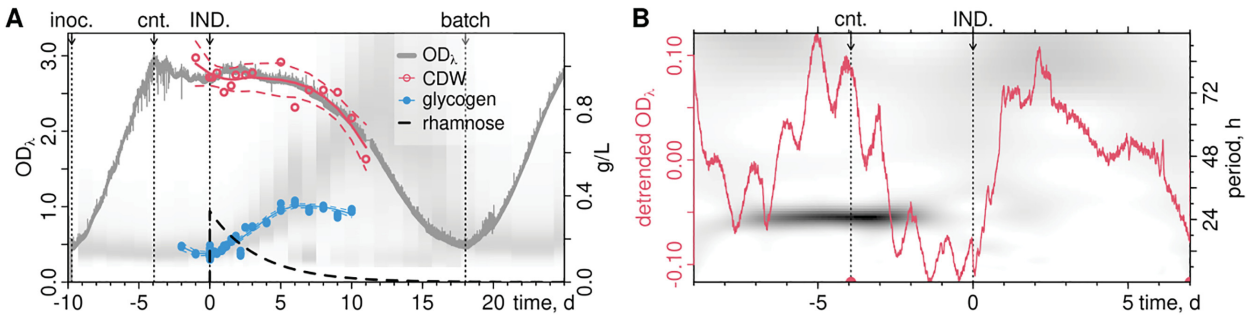


Figure 5. Pulsed induction in continuous culture. (**A**) Photobioreactor growth of the topA^{OX} strain (1 l BG11 medium, 0.5% CO_2, illumination ≈90 μmol m^{-2} m^{-1} per OD_{750}). Optical density was recorded online (OD_λ) and post-calibrated to offline OD_{750}. The arrows indicate `inoc.`: inoculation; `cnt.`: onset of continuous culture (rate $\phi = 0.01\ h^{-1}$); `IND.`: induction of *topA* by pulse-addition of rhamnose to 2 mM (0.33 g l^{-1}) at time 0 day; and `batch`: switch-off of dilution. The dashed black line shows the theoretical wash-out curve of rhamnose (g l^{-1}). Cell dry weight (CDW, g l^{-1}, red) and glycogen content (g l^{-1}, blue) were measured at the indicated times (points), and LOESS regressions are shown (solid lines) with 95% confidence intervals (dashed lines). The CASY-based cell volume distributions (Supplementary Figure S11A) are shown as a background in gray-scale for reference. (**B**) The detrended OD_λ signal (red line, Supplementary Figure S11D) shows a ≈24 h trend throughout batch phase and continuous culture before induction (IND.) A wavelet analysis of the dominant periods in the signal is shown as gray-scale background (right axis).

scription (Figure 1). The resulting phenotypes may therefore reflect such compensatory regulation. To investigate the direct effects of our manipulations, we selected the strain with the most pronounced phenotype, topA^{OX}, and studied the transient effects after induction in continuous culture. The culture was grown in continuous light at OD_{750} ≈ 2.7 and with a dilution rate $\phi \approx 0.24\ d^{-1}$ (Figure 5A, Supplementary Figure S11). After pulse-addition of the inducer rhamnose, the *topA* transcript abundance increased to ≈ 45x over its pre-induction level (Supplementary Figure S10F, G). Cell division was inhibited and cell volumes increased with similar kinetics as in the batch culture experiments. Glycogen content increased to ≈40 % of the CDW. After inducer wash-out, cells recovered to their pre-induction state. Appendix A provides a detailed record of these culture dynamics. The online OD signal (OD_λ) showed a subtle ≈24 h component which vanished after *topA* induction (Supplementary Figure S11D). Sustained circadian rhythms in constant light have been reported before (93). However, we sampled daily for OD_{750} and absorption spectra, and can not exclude that we inadvertently entrained the culture. Sampling in high temporal resolution may similarly have affected the disappearance of the signal after induction.

Dynamic transcriptome response in continuous culture. Samples for RNAseq analysis were taken at three different time scales, i.e., in highest resolution around induction (–35, 5, 20, 60 min), then over 3 days (4–8 h time steps), and three further samples until 26 days, covering the phases of volume increase and recovery. Coding gene transcript read counts were calculated, the resulting time series clustered (Figure 6A, Supplementary Figure S12) and clusters sorted based on the following analysis. The clusters were scanned for statistical enrichments with functional category annotations (Figure 6B, Supplementary Figures S14 and S15) and with clusterings from published experiments. Specifically, we tested for enrichments of (i) genes that responded coherently to stress conditions in the presence or absence of the gyrase inhibitor novobiocin (13), (ii) genes that were either upregulated or downregulated with increasing growth rate (94), and (iii) two diurnal (light/dark) time series (53,84) that were clustered with the same method (Supplementary Figure S17).

Over the first three days post-induction (Figure 6A), cluster 1 (red) transcripts were upregulated in waves with a ≈24 h pattern and afterwards returned to pre-induction levels. This cluster is enriched with genes encoding for ribosomal proteins and biosynthetic enzymes (Figure 6B), with genes that positively correlated with growth rate, and genes that peaked at dawn (Figure 6C, Supplementary Figure S14B). Cluster 2 (yellow) transcripts were downregulated in our experiment and comprise most photosynthesis genes, and overlap with gene groups that were downregulated in stress conditions, negatively correlated with growth rate, and peaked at noon. Cluster 3 (green) transcript abundances initially decreased and showed a weak ≈24 h pattern, opposite to the transcripts of cluster 1. It is enriched with genes peaking at noon or dusk. The transcript abundances of clusters 5 and 6 (blue and cyan) increased from 1 day post-induction, were enriched with genes that peak at night, with DNA replication and repair machinery, and with transposons. These clusters also contain most plasmid-encoded transcripts (Supplementary Figure S16), and were not upregulated in the endpoint measurement of the gyr^{kd} strains (Figure 6C). And finally, the largest cluster 4 (gray) comprises the genes with the weakest response to induction of *topA* overexpression. In summary, *topA* overexpression differentially affected gene cohorts that overlapped with genes whose transcript levels change over the diurnal cycle (84) and vary with growth rates (94). Diurnal cohorts that are expressed at night or at dawn were upregulated, while cohorts expressed at noon and dusk were downregulated.

Alignment of −10 and TSS with the structural code

Our intervention thus revealed gene groups that were also co-regulated in previous experiments. To analyze the underlying promoter structures we mapped the clustering onto transcription units (TU) (83) (Supplementary Figure S18), and calculated the nucleotide content around their transcription start sites (TSS). To avoid bias we only consid-

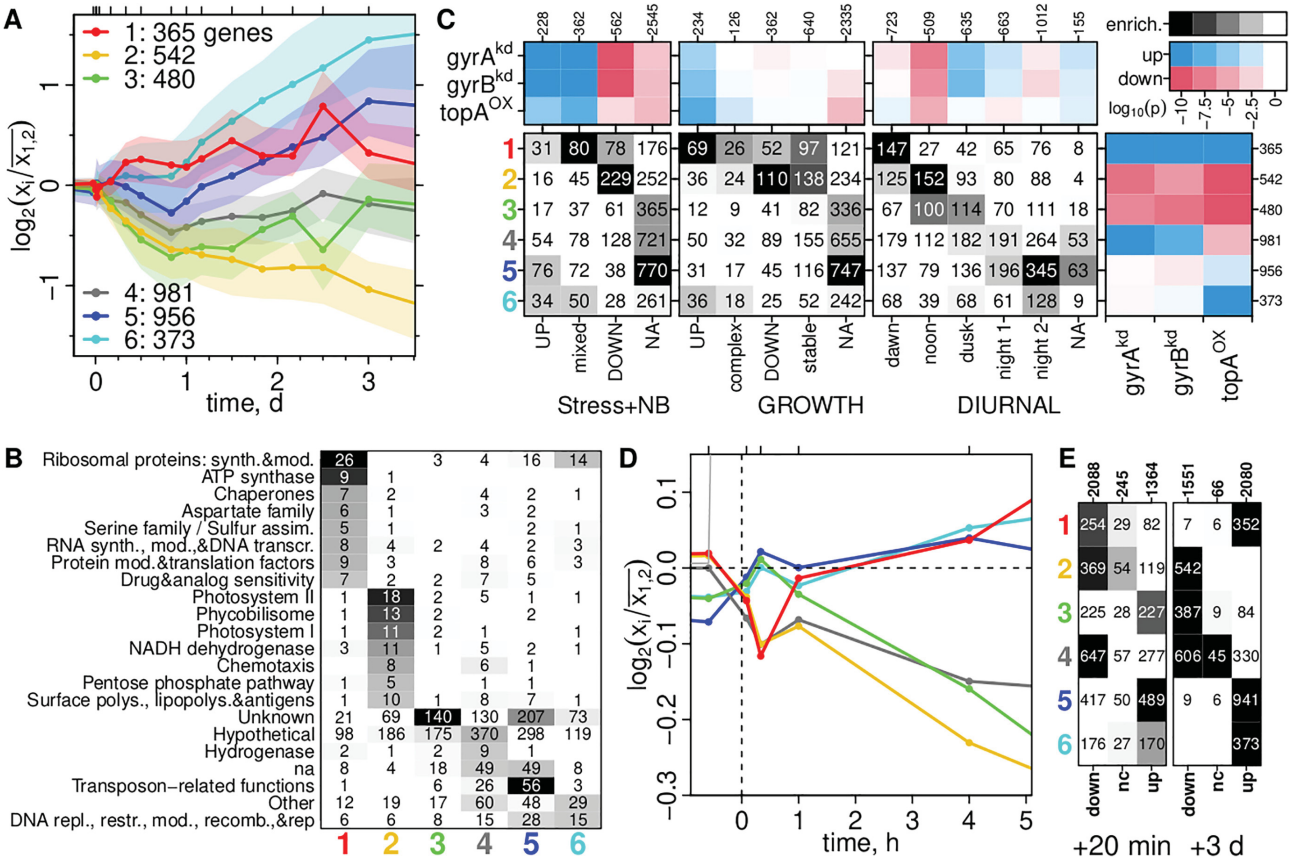


Figure 6. Cluster analysis of the transcriptome time series data. (**A**) Cluster medians of transcript abundances (solid lines), relative to the mean of two pre-induction samples. The transparent ranges indicate the 25%/75% quantiles; points and ticks on upper axis indicate the RNAseq sampling times. Cluster labels (1–6) and sizes (number of genes) are indicated in the legend. (**B**) Sorted enrichment profile of the six clusters with the CyanoBase 'categories' gene annotation. The numbers are the gene counts in each overlap, and gray scale indicates the statistical significance (enrichment) of these counts (black field: $p_{min} \leq 10^{-10}$; white text: $p_{txt} \leq 10^{-5}$). Only overlaps with $p_{sort} \leq 0.01$ are shown (full contingency table in Supplementary Figure S14A). (**C**) Enrichment profiles (gray scale as in (B)) with other published gene classifications (see text) and t-value profiles (red-blue scale, Supplementary Figure S13) of clusters in the end-point transcriptome experiments. Blue indicates upregulation ($t > 0$) and red downregulation ($t < 0$). (**D**) Cluster medians as in (A) but zoomed in on the first 5 h after induction. (**E**) Cluster enrichment profile (gray scale as in (B)) with genes upregulated (up), downregulated (down) or without change (nc), 5–20 min (left) or 2.5–3 days (right) after induction.

ered TU from the main genome for these analyses. As expected from many other bacterial species (10,12,14,16), the differential response to manipulation of supercoiling correlates with the G+C content of the coding region (Figure 7A). This is especially pronounced in the TU that were most upregulated or downregulated 20 min after induction, and in the typical direction, i.e., upregulated TU are A+T-rich and downregulated TU are G+C-rich (Supplementary Figure S19B). However, already 1 h post-induction the different clusters bifurcated, and one G+C-rich cluster (1, red) became upregulated, while one A+T-rich cluster (3, green) became downregulated (Figure 6D, E). Next, we focused on the core promoter (Figure 7B–D, Supplementary Figures S19–S24) to query for previously described supercoiling-sensitive structural features (32,52). This revealed a distinctive feature of *Synechocystis* promoters, namely, a strong coupling of the TSS with an A-tract-based structural DNA code present in most bacterial genomes (47,51), but specifically pronounced in polyploid cyanobacteria such as *Synechocystis* (53). A-tracts of length four show a clear helically phased enrichment with the maximal peak at the −10 bp region of the promoter (Supplementary Figure S20A). This A-tract pattern can be further decomposed into a helically phased enrichment of the AT2 dinucleotide motif (ApA, ApT, TpT) and a localized enrichment of the complementary TpA step just upstream of the −10 peak of AT2, and again at the TSS (Figure 7B, C, Supplementary Figures S20B, C, S21); i.e., spanning the region of single-stranded DNA (open bubble) in the transcription initiation complex. Each cluster showed significant deviations from this common structure. The TSS-associated peak of the TpA step is most pronounced in cluster 2 (yellow). Cluster 3 (green) shows the lowest AT2 peak at −10 bp but the highest peaks up to −50 bp, covering the σ factor-binding region. In contrast, cluster 1 (red) shows the highest AT2 peak at −10 bp, but significantly lower peaks directly upstream. Periodic enrichments further upstream may be out of phase due to variable distances from the TSS. Indeed, the autocorrelation function of concatenated promoter sequences shows comparable amplitudes in all clusters, and

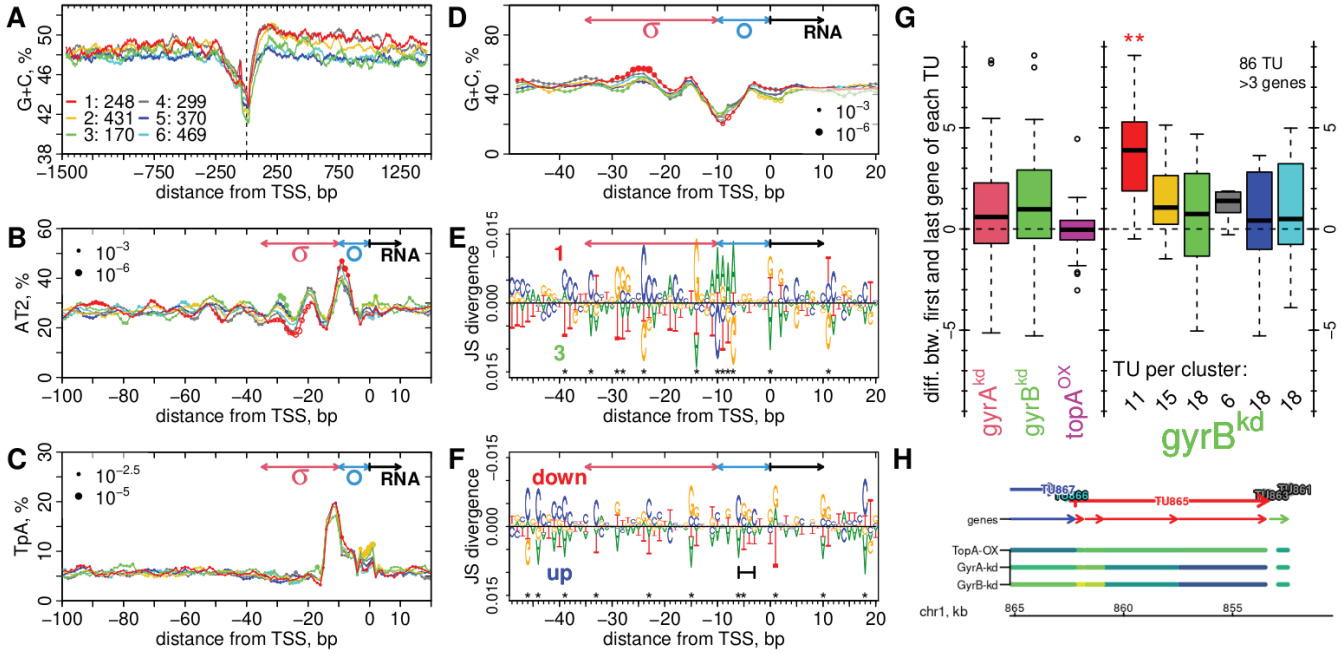


Figure 7. Promoter and Transcription Unit (TU) Structure. (**A**–**D**) Cluster nucleotide frequencies around transcription start sites (TSS) (Supplementary Figure S18); only TU on the main genome were considered and the legend in (A) provides the number of TU in each cluster. The G+C content in (A) was calculated in 66 bp windows at each position, all others in 5 bp windows. Point sizes (B–D) scale with $-\log_2(p)$ from local motif enrichment (filled points) and deprivation (open circles) tests, and the minimal p-values in each plot are indicated in the legends. The sigma factor binding region (σ, −35 to −10), the location of the open bubble (○, −10 to 0) and the transcript (RNA, from 0) are indicated. See Supplementary Figures S19–S23 for the full analysis. (**E**, **F**) The Jensen–Shannon (JS) divergence (Supplementary Figures S24–S25) between the position weight matrices of time series clusters 1 and 3 (E) and of immediate response clusters 'up' and 'down' (F); ⋆ indicates $p < 0.05$ (Supplementary Figure S24) (95). The short horizontal bar in (F) indicates the GC discriminator region −6 to −3. (**G**) Graded response along TU with ≥4 genes in the batch culture experiments in Figure 4. The y-axis shows the difference of the $\log_2$ fold changes between the first and last transcribed gene of each TU. The left panel shows all strains and the right panel the gyrB^{kd} strain and TUs by their cluster association. See Supplementary Figure S27 for all strains and all TU with ≥2 genes. (**H**) An example TU from cluster 1 (red) with a transcript abundance gradient in the gyr^{kd} strains but not in topA^{OX} strain. The genes on TU865 are, from 5' to 3', *rps20*, *tatD*, *rpoB* and *rpoC2*. The color scheme (viridis) in the strain tracks reflects the $\log_2$ fold-changes (Figure 4), where yellow indicates higher and blue lower expression than the control strain (EVC). The colors of genes and TU reflect their time series cluster association.

with higher periods (>11 bp) in upregulated and lower periods in downregulated clusters (Supplementary Figure S22).

The discriminator region and sigma factors

The deviations from this common structural pattern may be related to the differential immediate and adaptive responses of transcription to topA^{OX} induction. The GC content between the −10 element and the TSS, known as the discriminator region, partially determines positive and negative responses to both the regulatory metabolite ppGpp and to changes in DNA supercoiling (27,30,31,96,97). This pattern is consistently found in phylogenetically distant bacteria (32). We find two distinct nucleotide enrichment patterns in this region, one downstream and another one upstream of a conserved T at position −7 (Supplementary Figure S24A, B) which binds to a pocket of the σ^{70} factor of the *E. coli* RNA polymerase initiation complex (35,98,99). Firstly, the promoters that were upregulated 20 min post-induction show an enrichment of A+T nucleotides at −6 to −3 (Figure 7F, Supplementary Figure S19D,F, S24C). This is consistent with data from other bacterial species (32). However, it reflects the overall GC/AT bias of these promoters, extending beyond the core promoter and into the coding region. Secondly, cluster 1 (red) promoters are enriched in A between −7 and −11, upstream of the T at −7 (Figure 7D, E, Supplementary Figure S25).

Thus, the immediate response to topA^{OX} induction is largely consistent with responses observed throughout the bacterial phylogeny. The subsequent adaptive response likely reflects regulatory mechanisms specific to cyanobacteria or *Synechocystis*, and may involve specific sigma factors. We thus, investigated the expression patterns of the nine annotated sigma factors (100) (Supplementary Figure S26). The *sigA* transcript was downregulated quickly after topA^{OX} induction and was low in all endpoint measurements. The transcripts of group 2 sigma factors *sigB* and *sigC* were upregulated in all experiments but the time series shows that both are initially downregulated until 5 days post-induction. The group 2 sigma factors SigD and SigE are involved in circadian control, and their target genes partially overlap with those of SigA (101–103) (Supplementary Figure S26A). SigE activates sugar catabolic pathways during growth in light/dark conditions (104). Its transcript was down-regulated, but showed a diurnal pattern, slightly ahead (phase-advanced) of the diurnal pattern of cluster 3 (green) transcripts (Supplementary Figure S26C). The transcript of SigD was downregulated at all time points, but upregulated in the gyr^{kd} strains. And finally, the group 3 and 4 factors *sigH* and *sigI* are the only sigma factors that were

upregulated upon topA^{OX} induction: *sigI* transiently over the first three days, and *sigH* in three circadian steps, slightly phase-advanced of the circadian pattern of cluster 1 (red) transcripts (Supplementary Figure S26C).

Graded response along transcription units

Supercoiling does not only affect initiation but also elongation of transcription (105). Gyrase activity downstream of transcription units can resolve positive supercoiling that arises from the act of transcription itself (Figure 1B), and such sites are found, e.g., downstream of rRNA loci and highly transcribed operons in *E. coli* (21). Failure to remove downstream supercoiling leads to RNA polymerase stalling (19–21). Thus, we inspected the spatial fold-change patterns along TU in the batch culture RNAseq data (Figure 4) by analyzing the differences between the first and last transcribed gene of each multi-gene TU. Indeed, we find that the gyr^{kd} but not the topA^{OX} strains showed graded expression along TU (Figure 7G, H, Supplementary Figure S27). This mostly affected large (multi-gene) TU of the G+C-rich clusters 1, 2 and 4 (red, yellow, gray), is most pronounced in cluster 1, comprising of ribosomal protein genes, and is reminiscent of the graded effect at the rRNA loci (Figure 4A).

DISCUSSION

We manipulated the expression of gyrase and TopoI genes in *Synechocystis*, and showed that increased DNA relaxing (topA^{OX} strain) or decreased DNA supercoiling (gyr^{kd} strains) activity inhibits cell division and broadly affects physiology. Our data largely confirm the prevailing models of the role of DNA supercoiling homeostasis in bacteria for *Synechocystis* (Figure 1). We further demonstrate a direct coupling of *Synechocystis* promoters to helically phased A-tracts.

A toolbox for biotechnology and supercoiling research

Using the inducible dCas9-mediated CRISPR-interference system (64) we successfully repressed transcription of gyrase subunits *gyrA* and *gyrB*, or *gyrA* and *gyrB* simultaneously. Our tunable expression plasmid pSNDY (63) allowed us to over-express the native *topA*. All manipulations impacted pigment content, cell volume and ATP levels: pigments decreased whereas cell volume and ATP+ADP content increased. The most pronounced effects were observed for the strain topA^{OX}, which contained more than twice as much glycogen, comparable to the levels in nitrogen-starved cells (106,107). The sigma factor SigE activates glycogen degradation genes during the diurnal cycle (104). Its downregulation may thus underpin glycogen accumulation in the topA^{OX} strain. SigD, a diurnal counterpart of SigE (101), was downregulated in topA^{OX} but upregulated in the gyr^{kd} strains. This may underlie some of the differences between the phenotypes. The enlarged cell volume, in all strains, was further confirmed by flow cytometry and microscopy, which additionally revealed an increase in the fraction of 8-shaped cells, suggesting a block in cell division but not growth. Thus, we successfully redirected cellular resources by manipulation of DNA supercoiling, providing a promising platform for photoproduction. A combination of our constructs into a single strain, towards a fully synthetic control over the endogenous DNA supercoiling homeostasis, may allow optimization of growth and production phases in photobioreactors. The higher transcript abundances from strongly supercoiled plasmids in the topA^{OX} strain may prove specifically useful to boost expression of exogenous genes, as integration sites for most plasmids have been suggested recently (108). As a next step towards a biotechnological chassis organism, our manipulation of topoisomerase expression must be assessed on protein abundance level. Protein stability of the targeted topoisomerases will likely have to be modified, e.g. by inducible degron systems, to allow for a rapid switch of DNA supercoiling at an optimal point during a production phase.

The *gyrB* knockdown strains gyrB^{kd} and gyrAB^{kd} showed increased ATP+ADP content, and only the gyrB^{kd} strain showed (slightly) increased plasmid supercoiling. We did not further investigate these differences. They could be related to an additional function of GyrB, together with the second GyrA-like protein in *Synechocystis* (*sll1941*) and potentially as a decatenating topoisomerase (ParC/D, TopoIV) (89,90), or may stem from the CRISPRi off-target *sll1625*, a succinate dehydrogenase (91,92). The single knockdown gyrA^{kd} strain showed the weakest metabolic phenotype and is therefore best suited for future studies into the dynamic response to supercoiling in *Synechocystis*.

Evidence for the supercoiling homeostasis and the twin-domain models in *Synechocystis*

Overexpression of *topA* only transiently relaxed the plasmid DNA, and after ≈1 day, the plasmids became increasingly supercoiled. This overcompensation exemplifies the often counterintuitive consequences of manipulating a homeostatic feedback system. *In vitro*, hypernegative supercoiling of plasmids can be generated by gyrase and transcription (109). *In vivo*, it has been observed in a *topA*-deficient *E. coli* strain and depended on transcription (110,111). In our topA^{OX} strain, plasmid yields (per OD) and transcript abundances of plasmid-derived genes all increased with supercoiling. Transcript abundances of both gyrase subunits and the gyrase substrate, ATP, increased in parallel. The overexpression of *topA* may have triggered a positive feedback between plasmid transcription and/or replication and gyrase activity. Gyrase binding sites are frequently found in native plasmids and phage genomes (112) and such sites could contribute to this phenomenon.

Our other results are more consistent with previous observations. We observed compensatory upregulation of *topA* in the gyr^{kd} strains and of *gyrA* and *gyrB* in the topA^{OX} strain. Menzel and Gellert (1983) first suggested that transcription of the topoisomerase genes is under homeostatic control by negative feedback via the supercoiling status (2); and the same pattern is observed in many species across the bacterial phylogeny (2,6,12–15,113–115). Even the stronger response of *gyrA* than of *gyrB* (to a decrease of supercoiling) has been previously reported in *E. coli* (116). Likewise, the immediate genome-wide response to *topA* overexpression is consistent with reports

from many species (10–16): genes with G+C-rich coding regions were downregulated and A+T-rich upregulated 20 min post-induction. In the gyr^{kd} strains the G+C-rich TU clusters showed a graded response along TU, such that the downstream gene showed lower upregulation or stronger downregulation than the upstream gene. To date, there is no clear explanation for the correlation between the G+C content and the differential immediate response to DNA relaxation. G+C-rich DNA requires more energy for melting of the double helix, due to three instead of two hydrogen bonds per base pair. Indeed, the *in vivo* elongation rate was lower in G+C-rich genes of eukaryotes (117,118). In bacteria, elongation depends on downstream gyrase activity to avoid build-up of positive supercoiling and RNA polymerase stalling, especially at strongly transcribed loci such as RP and rRNA genes (19,21). This requirement could specifically explain the graded effect along G+C-rich TU in the gyrase knockdown strains which was most pronounced in cluster 1 (red), as well as the stronger downregulation of the downstream 23S than the upstream 16S rRNA at the rRNA loci. In summary, our data suggest that both the homeostatic feedback control of topoisomerase transcription (Figure 1A), and the twin-domain model of transcription-dependent supercoiling (Figure 1B) also hold in *Synechocystis*.

Helical phasing of the −10 elements and the TSS in *Synechocystis*

Already 1 h post-induction the transcriptional response diversified into at least six distinct groups of transcription units. Due to the quick compensatory reactions as well as the strong phenotype, we can not infer any causal models for this response. However, the six gene clusters overlapped with gene groups that were diurnally co-expressed (84,88) or responded differentially to growth rate (94). They may thus reflect physiologically relevant regulons (groups of TU with functionally interacting protein products). Their differential response correlated with significant deviations from a common promoter structure: a periodic enrichment of the AT2 motif, in-phase with the −10 element and the TSS. The AT2 motif is a minimal representation of short repeats of A and T nucleotides (A-tracts) without the TpA step (119,120). The TpA dinucleotide step is locally enriched just upstream of the AT2 peak at the −10 element (cf. 'TATA box'), and again at the TSS. TpA is structurally distinct and has been called a twist capacitor, since it can adopt both high and low twist states in molecular dynamics simulations, and thereby locally absorb torsional stress (121). Here, at −10 and the TSS, this property could facilitate open bubble formation of the transcription initiation complex. A-tracts have a narrower minor groove of the DNA double helix (46,122,123), providing binding sites for arginine residues in proteins that wrap DNA (124), notably: gyrase (21,125,126), or locally pinning DNA loop (plectoneme) formation (54). Their helically phased enrichment is observed throughout all domains of life (47,53), and specifically also upstream of bacterial promoters (46,50,52,54,127). However, only a few anecdotal observations reflect the direct coupling that we observe in *Synechocystis*. For example, four helically phased A-tracts alone served as a promoter, and the most downstream A-tract served as the TSS in artificial constructs (38). The phasing of the A-tracts relative to the −35 region determined the efficiency of a bacteriophage promoter (37). Kravatskaya *et al.* (2013) found that alignment of promoter sequences at the TSS facilitates the detection of AA+TT dinucleotide periodicities in supercoiling-sensitive *E. coli* promoters (52). To our knowledge, we provide the first observation of a direct alignment of the TSS and the −10 element with helical phased A-tracts on a genome-wide scale. It is possible that cyanobacterial RNA polymerases (128) and σ factors (102) rely more on such DNA structural features than the well-studied *E. coli* case. However, strong genome-wide A-tract periodicities in some cyanobacteria, incl. *Synechocystis*, correlated with a polyploid life style (53) and could also serve efficient packaging of the multiple genomes into plectonemic structures (48). The pattern we observe at promoters could thus merely reflect the proper integration of such a genome packaging code with promoters, similar to its embedding into the first and third codon positions in protein-coding regions (53). These explanations are not mutually exclusive, and evolution could yield A-tract-aligned promoters when these A-tracts are also beneficial for genome packaging.

Torsional strain and open bubble formation

Each time series cluster showed significant deviations from this common structure. Are these distinct patterns directly involved in the differential response to changes in DNA supercoiling? Several non-exclusive models how DNA supercoiling can affect transcription initiation have been proposed (5,29,32,39,40,42,52,129–131). A-tracts can locally stabilize DNA plectonemes, and such DNA loops can suppress *lac* operon promoters if positioned correctly (132,133). Notably, the dependence on a correct phasing of repressive motifs with the −35 element was stronger in *Synechocystis* than in *E. coli* (134). RNA polymerase can bind to the apical loop of a plectoneme and shifts this loop during transcript elongation, thereby avoiding rotation around the template (135). It was suggested that the RNA polymerase channels the torsional strain that is stored in the plectonemic structure into the opening of the DNA double helix between ca. −12 and +1 of the TSS (5,131), with differential supercoiling-dependence of A-tract periods shorter or longer than the DNA helical pitch (≈10.5 bp) (40,52). The auto-correlation analysis did reveal subtle differences in AT2 motif periods (Supplementary Figure S22), but a significance of these differences remains to be shown.

The sequence-dependent stability of the open bubble conformation of the RNA polymerase (open complex) determines the promoter's response to both ppGpp and DNA supercoiling (30–32,97). For example, the stability is affected by the GC content between the TSS and the −10 element, a region therefore known as the discriminator (27,96): a higher GC content correlates with both supercoiling-dependence and ppGpp repression of promoters. These differences are observed in a variety of bacterial species (32), including *S. elongatus*. We found a consistent pattern in the promoters affected 20 min after induction of $topA^{OX}$, specifically A+T are enriched from −6 to −3 bp of the TSS,

just downstream of a conserved T at position −7, in upregulated (relaxation-induced) promoters. In contrast, the time series cluster 1 (red) was enriched in A from −12 to −7. In *E. coli*, the T at −7 is flipped out of the helix and bound to a pocket of domain 2 of the σ^{70} factor during open complex formation (35,98,99). The discriminator is bound by the domain 1.2 of σ^{70} (96). All sigma factors that contain domain 1.2 (group 1 and group 2, (136)) were downregulated during the adaptive response. Only *sigH* and *sigI* were upregulated, and the latter with a circadian pattern. The *sigI* transcript was also upregulated during the dark phase of the diurnal cycle (84,137). The lack of domain 1.2 of these group 3 and 4 sigma factors may weaken the dependence of the promoters on supercoiling. And finally, differential enrichment of TpA at the TSS may point to a role of this twist capacitor dinucleotide during open complex formation (121).

DNA supercoiling and the diurnal program

Despite significant differences of the cyanobacterial core transcription infrastructure (102,138,139), ppGpp has very comparable consequences on transcription in cyanobacteria (59,60). Its increase is directly associated with the transcriptional shut-down during dark periods (60), and, during the light phase, it may modulate the diurnal transcription program (61). By inference from the roles of supercoiling and ppGpp in other bacteria, we can suggest a tentative model for the observed changes in gene expression upon *topA* induction or gyrase knockdown: in our constant light experiments ppGpp was likely low, while overexpression of *topA* shifts the supercoiling homeostasis and DNA structure towards the opposed physiological state, usually encountered during the dark phase. This combination, low ppGpp and low supercoiling, could reflect the dark/light transition during the diurnal cycle, and induce the expression of the dawn-specific cluster 1 (red), comprising of growth-relevant genes such as ribosomal proteins and the RNA polymerase. Indeed, the increase of translation-related transcript abundances started shortly before the actual onset of light in *Synechocystis* (84). In physiological context, strong transcription of this cluster would require downstream gyrase activity, and this transcription would lead to an overall increase in genomic supercoiling, according to the twin-domain model. This increase in genomic supercoiling in turn could be required to progress through the temporal expression program, and to initiate dawn-to-noon DNA replication (140). The promoters of clusters 2 and 3 (yellow and green), overlapping with noon-specific and dusk-specific cohorts, show a coherent helical phasing of the A-tract motif up to at least -60 bp which may specifically mediate sensitivity to the local or global level of supercoiling. In our experiments, increased TopoI or decreased gyrase activity would inhibit this transcription-dependent accumulation of supercoiling. The diurnal transcription program would be stuck in a dawn-like state, the genome would not be replicated and cell division blocked.

As an outlook, our strains should next be studied in diurnal conditions. A spatially resolved analysis of transcription along the genome (81), as well as DNA-structural footprinting methods, e.g., mapping of supercoiling-sensitive psoralen-binding sites (141), of gyrase-cleavage sites (21) or of the core transcription machinery (142) will provide an integrative picture of global regulatory mechanisms in a physiological context.

DATA AVAILABILITY

The clustering and time series data from the $topA^{OX}$ strain (both as raw abundances in TPM and as the log2 ratios to the mean of two pre-induction values, as plotted in this manuscript), and endpoint measurements (log2 ratio of abundances in the $gyrA^{kd}$, $gyrB^{kd}$ and $topA^{OX}$ strains to the EVC strain) are available as Supplemental Data File S1 (file `Datatable_S1.tsv`). The RNA-seq raw data have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-10949. The annotated sequence of the pSNDY plasmid, is available as Supplemental Data File S2 (genbank file `pSNDY_Prha_topA-6_119rhaS.gb`).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Nic Schmelling and Anika Wiegard for critical discussion of the manuscript, to Sam Meyer, Rick Gourse and Andrew Travers for helpful information regarding the publication history of the 'discriminator' and 'DNA curvature' in bacterial promoters; and to the anonymous reviewers for asking the right questions.

FUNDING

M.D., D.B. and T.B. were financially supported by the CLIB Competence Centre Biotechnology funded by the European Regional Development Fund ERDF (EFRE-0300096 and EFRE-0300095). Funding for E.P.H. and L.Y. is from Vetenskapsrådet (2016-06160). R.M. was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), grants AX 84/4-1 & STA 850/30-1 (COILseq). R.M., O.E. and I.M.A. were supported by the DFG under Germany's Excellence Strategy – EXC-2048/1–project ID 390686111 (CEPLAS). Funding for open access charge: Heinrich-Heine University Düsseldorf / CEPLAS.

Conflict of interest statement. None declared.

REFERENCES

1. Dorman,C. (2019) DNA supercoiling and transcription in bacteria: a two-way street. *BMC Mol. Cell Biol.*, **20**, 26.
2. Menzel,R. and Gellert,M. (1983) Regulation of the genes for *E. coli* DNA gyrase: homeostatic control of DNA supercoiling. *Cell*, **34**, 105–113.
3. Straney,R., Krah,R. and Menzel,R. (1994) Mutations in the −10 TATAAT sequence of the *gyrA* promoter affect both promoter strength and sensitivity to DNA supercoiling. *J. Bacteriol.*, **176**, 5999–6006.
4. Qi,H., Menzel,R. and Tse-Dinh,Y. (1997) Regulation of *Escherichia colitopA* gene transcription: involvement of a sigmaS-dependent promoter. *J. Mol. Biol.*, **267**, 481–489.

5. Unniraman,S. and Nagaraja,V. (2001) Axial distortion as a sensor of supercoil changes: a molecular model for the homeostatic regulation of DNA gyrase. *J. Genet.*, **80**, 119–124.
6. Dages,S., Dages,K., Zhi,X. and Leng,F. (2018) Inhibition of the *gyrA* promoter by transcription-coupled DNA supercoiling in *Escherichia coli*. *Sci. Rep.*, **8**, 14759.
7. Hsieh,L., Burger,R. and Drlica,K. (1991) Bacterial DNA supercoiling and [ATP]/[ADP]. Changes associated with a transition to anaerobic growth. *J. Mol. Biol.*, **219**, 443–450.
8. van Workum,M., van Dooren,S., Oldenburg,N., Molenaar,D., Jensen,P., Snoep,J. and Westerhoff,H. (1996) DNA supercoiling depends on the phosphorylation potential in *Escherichia coli*. *Mol. Microbiol.*, **20**, 351–360.
9. Snoep,J., van der Weijden,C., Andersen,H., Westerhoff,H. and Jensen,P. (2002) DNA supercoiling in *Escherichia coli* is under tight and subtle homeostatic control, involving gene-expression and metabolic regulation of both topoisomerase I and DNA gyrase. *Eur. J. Biochem.*, **269**, 1662–1669.
10. Peter,B., Arsuaga,J., Breier,A., Khodursky,A., Brown,P. and Cozzarelli,N. (2004) Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. *Genome Biol.*, **5**, R87.
11. Blot,N., Mavathur,R., Geertz,M., Travers,A. and Muskhelishvili,G. (2006) Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome. *EMBO Rep.*, **7**, 710–715.
12. Vijayan,V., Zuzow,R. and O'Shea,E. (2009) Oscillations in supercoiling drive circadian gene expression in cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 22564–22568.
13. Prakash,J., Sinetova,M., Zorina,A., Kupriyanova,E., Suzuki,I., Murata,N. and Los,D. (2009) DNA supercoiling regulates the stress-inducible expression of genes in the cyanobacterium *Synechocystis*. *Mol. Biosyst.*, **5**, 1904–1912.
14. Ferrandiz,M., Martin-Galiano,A., Schvartzman,J. and de la Campa,A. (2010) The genome of *Streptococcus pneumoniae* is organized in topology-reacting gene clusters. *Nucleic Acids Res.*, **38**, 3570–3581.
15. de la Campa,A., Ferrandiz,M., Martin-Galiano,A., Garcia,M. and Tirado-Velez,J. (2017) The Transcriptome of *Streptococcus pneumoniae* Induced by Local and Global Changes in Supercoiling. *Front. Microbiol.*, **8**, 1447.
16. Szafran,M., Gongerowska,M., Malecki,T., Elliot,M. and Jakimowicz,D. (2019) Transcriptional Response of *Streptomyces coelicolor* to Rapid Chromosome Relaxation or Long-Term Supercoiling Imbalance. *Front. Microbiol.*, **10**, 1605.
17. Liu,L. and Wang,J. (1987) Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 7024–7027.
18. Kim,S., Beltran,B., Irnov,I. and Jacobs-Wagner,C. (2019) Long-distance cooperative and antagonistic RNA polymerase dynamics via DNA supercoiling. *Cell*, **179**, 106–1196.
19. Rovinskiy,N., Agbleke,A., Chesnokova,O. and Higgins,N. (2019) Supercoil levels in *E. coli* and *Salmonella* chromosomes are regulated by the C-terminal 35(-)38 amino acids of GyrA. *Microorganisms*, **7**, 81 .
20. Ma,J., Bai,L. and Wang,M. (2013) Transcription under torsion. *Science*, **340**, 1580–1583.
21. Sutormin,D., Rubanova,N., Logacheva,M., Ghilarov,D. and Severinov,K. (2019) Single-nucleotide-resolution mapping of DNA gyrase cleavage sites across the *Escherichia coli* genome. *Nucleic Acids Res.*, **47**, 1373–1388.
22. Drolet,M., Broccoli,S., Rallu,F., Hraiky,C., Fortin,C., Masse,E. and Baaklini,I. (2003) The problem of hypernegative supercoiling and R-loop formation in transcription. *Front. Biosci.*, **8**, d210–d221.
23. Ahmed,W., Sala,C., Hegde,S., Jha,R., Cole,S. and Nagaraja,V. (2017) Transcription facilitated genome-wide recruitment of topoisomerase I and DNA gyrase. *PLoS Genet.*, **13**, e1006754.
24. Wu,H. and Fang,M. (2003) DNA supercoiling and transcription control: a model from the study of suppression of the leu-500 mutation in *Salmonella typhimuriumtopA*$^{-}$ strains. *Prog. Nucleic Acid Res. Mol. Biol.*, **73**, 43–68.
25. Zhi,X., Dages,S., Dages,K., Liu,Y., Hua,Z., Makemson,J. and Leng,F. (2017) Transient and dynamic DNA supercoiling potently stimulates the leu-500 promoter in *Escherichia coli*. *J. Biol. Chem.*, **292**, 14566–14575.
26. Sobetzko,P., Glinkowska,M., Travers,A. and Muskhelishvili,G. (2013) DNA thermodynamic stability and supercoil dynamics determine the gene expression program during the bacterial growth cycle. *Mol. Biosyst.*, **9**, 1643–1651.
27. Travers,A. (1980) Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. *J. Bacteriol.*, **141**, 973–976.
28. Ohlsen,K. and Gralla,J. (1992) Interrelated effects of DNA supercoiling, ppGpp, and low salt on melting within the *Escherichia coli* ribosomal RNA *rrnB* P1 promoter. *Mol. Microbiol.*, **6**, 2243–2251.
29. Conter,A., Menchon,C. and Gutierrez,C. (1997) Role of DNA supercoiling and RpoS sigma factor in the osmotic and growth phase-dependent induction of the gene *osmE* of *Escherichia coli* K12. *J. Mol. Biol.*, **273**, 75–83.
30. Figueroa-Bossi,N., Guerin,M., Rahmouni,R., Leng,M. and Bossi,L. (1998) The supercoiling sensitivity of a bacterial tRNA promoter parallels its responsiveness to stringent control. *EMBO J.*, **17**, 2359–2367.
31. Pemberton,I., Muskhelishvili,G., Travers,A. and Buckle,M. (2000) The G+C-rich discriminator region of the *tyrT* promoter antagonises the formation of stable preinitiation complexes. *J. Mol. Biol.*, **299**, 859–864.
32. Forquet,R., Pineau,M., Nasser,W., Reverchon,S. and Meyer,S. (2021) Role of the discriminator sequence in the supercoiling sensitivity of bacterial promoters. *mSystems*, **6**, e0097821.
33. Sudzinova,P., Kambova,M., Ramaniuk,O., Benda,M., Sanderova,H. and Krasny,L. (2021) Effects of DNA topology on transcription from rRNA promoters in *Bacillus subtilis*. *Microorganisms*, **9**, 87 .
34. Revyakin,A., Ebright,R. and Strick,T. (2004) Promoter unwinding and promoter clearance by RNA polymerase: detection by single-molecule DNA nanomanipulation. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 4776–4780.
35. Chen,J., Chiu,C., Gopalkrishnan,S., Chen,A., Olinares,P., Saecker,R., Winkelman,J., Maloney,M., Chait,B., Ross,W. *et al.* (2020) Stepwise promoter melting by bacterial RNA polymerase. *Mol. Cell*, **78**, 275–288.
36. Bossi,L. and Smith,D. (1984) Conformational change in the DNA associated with an unusual promoter mutation in a tRNA operon of *Salmonella*. *Cell*, **39**, 643–652.
37. McAllister,C. and Achberger,E. (1989) Rotational orientation of upstream curved DNA affects promoter function in *Bacillus subtilis*. *J. Biol. Chem.*, **264**, 10451–10456.
38. Collis,C., Molloy,P., Both,G. and Drew,H. (1989) Influence of the sequence-dependent flexure of DNA on transcription in *E. coli*. *Nucleic Acids Res.*, **17**, 9447–9468.
39. Figueroa,N., Wills,N. and Bossi,L. (1991) Common sequence determinants of the response of a prokaryotic promoter to DNA bending and supercoiling. *EMBO J.*, **10**, 941–949.
40. Hirota,Y. and Ohyama,T. (1995) Adjacent upstream superhelical writhe influences an *Escherichia coli* promoter as measured by *in vivo* strength and *in vitro* open complex formation. *J. Mol. Biol.*, **254**, 566–578.
41. Espinosa-Urgel,M. and Tormo,A. (1993) Sigma s-dependent promoters in *Escherichia coli* are located in DNA regions with intrinsic curvature. *Nucleic Acids Res.*, **21**, 3667–3670.
42. Perez-Martin,J. and de Lorenzo,V. (1997) Clues and consequences of DNA bending in transcription. *Annu. Rev. Microbiol.*, **51**, 593–628.
43. Asayama,M., Kato,H., Shibato,J., Shirai,M. and Ohyama,T. (2002) The curved DNA structure in the 5'-upstream region of the light-responsive genes: its universality, binding factor and function for cyanobacterial *psbA* transcription. *Nucleic Acids Res.*, **30**, 4658–4666.
44. Agrawal,G., Asayama,M. and Shirai,M. (2003) Two distinct curved DNAs upstream of the light-responsive *psbA* gene in a cyanobacterium. *Biosci. Biotechnol. Biochem.*, **67**, 1817–1821.
45. Petersen,L., Larsen,T., Ussery,D., On,S. and Krogh,A. (2003) RpoD promoters in *Campylobacter jejuni* exhibit a strong periodic signal instead of a −35 box. *J. Mol. Biol.*, **326**, 1361–1372.
46. Haran,T. and Mohanty,U. (2009) The unique structure of A-tracts and intrinsic DNA bending. *Q. Rev. Biophys.*, **42**, 41–81.
47. Schieg,P. and Herzel,H. (2004) Periodicities of 10-11bp as indicators of the supercoiled state of genomic DNA. *J. Mol. Biol.*, **343**, 891–901.
48. Tolstorukov,M., Virnik,K., Adhya,S. and Zhurkin,V. (2005) A-tract clusters may facilitate DNA packaging in bacterial nucleoid. *Nucleic Acids Res.*, **33**, 3907–3918.

49. Olivares-Zavaleta,N., Jauregui,R. and Merino,E. (2006) Genome analysis of *Escherichia coli* promoter sequences evidences that DNA static curvature plays a more important role in gene transcription than has previously been anticipated. *Genomics*, **87**, 329–337.
50. Nov Klaiman,T., Hosid,S. and Bolshoy,A. (2009) Upstream curved sequences in *E. coli* are related to the regulation of transcription initiation. *Comput. Biol. Chem.*, **33**, 275–282.
51. Mrazek,J. (2010) Comparative analysis of sequence periodicity among prokaryotic genomes points to differences in nucleoid structure and a relationship to gene expression. *J. Bacteriol.*, **192**, 3763–3772.
52. Kravatskaya,G., Chechetkin,V., Kravatsky,Y. and Tumanyan,V. (2013) Structural attributes of nucleotide sequences in promoter regions of supercoiling-sensitive genes: how to relate microarray expression data with genomic sequences. *Genomics*, **101**, 1–11.
53. Lehmann,R., Machné,R. and Herzel,H. (2014) The structural code of cyanobacterial genomes. *Nucleic Acids Res.*, **42**, 8873–8883.
54. Kim,S., Ganji,M., Kim,E., van der Torre,J., Abbondanzieri,E. and Dekker,C. (2018) DNA sequence encodes the position of DNA supercoils. *eLife*, **7**, e36557.
55. Salvador,M., Klein,U. and Bogorad,L. (1998) Endogenous fluctuations of DNA topology in the chloroplast of *Chlamydomonas reinhardtii*. *Mol. Cell Biol.*, **18**, 7235–7242.
56. Smith,R. and Williams,S. (2006) Circadian rhythms in gene transcription imparted by chromosome compaction in the cyanobacterium *Synechococcus elongatus*. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 8564–8569.
57. Mori,T. and Johnson,C. (2001) Circadian programming in cyanobacteria. *Semin. Cell Dev. Biol.*, **12**, 271–278.
58. Woelfle,M., Xu,Y., Qin,X. and Johnson,C. (2007) Circadian rhythms of superhelical status of DNA in cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 18819–18824.
59. Doolittle,W. (1979) The cyanobacterial genome, its expression, and the control of that expression. *Adv. Microb. Physiol.*, **20**, 1–102.
60. Hood,R., Higgins,S., Flamholz,A., Nichols,R. and Savage,D. (2016) The stringent response regulates adaptation to darkness in the cyanobacterium *Synechococcus elongatus*. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E4867–E4876.
61. Puszynska,A. and O'Shea,E. (2017) ppGpp controls global gene expression in light and in darkness in *S. elongatus*. *Cell Rep.*, **21**, 3155–3165.
62. Los,D. (2004) The effect of low-temperature-induced DNA supercoiling on the expression of the desaturase genes in *Synechocystis*. *Cell Mol. Biol. (Noisy-le-grand)*, **50**, 605–612.
63. Behle,A., Saake,P., Germann,A., Dienst,D. and Axmann,I. (2020) Comparative dose-response analysis of inducible promoters in cyanobacteria. *ACS Synth. Biol.*, **9**, 843–855.
64. Yao,L., Cengic,I., Anfelt,J. and Hudson,E. (2016) Multiple gene repression in cyanobacteria using CRISPRi. *ACS Synth. Biol.*, **5**, 207–212.
65. Labun,K., Montague,T., Krause,M., Torres Cleuren,Y., Tjeldnes,H. and Valen,E. (2019) CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.*, **47**, W171–W174.
66. Cui,L., Vigouroux,A., Rousset,F., Varet,H., Khanna,V. and Bikard,D. (2018) A CRISPRi screen in *E. coli* reveals sequence-specific toxicity of dCas9. *Nat. Commun.*, **9**, 1912.
67. Rippka,R., Deruelles,J., Waterbury,J. B., Herdman,M. and Stanier,R. Y. (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J. Gen. Microbiol.*, **111**, 1–61.
68. Vischer,N., Verheul,J., Postma,M., van den Berg van Saparoea,B., Galli,E., Natale,P., Gerdes,K., Luirink,J., Vollmer,W., Vicente,M. and den Blaauwen,T. (2015) Cell age dependent concentration of *Escherichia coli* divisome proteins analyzed with ImageJ and ObjectJ. *Front. Microbiol.*, **6**, 586.
69. Nelson,G., Boehm,U., Bagley,S., Bajcsy,P., Bischof,J., Brown,C., Dauphin,A., Dobbie,I., Eriksson,J., Faklaris,O. *et al.* (2021) QUAREP-LiMi: A community-driven initiative to establish guidelines for quality assessment and reproducibility for instruments and images in light microscopy. *J. Microsc.*, **284**, 56–73.
70. Hahne,F., Le Meur,N., Brinkman,R., Ellis,B., Haaland,P., Sarkar,D., Spidlen,J., Strain,E. and Gentleman,R. (2009) flowCore: a Bioconductor package for high throughput flow cytometry. *BMC Bioinformatics*, **10**, 106.
71. Rüdiger,S., Rediger,A., Kölsch,A., Dienst,D., Axmann,I. M. and Machné,R. (2021) Plasmid supercoiling decreases during the dark phase in cyanobacteria: a clarification of the interpretation of chloroquine-agarose gels. bioRxiv doi: https://doi.org/10.1101/2021.07.26.453679, 26 July 2021, preprint: not peer reviewed.
72. Pinto,F., Thapper,A., Sontheim,W. and Lindblad,P. (2009) Analysis of current and alternative phenol based RNA extraction methodologies for cyanobacteria. *BMC Mol. Biol.*, **10**, 79.
73. Livak,K. and Schmittgen,T. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. *Methods*, **25**, 402–408.
74. Pinto,F., Pacheco,C., Ferreira,D., Moradas-Ferreira,P. and Tamagnini,P. (2012) Selection of suitable reference genes for RT-qPCR analyses in cyanobacteria. *PLoS One*, **7**, e34983.
75. Foley,J. (2021) bioanalyzeR: analysis of agilent electrophoresis data. R package version 0.7.3.
76. Bolger,A., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
77. Langmead,B. and Salzberg,S. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
78. Love,M., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
79. Hilker,R., Stadermann,K., Schwengers,O., Anisiforov,E., Jaenicke,S., Weisshaar,B., Zimmermann,T. and Goesmann,A. (2016) ReadXplorer 2 – detailed read mapping analysis and visualization from one single source. *Bioinformatics*, **32**, 3702–3708.
80. Machné,R. and Murray,D. (2012) The yin and yang of yeast transcription: elements of a global feedback system between metabolism and chromatin. *PLoS One*, **7**, e37906.
81. Machné,R., Murray,D. and Stadler,P. (2017) Similarity-based segmentation of multi-dimensional signals. *Sci. Rep.*, **7**, 12355.
82. Lo,K., Hahne,F., Brinkman,R. and Gottardo,R. (2009) flowClust: a Bioconductor package for automated gating of flow cytometry data. *BMC Bioinformatics*, **10**, 145.
83. Kopf,M., Klahn,S., Scholz,I., Matthiessen,J., Hess,W. and Voss,B. (2014) Comparative analysis of the primary transcriptome of *Synechocystis* sp. PCC 6803. *DNA Res.*, **21**, 527–539.
84. Saha,R., Liu,D., Hoynes-O'Connor,A., Liberton,M., Yu,J., Bhattacharyya-Pakrasi,M., Balassy,A., Zhang,F., Moon,T., Maranas,C. and Pakrasi,H. (2016) Diurnal regulation of cellular processes in the cyanobacterium *Synechocystis* sp. strain PCC 6803: insights from transcriptomic, fluxomic, and physiological analyses. *mBio*, **7**, e00464-16.
85. Fujisawa,T., Narikawa,R., Maeda,S., Watanabe,S., Kanesaki,Y., Kobayashi,K., Nomata,J., Hanaoka,M., Watanabe,M., Ehira,S. *et al.* (2017) CyanoBase: a large-scale update on its 20th anniversary. *Nucleic Acids Res.*, **45**, D551–D554.
86. UniProt Consortium (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.*, **49**, D480–D489.
87. Rodgers,M. and Woodson,S. (2021) A roadmap for rRNA folding and assembly during transcription. *Trends Biochem. Sci.*, **46**, 889–901.
88. Lehmann,R., Machné,R., Georg,J., Benary,M., Axmann,I. M. and Steuer,R. (2013) How cyanobacteria pose new problems to old methods: challenges in microarray time series analysis. *BMC Bioinformatics*, **14**, 133.
89. Mrazek,J., Bhaya,D., Grossman,A. and Karlin,S. (2001) Highly expressed and alien genes of the *Synechocystis* genome. *Nucleic Acids Res.*, **29**, 1590–1601.
90. Gärtner,K., Klähn,S., Watanabe,S., Mikkat,S., Scholz,I., Hess,W. and Hagemann,M. (2019) Cytosine N4-methylation via M.Ssp6803II is involved in the regulation of transcription, fine-tuning of DNA replication and DNA repair in the cyanobacterium *Synechocystis* sp. PCC 6803. *Front. Microbiol.*, **10**, 1233.
91. Cooley,J., Howitt,C. and Vermaas,W. (2000) Succinate:quinol oxidoreductases in the cyanobacterium *Synechocystis* sp. strain PCC 6803: presence and function in metabolism and electron transport. *J. Bacteriol.*, **182**, 714–722.

92. Mock,M., Schmid,A. and Buhler,K. (2020) Directed reaction engineering boosts succinate formation of *Synechocystis* sp. PCC 6803_Δ*sll1625*. *Biotechnol. J.*, **15**, e2000127.
93. van Alphen,P. and Hellingwerf,K. (2015) Sustained circadian rhythms in continuous light in *Synechocystis* sp. PCC6803 growing in a well-controlled photobioreactor. *PLoS One*, **10**, e0127715.
94. Zavrel,T., Faizi,M., Loureiro,C., Poschmann,G., Stuhler,K., Sinetova,M., Zorina,A., Steuer,R. and Cerveny,J. (2019) Quantitative insights into the cyanobacterial cell economy. *eLife*, **8**, e42508.
95. Nettling,M., Treutler,H., Grau,J., Keilwagen,J., Posch,S. and Grosse,I. (2015) DiffLogo: a comparative visualization of sequence motifs. *BMC Bioinformatics*, **16**, 387.
96. Haugen,S., Berkmen,M., Ross,W., Gaal,T., Ward,C. and Gourse,R. (2006) rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. *Cell*, **125**, 1069–1082.
97. Gummesson,B., Lovmar,M. and Nystrom,T. (2013) A proximal promoter element required for positive transcriptional control by guanosine tetraphosphate and DksA protein during the stringent response. *J. Biol. Chem.*, **288**, 21055–21064.
98. Shultzaberger,R., Chen,Z., Lewis,K. and Schneider,T. (2007) Anatomy of *Escherichia coli* sigma70 promoters. *Nucleic Acids Res.*, **35**, 771–788.
99. Feklistov,A. and Darst,S. (2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell*, **147**, 1257–1269.
100. Imamura,S., Yoshihara,S., Nakano,S., Shiozaki,N., Yamada,A., Tanaka,K., Takahashi,H., Asayama,M. and Shirai,M. (2003) Purification, characterization, and gene expression of all sigma factors of RNA polymerase in a cyanobacterium. *J. Mol. Biol.*, **325**, 857–872.
101. Imamura,S., Asayama,M., Takahashi,H., Tanaka,K., Takahashi,H. and Shirai,M. (2003) Antagonistic dark/light-induced SigB/SigD, group 2 sigma factors, expression through redox potential and their roles in cyanobacteria. *FEBS Lett.*, **554**, 357–362.
102. Koskinen,S., Hakkila,K., Kurkela,J., Tyystjarvi,E. and Tyystjarvi,T. (2018) Inactivation of group 2 sigma factors upregulates production of transcription and translation machineries in the cyanobacterium *Synechocystis* sp. PCC 6803. *Sci. Rep.*, **8**, 10305.
103. Kariyazono,R. and Osanai,T. (2022) Identification of the genome-wide distribution of cyanobacterial group-2 sigma factor SigE, accountable for its regulon. *Plant J.*, **110**, 548–561.
104. Osanai,T., Kanesaki,Y., Nakano,T., Takahashi,H., Asayama,M., Shirai,M., Kanehisa,M., Suzuki,I., Murata,N. and Tanaka,K. (2005) Positive regulation of sugar catabolic pathways in the cyanobacterium *Synechocystis* sp. PCC 6803 by the group 2 sigma factor sigE. *J. Biol. Chem.*, **280**, 30653–30659.
105. Rovinskiy,N., Agbleke,A., Chesnokova,O., Pang,Z. and Higgins,N. (2012) Rates of gyrase supercoiling and transcription elongation control supercoil density in a bacterial chromosome. *PLoS Genet.*, **8**, e1002845.
106. Hasunuma,T., Kikuyama,F., Matsuda,M., Aikawa,S., Izumi,Y. and Kondo,A. (2013) Dynamic metabolic profiling of cyanobacterial glycogen biosynthesis under conditions of nitrate depletion. *J. Exp. Bot.*, **64**, 2943–2954.
107. Mollers,K., Cannella,D., Jorgensen,H. and Frigaard,N. (2014) Cyanobacterial biomass as carbohydrate and nutrient feedstock for bioethanol production by yeast fermentation. *Biotechnol. Biofuels.*, **7**, 64.
108. Nagy,C., Thiel,K., Mulaku,E., Mustila,H., Tamagnini,P., Aro,E., Pacheco,C. and Kallio,P. (2021) Comparison of alternative integration sites in the chromosome and the native plasmids of the cyanobacterium *Synechocystis* sp. PCC 6803 in respect to expression efficiency and copy number. *Microb. Cell Fact.*, **20**, 130.
109. Drolet,M., Bi,X. and Liu,L. (1994) Hypernegative supercoiling of the DNA template during transcription elongation *in vitro*. *J. Biol. Chem.*, **269**, 2068–2074.
110. Samul,R. and Leng,F. (2007) Transcription-coupled hypernegative supercoiling of plasmid DNA by T7 RNA polymerase in *Escherichia coli* topoisomerase I-deficient strains. *J. Mol. Biol.*, **374**, 925–935.
111. Zhi,X. and Leng,F. (2013) Dependence of transcription-coupled DNA supercoiling on promoter strength in *Escherichia coli* topoisomerase I deficient strains. *Gene*, **514**, 82–90.
112. Oram,M., Howells,A., Maxwell,A. and Pato,M. (2003) A biochemical analysis of the interaction of DNA gyrase with the bacteriophage Mu, pSC101 and pBR322 strong gyrase sites: the role of DNA sequence in modulating gyrase supercoiling and biological activity. *Mol. Microbiol.*, **50**, 333–347.
113. Tse-Dinh,Y. and Beran,R. (1988) Multiple promoters for transcription of the *Escherichia coli* DNA topoisomerase I gene and their regulation by DNA supercoiling. *J. Mol. Biol.*, **202**, 735–742.
114. Unniraman,S. and Nagaraja,V. (1999) Regulation of DNA gyrase operon in *Mycobacterium smegmatis*: a distinct mechanism of relaxation stimulated transcription. *Genes. Cells*, **4**, 697–706.
115. Ahmed,W., Menon,S., Karthik,P. and Nagaraja,V. (2016) Autoregulation of topoisomerase I expression by supercoiling sensitive transcription. *Nucleic Acids Res.*, **44**, 1541–1552.
116. Neumann,S. and Quinones,A. (1997) Discoordinate gene expression of *gyrA* and *gyrB* in response to DNA gyrase inhibition in *Escherichia coli*. *J. Basic Microbiol.*, **37**, 53–69.
117. Jonkers,I., Kwak,H. and Lis,J. (2014) Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife*, **3**, e02407.
118. Veloso,A., Kirkconnell,K., Magnuson,B., Biewen,B., Paulsen,M., Wilson,T. and Ljungman,M. (2014) Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Res.*, **24**, 896–905.
119. Leroy,J., Charretier,E., Kochoyan,M. and Gueron,M. (1988) Evidence from base-pair kinetics for two types of adenine tract structures in solution: their relation to DNA curvature. *Biochemistry*, **27**, 8894–8898.
120. Stefl,R., Wu,H., Ravindranathan,S., Sklenar,V. and Feigon,J. (2004) DNA A-tract bending in three dimensions: solving the dA4T4 vs. dT4A4 conundrum. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 1177–1182.
121. Reymer,A., Zakrzewska,K. and Lavery,R. (2018) Sequence-dependent response of DNA to torsional stress: a potential biological regulation mechanism. *Nucleic Acids Res.*, **46**, 1684–1694.
122. Wu,H. and Crothers,D. (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature*, **308**, 509–513.
123. Calladine,C., Drew,H. and McCall,M. (1988) The intrinsic curvature of DNA in solution. *J. Mol. Biol.*, **201**, 127–137.
124. Rohs,R., West,S., Sosinsky,A., Liu,P., Mann,R. and Honig,B. (2009) The role of DNA shape in protein-DNA recognition. *Nature*, **461**, 1248–1253.
125. Wahle,E. and Kornberg,A. (1988) The partition locus of plasmid pSC101 is a specific binding site for DNA gyrase. *EMBO J.*, **7**, 1889–1895.
126. Oram,M., Travers,A., Howells,A., Maxwell,A. and Pato,M. (2006) Dissection of the bacteriophage Mu strong gyrase site (SGS): significance of the SGS right arm in Mu biology and DNA gyrase mechanism. *J. Bacteriol.*, **188**, 619–632.
127. Asayama,M. and Ohyama,T. (2003) Curved DNA and prokaryotic promoters: a mechanism for activation of transcription. In: *Madame Curie Bioscience Database [Internet]. Landes Bioscience*. Austin (TX), USA.
128. Riaz-Bradley,A., James,K. and Yuzenkova,Y. (2020) High intrinsic hydrolytic activity of cyanobacterial RNA polymerase compensates for the absence of transcription proofreading factors. *Nucleic Acids Res.*, **48**, 1341–1352.
129. Kusano,S., Ding,Q., Fujita,N. and Ishihama,A. (1996) Promoter selectivity of *Escherichia coli* RNA polymerase E sigma 70 and E sigma 38 holoenzymes. Effect of DNA supercoiling. *J. Biol. Chem.*, **271**, 1998–2004.
130. Opel,M., Aeling,K., Holmes,W., Johnson,R., Benham,C. and Hatfield,G. (2004) Activation of transcription initiation from a stable RNA promoter by a Fis protein-mediated DNA structural transmission mechanism. *Mol. Microbiol.*, **53**, 665–674.
131. Travers,A. and Muskhelishvili,G. (2005) DNA supercoiling - a global transcriptional regulator for enterobacterial growth?. *Nat. Rev. Microbiol.*, **3**, 157–169.
132. Becker,N., Greiner,A., Peters,J. and Maher,L. 3rd (2014) Bacterial promoter repression by DNA looping without protein-protein binding competition. *Nucleic Acids Res.*, **42**, 5495–5504.
133. Fulcrand,G., Dages,S., Zhi,X., Chapagain,P., Gerstman,B., Dunlap,D. and Leng,F. (2016) DNA supercoiling, a critical signal

regulating the basal expression of the *lac* operon in *Escherichia coli*. *Sci. Rep.*, **6**, 19243.
134. Camsund,D., Heidorn,T. and Lindblad,P. (2014) Design and analysis of LacI-repressed promoters and DNA-looping in a cyanobacterium. *J. Biol. Eng.*, **8**, 4.
135. ten Heggeler-Bordier,B., Wahli,W., Adrian,M., Stasiak,A. and Dubochet,J. (1992) The apical localization of transcribing RNA polymerases on supercoiled DNA prevents their rotation around the template. *EMBO J.*, **11**, 667–672.
136. Imamura,S. and Asayama,M. (2009) Sigma factors for cyanobacterial transcription. *Gene. Regul. Syst. Biol.*, **3**, 65–87.
137. Srivastava,A., Summers,M. and Sobotka,R. (2020) Cyanobacterial sigma factors: Current and future applications for biotechnological advances. *Biotechnol. Adv.*, **40**, 107517.
138. Xie,W., Jager,K. and Potts,M. (1989) Cyanobacterial RNA polymerase genes *rpoC1* and *rpoC2* correspond to *rpoC* of *Escherichia coli*. *J. Bacteriol.*, **171**, 1967–1973.
139. Riaz-Bradley,A. (2019) Transcription in cyanobacteria: a distinctive machinery and putative mechanisms. *Biochem. Soc. Trans.*, **47**, 679–689.
140. Werner,A., Broeckling,C., Prasad,A. and Peebles,C. (2019) A comprehensive time-course metabolite profiling of the model cyanobacterium *Synechocystis* sp. PCC 6803 under diurnal light:dark cycles. *Plant J.*, **99**, 379–388.
141. Teves,S. and Henikoff,S. (2014) Transcription-generated torsional stress destabilizes nucleosomes. *Nat. Struct. Mol. Biol.*, **21**, 88–94.
142. John,J., Jabbar,J., Badjatia,N., Rossi,M., Lai,W. and Pugh,B. (2022) Genome-wide promoter assembly in *E. coli* measured at single-base resolution. *Genome Res.*, **32**, 878–892.

Supplementary Information for Behle, Dietsch, *et al.* (2022): Manipulation of Topoisomerase Expression Inhibits Cell Division but not Growth, and Reveals A Distinctive Promoter Structure in *Synechocystis*.

Strain name	Chromosomal genotype	Plasmid
EVC	P_{L22}:dCas9	pSNDY (EVC)
gyrA^{kd}	P_{L22}:dCas9; P_{L22}:sgRNA$_{gyrA}$	pSNDY (EVC)
gyrB^{kd}	P_{L22}:dCas9; P_{L22}:sgRNA$_{gyrB}$	pSNDY (EVC)
gyrAB^{kd}	P_{L22}:dCas9; P_{L22}:sgRNA$_{gyrA}$; P_{L22}:sgRNA$_{gyrB}$	pSNDY (EVC)
topA^{kd}	P_{L22}:dCas9; P_{L22}:sgRNA$_{topA}$	pSNDY (EVC)
topA^{OX}	P_{L22}:dCas9	pSNDY P_{J23119}:*rhaS*; P_{rha}:*topA*

Table S1. Construction of strains investigated in this work. The parental strain for all strains listed here (*Synechocystis* sp. PCC 6803 encoding aTc-inducible dCas9) is based on a strain originally obtained from Martin Fulda (Göttingen, Germany) and engineered for CRISPRi by Yao *et al.* [1]. It contains a TetR cassette, as well as dCas9 under the promoter PL_{22}, inducible with anhydrotetracycline (aTc), at the genomic insertion site *psbA1*. The sgRNA sequences (Tab. S2) were constructed *via* overlap extension PCR and integrated into the vector designed by Yao *et al.* [1] (Addgene ID 73224), which inserts into the *slr0230* site of the *Synechocystis* genome. The sgRNA plasmids were integrated *via* transformation. Briefly, 10 mL of exponentially grown culture was concentrated to 250 µL, 1 µg–2 µg of pure plasmid was added and the mixture was incubated up to 5 h before plating the entire mixture on BG11 plates. After drying the plates, agar was underlaid with 300 µL of 1 mg mL^{-1} kanamycin stock using a sterile spatula, thereby forming a diffusion gradient. After 1-2 weeks of incubation at 30 °C with the lid facing upward, isolated green colonies were carefully transferred to a fresh plate. Over time, positive clones were gradually shifted to higher concentrations of kanamycin (4, 8, 12, 20, 40 µg mL^{-1} final concentration in the plate). Complete segregation of mutants was ensured *via* colony PCR. For rhamnose-inducible overexpression, the coding sequence of *slr2058* (*topA*) was integrated into pSHDY P_{J23119}:*rhaS*; P_{rha}:*mVenus* (Addgene ID 137662) [2] in place of *mVenus via* Gibson assembly. Both this new construct and pSHDY (Addgene ID 137661) were further modified by exchanging the spectinomycin resistance cassette with the nourseothricin cassette, resulting in pSNDY P_{J23119}:*rhaS*; P_{rha}:*topA* and pSNDY (EVC), respectively. Replicative vectors were introduced into the dCas9 background strain *via* conjugation as described (`dx.doi.org/10.17504/protocols.io.ftpbnmn`). Clones were selected using nourseothricin (Jena Bioscience, #AB-102L) at a final concentration of 50 µg mL^{-1} and verified *via* colony PCR.

Gene	Protospacer Sequence	Predicted Off-Targets
gyrA	TCAGTCATGCAATTACTCCA	*ssr3154*, *slr1560*
gyrB	ACTCCAAAATCAGGCTGAGCTT	*sll1625*, *slr0896*
topA	GATAGTGCGGGCTTTAGTGG	*sll1660*

Table S2. sgRNA Protospacer Sequences. The target-specific parts (protospacer) of the sgRNA for CRISPRi-based knockdown strains (Tab. S1) were designed using CHOPCHOP [3]. and potential off-targets analyzed by the Cas-OFFinder [4], allowing ≤ 2 mismatches or DNA/RNA bulges. Only the *topA* sgRNA had had one additional target (in gene *sll1660*), but this strain was not further analyzed beyond Figure 2. Furthermore, Cui *et al.* [5] found that a 9-nt identity of an an off-target site to the sgRNA seed sequence (seed sequence = first 12 nt) can result in significant repression, if an NGG PAM site is present and that the sgRNA anneals to the coding strand. We searched for such potential off-target sites for the three sgRNAs used here. The *topA* sgRNA does not meet this criteria for any other sites. The *gyrB* sgRNA meets this criteria for two off-target sites, *sll1625* and *slr0896*. The *gyrA* sgRNA meets this criteria for two off-target sites in the genome, *ssr3154* and *slr1560*. These off-targets were included in the RNAseq analysis in Figure 4D.

Gene	L_A	T_m	Primer efficiency	Direction	Sequence
gyrA	102	82.9℃	104.16%	forward	GAACTTTGGCTCCGTGGATAA
				reverse	GCCTCAATGTCCCGCAATAA
gyrB	107	78.9℃	100.32%	forward	TGCCCGTAAGCGCAATAA
				reverse	ATTCTGGGTCCGGTACTTTAAC
topA	105	80.6℃	99.43%	forward	AGACCGGGAAGGAGAAAGTA
				reverse	CGAATGGCTTCCTGGGTAAT
rpoA	96	80.4℃	97.86%	forward	CCATGAGTTCGCCACTATTCT
				reverse	GGCTGATCGGTGTAGCTTT
rnpB	93	82.04 ℃	98.29%	forward	AGAGGTACTGGCTCGGTAAA
				reverse	TCAAGCGGTTCCACCAATC

Table S3. RT-qPCR Primers. Primers used for RT-qPCR of the indicated genes were designed using the IDT PrimerQuest tool (`https://eu.idtdna.com/pages/tools/primerquest`). L_A: amplicon length, T_m: melting temperature

Name	RefSeq ID
genome	NC_000911
pCA2.4_M	CP003270
pCB2.4_M	CP003271
pCC5.2_M	CP003272
pSYSM	NC_005229
pSYSA	NC_005230
pSYSG	NC_005231
pSYSX	NC_005232
pSNDY P_{J23119}:*rhaS*; P_{rha}:*topA*	file `pSNDY_Prha_topA-6_119rhaS.gb`

Table S4. Genome and Plasmid Sequences for RNA-seq Mapping. RefSeq IDs of the genome and plasmid sequences of *Synechocystis* used for mapping of the RNA-seq reads. Construction of pSNDY is described in Table S1 and the full sequence is available as Supplemental Data File S2 (genbank file `pSNDY_Prha_topA-6_119rhaS.gb`).

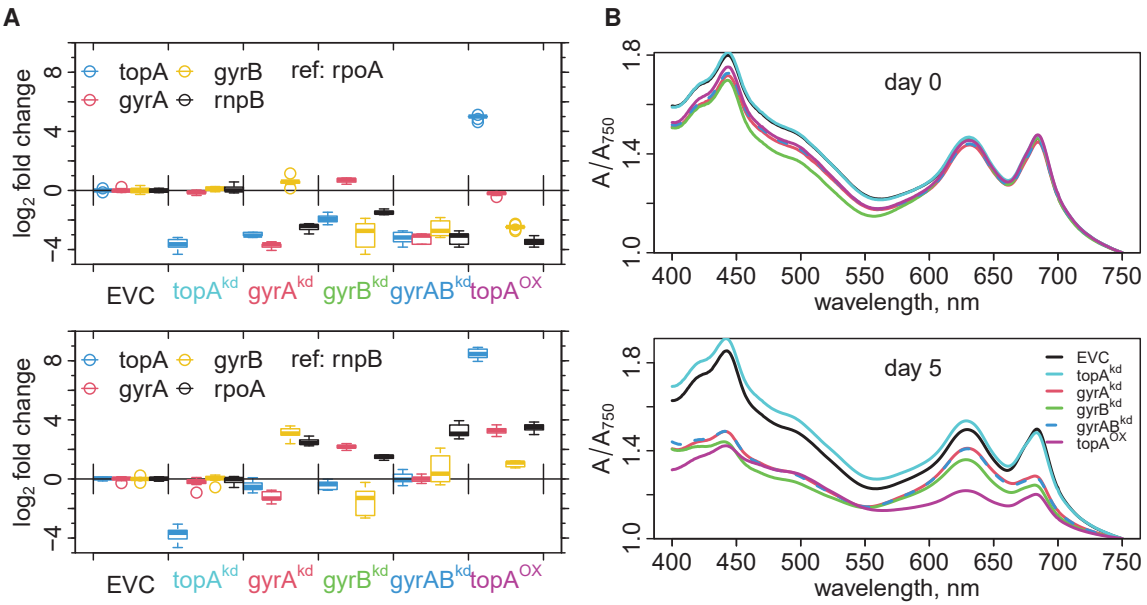


Figure S1. Batch Culture Endpoint Measurements. See Figure 2 for details. **A:** RT-qPCR results using *rpoA* (top panel) or *rnpB* (bottom) as reference genes [6]. Boxplots of 9 technical replicates (3 samples, each measured 3x) . **B:** Absorption spectra at inoculation and harvest times.

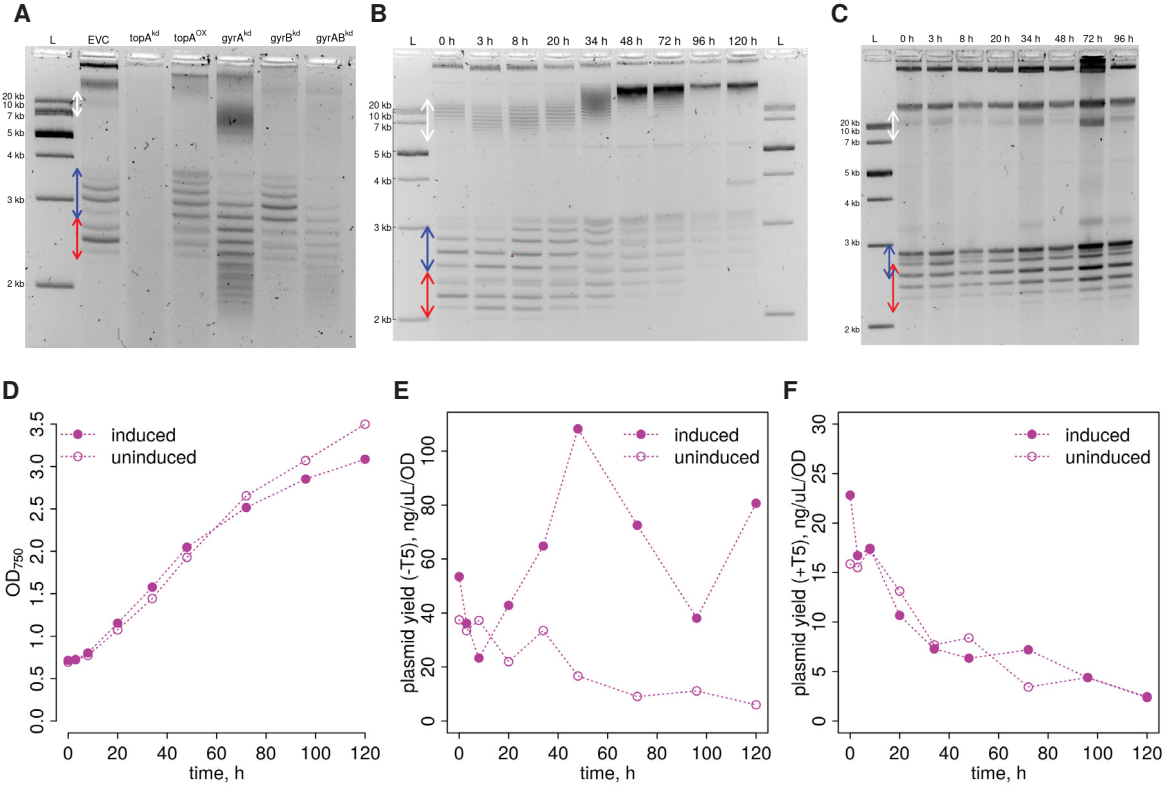
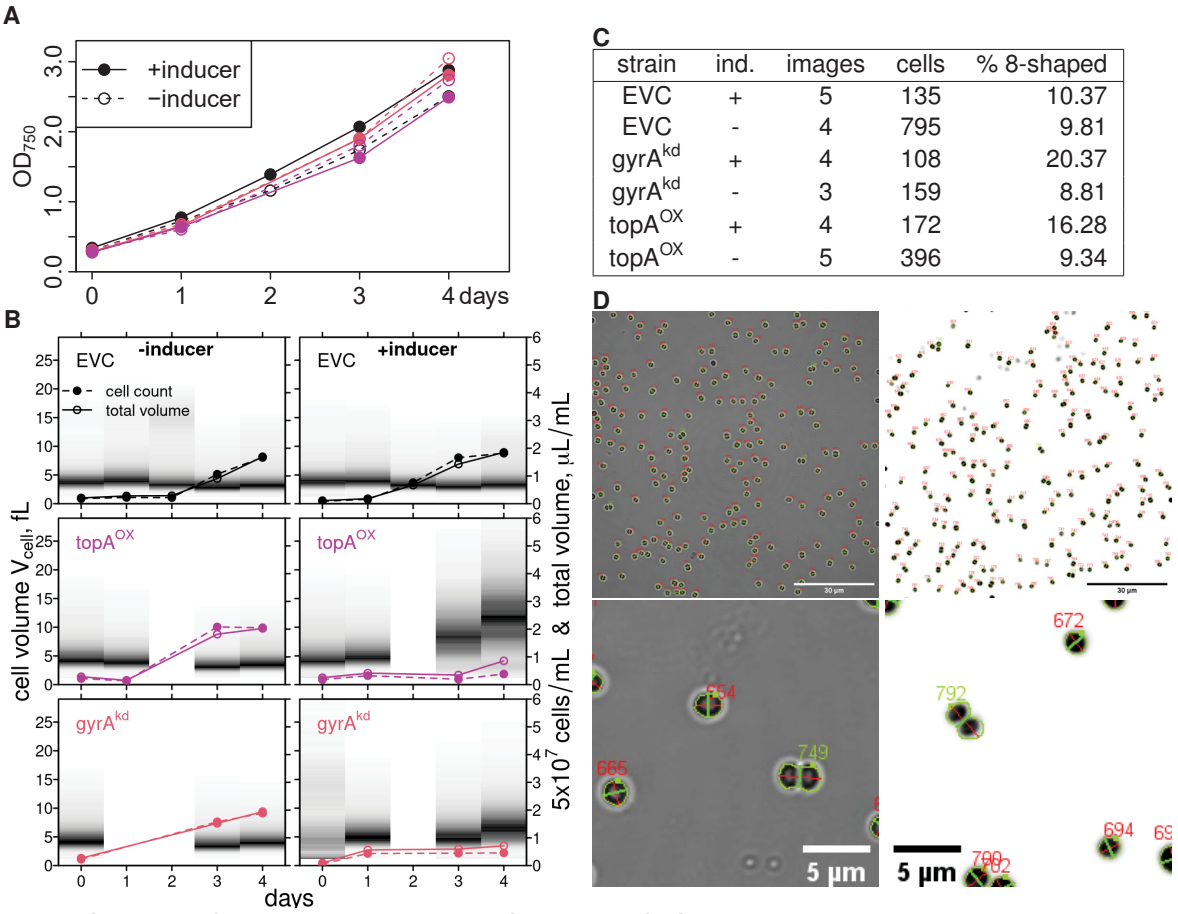


Figure S2. Plasmid Supercoiling Gels. **A–C**: Chloroquine-agarose gels 1.2 % agarose, 0.5x TBE and 20 µg mL^{-1} CQ, 1.8 V cm^{-1}) of plasmids extracted from all strains at harvest time (A) of the experiment shown in Figure 2, or as a time series (see D–F) of the topA^{OX} strain with (B), or without the inducer (C). At 20 µg mL^{-1} CQ, originally more relaxed (rel.) plasmids migrate further (higher migration distance) than more negatively supercoiled plasmids [7]. Two distinct plasmid topoisomer distributions can be distinguished in the EVC in gel (A), run for 20 h. We assume the less far migrated bands to correspond to the larger plasmid pCA2.4_M (2378 bp, blue arrows), and the further migrated bands to the smaller plasmid pCB2.4_M (2345 bp, red arrows). The size of a larger plasmid, pCC5.2_M (5214 bp) fits to a series of topoisomer bands indicated by white arrows. Bands above that likely stem from our pSNDY and or larger endogenous plasmids of *Synechocystis*. The gel of the induced culture (B) was run for a longer time (22 h) to get a better separation of the pCC5.2_M topoisomers. The gel in (C) was only run for 18 h. Baseline-corrected electropherograms of the gels in (A) and (B) are shown in Figure 2B and 2C of the main manuscript, respectively. **D:** Growth of the topA^{OX} strain, with or without induction with 1 mM rhamnose at time 0 h. For both conditions, a starter culture was split into 9 cultures at 0 h, each harvested at the indicated time points for OD_{750} measurement and plasmid extraction (used for electrophoresis shown in (B) and (C)). **E & F**: Yields of plasmid extraction over time, each normalized to the OD_{750} (D), and before (E) and after (F) treatment with the T5 exonuclease to remove all non closed circular DNA.



strain	ind.	images	cells	% 8-shaped
EVC	+	5	135	10.37
EVC	-	4	795	9.81
$gyrA^{kd}$	+	4	108	20.37
$gyrA^{kd}$	-	3	159	8.81
$topA^{OX}$	+	4	172	16.28
$topA^{OX}$	-	5	396	9.34

Figure S3. Batch Cultures for Microscopy. OD (**A**) and CASY cell counts and size distributions (**B**) from cultures of the indicated strains in the presence or absence of the inducer. Samples for microscopy were taken on day 4. **C:** Number of microscopy images and individual cells analyzed with the `Coli-Inspector` [8] (results in Fig. S5) and manual counts of the fraction of 8-shaped cells. The example cells in Figure S4 were also chosen from these images. **D:** Example image illustrating the semi-automated counting procedure. Each cell recognized by the `Coli-Inspector` is characterized by a length (thin red line) and width (thicker green line) parameter, where the shorter of both is always the width. For cell dimension measurements (Fig. S5A-C) the channel for chlorophyll fluorescence was used (right). Because of variations of focal planes those measurements were more reliable than those of the bright field channel (left). The fractions of 8-shaped cells (Fig. S5D, E) were manually counted in the brightfield images.

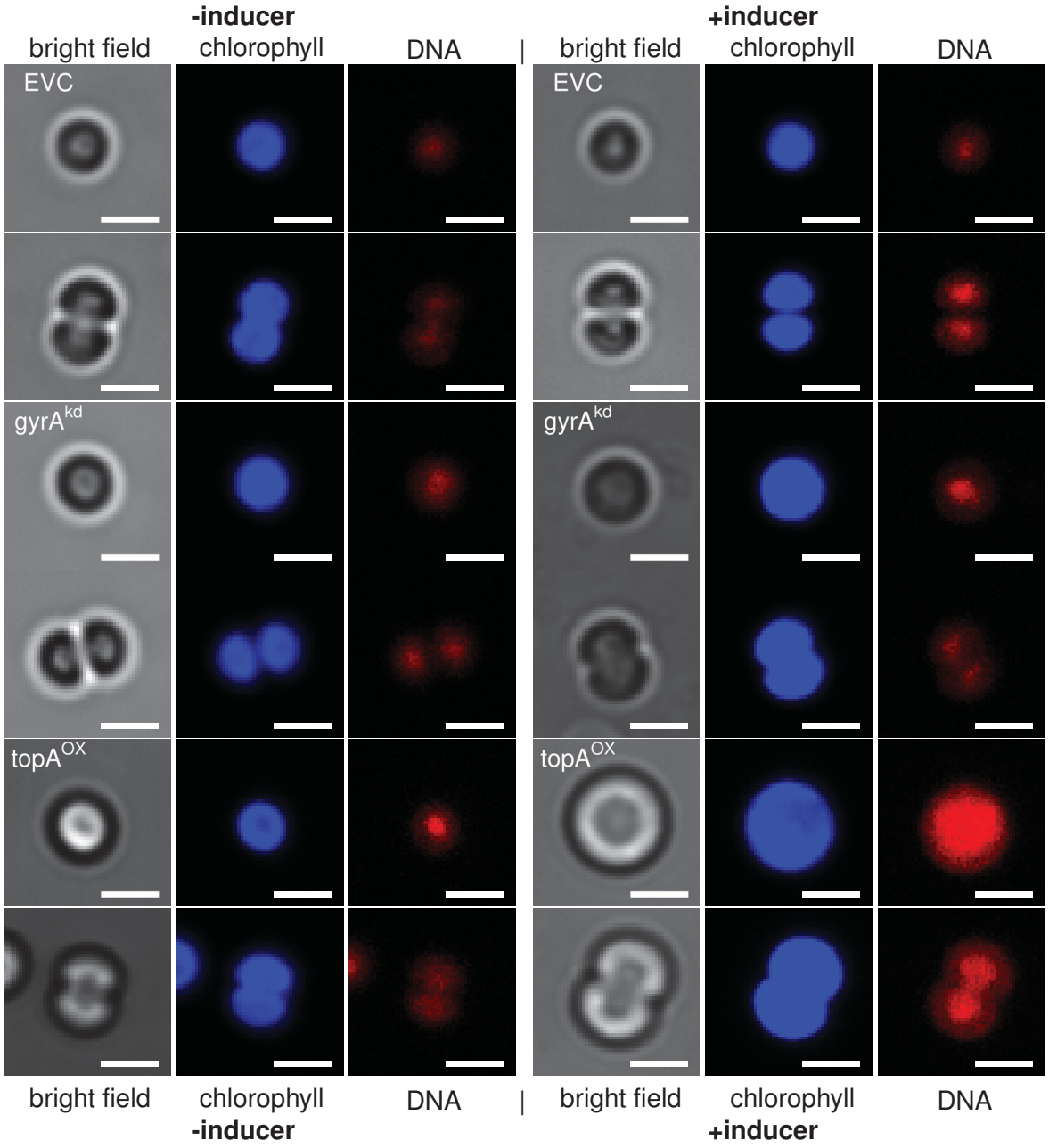


Figure S4. Batch Culture Endpoint Measurements for Microscopy. Zoom on typical single cells from the microscopy images of the cultures shown in Figure S3. The size bar length is 2 µm. From each strain and condition (± inducer) a typical round cell and a typical 8-shaped cell in division was chosen, and its brightfield (left panels) and chlorophyll- (blue, middle) and HOECHST DNA stain-specific (red, right) fluorescence channels are shown. Merged versions of these images, showing both fluorescence channels for the $topA^{OX}$ and $gyrA^{kd}$ strains with and without inducer, are shown in Figure 3.

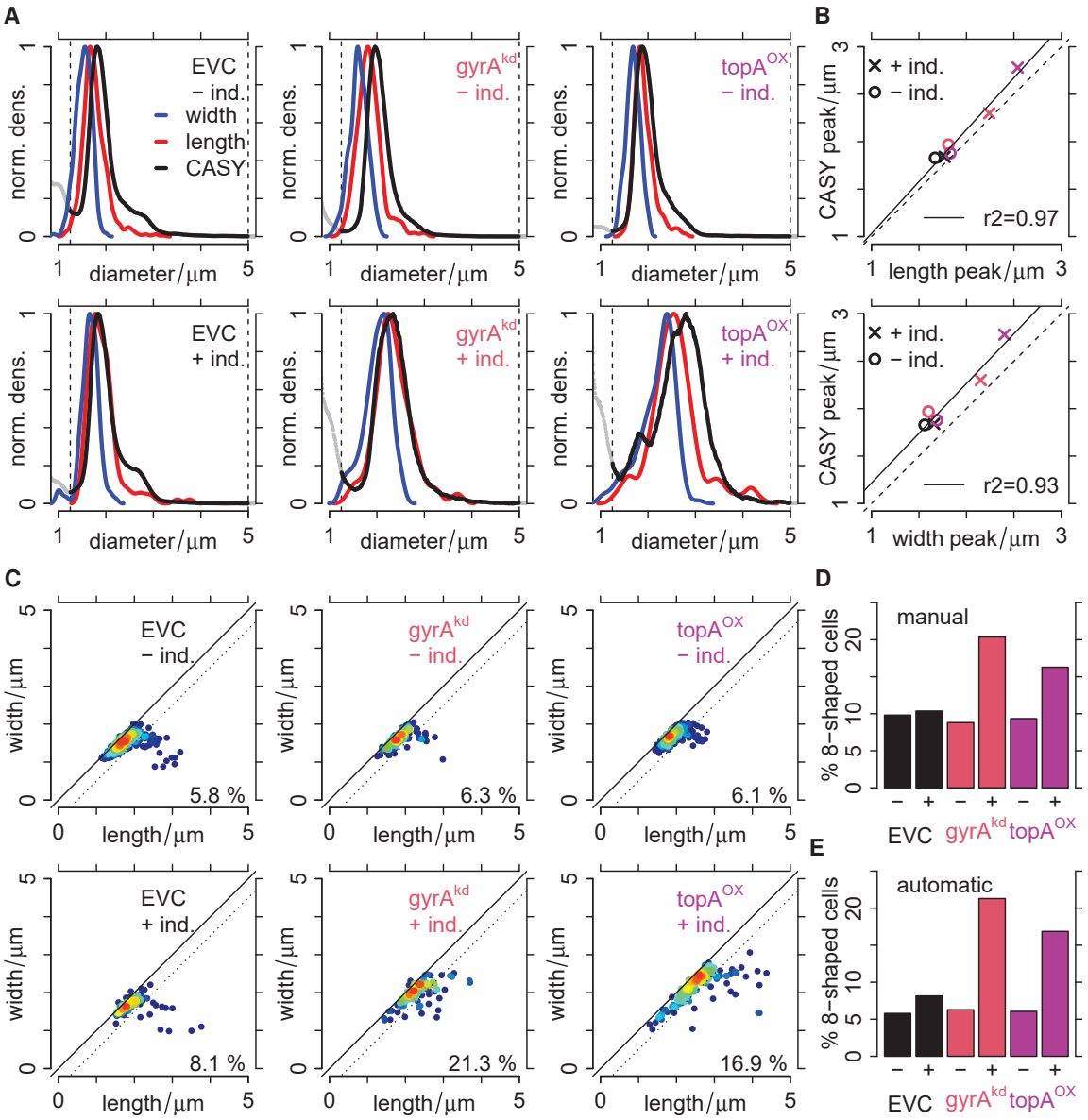


Figure S5. Microscopy Statistics and CASY Calibration. **A:** Density distributions of the cell diameters (black lines) measured by the CASY cell counter and cell lengths and widths (blue and red lines) determined by the `Coli-Inspector` [8] from microscopy images (Fig. S3C,D). All distributions were normalized to a maximum of 1. The dashed lines are the cut-off used for CASY data. **B**: Comparison of peak cell diameters as measured by the CASY cell counter and from microscopy images. The dashed line is the diagonal, and the solid line is a linear regression line, with the r^2 value shown in the bottom right legend. For microscopy, the peaks of the distributions were directly obtained from the density distributions shown in (A), and for CASY we used the peaks reported by the CASY. The CASY-based peak diameters are slightly larger than microscopy-based cell length and width peaks. **C:** cell-wise comparison of width and length estimation by the `Coli-Inspector` (with manual adjustments). For round cells width and length correspond well, for 8-shaped cells (in division) the width negatively correlates with length, since cells further into division are longer but the width, corresponding to the division plane, is shorter. The dotted line is the diagonal $-0.5\,\mu m$, and the cell fraction below this line, with width $<$ length $- 0.5\,\mu m$, was counted (bottom right text). **D & E:** Percent 8-shaped cells of all counted cells. 8-shaped cells were manually counted (**D**, see table in Fig. S3C) or determined automatically (**E**) by a difference in widths and lengths reported by the `Coli-Inspector` (see dotted line and % in (C)).

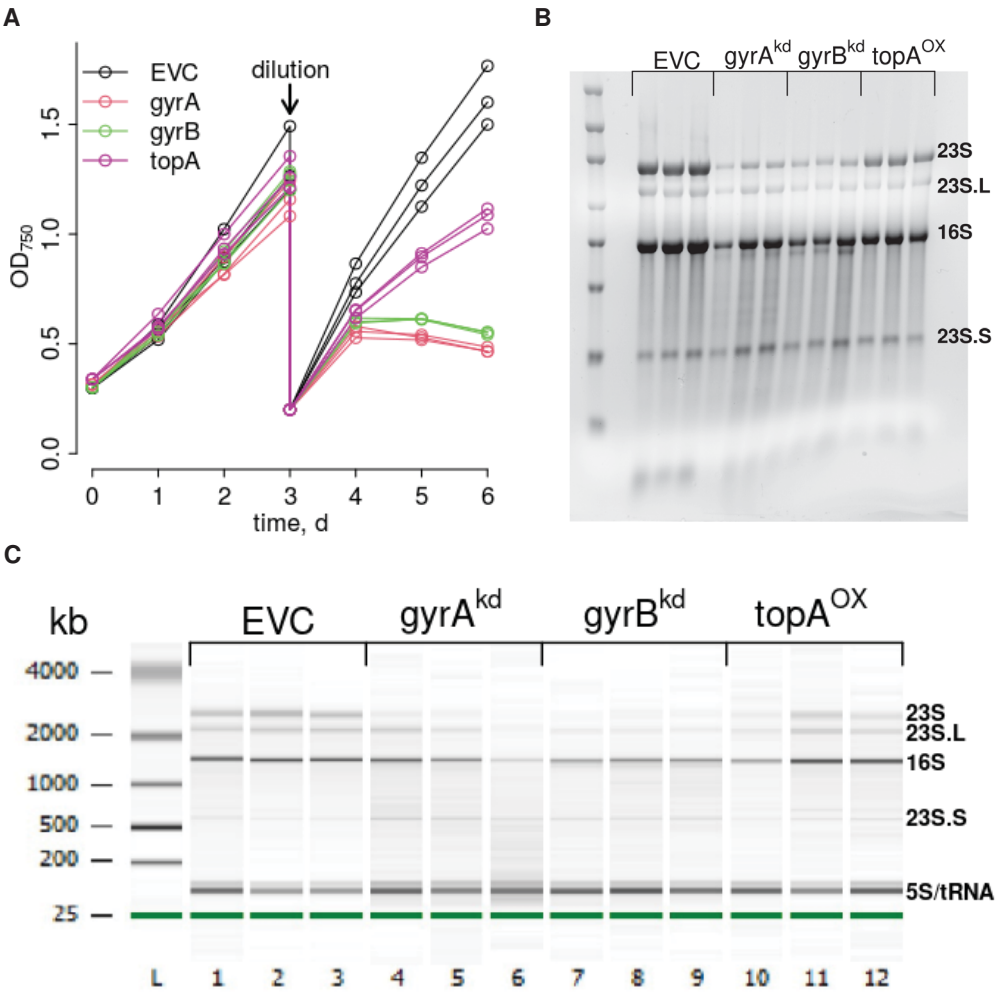


Figure S6. RNA Extraction for RNA-seq Experiment. **A:** Growth of triplicate cultures (split upon induction at 0 d).**B & C:** Total RNA compositions were analyzed by formaldehyde-agarose gel electrophoresis (B, 500 ng RNA per well) and by capillary gel electrophoresis (Agilent Bioanalyzer); the electropherograms of the pseudo-gel figure in (C, from the Bioanalyzer report) samples were analyzed for Figures 4A, S8 and S9.

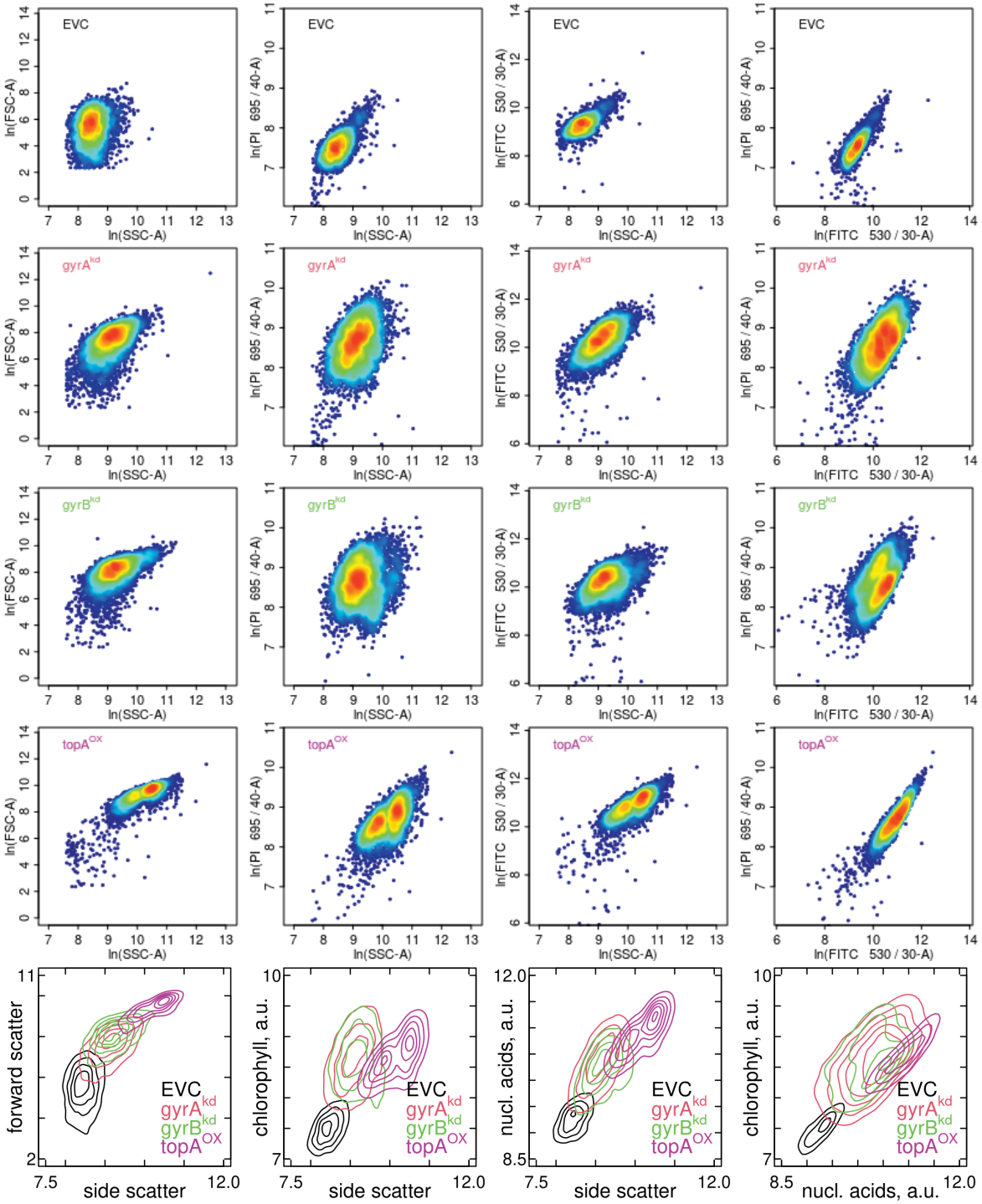


Figure S7. Flow Cytometry. Cells from the cultures used for RNA-seq and total RNA analysis (Fig. 4) were fixed in para-formaldehyde and stained with Syto-9, a nucleic acid fluorescence marker, and analyzed by flow cytometry. The data were gated by the side scatter signal (SSC-A> 2000) and the forward scatter signal (FSC-A> 10) to filter debris and background signals. Forward scatter (FSC-A) is proportional to cell size, side scatter (SSC-A) reflects cytoplasmic granularity and morphology; the FITC fluorescence channel (530/30 nm) excites the Syto-9 stain, and the PI channel (695/40 nm) excites chlorophyll. The natural logarithm $(\ln)$ of all data was plotted. Colors reflect local density (red: high, blue: low). The bottom panels show a zoom into the data, and local densities are displayed as contour lines, merged for all strains.

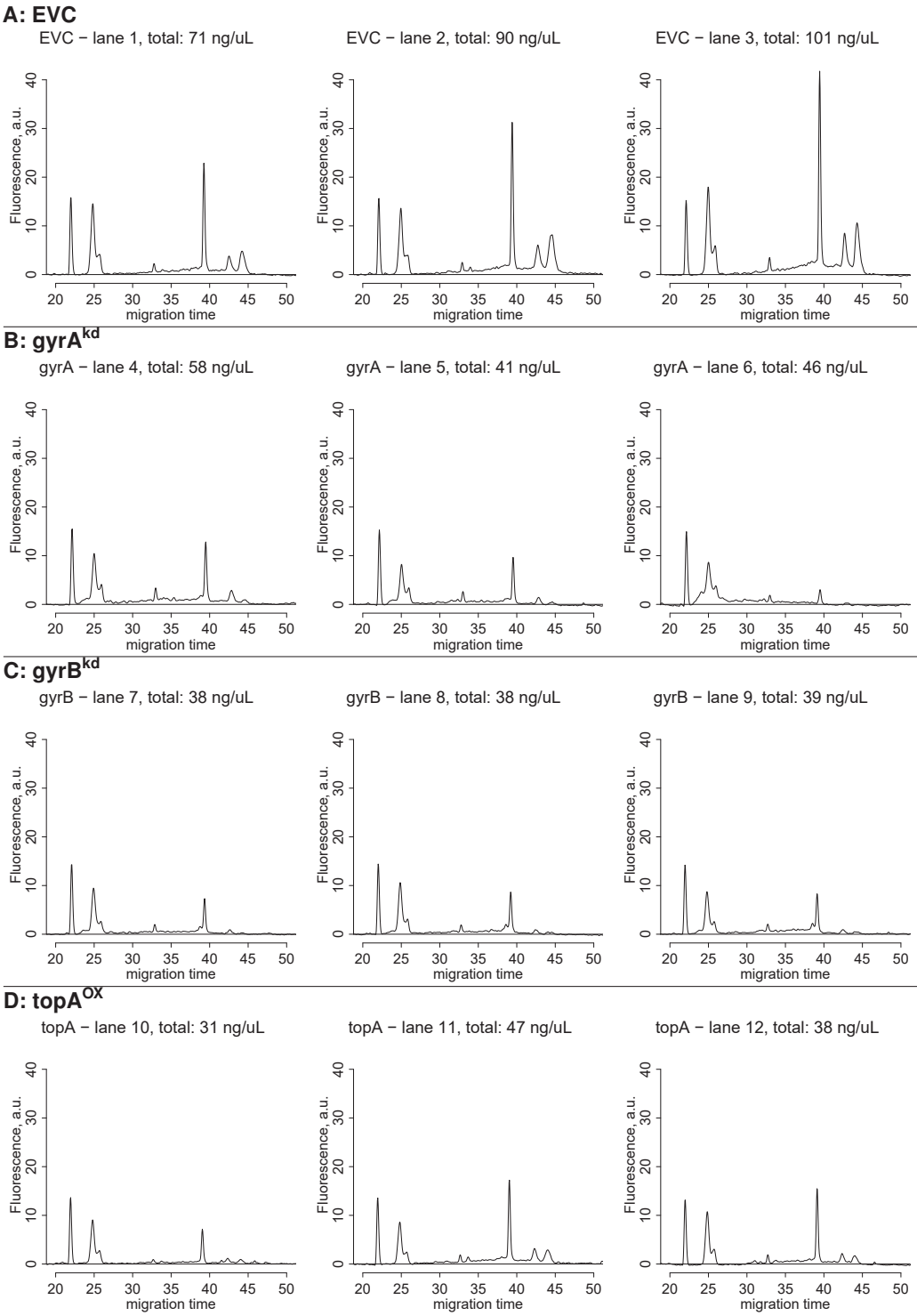


Figure S8. Raw RNA Electropherograms. Electropherograms of the capillary gel electrophoresis (Fig. S6C), exported as XML files from the 2100 Bioanalyzer software, and imported to R with the `bioanalyzeR` R package (v 0.9.1, `https://github.com/jwfoley/bioanalyzeR`) [9]. **A–D:** are each the triplicate samples for the indicated strains. Samples are the same as shown on the formaldehyde-agarose gel in Figure S6B and subsequently used for RNAseq analysis (Fig. 4B,C). The locations of rRNA peaks are indicated; r23S.S and r23S.L are short and long fragments of the 23S rRNA typically seen in *Synechocystis*, the other peaks were assigned to 5S, 16S and (full length) 23S rRNA. The total RNA concentration in the samples, as provided by the 2100 Bioanalyzer software report, are indicated on top of each plot, and these numbers were used for normalization of the electropherograms for Figures S9 and 4A.

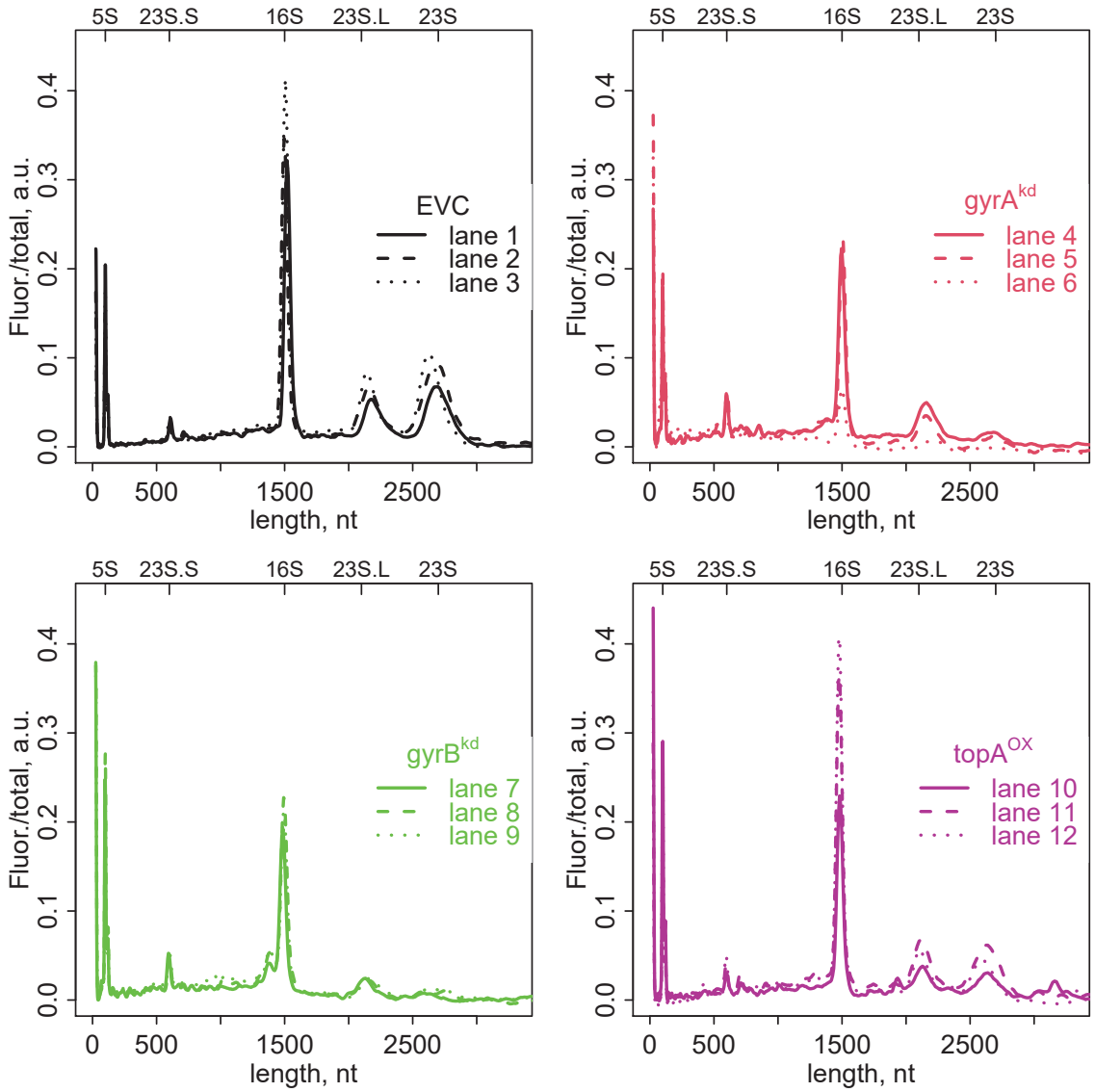


Figure S9. rRNA Relative Abundances. The RNA electropherograms in Figure S8 were each normalized by the total RNA content of the sample, as reported by the Bioanalyzer 2100 software, and plotted by replicate groups. The mean of each replicate group was calculated and plotted for Figure 4A.

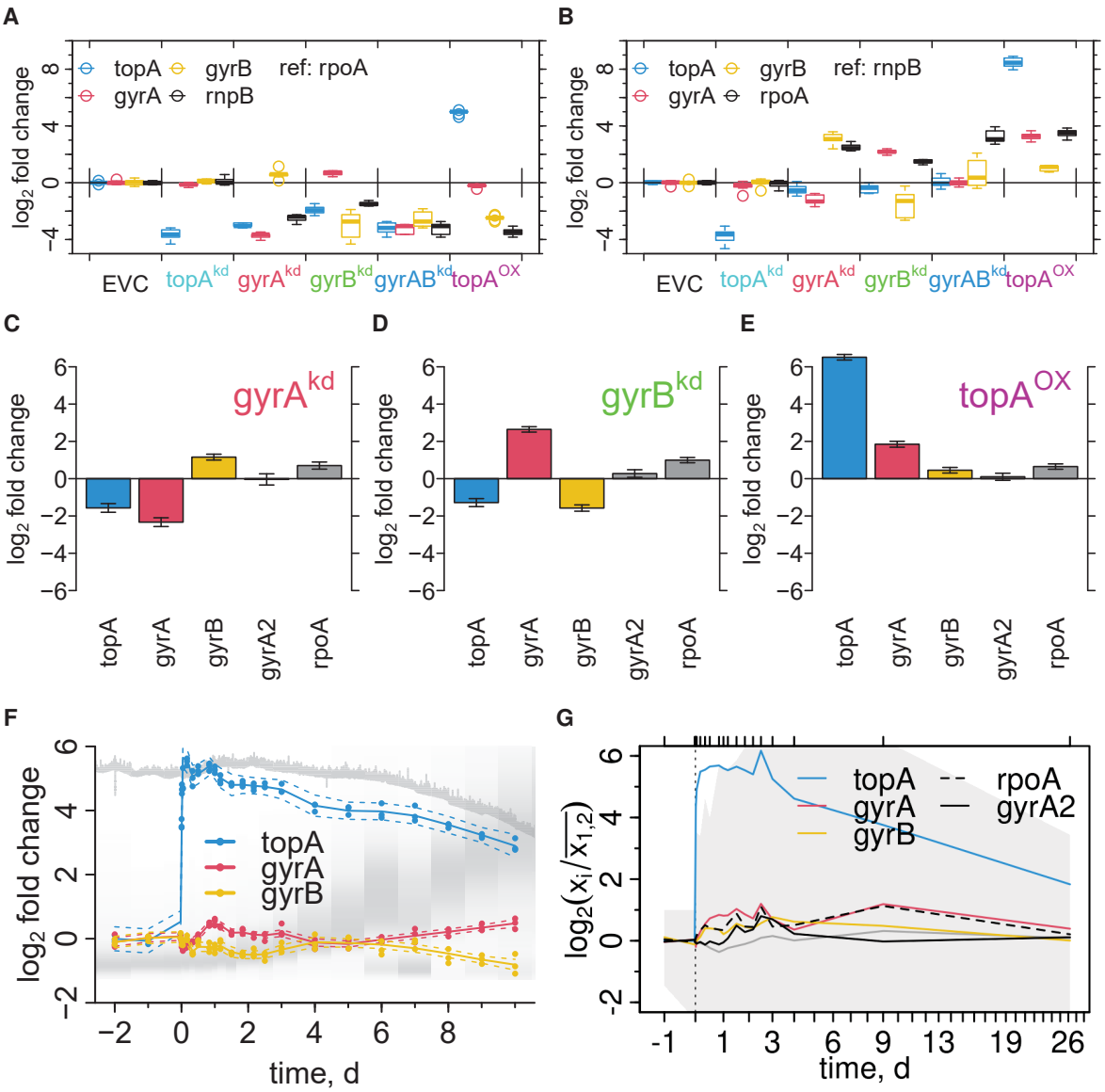


Figure S10. Homeostatic Regulation of Topoisomerase Genes: RT-qPCR vs. RNA-seq. **A and B:** RT-qPCR results from the endpoint experiments, reproduced from Figure S1A for comparison. **C-E:** Transcript abundance of topoisomerase genes and the reference gene *rpoA* used in qPCR (A,B) in the RNAseq data (Fig. 4, S6). *gyrA2* denotes the gene *sll1941*, a homolog of *gyrA*. **F-G:** The log2 fold change over the mean of the two pre-induction samples as measured by RT-qPCR (F) and RNAseq (G). RNAseq data shows that the expression of the reference gene *rpoA* increased with a periodic pattern and at a slightly higher fold change than *gyrB*, while *gyrA* stayed above or at similar levels as the reference gene. Considering this normalization effect (cf. [10]) the apparent downregulation of *gyrB* in (F) is an artifact of the reference gene, and the RT-qPCR and RNA-seq time series data are consistent.

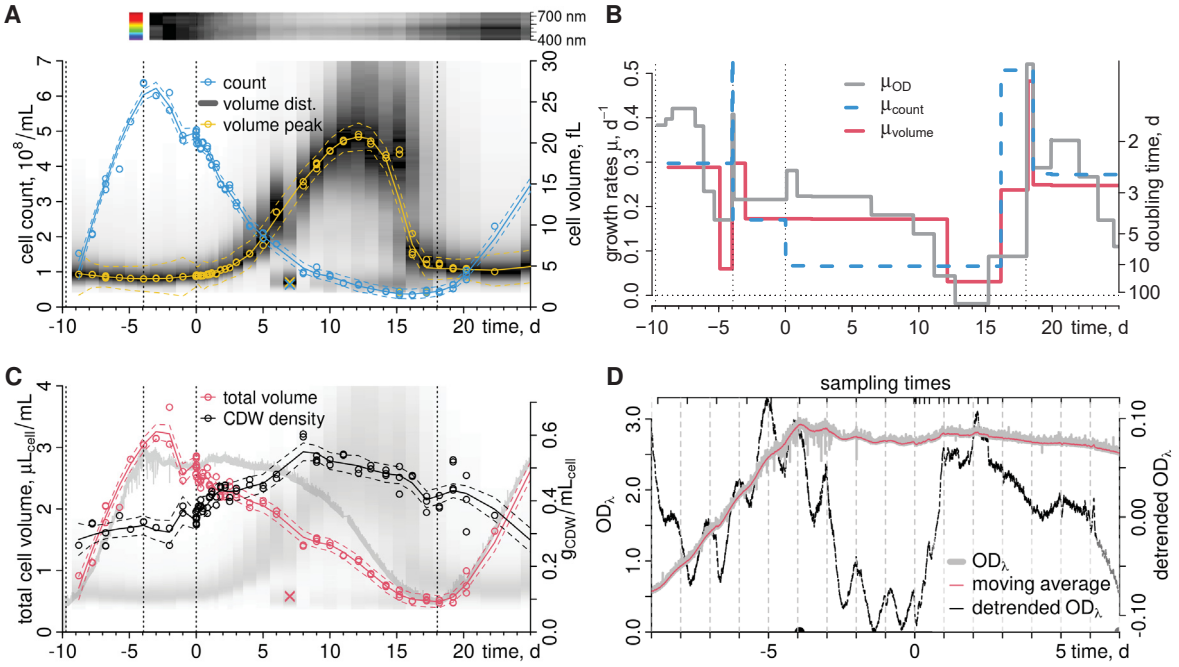


Figure S11. Pulsed Induction in Continuous Culture. Detailed biomass and growth rate data for Figure 5. **A:** Cell numbers (blue points) and volume distributions (gray scale) were recorded daily, and at higher resolution after induction, with the CASY cell counter. The peaks of the cell volume distributions are shown as yellow points. LOESS regressions with 95% confidence intervals are shown as lines. One outlier of the CASY measurement (x) was due to cell lysis during a washing step and was not included for regression. The top panel shows daily recorded spectra in gray-scale, where black indicates higher and white lower absorption. The underlying spectra were consistent with those of the batch culture shown in Figure S1.. **B:** The total cell volume (V_{total}) was calculated by integrating the single cell volume distributions in (B), and the CDW density were calculated by dividing the OD_λ signal, calibrated to the CDW measurements (A, Fig. A2C), by V_{total}. **C:** growth rates μ were calculated by local (piecewise) linear regressions of the OD_λ (A), and cell count (B) and total cell volume (D) measurements and subtraction of the culture dilution rate (Fig. A3). **D:** Detrending of the OD_λ signal for wavelet analysis (Fig. 5B). A moving average (red) of the raw signal by the OD probe (gray, time resolution: 30 s) was calculated using a window size of 30 min. The signal was detrended using the `detrend` function of the `pracma` R package (v2.3.6), where the end of the batch phase was used as the breakpoint (half bullet point on the bottom x-axis at ≈ -4 d). This subtracts two linear least-squares fits (before and after the breakpoint) from the data. This detrended signal was then used for Wavelet analysis with the `analyze.wavelet` function of `WaveletComp` R package (v1.1) for periods 0.24 h–96 h, and with `loess.span=0`, and `dj=0.05` and the mean time resolution of `dt=30 s`. The dashed vertical lines indicate full days around induction as a 24 h reference. Sampling daily before and in higher resolution after induction (time 0 h) may have induced or removed the $\approx$ 24 h pattern, and sampling times are indicated on the top axis. Note, that daily sampling times shifted while the detrended OD peaks remained at constant times.

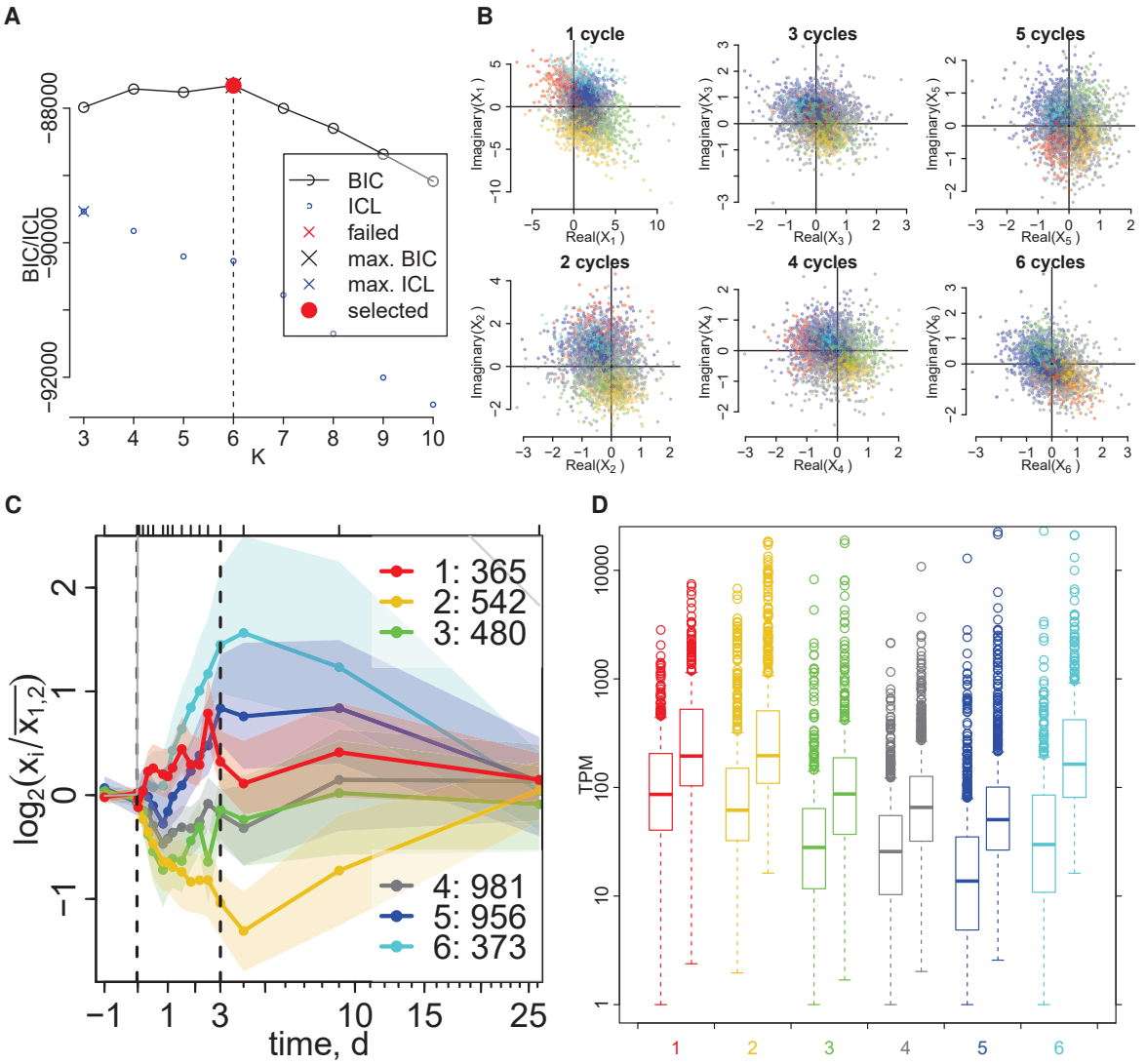


Figure S12. Clustering & Total Read-Count Distribution. **A:** Bayesian Information Content (BIC) as reported by `flowClust` for clustering of selected scaled components $X'_{k=1,\ldots,6}$ of the Discrete Fourier Transform (DFT) of the `arcsinh`-transformed TPM data over varying number of cluster centers (K). The maximal BIC was reached for a classification into $K = 6$ distinct clusters (co-expression cohorts). This clustering was chosen for further analysis. **B:** real and imaginary parts of the DFT that were used for clustering (R package `flowClust` [11]). Colors already indicate the final cluster assignments of each transcript at $K = 6$ (A). **C:** Cluster medians (solid lines) of the relative transcript abundances (rel. abund.). For each transcript the $\log_2$ of the ratio of read-counts at time points i to mean of the two samples before induction ($i = 1, 2$, at -1 d and -1 h) was calculated (points indicate the sampling time points i). The transparent ranges indicate the 25% and 75% quantiles of each cluster. Only the time points within to two vertical lines were used for clustering. **D:** Cluster-wise distributions (boxplots) of minimal (left) and maximal (right) read-count values (TPM) of each transcripts.

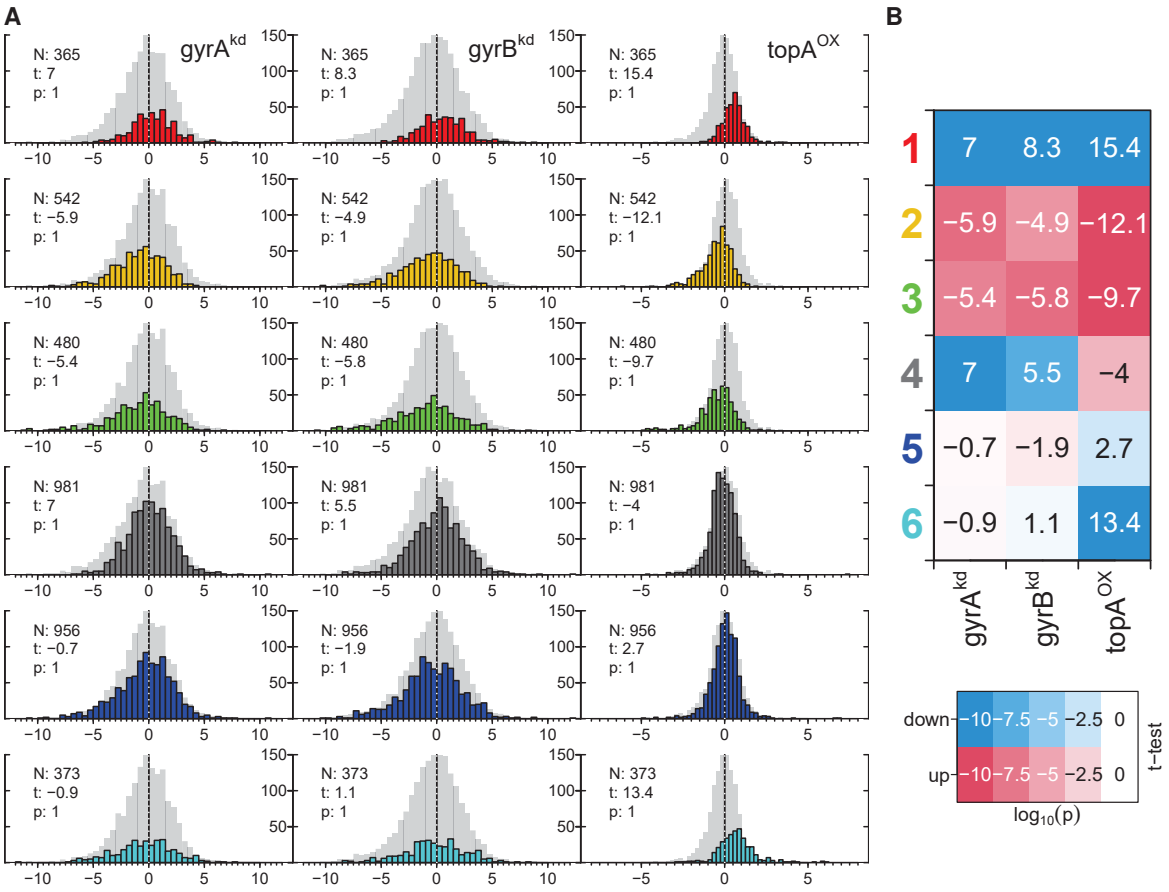


Figure S13. Co-Expression Cohorts in the Endpoint RNA-seq and Construction of t-Test Profiles. A: Distributions of the $\log_2$ fold-change of transcript abundances in the three strain endpoint experiment for each of the co-expression clusters derived from the topA^{OX} time series data. The gray background shows the distribution of all other transcripts. The y-axis are the counts for the colored distributions, while the gray background distributions are densities (without axis). For each cluster a t-test was performed (base R function `t.test`) against all transcripts not in the cluster, and the cluster sizes n, and the t-values and the p-values from each test are shown in each plot. The total number of transcripts with expression values were 3676 for gyrA^{kd} and gyrB^{kd}, and 3680 for topA^{OX}. **B:** A t-test profile plot is constructed from the t-test results in (A). A negative t-value indicates that the tested cluster transcripts have a lower mean abundance than all other transcripts and this is indicated by a red color field, the rounded t-value is shown in the fields; blue indicates a positive t-value and higher mean abundance. The p-value is converted to a transparency value for the red and blue colors (along a color palette from red/blue to white), such that the full color is reached for $p \leq p_{\text{min}}$, and for higher p-values the transparency scales with $\log_2(p)$. Both, for visibility of the text and to indicate an additional p-value cut-off the text (t-values) is plotted in white if $p \leq p_{\text{text}}$. The bottom legend shows 5 p-values (text: $\log_{10}(p)$) and the resulting field and text colors. Here $p_{\text{min}} = 10^{-10}$ and $p_{\text{text}} = 10^{-5}$.

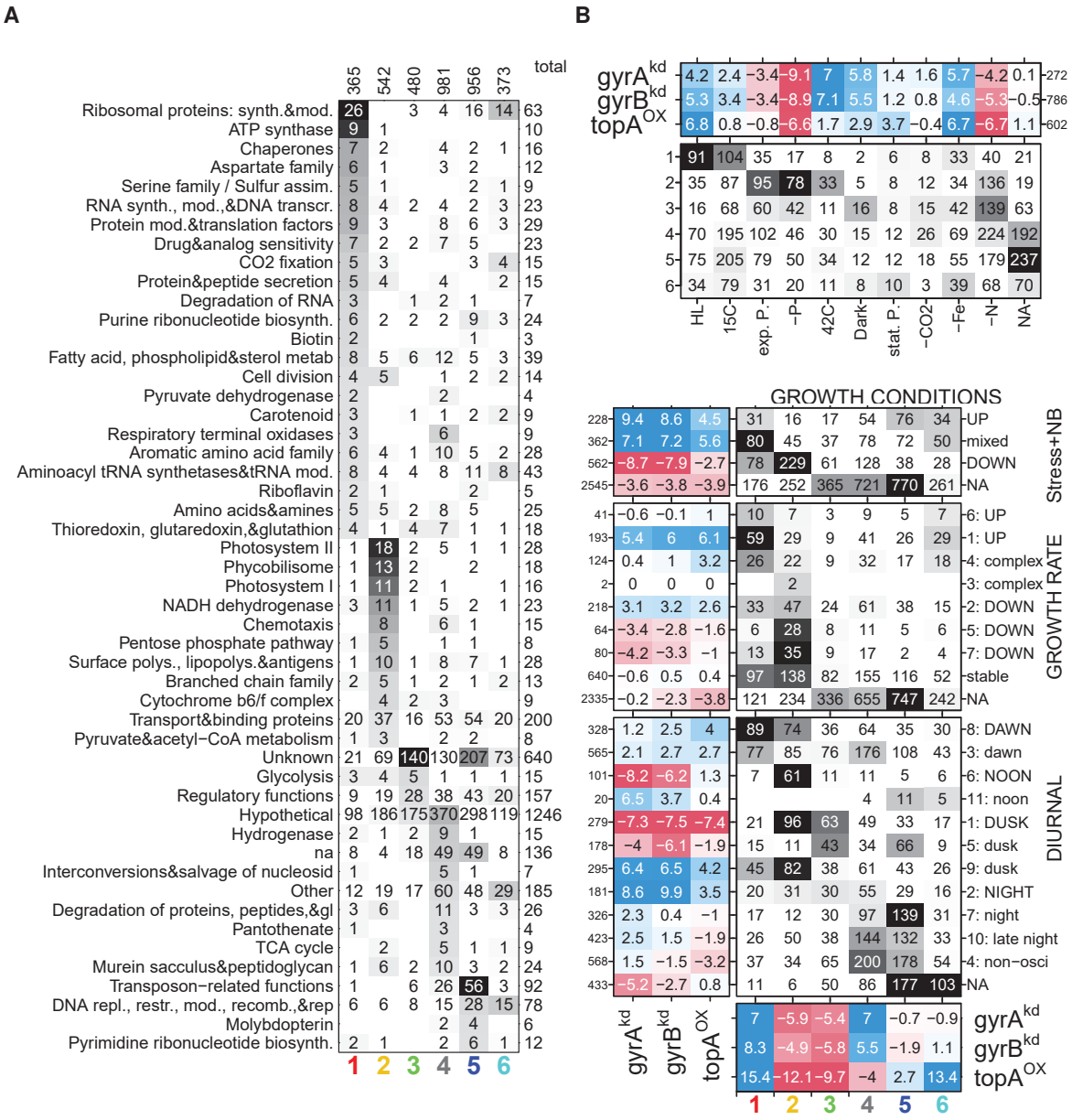


Figure S14. `CyanoBase` **Category Analysis of Co-Expressed Cohorts**. **A:** Sorted enrichment profile of functional category annotations as in Figure 6B (colored with $p_{\text{min}} = 10^{-10}$ and $p_{\text{text}} = 10^{-5}$) but sorted at $p_{\text{sort}} = 0.1$. All categories below the red line had only $p > p_{\text{sort}}$ and are unsorted. Some abbreviations of the original annotation terms are used for readability of the plot: synth. - synthesis, mod. - modification, repl. - replication, transcr. - transcription, recomb. - recombination, restr. - restriction, s. - saccharides, assim. - assimilation, & - and. **B:** Overlap enrichment and t-test profiles with clusterings as Figure 6C but for additional gene classifications from other publications; from top to bottom: experimental `GROWTH CONDITIONS` with maximal expression of transcription units from Kopf *et al.* [12], stress and novobiocin (`Stress + NB`) treatment (same as in Fig. 6C) by Prakash *et al.* [13], the original non-collapsed clustering of protein abundance level response to `GROWTH RATE` by Zavrel *et al.* [14], and a clustering of a `DIURNAL` transcriptome data set from the supplemental material by Lehmann *et al.* [15]. Numbers in the t-test fields (red and blue color scales) are the t statistic.

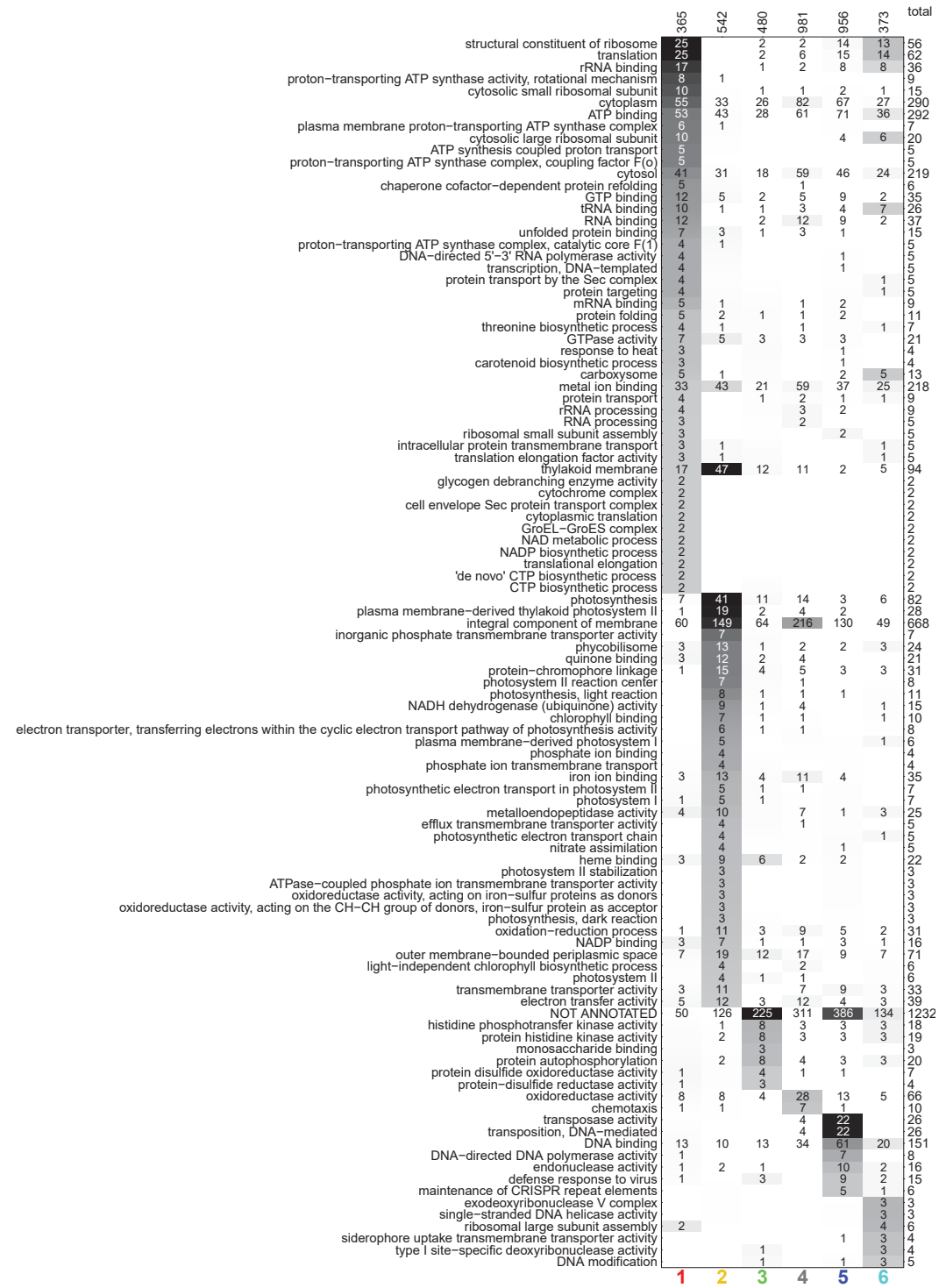


Figure S15. GO Analysis of Co-Expressed Cohorts. Sorted enrichment profile as in Figures 6B and S14A (colored with $p_{\text{min}} = 10^{-10}$ and $p_{\text{text}} = 10^{-5}$) but for Gene Ontology (GO) terms, downloaded from the `UniProt` database (2021-03-20, `organism:1111708`). Rows are cut and sorted along columns at $p_{\text{sort}} = 0.01$.

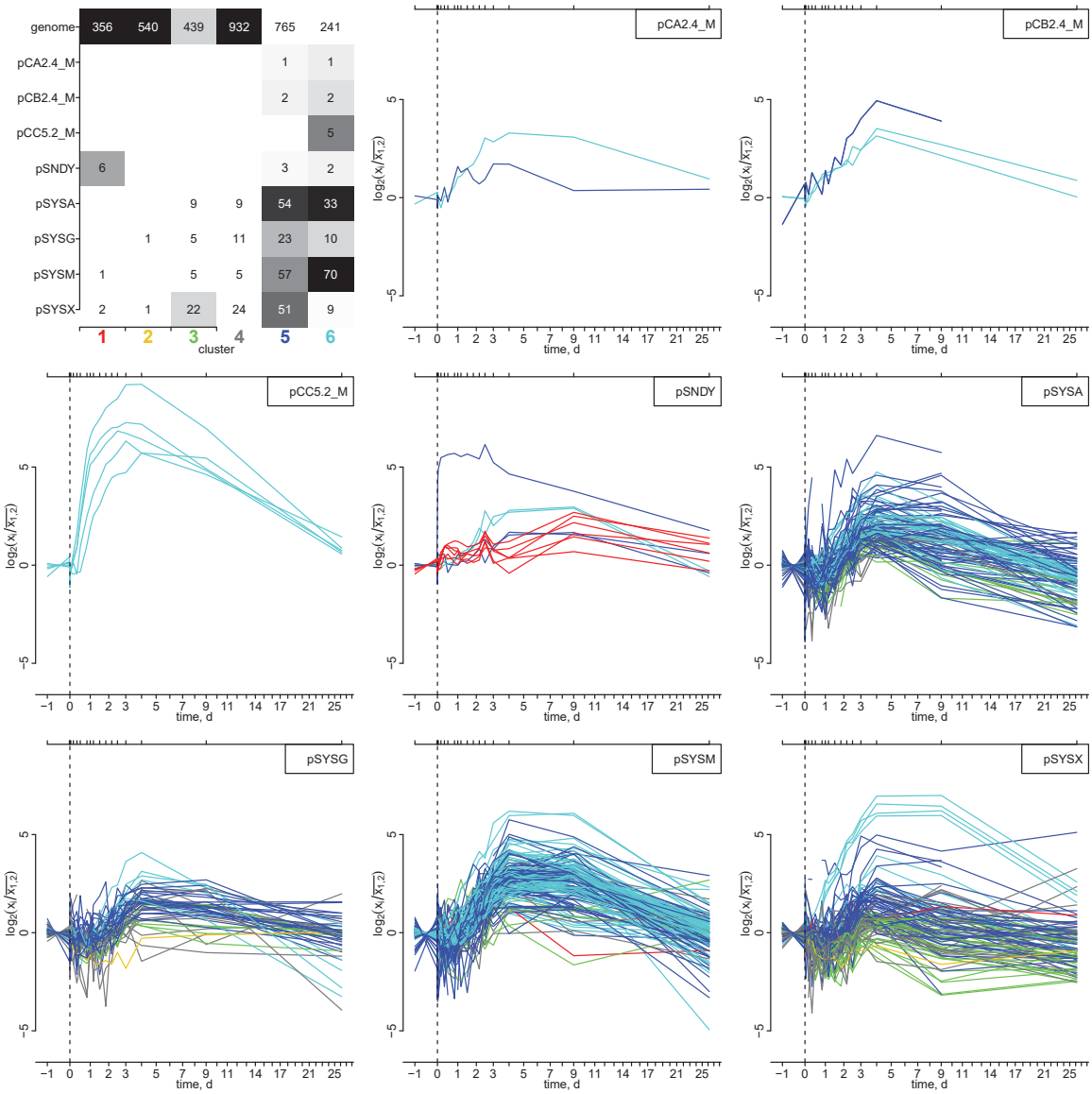


Figure S16. Transcriptome Time Series - Plasmids. Top left panel: Enrichment profile of time series clusters with the locations on the chromosome, one of the seven endogenous plasmids, or our construct pSNDY [2] (Table S1). All other panels show the temporal transcript abundance profiles for the coding genes of each plasmid (see top right legends for plasmid names); each transcript is colored according to its cluster label. Missing values stem from 0 read-counts in the raw data.

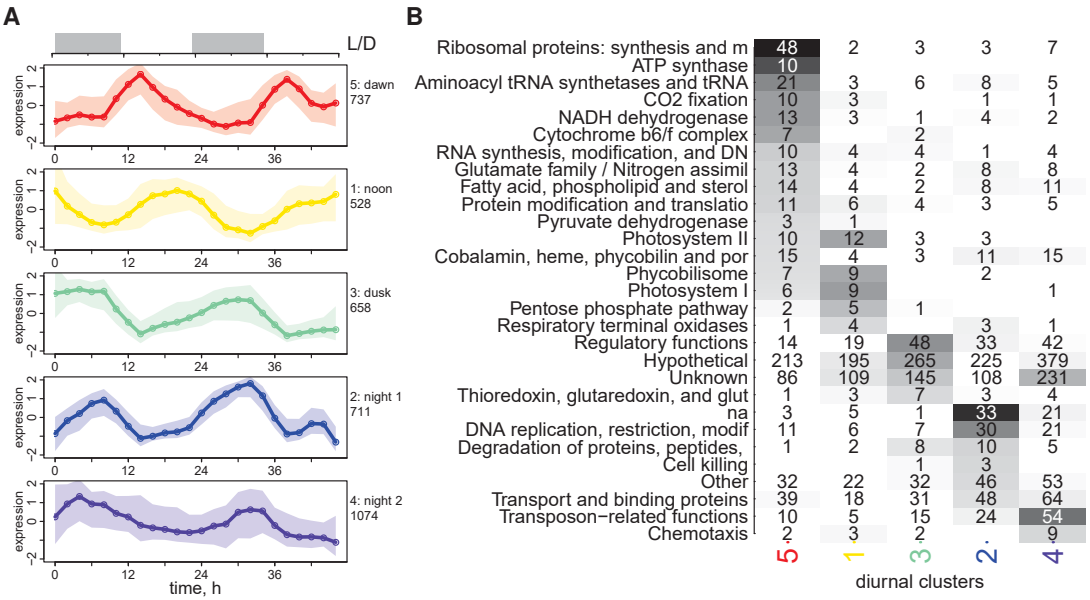


Figure S17. Diurnal Co-Expression Cohorts. Clustering of diurnal transcriptome data by Saha *et al.* [16] into 5 co-expression cohorts, using the Fourier transformed data and the `flowClust` algorithm. The maximal BIC clustering at $K = 5$ clusters was used for analysis. **A:** Cluster medians of the normalized (to mean 0) expression values with an additional moving average over 3 samples. Transparent ranges show the 10% and 90% quantiles, *i.e.* they encompass 80% of all values in a cluster. Cluster labels and sizes (number of genes) are indicated on the right y-axis. The gray and white bars on the top indicate dark and light phases of the experiment. **B:** Enrichment profiles of co-expressed cohorts with `CyanoBase` functional categories as for Figure 6B (colored with $p_{\text{min}} = 10^{-10}$ and $p_{\text{text}} = 10^{-5}$), but cut and sorted at $p_{\text{sort}} = 0.05$.

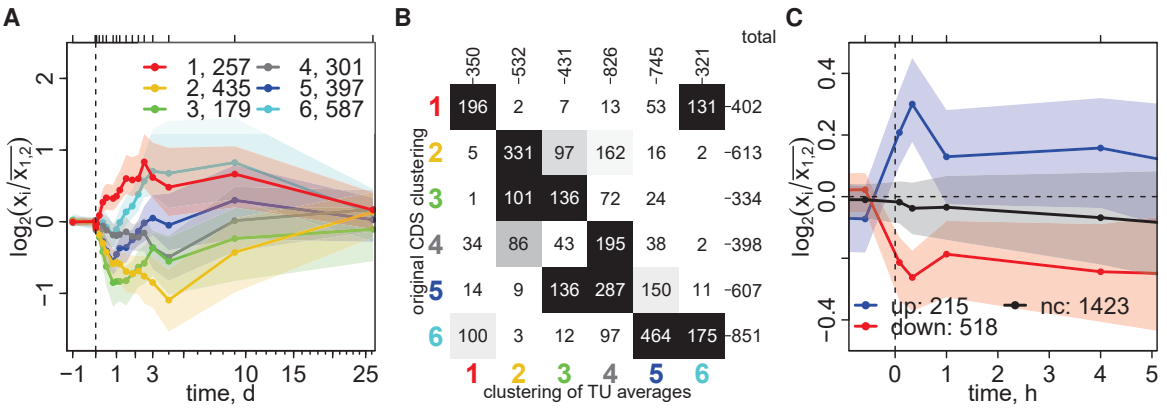


Figure S18. Transcription Unit Clustering. A: Clustering of transcription units (TU) defined by [12]. Average expression was calculated for all TU from the expression of coding genes they encompass (*via* the "Sense.tags" column of the original data set), and the resulting TU time-series was clustered by k-means, using cluster centers from the CDS clustering (Fig. 6, S12) and identical time-series processing. The number of TU in each cluster is indicated in the legend. The lines and dots are the medians and the ranges show the 25 % and 75 % quantiles of each cluster. **B:** enrichment profile of the original CDS clustering (y-axis) with the TU-based re-clustering; colored with $p_{\text{min}} = 10^{-10}$ and $p_{\text{text}} = 10^{-5}$ and with the original order. **C:** As (A) but for immediate response clusters. All TU were classified by the log2 ratio of the means of the two post-induction time points to the means the two pre-induction time points ($\log_2(\overline{x_{3,4}}/\overline{x_{1,2}})$); up: > 0.15, down: < -0.15, and nc: all others. The number of TU in each class is indicated in the legend.

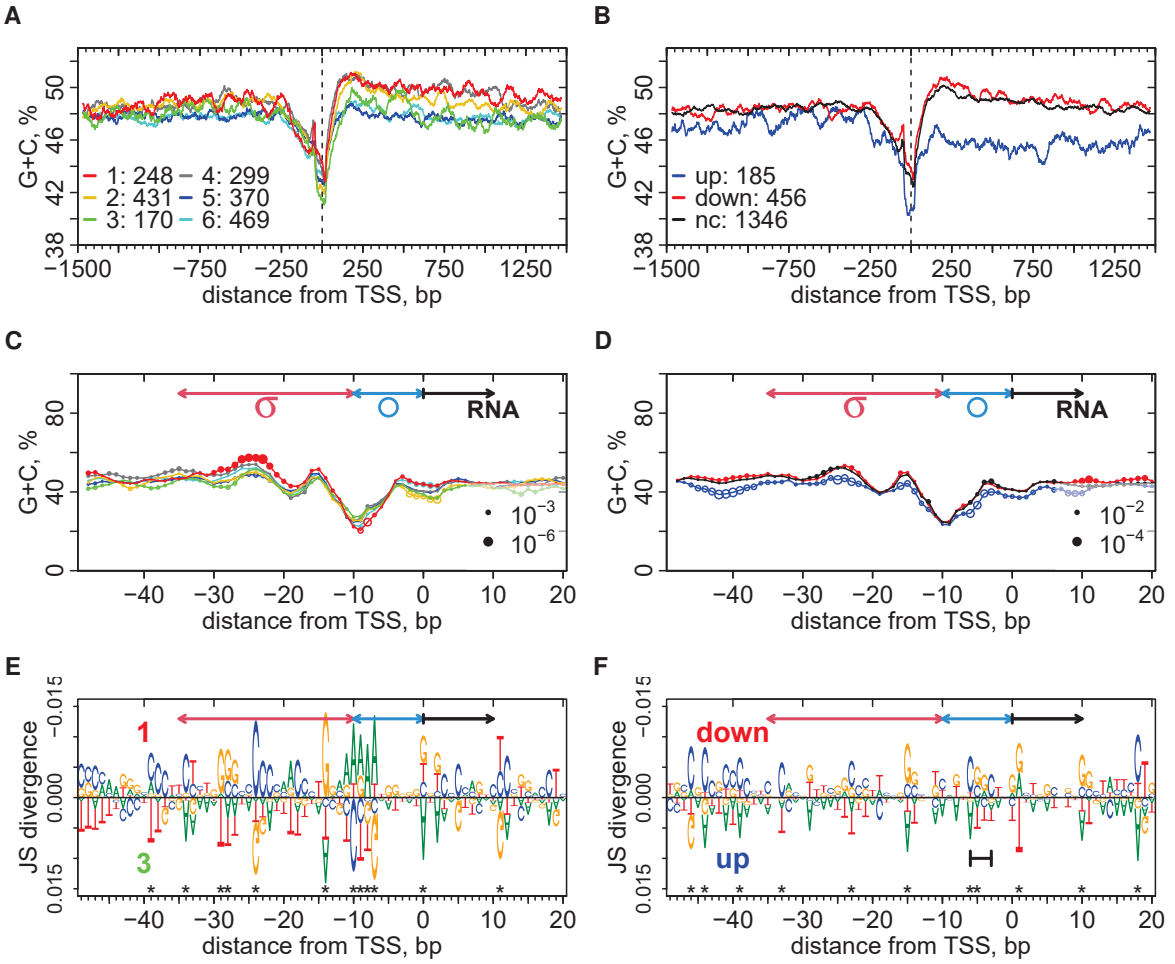


Figure S19. Long Range G+C Content and the Discriminator Region. The G+C frequencies of clustered TU (Fig. S18) aligned at their transcription start sites (TSS) were calculated in 66 bp bins (A and B) or 5 bp bins (C and D) around the TSS, sequence logos (E and F) were calculated without binning. **A & C:** Nucleotide frequency profiles for the differential response clusters shown in Figure S18A. (A) is identical to Figure 7A, and reproduced here for comparison. **B & D:** same as (A and C) but for the immediate response clusters (20 min post-induction) described in Figure S18C. The cluster legends in (A) and (B) provide the number of TUs in each cluster. Note, that only TU from the main genome were considered for this analysis, therefore the numbers are lower than those reported in Figure S18. The point sizes in (B) and (D) scale with $-\log_2(p)$, where p are the minimal p-values of two-sided cumulative hypergeometric distribution tests of the total counts in each cluster. Filled points indicate significant enrichment and open circles indicate significant deprivation of the motif count (here simply G or C on the forward strand) in the respective cluster. The dot size for the minimal and half-minimal p-values are indicated in the bottom-right legends. **E & F:** The Jensen-Shannon (JS) divergence [17] between the position weight matrices of the indicated promoter clusters; as in Figure 7E-F, reproduced here for direct visual comparison with G+C frequency profiles (B-D). The short horizontal bar in (F) indicates the GC discriminator region -6 to -3. See S24-S25 for details.

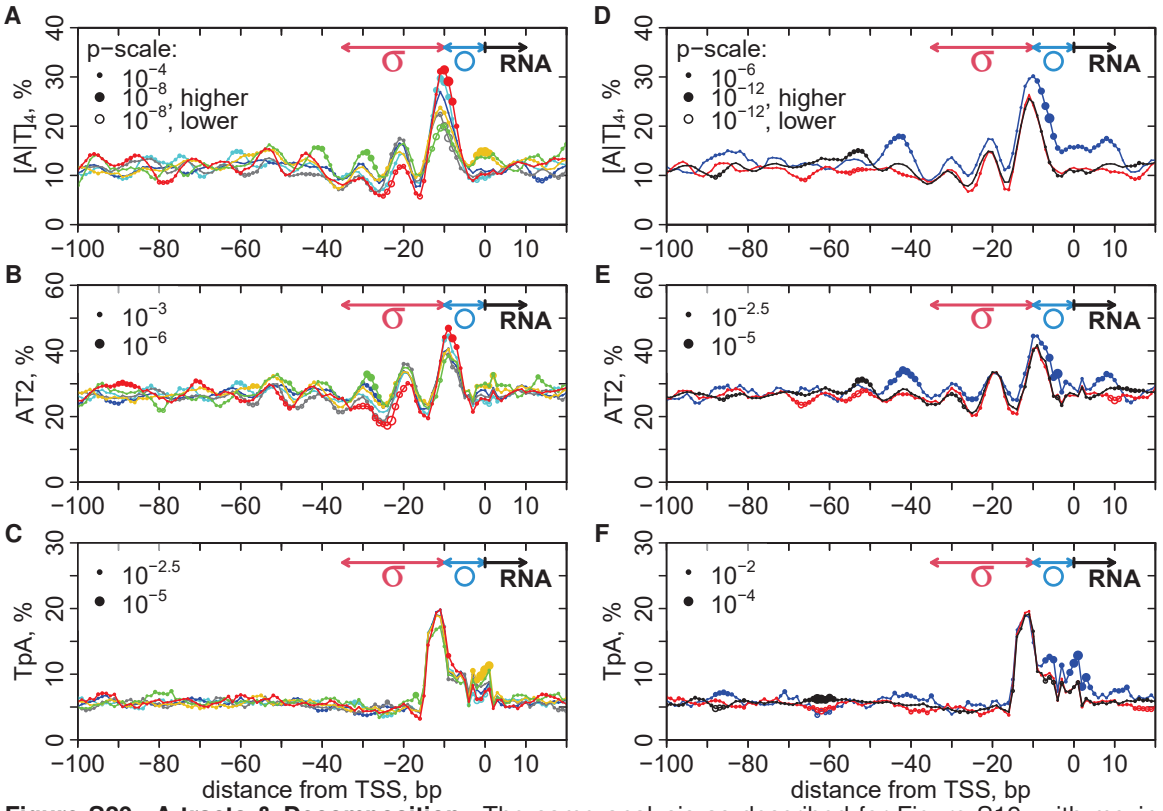


Figure S20. A-tracts & Decomposition. The same analysis as described for Figure S19, with moving averages over 5 bp, and for different DNA motifs: "$[A|T]_4$" are repeats of A and T nucleotides of length 4 (A-tracts). The "AT2" motif are dinucleotides ApA, ApT or TpA, and "TpA" is the dinucleotide "TpA". **A–C** are frequency profiles for the differential response TU clusters (Fig. S18A). **D–F** are frequency profiles for the immediate response clusters (Fig. S18C).

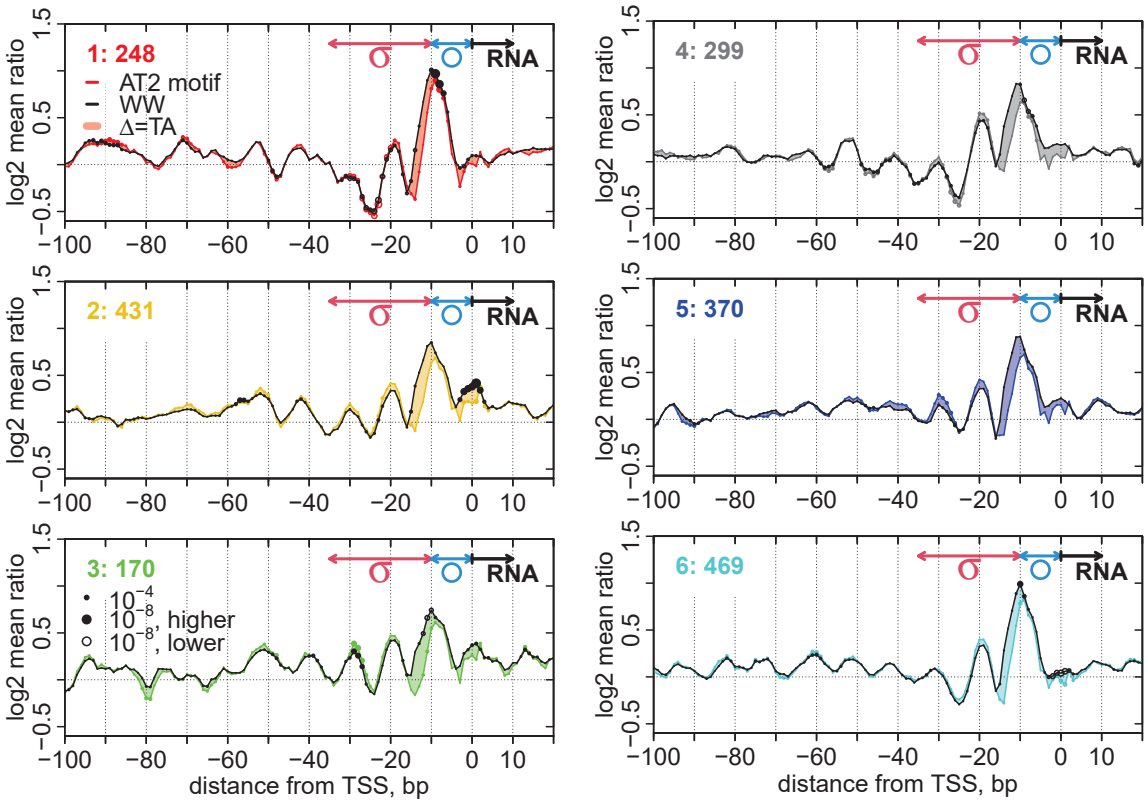


Figure S21. Normalized WW and AT2 Motif Frequencies. Direct visualization of the spatial relation of periodic enrichment of the AT2 motif and localized enrichment of the TpA step. The log2 mean ratios of the WW (black) and the AT2 dinucleotide motifs (colored by cluster) were calculated as described for Figure 7B and then normalized by dividing each position by the frequency in the total analyzed window. The TpA step is the difference between the WW and the AT2 motif frequencies and indicated by shade.

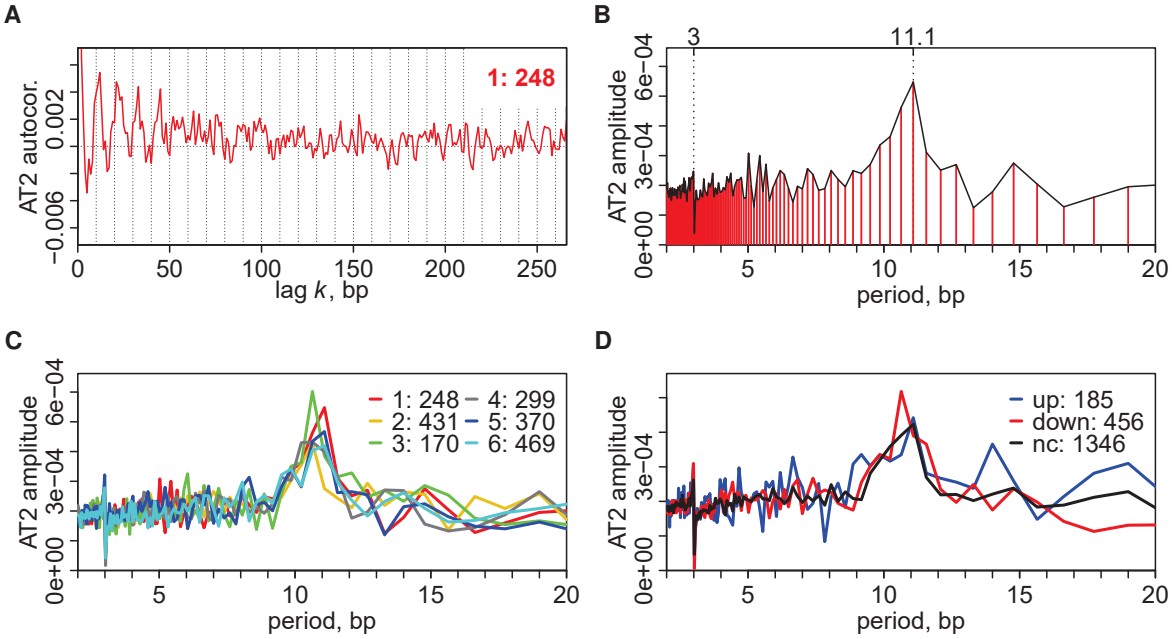


Figure S22. Autocorrelation Analysis of AT2 Motif Frequencies in Upstream Regions. A: The autocorrelation function of AT2 motif frequencies in concatenated promoter sequences (−150 bp–+15 bp around the TSS), here exemplary for cluster 1, was calculated after Schieg and Herzel [18] and as described by Lehmann, Machné and Herzel [15], but without removal of the first 30 positions which controls for periodic coding regions by alpha helices. **B:** Power spectrum of the autocorrelation functions in (A). **C & D:** Power spectra of the AT2 autocorrelation function is described in (A and B) but for all differential response clusters (C, Fig. S18A) and for all immediate response clusters (D, Fig. S18C).

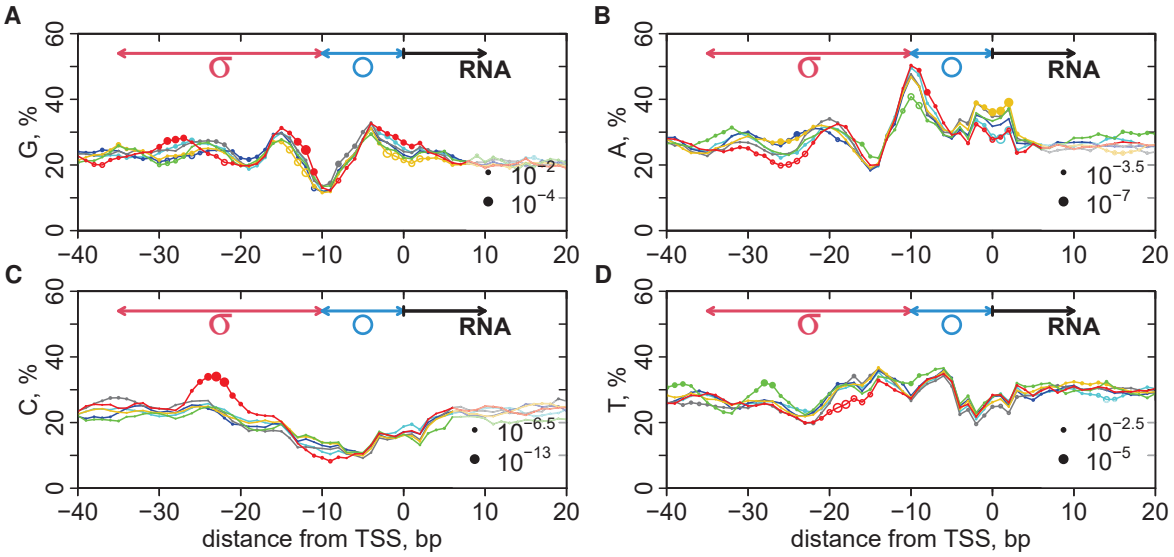


Figure S23. Single Nucleotide Frequencies. The same analysis as described for Figure S19, with moving averages over 5 bp, and for single nucleotides.

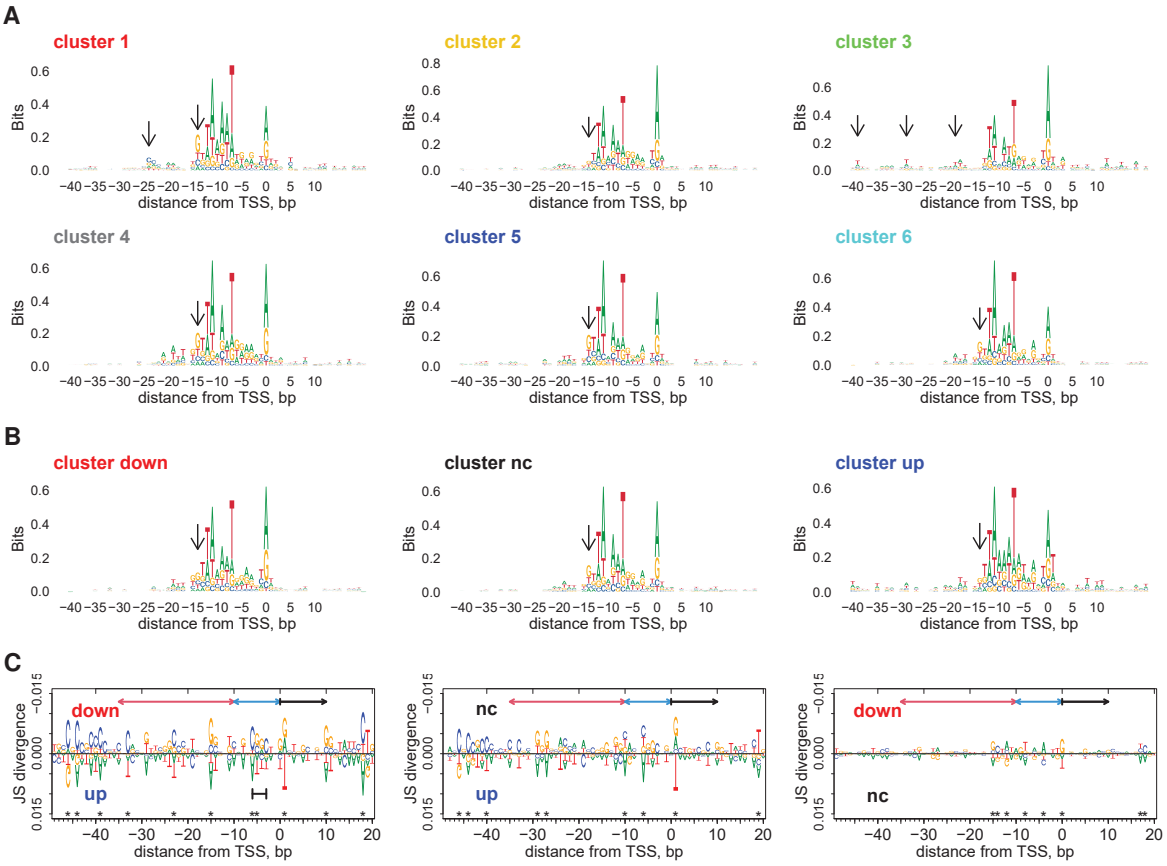


Figure S24. Sequence Logos at the Transcription Start Site. Sequence logos [19] were generated with `ggseqlogo` (`https://CRAN.R-project.org/package=ggseqlogo`) for the time series clusters 1-6 (**A**) and the early response clusters (**B**) from Figures S18A and C. The Jensen-Shannon (JS) divergence (**C**) between the early response clusters in (B) was calculated with `DiffLogo`, where $*$ indicate significance at $p < 0.05$ [17]. In all plots the TSS is 0, thus +1 should be added to get conventional nomenclature, where TSS is +1. **A-B:** All clusters showed overall similar sequence properties, reflecting previous results [20–24]: an enrichment of T and A at -12 and -11, and T at -7. The former (-12/-11) mark the first nucleotides of the open DNA bubble during initiation of transcription. The latter (-7) binds to a specific pocket in the σ^{70} factor of *E. coli* during open bubble formation [25–27]. A G at -14 ("extended -10") is enriched in all clusters except the downregulated clusters 2 (yellow) and 3 (green). This was also observed in promoters of genes upregulated in the ΔsigBCDE strain [23], and was stronger in SigA-bound than in SigE-bound TU [24]. Only cluster 1 (red) had a weak enrichment of C at -22 to -24, and only cluster 3 (green) showed weak enrichments of T at ca. -29, -39 and -49, i.e., in helically phased distances. A T at ca. -30 was also observed in genes downregulated in a ΔrpoZ strain [22]. **C:** The difference logo confirms a lower GC-content in the discriminator region, (marked by horizontal bars from -6 to -3), in the promoters that were upregulated ("up") immediately after topA^{OX} induction than in downregulated promoters ("down"), but shows similar enrichments throughout the core promoter.

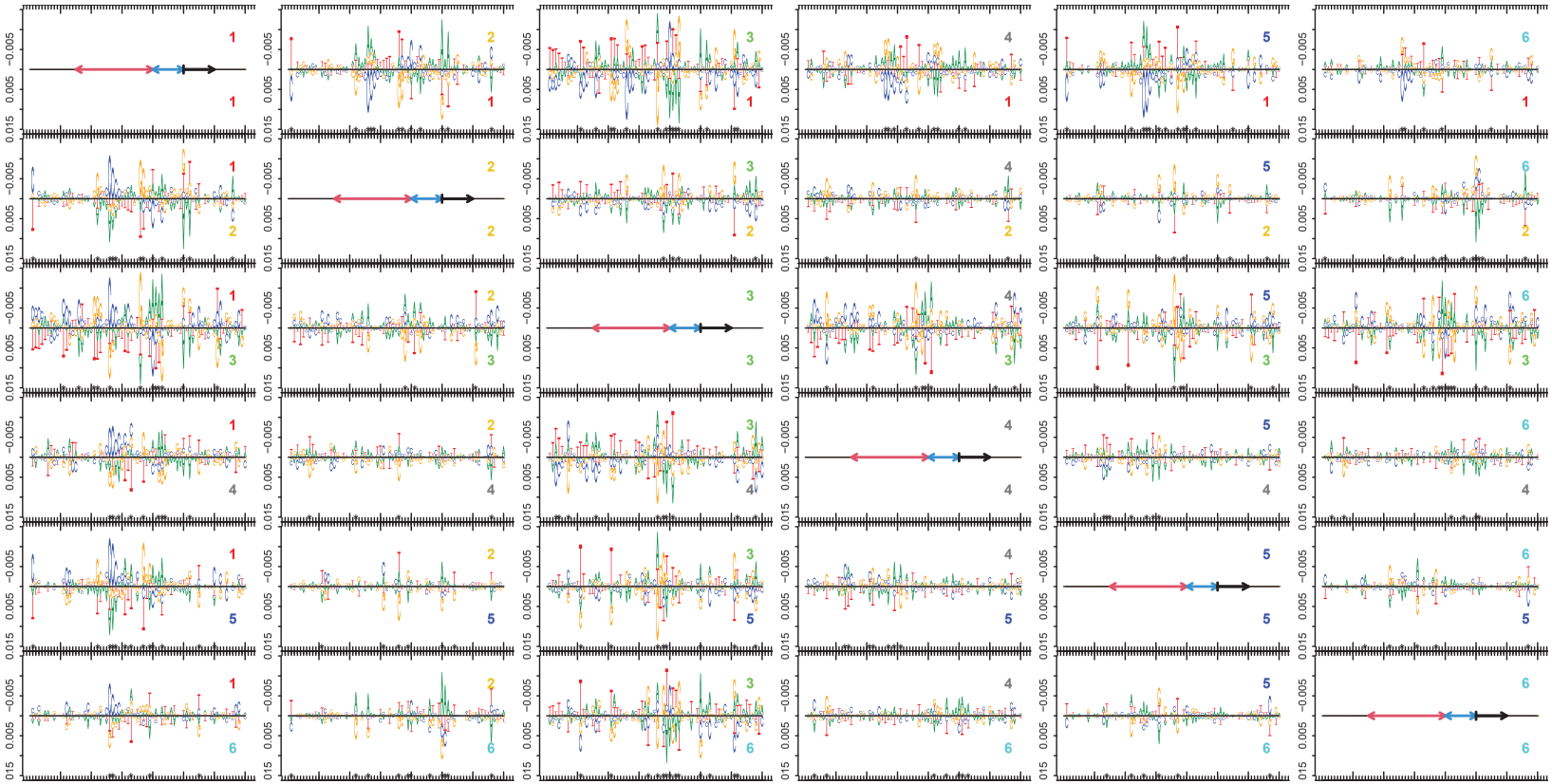


Figure S25. Difference logos for all pairs of time series clusters, see Figure S24A for details. The arrows indicate the -35, -10 and TSS (0) positions, as in all other promoter plots. The difference logos confirm the enrichment of C between -20 and -25 and A between -10 and -7 of cluster 1 promoters (Fig. S23), and the helically phased enrichment of A-tracts in cluster 3 (Fig. S20A).

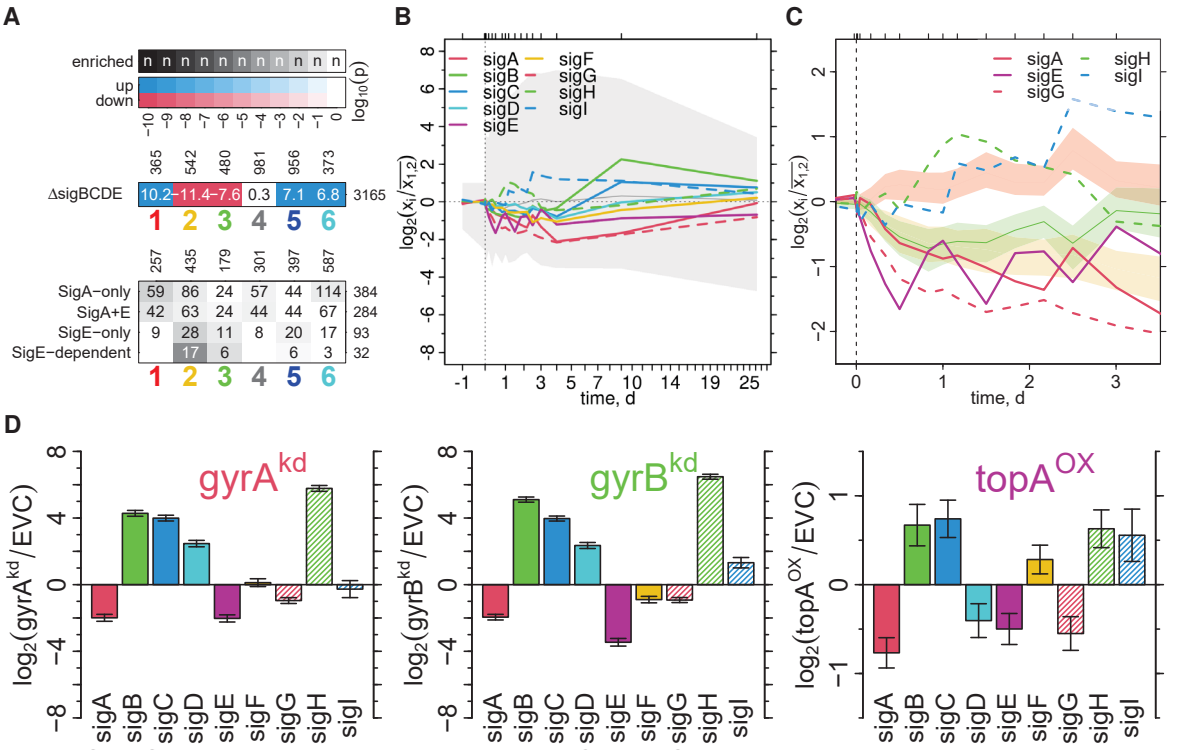


Figure S26. Sigma Factor Expression and Target Genes. *Synechocystis* carries 9 different sigma factors. SigA is essential and activates the transcription and translation machineries during autotrophic growth conditions [28]. The target genes of group 2 sigma factors (SigBCDE) overlap with SigA target genes and are involved in transcriptional remodeling during dark/light transitions and rapid respond to various adverse conditions [23, 24, 29, 30]. SigD and SigE may be involved in circadian control of transcription, since their transcript and protein abundances showed diurnal patterns [30–32]; and SigE activates transcription of sugar catabolic pathways, incl. glycogen breakdown [33]. The functions of group 3 sigma factors (SigFGHI) are not well studied. They lack the domain 1.2 [28] which in *E. coli* is known to mediate differential sensitivity of promoters to ppGpp [34]. SigH is upregulated and SigG downregulated during heat stress [35], and SigH is involved in acid acclimation [36]. SigI is upregulated during the dark phase of the diurnal cycle [16, 37]. **A:** T-test profile of cluster expression (by single genes, as in Fig. 6A) in the ΔsigBCDE strain [23] (blue/red color scale) and cluster enrichment profile (black/white color scale) of the TU clusters (Fig. S18) with SigA-bound, SigE-bound and SigE-dependent TU (column "TSSs within 100bp of peak summit" of Table S6, and Table 1 of [24]). Total numbers in each group are shown on the top and right axes and colors scale with the p-values as indicated in the legend. **B & C:** Time-series of normalized transcript abundances (Fig. 6, S12) for all nine annotated sigma factors of *Synechocystis* [28], for the full time series (B) and as a zoom on the first 3 days post-induction for selected sigma factors and with 25%/75% quantile ranges of clusters 1–3 shown as transparent ranges (C). **D:** Normalized abundances of the sigma factor transcripts in the endpoint measurements of the gyrA^{kd}, gyrB^{kd} and topA^{OX} strains (Fig. 4). **Summary:** SigA-bound TU are slightly enriched in the upregulated cluster 1 (A, bottom), yet the *sigA* transcript is downregulated in all conditions (B-D). Transcripts of the group 2 factors *sigB* and *sigC* are upregulated in all conditions (D), but only as a late adaptive response in the topA^{OX} time series (B); *sigD* is upregulated only in the gyr^{kd} strains, and *sigE* is strongly downregulated in all conditions (D). The time series (B-C) reveals a notable circadian pattern in *sigE* downregulation, preceding the cluster 3 peaks, while *sigH* is transiently upregulated during day 1, accompanied by an upregulation of *sigI* in three circadian steps, preceding the circadian peaks of cluster 1 transcripts. Thus, the response of sigma factors in our experiments may be related to both the glycogen-enriched phenotype (*sigE*, [33]) and the circadian pattern of gene expression after topA^{OX} induction. However, more data, e.g., chromatin-immunoprecipitation based binding studies, are required to test the relevance of these pattern.

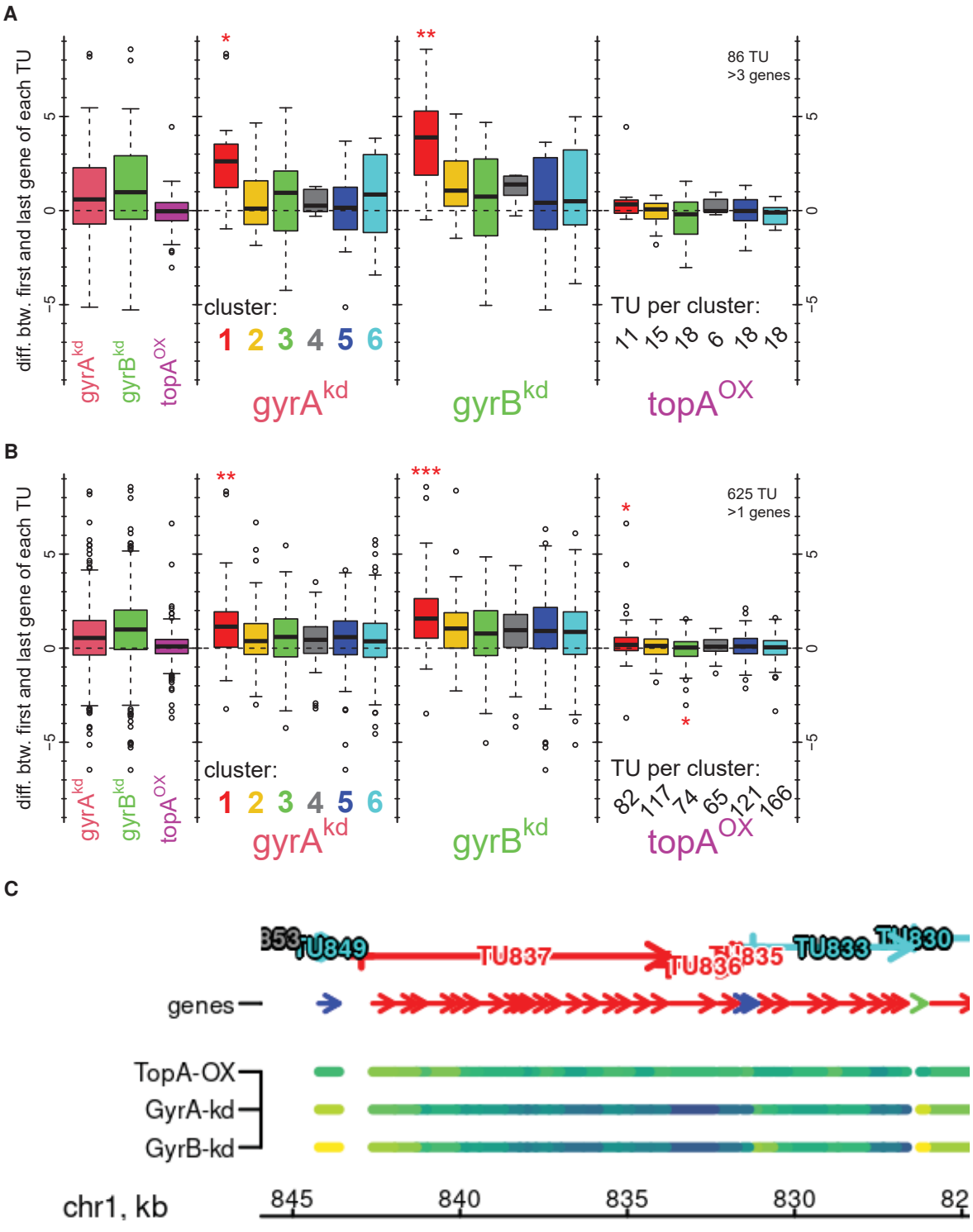


Figure S27. Graded Response Along Transcription Units. A & B: As Figure 7G but for all strains and all TU with ≥ 4 genes (A) or for all TU with ≥ 2 genes (B). Red stars above or below the boxplots indicate significance (*: $p < 0.05$, **: $p < 0.01$ and ***: $p < 0.001$) in two-sided t-tests of each cluster vs. all other clusters. **C:** Example TUs coding for ribosomal proteins. TU and coding genes are colored by their cluster labels. The transcript abundance levels of coding genes relative to the empty vector control in the endpoint experiments are color-coded by the `viridis` color scheme such that blue are lower and yellow are higher values. The values become progressively lower along the TU, in 5' to 3' direction, the gyrA^{kd} and gyrB^{kd} strains but not in the topA^{OX} strain.

Appendix A: Reactor Dynamics

Transient Increase in Cell Volume and Density. To study the dynamic response to transient *topA* induction, the topA^{OX} strain was grown in a Lambda Minifor bioreactor (Fig. A1) with continuous (online) monitoring of turbidity (OD_λ, Fig . A2A,B). Continuous culture dilution was initiated at $OD_\lambda \approx 2.9$ and with dilution rate $\phi \approx 0.24\,\mathrm{d}^{-1}$. The culture stabilized around $OD_\lambda \approx 2.7$. Notably, a subtle $\approx 24\,\mathrm{h}$ pattern of OD_λ was observed in both batch and pre-induction continuous growth phases. Then rhamnose was injected to $2\,\mathrm{mM}$ to induce overexpression of *topA*. The *topA* transcript was upregulated to ≈45-fold over the pre-induction level within $4\,\mathrm{h}$, as measured by RT-qPCR and confirmed by RNA-seq (Fig. S10) and decreased slowly over the course of the experiment. The OD_λ initially increased for $1\,\mathrm{d}$ post-induction, then slowly decreased. Cell dry weight (CDW) measurements were noisy but matched the OD_λ signal over the sampled period (Fig. 5A, A2C). In contrast, cell numbers started to decrease immediately, and cell volumes increased (Fig. S11A). We calculated growth rates of OD_λ, cell numbers and the total cell volume (Fig. S11C, A3). Cell division was not completely blocked but severely reduced to a division time of ≈10 d ($\mu_{\mathrm{count}} \approx 0.07\,\mathrm{d}^{-1}$). Total cell volume growth was much less affected and remained stable ($\mu_{\mathrm{volume}} \approx 0.18\,\mathrm{d}^{-1}$) throughout continuous culture operation until $12\,\mathrm{d}$ post-induction. Thus, artificial *topA* overexpression blocked cell division but not cell volume growth. OD_λ growth remained highest ($\mu_{\mathrm{OD}} \approx 0.23\,\mathrm{d}^{-1}$) and stable over the first 5 d–6 d. In parallel, glycogen content increased to about $35\,\%$–$40\,\%$ of the CDW (Fig. 5A). We further noticed that sampled cells started to sediment much faster, indicating increased intracellular density. By calibrating the OD_λ signal to the CDW measurements (Fig. A2C) and dividing by the total cell volume we can estimate a CDW density and this value also increased over time from 0.3 to $0.5\,\mathrm{g_{DCW}/mL_{cell}}$ (Fig. S11B). This range is consistent with data from *E. coli* [38, 39]. However, the CDW per OD_{750} was relatively lower for the enlarged strains in the endpoint measurement (Fig. 2A), and thus, the calibration to OD_λ may overestimate true CDW density. The enlarged and denser cells also became increasingly fragile: in the CASY cell counter data a small population of varying intensity appeared at $<2\,\mathrm{fL}$. This peak was highest at 7 d (outlier x in Fig. S11A), where cells were lysed during centrifugation in a washing step. The washing step was skipped thereafter, and the peak of small cells (dead or fragmented) remained small but increased towards the end of the continuous culture. Maximal cell volumes >20 fL were reached 10 d–15 d post-induction. From day 14 a population of smaller cells, $\approx 7.5\,\mathrm{fL}$, appeared. On 16 d this population was the majority, and cell volume further decreased to 5 fL. Cell pigmentation recovered and the culture appeared greener again. We then switched off dilution, and the culture resumed growth, although at lower growth rates than pre-induction.

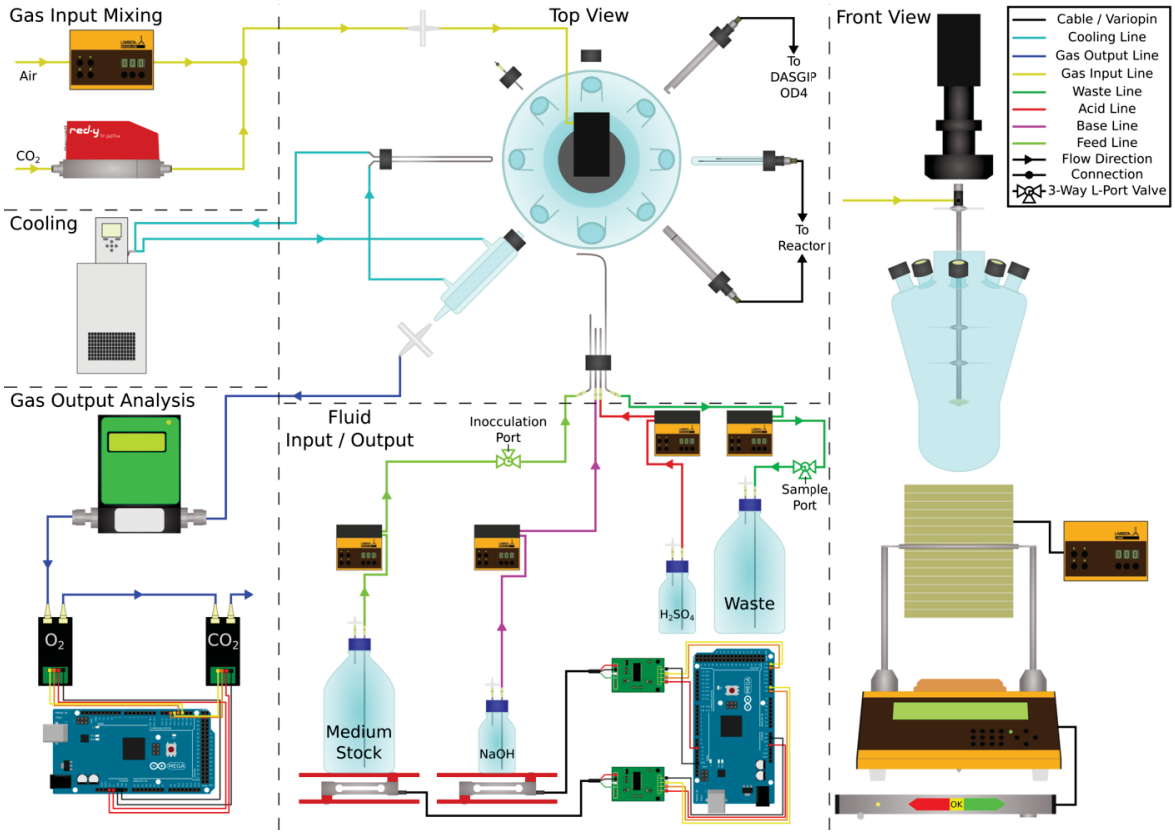


Figure A1. Photobioreactor Setup. A schematic overview of the cultivation setup showing the Lambda Minifor bioreactor in front and top-down views alongside its external components and custom expansions. The gas input mixture is generated by a Lambda MASSFLOW 5000 gas flow controller and a Voegtlin red-y smart controller, which regulate the flow of compressed air and CO_2 respectively. This input gas mixture is then introduced into the cultivation vessel via the sparger at the end of the agitation unit. The offgas condenser as well as the reactor's cooling finger are part of a water cooling circuit which is regulated by a Lauda Eco Silver thermostat set to $16\,^{\circ}C$. An Aalborg Massflow Meter monitors the flow rate of the culture's offgas before it is lead through a custom microcontroller-based gas sensor array in order to evaluate its O_2 and CO_2 content. The reactor actively regulates the culture's pH and temperature values by controlling its heating compartment as well as the Lambda Preciflow peristaltic pumps which are attached to NaOH and H_2SO_4 stock bottles, each $0.5\,M$. Additional culture parameters are monitored by a dissolved O_2 probe attached to the reactor and an OD4 probe connected to a DASGIP OD4 device. An additional set of peristaltic pumps is attached to the culture's medium stock and waste containers in order to control the reactor's volume and medium turnover. The reactor weighting module enables the system to operate under chemostat conditions. This is achieved by manually configuring the medium feed peristaltic pump at a constant speed in order to achieve a desired medium turnover rate while automatically regulating the waste pump speed to keep the total reactor weight constant. Additionally, a custom microcontroller-based scale setup is monitoring the weight of both the medium and NaOH stock bottles, which allows for the calculation of medium and base pump rates from the recorded data. The culture's illumination is provided by the Lambda LUMO modules, an LED strip fitted around the cultivation vessel.

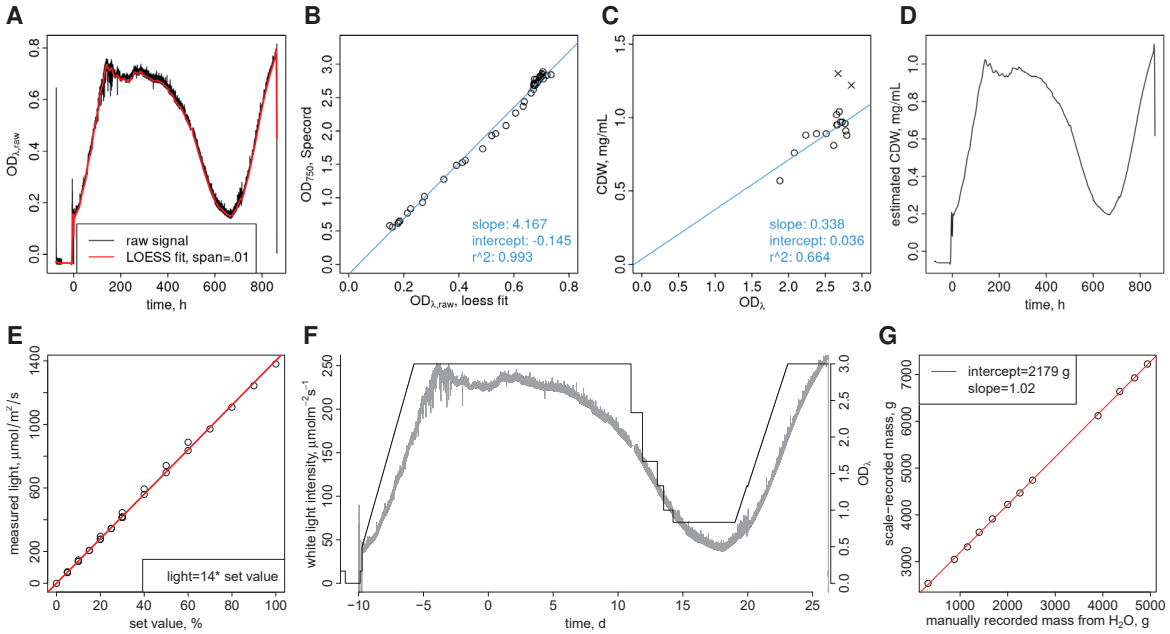


Figure A2. Calibrations. **A:** LOESS regression (R `loess`) of the raw signal (resolution ca. 1 sec) from the DASGIP OD4 module ($OD_{\lambda,raw}$). **B:** calibration of the LOESS fit of the $OD_{\lambda,raw}$ signal to offline $OD_{750\,nm}$ by linear regression (R *lm*). The calibrated signal is used throughout the document and denoted OD_{λ}. **C:** calibration cell dry weight (CDW) to the OD_{λ} signal. Data points marked by X were removed as outliers. **D:** the LOESS fit of the OD4 signal was then used to estimate CDW for all time points. **E:** calibration of the Lambda LUMO light module with a Licor light meter (LI-250A) with a spherical sensor bulb (LI-193). **F:** time-series of set and calibrated (white) light intensities (black line, left y-axis) compared to the OD_{λ} time-series (gray line, right axis). The light intensity was manually adjusted to avoid high-light stress in the culture during biomass decrease: light was initially increased as a ramp from 42 to 250 photons, then kept constant, and manually decreased to maintain light intensity approximately at $\sim$90 µmol m^{-2} s^{-1} per OD_{750}. After the switch to batch culture light was again increased from 70 to 250 µmol m^{-2} s^{-1}. **G:** The Arduino-based scales where calibrated prior to the experiment (not shown). During the experiment the liquid level on the 5L feed bottle was marked regularly, and the mass of water filled to these marked was recorded on a benchtop scale (Kern) after the experiment to test consistent performance. The recorded mass was reproduced sufficiently well (red line: linear regression): the intercept of the linear regression corresponds to the mass of the empty feed bottle and the slope was ≈ 1. Since the manual marks on the bottle are more error prone than the pre-calibration, we did not re-calibrate the data but relied on the recorded mass for calculation of the dilution rate.

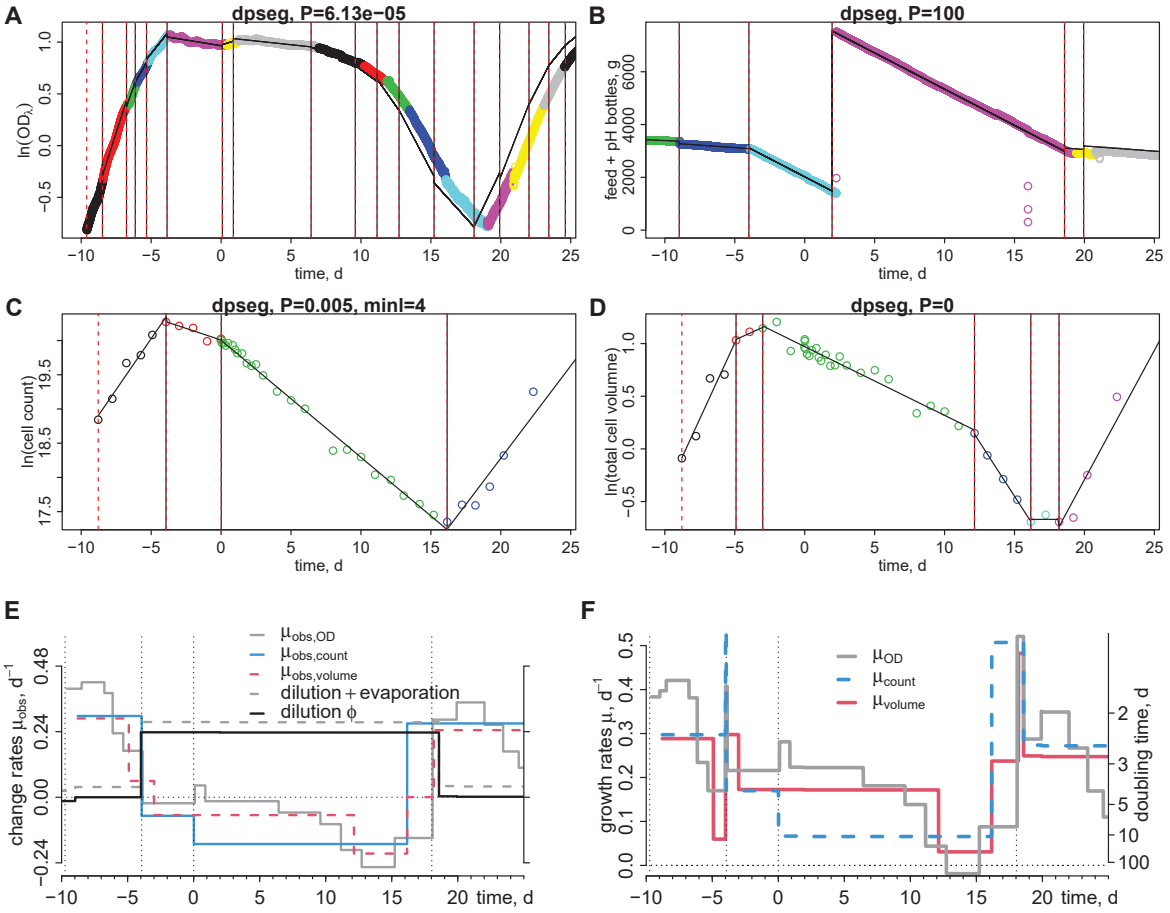


Figure A3. Calculation of Dilution and Growth Rates. All rates were calculated from the slopes of measured data (or of their natural logarithms as indicated) using piecewise linear segmentation with the R package `dpseg`. The plots in **A-D** were generated by `dpseg` and the vertical lines indicate borders of the piecewise segments, and the used penalty parameter P is shown in the plot title on the top axis. The minimal segment length parameter `minl` was only used in (C). **A**: the calibrated OD$_\lambda$ signal (1 sec resolution) was smoothed with a moving average and window size 15 and interpolated at 300 sec intervals. **B**: sum of the recorded weights of medium feed and pH control bottle weight; outliers (faulty measurements or bottle changes) were removed and data interpolated at 300 sec intervals. **C:** the total cell count for each CASY measurement, single measurements and means of technical duplicates. **D:** the total cell volume, calculated as the integral of the single cell volume distribution, for each CASY measurement, single measurements and means of technical duplicates. **E:** Observed rates. The (negative) slopes of the summed bottle weight changes (B) reflect the amounts added to the reactor culture by the Lambda reactor mass control system, assuming 1 g/mL density. The total culture dilution rate (dashed gray line, "dilution + evaporation") is obtained by division by the culture volume ($V_\ell = 1$ L). The liquid loss by evaporation is seen at times before onset of continuous culture (time -4 d) and is subtracted to obtain the actual dilution rate ϕ (black line). The slopes of the change of the natural logarithms of the OD$_\lambda$ signal (A), the total integrated cell volume (B), and the cell counts (C) are the observed change rates $\mu_{obs,OD}$ (gray line), $\mu_{obs,volume}$ (red line) and $\mu_{obs,count}$ (blue line), respectively. **F:** The culture growth rates μ_{OD} (gray line) and μ_{count} (blue line) and μ_{volume} (red line) were calculated as the difference between observed change rates and the culture dilution rate: $\mu = \mu_{obs} - \phi$.

References

1. Yao, L., Cengic, I., Anfelt, J., and Hudson, E. (Mar, 2016) Multiple Gene Repression in Cyanobacteria Using CRISPRi.. *ACS Synth Biol,* **5**(3), 207–212.
2. Behle, A., Saake, P., Germann, A., Dienst, D., and Axmann, I. (Apr, 2020) Comparative Dose-Response Analysis of Inducible Promoters in Cyanobacteria.. *ACS Synth Biol,* **9**(4), 843–855.
3. Labun, K., Montague, T., Krause, M., Torres Cleuren, Y., Tjeldnes, H., and Valen, E. (Jul, 2019) CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing.. *Nucleic Acids Res,* **47**(W1), W171–W174.
4. Bae, S., Park, J., and Kim, J. (May, 2014) Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases.. *Bioinformatics,* **30**(10), 1473–1475.
5. Cui, L., Vigouroux, A., Rousset, F., Varet, H., Khanna, V., and Bikard, D. (May, 2018) A CRISPRi screen in *E. coli* reveals sequence-specific toxicity of dCas9.. *Nat Commun,* **9**(1), 1912.
6. Pinto, F., Pacheco, C., Ferreira, D., Moradas-Ferreira, P., and Tamagnini, P. (2012) Selection of suitable reference genes for RT-qPCR analyses in cyanobacteria.. *PLoS One,* **7**(4), e34983.
7. Rüdiger, S., Rediger, A., Kölsch, A., Dienst, D., Axmann, I. M., and Machné, R. (2021) Plasmid supercoiling decreases during the dark phase in cyanobacteria: a clarification of the interpretation of chloroquine-agarose gels.. *bioRxiv,*.
8. Vischer, N., Verheul, J., Postma, M., van den Berg van Saparoea, B., Galli, E., Natale, P., Gerdes, K., Luirink, J., Vollmer, W., Vicente, M., and den Blaauwen, T. (2015) Cell age dependent concentration of *Escherichia coli* divisome proteins analyzed with ImageJ and ObjectJ.. *Front Microbiol,* **6**, 586.
9. Foley, J. bioanalyzeR: Analysis of Agilent electrophoresis data (2021) R package version 0.7.3.
10. Lehmann, R., Machné, R., Georg, J., Benary, M., Axmann, I. M., and Steuer, R. (Apr, 2013) How cyanobacteria pose new problems to old methods: challenges in microarray time series analysis.. *BMC Bioinformatics,* **14**(1), 133.
11. Lo, K., Hahne, F., Brinkman, R., and Gottardo, R. (2009) flowClust: a Bioconductor package for automated gating of flow cytometry data.. *BMC Bioinformatics,* **10**, 145.
12. Kopf, M., Klahn, S., Scholz, I., Matthiessen, J., Hess, W., and Voss, B. (Oct, 2014) Comparative analysis of the primary transcriptome of *Synechocystis* sp. PCC 6803.. *DNA Res,* **21**(5), 527–539.
13. Prakash, J., Sinetova, M., Zorina, A., Kupriyanova, E., Suzuki, I., Murata, N., and Los, D. (Dec, 2009) DNA supercoiling regulates the stress-inducible expression of genes in the cyanobacterium *Synechocystis*.. *Mol Biosyst,* **5**(12), 1904–1912.
14. Zavrel, T., Faizi, M., Loureiro, C., Poschmann, G., Stuhler, K., Sinetova, M., Zorina, A., Steuer, R., and Cerveny, J. (Feb, 2019) Quantitative insights into the cyanobacterial cell economy.. *eLife,* **8**, e42508.
15. Lehmann, R., Machné, R., and Herzel, H. (Oct, 2014) The structural code of cyanobacterial genomes.. *Nucleic Acids Res,* **42**(14), 8873–8883.
16. Saha, R., Liu, D., Hoynes-O'Connor, A., Liberton, M., Yu, J., Bhattacharyya-Pakrasi, M., Balassy, A., Zhang, F., Moon, T., Maranas, C., and Pakrasi, H. (2016) Diurnal Regulation of Cellular Processes in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803: Insights from Transcriptomic, Fluxomic, and Physiological Analyses.. *mBio,* **7**(3), e00464–16.
17. Nettling, M., Treutler, H., Grau, J., Keilwagen, J., Posch, S., and Grosse, I. (Nov, 2015) DiffLogo: a comparative visualization of sequence motifs.. *BMC Bioinformatics,* **16**, 387.
18. Schieg, P. and Herzel, H. (Oct, 2004) Periodicities of 10-11bp as indicators of the supercoiled state of genomic DNA.. *J Mol Biol,* **343**(4), 891–901.
19. Schneider, T. and Stephens, R. (Oct, 1990) Sequence logos: a new way to display consensus sequences.. *Nucleic Acids Res,* **18**(20), 6097–6100.
20. Vogel, J., Axmann, I., Herzel, H., and Hess, W. (Jun, 2003) Experimental and computational analysis of transcriptional start sites in the cyanobacterium *Prochlorococcus* MED4.. *Nucleic Acids Res,* **31**(11), 2890–2899.
21. Mitschke, J., Georg, J., Scholz, I., Sharma, C., Dienst, D., Bantscheff, J., Voss, B., Steglich, C., Wilde, A., Vogel, J., and Hess, W. (Feb, 2011) An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis sp.* PCC6803.. *Proc Natl Acad Sci U S A,* **108**(5), 2124–2129.
22. Gunnelius, L., Hakkila, K., Kurkela, J., Wada, H., Tyystjarvi, E., and Tyystjarvi, T. (Apr, 2014) The omega subunit of the RNA polymerase core directs transcription efficiency in cyanobacteria.. *Nucleic Acids Res,* **42**(7), 4606–4614.
23. Koskinen, S., Hakkila, K., Kurkela, J., Tyystjarvi, E., and Tyystjarvi, T. (Jul, 2018) Inactivation of group 2 sigma factors upregulates production of transcription and translation machineries in the cyanobacterium *Synechocystis* sp. PCC 6803.. *Sci Rep,* **8**(1), 10305.
24. Kariyazono, R. and Osanai, T. (Apr, 2022) Identification of the genome-wide distribution of cyanobacterial group-2 sigma factor SigE, accountable for its regulon.. *Plant J,* **110**(2), 548–561.
25. Shultzaberger, R., Chen, Z., Lewis, K., and Schneider, T. (2007) Anatomy of *Escherichia coli* sigma70 promoters.. *Nucleic Acids Res,* **35**(3), 771–788.
26. Feklistov, A. and Darst, S. (Dec, 2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit.. *Cell,* **147**(6), 1257–1269.
27. Chen, J., Chiu, C., Gopalkrishnan, S., Chen, A., Olinares, P., Saecker, R., Winkelman, J., Maloney, M., Chait, B., Ross, W., Gourse, R., Campbell, E., and Darst, S. (Apr, 2020) Stepwise Promoter Melting by Bacterial RNA Polymerase.. *Mol Cell,* **78**(2), 275–288.e6.

28. Imamura, S. and Asayama, M. (2009) Sigma factors for cyanobacterial transcription.. *Gene Regul Syst Bio,* **3**, 65–87.
29. Imamura, S., Yoshihara, S., Nakano, S., Shiozaki, N., Yamada, A., Tanaka, K., Takahashi, H., Asayama, M., and Shirai, M. (Jan, 2003) Purification, characterization, and gene expression of all sigma factors of RNA polymerase in a cyanobacterium.. *J Mol Biol,* **325**(5), 857–872.
30. Imamura, S., Asayama, M., Takahashi, H., Tanaka, K., Takahashi, H., and Shirai, M. (Nov, 2003) Antagonistic dark/light-induced SigB/SigD, group 2 sigma factors, expression through redox potential and their roles in cyanobacteria.. *FEBS Lett,* **554**(3), 357–362.
31. Kucho, K., Okamoto, K., Tsuchiya, Y., Nomura, S., Nango, M., Kanehisa, M., and Ishiura, M. (Mar, 2005) Global analysis of circadian expression in the cyanobacterium *Synechocystis* sp. strain PCC 6803.. *J Bacteriol,* **187**(6), 2190–2199.
32. Yoshimura, T., Imamura, S., Tanaka, K., Shirai, M., and Asayama, M. (Apr, 2007) Cooperation of group 2 sigma factors, SigD and SigE for light-induced transcription in the cyanobacterium *Synechocystis* sp. PCC 6803.. *FEBS Lett,* **581**(7), 1495–1500.
33. Osanai, T., Kanesaki, Y., Nakano, T., Takahashi, H., Asayama, M., Shirai, M., Kanehisa, M., Suzuki, I., Murata, N., and Tanaka, K. (Sep, 2005) Positive regulation of sugar catabolic pathways in the cyanobacterium *Synechocystis* sp. PCC 6803 by the group 2 sigma factor sigE.. *J Biol Chem,* **280**(35), 30653–30659.
34. Haugen, S., Berkmen, M., Ross, W., Gaal, T., Ward, C., and Gourse, R. (Jun, 2006) rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase.. *Cell,* **125**(6), 1069–1082.
35. Huckauf, J., Nomura, C., Forchhammer, K., and Hagemann, M. (Nov, 2000) Stress responses of *Synechocystis* sp. strain PCC 6803 mutants impaired in genes encoding putative alternative sigma factors.. *Microbiology (Reading),* **146 (Pt 11)**, 2877–2889.
36. Zhang, X., Chen, G., Qin, C., Wang, Y., and Wei, D. (Nov, 2012) Slr0643, an S2P homologue, is essential for acid acclimation in the cyanobacterium *Synechocystis* sp. PCC 6803.. *Microbiology (Reading),* **158**(Pt 11), 2765–2780.
37. Srivastava, A., Summers, M., and Sobotka, R. (May, 2020) Cyanobacterial sigma factors: Current and future applications for biotechnological advances.. *Biotechnol Adv,* **40**, 107517.
38. Zimmerman, S. and Trach, S. (Dec, 1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*.. *J Mol Biol,* **222**(3), 599–620.
39. Oldewurtel, E., Kitahara, Y., and van Teeffelen, S. (Aug, 2021) Robust surface-to-mass coupling and turgor-dependent cell width determine bacterial dry-mass density.. *Proc Natl Acad Sci U S A,* **118**(32), e2021416118.

6 Chapter 6
Conclusion and outlook

Homo sapiens, the only surviving species of the human genus, has existed on Earth for about 300 thousand years. Due to the looming climate collapse, depletion of finite resources and economic systems based on petrochemicals, we are not only threatening humanity and its culture but are also presented with the sixth great mass extinction, if we do not adopt a new way of life. Authorities should acknowledge the importance and potential of research on reducing fossil fuel consumption and optimizing industrial processes. Microbial production hosts have been identified as a sustainable source for the synthesis of raw materials useful for the chemical and pharmaceutical industries and can be an important building block towards finding solutions. Therefore, sustainable production of terpenoids in microbial hosts is an opportunity in today's biotechnological research which provides access to biochemical diversity. Cyanobacteria present themselves as a particularly attractive platform for circular bioeconomy. While production in established heterotroph platform organisms has highlighted the general promise of microbial chassis organisms, sustainable bioproduction using light and CO_2 will require the development of robust cyanobacterial host organisms. Engineering of the production pathway is necessary, but it is also necessary to change other routes in order to reroute carbon from growth to product in order to reach production levels that are commercially viable. The ability to employ CRISPRi to inhibit transcription of a crucial route, in this case, the pathway leading to the synthesis of carotenoids presents an important proof-of-concept. Simple knock-out mutation would not have made this conceivable, but it demonstrates the range of what is theoretically feasible outside of essentiality. Future engineering of photoproduction in cyanobacteria will be guided by metabolomics-based methodologies in addition to *in silico* flux analysis, enabling the creation of more metabolically balanced and durable strains. Repression of the *crtE'* transcript revealed that there was a delicate balance between reduced carotenoid content and *Synechocystis* viability. Fine tuning of carotenoid pathway repression is required for more efficient production of valencene. Ideally, the knowledge gained from the *in silico* analysis to improve squalene production can also be used to improve valencene production. In the case of squalene improvement, mainly the reaction steps in the MEP synthesis pathway and in parts of the Calvin-Benson cycle and lower glycolysis were used. By combining this improvement and the adapted terpenoid synthesis pathway, valencene production could be further optimized. At the same time, *in silico* analysis allows prediction of an optimal flux through *CrtE*, in which viability and production are in an ideal balance.

Numerous previous works in highly various organisms have been performed with similar results. *Synechocystis* sp. PCC 6803 has been subjected to genetic alterations involving the deletion of the *shc* gene in previous works already, resulting in a squalene titer of 0.67 $mg \cdot OD^{-1} \cdot L^{-1}$ [27]. Introducing an additional Squalene Synthase from *A. thaliana* led to an improved titer of 5.1 $mg \cdot L^{-1}$ [110], a result which was replicated by another group with an algal squalene synthase[111]. Building on these works and the deletion of *shc*, further modifications such as the deletion of sqs and dCas9-mediated knock-down of *crtE* coupled with the overexpression of CnVS and *ispA*, were employed for the production of valencene in the same host in this work, reaching titers of 17.6 $mg \cdot L^{-1}$ [112]. Almost simultaneously, a different group was able to show biosynthesis of valencene in *Synechocystis* with similar strategies[113]. The same genes (CnVS and *ispA*) were used for overexpression, resulting in a slightly lower titer of 9.6 $mg \cdot L^{-1}$ compared to titers achieved in a comparable strain in this work (12.5 $mg \cdot L^{-1}$), the slight

difference likely resulting from the lack of the *ΔshcΔsqs* deletion in the cited work. The similar titers achieved in these two independent studies highlight the robust methods and evaluation of the strains constructed in this work.

Table 1: Overview of product yields for squalene and valencene produced in different organisms as found in literature.

Organism	Product	Modifications	Titer	Ref.
Synechocystis sp. PCC 6803	Squalene	Δ*shc*; expression of BSS	5.1 mg·L^{-1}	111
Synechocystis sp. PCC 6803	Squalene	Δ*shc*	0.67 mg·OD^{-1}·L^{-1}	27
Synechocystis sp. PCC 6803	Squalene	Δ*shc*	5.1 mg·L^{-1}	110
Synechocystis sp. PCC 6803	Valencene	Overexpression of CnVS and *ispA*	9.6 mg·L^{-1}	113
Synechocystis sp. PCC 6803	Valencene	Δ*sqs,* Δ*shc;* Overexpression of CnVS, ispA; CRISPRi of crtE	17.6 mg·L^{-1}	112
Synechococcus sp. 7942	Squalene	Expression of SQS, *idi, dxs, ispA*	4.98 mg·OD^{-1}·L^{-1}	114
Escherichia coli	Squalene	Δ*pgi*, Δ*menA*, expression of *hsqs, zwf, pgi, idi, ispA, dxs, udhA*	52.1 mg·L^{-1}	115
Rhodobacter capsulatus	Squalene	Expression of A. *thaliana* SQS1	8.24 mg·L^{-1}	110
Rhodobacter capsulatus	Valencene	Expression of CnVS, ispA, MVA cluster	18 mg·L^{-1}	116
Corynebacterium glutamicum	Valencene	Expression of ERG20, ispA, CnVS; ΔcrtE, ΔidsA	2.41 mg·L^{-1}	117
Saccharomyces cerevisiae	Valencene	Overexpression of *erg10*, *erg13*, *tHMG1*, *erg12*, *erg8*, *erg19*, *erg20* and *idi1, Δrox1*, down-regulated erg9, *Δbts1*, *Δdpp1*, *Δlpp*	539 mg·L^{-1}	118
Yarrowia lipolytica ATCC 20460	Valencene	Overexpression of HMG1, ERG12, ACL1, SeACS, IDI, ERG20, CnVS downregulation of SQS	113.9 mg·L^{-1}	119
Ustilago maydis	Valencene	Expression of CnVS, *crtB*; Δ*car2*	5.5 mg·L^{-1}	120

Synechococcus sp. PCC 7942, another cyanobacterial strain, exhibits the production of squalene through the expression of key biosynthetic genes, resulting in a titer of 4.98 mg OD^{-1}·L^{-1} [114]. *Escherichia coli*, a well-established microbial workhorse, achieves a notable squalene titer of 52.1 mg/L through the deletion of *pgi* and *menA* genes, coupled with the expression of an array of biosynthetic genes[115].

Rhodobacter, a photosynthetic bacterium, exhibits a Squalene titer of 8.24 mg·L^{-1} through the expression of the *Arabidopsis thaliana* SQS1 gene[110]. Additionally, valencene production in *Rhodobacter* is achieved through the expression of CnVS, *ispA*, and the MVA cluster, yielding a titer of 18 mg·L^{-1} [116].

Next to work in photosynthetic chassis organisms, valencene and squalene were also successfully produced in various heterotrophs, some of which are established platform organisms in industrial biotechnology.

Valencene production was demonstrated in *Corynebacterium glutamicum*, a gram-positive bacterium, through the expression of ERG20, *ispA*, and CnVS genes, accompanied by the deletion of *crtE* and *idsA* genes, resulting in a titer of 2.41 mg·L^{-1} [117].

Ustilago maydis, a fungus originally isolated as a pathogen from maize plants, but now a relevant chassis in biotechnological applications, serves as a host for valencene production through the expression of CnVS and *crtB* genes, along with the deletion of the *car2* gene, yielding a titer of 5.5 mg·L^{-1} [120].

Significantly higher titers of valencene have been achieved in the two yeasts *Yarrowia lipolytica* and *Saccharomyces cerevisiae*, both of which serve as biotechnological platform organisms in

the industry. With 113.9 and 539 mg·L^{-1}, respectively[118,119], they are capable of achieving 10-50 fold higher titers. It should be noted that in both cases, the cultivation conditions and volumes differ significantly from, e.g., conditions used in this work. For *S. cerevisiae*, a cultivation volume of 3L was chosen. Cells were cultivated using fed-batch fermentation, for a total of 135h, supplementing excess glucose semi-continuously. On a biomass-specific level, the same strain achieving 539 mg·L^{-1} achieved 22.7 mg/DCW, which is close to modified *Synechocystis* from this work (19 mg/DCW). Similarly, *Y. lipolytica* reached more than a 5-fold higher titer compared to *Synechocystis*, but only close to 4 mg valencene per DCW, due to the high cell density reached. This highlights the need for high-volume photobioreactors tailored to the needs of cyanobacteria, in order to achieve similar biomass. It should also be noted that there is a trade-off between high biomass and light availability for the individual cell, which may lead to lower productivity at high biomass concentrations, due to light limitation and reduced photosynthetic activity. For this reason, alternative cyanobacterial strains with shorter doubling times and high light tolerance are being explored as more suitable biotechnological chassis[121,122].

Next to the modification of central carbon and terpenoid biosynthesis pathways also explored in varous cited literature (*Table 1*), a logical next step in strain engineering would be the production of the sesquiterpenoid nootkatone, an aromatic compound found in grapefruits which is an effective repellent or insecticide against mosquitoes. The main challenge in the synthesis of nootkatone so far has been the low catalytic efficiency of the membrane-anchored cytochrome P450/P450 reductase system[123]. In *C. nootkatensis,* the biosynthesis of nootkatone starts at FPP, which is converted to valencene by valencene synthase. This is followed by the regioselective oxidation of valencene to nootkatol by cytochrome P450 coupled with P450 reductase. Nootkatone is finally synthesized by the dehydrogenase-catalyzed oxidation of nootkatol. The pronounced membrane system of cyanobacteria *could* turn out to be particularly suitable for the membrane-anchored enzymes. In addition, *Synechocystis* contains not only endogenous CP450, for which oxygen and NADPH are readily available via photosynthetic activity, but also native alcohol hydrogenases, which are required for the final oxidation.

As mentioned before, there is a trade-off between cell density and light availability in cyanobacterial cultivation. An additional trade-off must be considered for producing strains – the trade-off between accumulation of biomass and desired product[88]. In this case, a two-phase cultivation system could be beneficial for cultivating the valencene-producing strain (*Figure 3*). It is ideal to grow the organism to the best possible cell density, taking into account factors like self-shading and gas exchange, and then inducing the valencene production cassette and reducing carotenoid production via *crtE*. Thus, in the growth phase, the organism has enough capacity to build up its carotenoid pool and in the production phase, in which growth is undesirable, the free precursor molecules of the terpene metabolism can be put into the production of valencene.

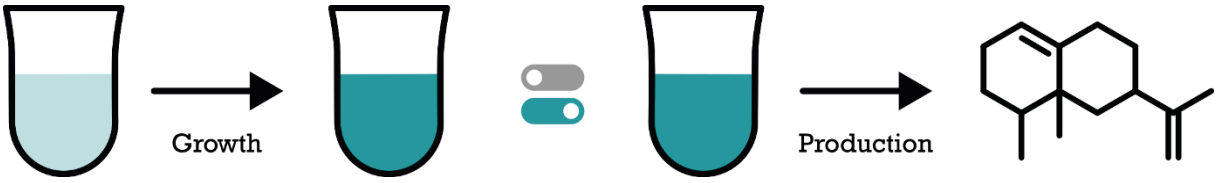


Figure 3: Scheme of a proposed switchable two-phase cultivation system. Left: Growth phase. Right: Production phase.

In this work it could be shown that by manipulating supercoiling, an increase in glycogen and ATP could be achieved. However, the simultaneous reduction of pigments that are required for photosynthesis entails a strong loss of viability. A quickly switchable system is required in which growth and production phases can be changed at short notice. The breakdown of the overproduced topoisomerase I by a protein degradation tag could be a possible solution here. In a finely tuned system, the next step would be to assess whether the additional energy carriers ATP and energy storage glycogen can be converted into a product before they are enriched. The formation of a fluorescent protein is a good way to quickly measure a positive effect on the protein level, while valencene production would confirm a positive effect on the metabolic level. The amount of reduction equivalents such as NADPH was not investigated in these strains. Since they are an important requirement for many production pathways, a redox balance is desirable. Computational models developed for cyanobacterial will further aid in understanding and engineering the metabolism of cyanobacterial chassis.

In a recent publication, a novel method for decoupling growth and production in *E. coli* was introduced by eliminating oriC from its genome[102]. While the current design relies on temperature elevation as the switch trigger, challenges may arise in large-scale fermentations due to inefficient distribution of equal temperature across the vessel. The operation is dependent on a serine recombinase sourced from bacteriophage phiC31. The expression of this recombinase is regulated by the temperature-sensitive cI857 repressor derived from phage lambda. In cells that undergo switching, the expression of the reporter protein persists even after growth has stopped, resulting in protein levels up to 5 times higher compared to non-switching cells. Alternative regulatory mechanisms, such as small molecules, could be explored to control serine recombinase expression, provided it remains tightly repressed under non-inducing conditions.
The findings underscore how cells can sustain metabolic activity and continue protein synthesis after stopping growth. In the context of bioproduction, the growth-decoupling system holds potential when combined with strategies that enhance product synthesis, such as the modulation of specific metabolic pathways, substrates, or cofactor availability. Instead of being a substitute, arresting growth emerges as a complementary approach to enhance product titers, yield, and overall productivity.
Historically, reprogramming microorganisms for novel functions has heavily relied on specialized molecular biology tools addressing genetic regulation. Recent advancements in pathway engineering, as exemplified by the FENIX device discussed in cited literature[124], explore a previously unexplored feature – the constitutive degradation of a target protein within a pathway. This degradation is triggered by a user-controlled, cost-effective inducer. The FENIX system not only finds application in metabolic engineering for biopolymer accumulation in recombinant *E. coli* strains, but also enables complex pathway engineering with external control over the production of multiple proteins in different metabolic network domains. Although the study focuses on intracellular PHB accumulation, a macromolecule also present in multiple cyanobacterial species[125], the system's versatility extends to increasing biosynthesis of extracellular products, particularly those challenging to control at the gene expression level.
Efficient bioproduction demands precise fine-tuning of cell growth and biosynthesis. A GCE-based orthogonal protein translation system, introduced in[126], into genetically recoded *E. coli*

and *B. subtilis* allows the incorporation of non-canonical amino acids. This system, combined with amber stop codon insertion, was successfully applied in proof-of-concept metabolic engineering for increased GlcNAc and NeuAc production in *E. coli* and *B. subtilis*, respectively. The combination of GCE and genetically recoded bacteria demonstrates potential for the precise regulation of cellular metabolism. The incorporation of the highly efficient ncAA OMeY tool, the first developed for *B. subtilis*, expands ncAA-incorporated protein production to this bacterium, leveraging its strong protein secretion capability. Given the pathway-independence and inactivity of the GCE-CGBBE strategy without naturally occurring ncAAs, it exhibits generic and robust features applicable to other crucial biochemical productions in engineered bacteria. Our strategy implies a global reprogramming approach resulting in a growth stop, metabolic rewiring (glycogen, ATP, pigments) and genetic reprogramming on a transcriptional level. It could be hypothesized that such a strain could be used for the production of valuable compounds by making use of a globally reprogrammed cell with reduced flux towards biomass. Further work is required to efficiently redirect the energy and carbon set free in the form of ATP and glycogen, while maintaining a sufficient level of cell viability.

Finally, the optimization of growth and production phases in photobioreactors may be possible by combining the two genetic engineering approaches – computer-aided metabolic engineering via CRISPRi technology and manipulation of DNA-topology - into a single strain and working towards a fully synthetic control over the endogenous DNA supercoiling equilibrium.

7 Chapter 7 References

7 References

1. Angermayr SA, Hellingwerf KJ, Lindblad P, Teixeira de Mattos MJ. Energy biotechnology with cyanobacteria. *Curr Opin Biotechnol*. 2009;20(3):257-263. doi:10.1016/j.copbio.2009.05.011

2. Heidorn T, Camsund D, Huang HH, et al. Synthetic Biology in Cyanobacteria. In: *Methods in Enzymology*. Vol 497. Elsevier; 2011:539-579. doi:10.1016/B978-0-12-385075-1.00024-X

3. Stal LJ. Nitrogen Fixation in Cyanobacteria. In: John Wiley & Sons, Ltd, ed. *eLS*. 1st ed. Wiley; 2015:1-9. doi:10.1002/9780470015902.a0021159.pub2

4. Lau NS, Matsui M, Abdullah AAA. Cyanobacteria: Photoautotrophic Microbial Factories for the Sustainable Synthesis of Industrial Products. *BioMed Res Int*. 2015;2015:1-9. doi:10.1155/2015/754934

5. Huang HH, Camsund D, Lindblad P, Heidorn T. Design and characterization of molecular tools for a Synthetic Biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Res*. 2010;38(8):2577-2593. doi:10.1093/nar/gkq164

6. Savakis P, Hellingwerf KJ. Engineering cyanobacteria for direct biofuel production from CO2. *Curr Opin Biotechnol*. 2015;33:8-14. doi:10.1016/j.copbio.2014.09.007

7. Ashour M, Wink M, Gershenzon J. Biochemistry of Terpenoids: Monoterpenes, Sesquiterpenes and Diterpenes. In: Wink M, ed. *Biochemistry of Plant Secondary Metabolism*. Wiley-Blackwell; 2010:258-303. doi:10.1002/9781444320503.ch5

8. Dismukes GC, Carrieri D, Bennette N, Ananyev GM, Posewitz MC. Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. *Curr Opin Biotechnol*. 2008;19(3):235-240. doi:10.1016/j.copbio.2008.05.007

9. Lagi M, Bar-Yam Y, Bertrand KZ, Bar-Yam Y. Accurate market price formation model with both supply-demand and trend-following for global food prices providing policy recommendations. *Proc Natl Acad Sci*. 2015;112(45). doi:10.1073/pnas.1413108112

10. Armstrong GA. Eubacteria show their true colors: genetics of carotenoid pigment biosynthesis from microbes to plants. *J Bacteriol*. 1994;176(16):4795-4802. doi:10.1128/jb.176.16.4795-4802.1994

11. Zhang L, Selão TT, Selstam E, Norling B. Subcellular Localization of Carotenoid Biosynthesis in Synechocystis sp. PCC 6803. *PloS One*. 2015;10(6):e0130904. doi:10.1371/journal.pone.0130904

12. Agger SA, Lopez-Gallego F, Hoye TR, Schmidt-Dannert C. Identification of Sesquiterpene Synthases from *Nostoc punctiforme* PCC 73102 and *Nostoc* sp. Strain PCC 7120. *J Bacteriol*. 2008;190(18):6084-6096. doi:10.1128/JB.00759-08

13. Jensen K, Møller BL. Plant NADPH-cytochrome P450 oxidoreductases. *Phytochemistry*. 2010;71(2-3):132-141. doi:10.1016/j.phytochem.2009.10.017

14. Lindberg P, Park S, Melis A. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using Synechocystis as the model organism. *Metab Eng*. 2010;12(1):70-79. doi:10.1016/j.ymben.2009.10.001

7 Chapter 7
References

15. Hoshino Y, Gaucher EA. On the Origin of Isoprenoid Biosynthesis. McInerney J, ed. *Mol Biol Evol*. 2018;35(9):2185-2197. doi:10.1093/molbev/msy120

16. Opitz S, Nes WD, Gershenzon J. Both methylerythritol phosphate and mevalonate pathways contribute to biosynthesis of each of the major isoprenoid classes in young cotton seedlings. *Phytochemistry*. 2014;98:110-119. doi:10.1016/j.phytochem.2013.11.010

17. Heath RJ, Rock CO. The Claisen condensation in biology. *Nat Prod Rep*. 2002;19(5):581-596. doi:10.1039/b110221b

18. Ajikumar PK, Tyo K, Carlsen S, Mucha O, Phon TH, Stephanopoulos G. Terpenoids: Opportunities for Biosynthesis of Natural Product Drugs Using Engineered Microorganisms. *Mol Pharm*. 2008;5(2):167-190. doi:10.1021/mp700151b

19. Miziorko HM. Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch Biochem Biophys*. 2011;505(2):131-143. doi:10.1016/j.abb.2010.09.028

20. Pérez-Gil J, Rodríguez-Concepción M. Metabolic plasticity for isoprenoid biosynthesis in bacteria. *Biochem J*. 2013;452(1):19-25. doi:10.1042/BJ20121899

21. Wang B, Eckert C, Maness PC, Yu J. A Genetic Toolbox for Modulating the Expression of Heterologous Genes in the Cyanobacterium *Synechocystis* sp. PCC 6803. *ACS Synth Biol*. 2018;7(1):276-286. doi:10.1021/acssynbio.7b00297

22. Gruchattka E, Hädicke O, Klamt S, Schütz V, Kayser O. In silico profiling of Escherichia coli and Saccharomyces cerevisiae as terpenoid factories. *Microb Cell Factories*. 2013;12(1):84. doi:10.1186/1475-2859-12-84

23. Lv X, Gu J, Wang F, et al. Combinatorial pathway optimization in *Escherichia coli* by directed co-evolution of rate-limiting enzymes and modular pathway engineering: Combinatorial Pathway Optimization in *Escherichia coli*. *Biotechnol Bioeng*. 2016;113(12):2661-2669. doi:10.1002/bit.26034

24. Pattanaik B, Lindberg P. Terpenoids and Their Biosynthesis in Cyanobacteria. *Life*. 2015;5(4):269-293. doi:10.3390/life5010269

25. Lin PC, Saha R, Zhang F, Pakrasi HB. Metabolic engineering of the pentose phosphate pathway for enhanced limonene production in the cyanobacterium Synechocysti s sp. PCC 6803. *Sci Rep*. 2017;7(1). doi:10.1038/s41598-017-17831-y

26. Melis A. Carbon partitioning in photosynthesis. *Curr Opin Chem Biol*. 2013;17(3):453-456. doi:10.1016/j.cbpa.2013.03.010

27. Englund E, Pattanaik B, Ubhayasekera SJK, Stensjö K, Bergquist J, Lindberg P. Production of Squalene in Synechocystis sp. PCC 6803. Neilan B, ed. *PLoS ONE*. 2014;9(3):e90270. doi:10.1371/journal.pone.0090270

28. Abe I. Enzymatic synthesis of cyclic triterpenes. *Nat Prod Rep*. 2007;24(6):1311-1331. doi:10.1039/b616857b

7 Chapter 7
References

29. Maeda H, Sakuragi Y, Bryant DA, DellaPenna D. Tocopherols Protect *Synechocystis* sp. Strain PCC 6803 from Lipid Peroxidation. *Plant Physiol*. 2005;138(3):1422-1435. doi:10.1104/pp.105.061135

30. Yang Y, Yin C, Li W, Xu X. α-Tocopherol Is Essential for Acquired Chill-Light Tolerance in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803. *J Bacteriol*. 2008;190(5):1554-1560. doi:10.1128/JB.01577-07

31. Collakova E, DellaPenna D. Isolation and Functional Analysis of Homogentisate Phytyltransferase from *Synechocystis* sp. PCC 6803 and Arabidopsis. *Plant Physiol*. 2001;127(3):1113-1124. doi:10.1104/pp.010421

32. Schledz M, Seidler A, Beyer P, Neuhaus G. A novel phytyltransferase from *Synechocystis* sp. PCC 6803 involved in tocopherol biosynthesis. *FEBS Lett*. 2001;499(1-2):15-20. doi:10.1016/S0014-5793(01)02508-X

33. Fatihi A, Latimer S, Schmollinger S, et al. A Dedicated Type II NADPH Dehydrogenase Performs the Penultimate Step in the Biosynthesis of Vitamin K1 in S *ynechocystis* and Arabidopsis. *Plant Cell*. 2015;27(6):1730-1741. doi:10.1105/tpc.15.00103

34. Vavilin D, Vermaas W. Continuous chlorophyll degradation accompanied by chlorophyllide and phytol reutilization for chlorophyll synthesis in Synechocystis sp. PCC 6803. *Biochim Biophys Acta BBA - Bioenerg*. 2007;1767(7):920-929. doi:10.1016/j.bbabio.2007.03.010

35. Gershenzon J, Dudareva N. The function of terpene natural products in the natural world. *Nat Chem Biol*. 2007;3(7):408-414. doi:10.1038/nchembio.2007.5

36. Langenheim JH. Higher plant terpenoids: A phytocentric overview of their ecological roles. *J Chem Ecol*. 1994;20(6):1223-1280. doi:10.1007/BF02059809

37. Jean-Claude Caissard. Secretion mechanisms of volatile organic compounds in specialized cells of aromatic plants. *Recent Res Dev Cell Biol*. 2004;(2):1-15.

38. Caruthers JM, Kang I, Rynkiewicz MJ, Cane DE, Christianson DW. Crystal Structure Determination of Aristolochene Synthase from the Blue Cheese Mold, Penicillium roqueforti*. *J Biol Chem*. 2000;275(33):25533-25539. doi:10.1074/jbc.M000433200

39. Ye L, Zhang C, Bi C, Li Q, Zhang X. Combinatory optimization of chromosomal integrated mevalonate pathway for β-carotene production in Escherichia coli. *Microb Cell Factories*. 2016;15(1):202. doi:10.1186/s12934-016-0607-3

40. Oku H, Mutanda I, Inafuku M. Molecular characteristics of isoprene synthase and its control effects on isoprene emissions from tropical trees. *J Plant Res*. Published online November 11, 2022. doi:10.1007/s10265-022-01418-4

41. Zhou F, Pichersky E. The complete functional characterisation of the terpene synthase family in tomato. *New Phytol*. 2020;226(5):1341-1360. doi:10.1111/nph.16431

42. Singsaas EL, Lerdau M, Winter K, Sharkey TD. Isoprene Increases Thermotolerance of Isoprene-Emitting Species. *Plant Physiol*. 1997;115(4):1413-1420. doi:10.1104/pp.115.4.1413

7 Chapter 7
References

43. Banthorpe DV, Charlwood BV, Francis MJO. Biosynthesis of monoterpenes. *Chem Rev*. 1972;72(2):115-155. doi:10.1021/cr60276a002

44. Schewe H, Mirata MA, Holtmann D, Schrader J. Biooxidation of monoterpenes with bacterial monooxygenases. *Process Biochem*. 2011;46(10):1885-1899. doi:10.1016/j.procbio.2011.06.010

45. Silva RC e, Costa JS da, Figueiredo RO de, et al. Monoterpenes and Sesquiterpenes of Essential Oils from Psidium Species and Their Biological Properties. *Molecules*. 2021;26(4):965. doi:10.3390/molecules26040965

46. Izaguirre G, Hwang CJ, Krasner SW, McGuire MJ. Geosmin and 2-Methylisoborneol from Cyanobacteria in Three Water Supply Systems. *Appl Environ Microbiol*. 1982;43(3):708-714. doi:10.1128/aem.43.3.708-714.1982

47. Zerbe P, Bohlmann J. Plant diterpene synthases: exploring modularity and metabolic diversity for bioengineering. *Trends Biotechnol*. 2015;33(7):419-428. doi:10.1016/j.tibtech.2015.04.006

48. Nakashima T, Inoue T, Oka A, Nishino T, Osumi T, Hata S. Cloning, expression, and characterization of cDNAs encoding Arabidopsis thaliana squalene synthase. *Proc Natl Acad Sci*. 1995;92(6):2328-2332. doi:10.1073/pnas.92.6.2328

49. Umeno D, Tobias AV, Arnold FH. Evolution of the C_{30} Carotenoid Synthase CrtM for Function in a C_{40} Pathway. *J Bacteriol*. 2002;184(23):6690-6699. doi:10.1128/JB.184.23.6690-6699.2002

50. Kirby J, Keasling JD. Biosynthesis of Plant Isoprenoids: Perspectives for Microbial Engineering. *Annu Rev Plant Biol*. 2009;60(1):335-355. doi:10.1146/annurev.arplant.043008.091955

51. Lin PC, Pakrasi HB. Engineering cyanobacteria for production of terpenoids. *Planta*. 2019;249(1):145-154. doi:10.1007/s00425-018-3047-y

52. Tracy NI, Chen D, Crunkleton DW, Price GL. Hydrogenated monoterpenes as diesel fuel additives. *Fuel*. 2009;88(11):2238-2240. doi:10.1016/j.fuel.2009.02.002

53. Chuck CJ, Donnelly J. The compatibility of potential bioderived fuels with Jet A-1 aviation kerosene. *Appl Energy*. 2014;118:83-91. doi:10.1016/j.apenergy.2013.12.019

54. Paddon CJ, Keasling JD. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat Rev Microbiol*. 2014;12(5):355-367. doi:10.1038/nrmicro3240

55. Tsuruta H, Paddon CJ, Eng D, et al. High-Level Production of Amorpha-4,11-Diene, a Precursor of the Antimalarial Agent Artemisinin, in Escherichia coli. Gregson A, ed. *PLoS ONE*. 2009;4(2):e4489. doi:10.1371/journal.pone.0004489

56. Solomon S, Plattner GK, Knutti R, Friedlingstein P. Irreversible climate change due to carbon dioxide emissions. *Proc Natl Acad Sci*. 2009;106(6):1704-1709. doi:10.1073/pnas.0812721106

7 Chapter 7
References

57. Megarity CF. Engineering enzyme catalysis: an inverse approach. *Biosci Rep*. 2019;39(2):BSR20181107. doi:10.1042/BSR20181107

58. Jones WP, Kinghorn AD. Extraction of Plant Secondary Metabolites. In: Sarker SD, Nahar L, eds. *Natural Products Isolation*. Vol 864. Methods in Molecular Biology. Humana Press; 2012:341-366. doi:10.1007/978-1-61779-624-1_13

59. Leonhardt RH, Berger RG. Nootkatone. In: Schrader J, Bohlmann J, eds. *Biotechnology of Isoprenoids*. Vol 148. Advances in Biochemical Engineering/Biotechnology. Springer International Publishing; 2014:391-404. doi:10.1007/10_2014_279

60. Ehrhardt DW, Frommer WB. New Technologies for 21st Century Plant Science. *Plant Cell*. 2012;24(2):374-394. doi:10.1105/tpc.111.093302

61. Sharkey TD, Yeh S. I SOPRENE E MISSION FROM P LANTS. *Annu Rev Plant Physiol Plant Mol Biol*. 2001;52(1):407-436. doi:10.1146/annurev.arplant.52.1.407

62. Melis A. Photosynthesis-to-fuels: from sunlight to hydrogen, isoprene, and botryococcene production. *Energy Env Sci*. 2012;5(2):5531-5539. doi:10.1039/C1EE02514G

63. Peralta-Yahya PP, Keasling JD. Advanced biofuel production in microbes. *Biotechnol J*. 2010;5(2):147-162. doi:10.1002/biot.200900220

64. Moser S, Pichler H. Identifying and engineering the ideal microbial terpenoid production host. *Appl Microbiol Biotechnol*. 2019;103(14):5501-5516. doi:10.1007/s00253-019-09892-y

65. Lassen LM, Nielsen AZ, Ziersen B, Gnanasekaran T, Møller BL, Jensen PE. Redirecting Photosynthetic Electron Flow into Light-Driven Synthesis of Alternative Products Including High-Value Bioactive Natural Compounds. *ACS Synth Biol*. 2014;3(1):1-12. doi:10.1021/sb400136f

66. Newman DM, Sake CL, Metcalf AJ, Davies FK, Boyle NR. *Characterizing Photosynthetic Biofuel Production: Isotopically Non-Stationary* ^{13} *C Metabolic Flux Analysis (INST-* ^{13} *CMFA) on Limonene Producing* Synechococcus *Sp. PCC 7002*. Bioengineering; 2022. doi:10.1101/2022.03.30.486112

67. Mills LA, McCormick AJ, Lea-Smith DJ. Current knowledge and recent advances in understanding metabolism of the model cyanobacterium *Synechocystis* sp. PCC 6803. *Biosci Rep*. 2020;40(4):BSR20193325. doi:10.1042/BSR20193325

68. Torrado A, Connabeer HM, Röttig A, et al. Directing cyanobacterial photosynthesis in a cytochrome *c* oxidase mutant using a heterologous electron sink. *Plant Physiol*. 2022;189(4):2554-2566. doi:10.1093/plphys/kiac203

69. Herold RA, Bryan SJ. *Engineered Terpenoid Production in* Synechocystis *Sp. PCC 6803 Under Different Growth Conditions*. Bioengineering; 2020. doi:10.1101/2020.08.19.256149

70. Sebesta J, Peebles CAM. Improving heterologous protein expression in Synechocystis sp. PCC 6803 for alpha-bisabolene production. *Metab Eng Commun*. 2020;10:e00117. doi:10.1016/j.mec.2019.e00117

7 Chapter 7
References

71. Hickman JW, Kotovic KM, Miller C, et al. Glycogen synthesis is a required component of the nitrogen stress response in Synechococcus elongatus PCC 7942. *Algal Res*. 2013;2(2):98-106. doi:10.1016/j.algal.2013.01.008

72. van der Woude AD, Angermayr SA, Puthan Veetil V, Osnato A, Hellingwerf KJ. Carbon sink removal: Increased photosynthetic production of lactic acid by Synechocystis sp. PCC6803 in a glycogen storage mutant. *J Biotechnol*. 2014;184:100-102. doi:10.1016/j.jbiotec.2014.04.029

73. Choi SY, Lee HJ, Choi J, et al. Photosynthetic conversion of CO2 to farnesyl diphosphate-derived phytochemicals (amorpha-4,11-diene and squalene) by engineered cyanobacteria. *Biotechnol Biofuels*. 2016;9(1). doi:10.1186/s13068-016-0617-8

74. Tóth TN, Chukhutsina V, Domonkos I, et al. Carotenoids are essential for the assembly of cyanobacterial photosynthetic complexes. *Biochim Biophys Acta BBA - Bioenerg*. 2015;1847(10):1153-1165. doi:10.1016/j.bbabio.2015.05.020

75. Antoniewicz MR. A guide to metabolic flux analysis in metabolic engineering: Methods, tools and applications. *Metab Eng*. 2021;63:2-12. doi:10.1016/j.ymben.2020.11.002

76. Hidalgo Martinez D, Betterle N, Melis A. Phycocyanin Fusion Constructs for Heterologous Protein Expression Accumulate as Functional Heterohexameric Complexes in Cyanobacteria. *ACS Synth Biol*. 2022;11(3):1152-1166. doi:10.1021/acssynbio.1c00449

77. Kang DH, Ko SC, Heo YB, Lee HJ, Woo HM. RoboMoClo: A Robotics-Assisted Modular Cloning Framework for Multiple Gene Assembly in Biofoundry. *ACS Synth Biol*. 2022;11(3):1336-1348. doi:10.1021/acssynbio.1c00628

78. Gao X, Gao F, Liu D, Zhang H, Nie X, Yang C. Engineering the methylerythritol phosphate pathway in cyanobacteria for photosynthetic isoprene production from CO_2. *Energy Environ Sci*. 2016;9(4):1400-1411. doi:10.1039/C5EE03102H

79. Lv X, Xue H, Qin L, Li C. Transporter Engineering in Microbial Cell Factory Boosts Biomanufacturing Capacity. *BioDesign Res*. 2022;2022:1-8. doi:10.34133/2022/9871087

80. Zhou J, Lin X, Liu S, et al. Effects of Compound Elicitors on the Biosynthesis of Triterpenoids and Activity of Defense Enzymes from Inonotus hispidus (Basidiomycetes). *Molecules*. 2022;27(9):2618. doi:10.3390/molecules27092618

81. Liu Y, Wang Z, Cui Z, Qi Q, Hou J. Progress and perspectives for microbial production of farnesene. *Bioresour Technol*. 2022;347:126682. doi:10.1016/j.biortech.2022.126682

82. Chen X, Huang C, Li K, et al. Recent advances in biosynthesis and pharmacology of β-elemene. *Phytochem Rev*. Published online August 18, 2022. doi:10.1007/s11101-022-09833-0

83. Cheng T, Wang L, Sun C, Xie C. Optimizing the downstream MVA pathway using a combination optimization strategy to increase lycopene yield in Escherichia coli. *Microb Cell Factories*. 2022;21(1):121. doi:10.1186/s12934-022-01843-z

7 Chapter 7
References

84. LI Z. Role of rate-limiting step of mevalonate pathway in improving lycopene production in Escherichia coli. *Editor Dep Chin J Bioeng*. 2020;Journal of Bioengineering(36(1)):77-89. doi:10.13345/j.cjb.190189

85. Noushahi HA, Khan AH, Noushahi UF, et al. Biosynthetic pathways of triterpenoids and strategies to improve their Biosynthetic Efficiency. *Plant Growth Regul*. 2022;97(3):439-454. doi:10.1007/s10725-022-00818-9

86. Meyer AJ, Segall-Shapiro TH, Glassey E, Zhang J, Voigt CA. Escherichia coli "Marionette" strains with 12 highly optimized small-molecule sensors. *Nat Chem Biol*. 2019;15(2):196-204. doi:10.1038/s41589-018-0168-3

87. Raman K, Chandra N. Flux balance analysis of biological systems: applications and challenges. *Brief Bioinform*. 2009;10(4):435-449. doi:10.1093/bib/bbp011

88. Knoop H, Steuer R. A Computational Analysis of Stoichiometric Constraints and Trade-Offs in Cyanobacterial Biofuel Production. *Front Bioeng Biotechnol*. 2015;3. doi:10.3389/fbioe.2015.00047

89. Chaves JE, Romero PR, Kirst H, Melis A. Role of isopentenyl-diphosphate isomerase in heterologous cyanobacterial (Synechocystis) isoprene production. *Photosynth Res*. 2016;130(1-3):517-527. doi:10.1007/s11120-016-0293-3

90. Park SY, Binkley RM, Kim WJ, Lee MH, Lee SY. Metabolic engineering of Escherichia coli for high-level astaxanthin production with high productivity. *Metab Eng*. 2018;49:105-115. doi:10.1016/j.ymben.2018.08.002

91. Yamada Y, Kuzuyama T, Komatsu M, et al. Terpene synthases are widely distributed in bacteria. *Proc Natl Acad Sci U S A*. 2015;112(3):857-862. doi:10.1073/pnas.1422108112

92. Kwak S, Crook N, Yoneda A, et al. Functional mining of novel terpene synthases from metagenomes. *Biotechnol Biofuels Bioprod*. 2022;15(1):104. doi:10.1186/s13068-022-02189-9

93. Greenhagen BT, O'Maille PE, Noel JP, Chappell J. Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. *Proc Natl Acad Sci*. 2006;103(26):9826-9831. doi:10.1073/pnas.0601605103

94. Goldsmith M, Tawfik DS. Directed enzyme evolution: beyond the low-hanging fruit. *Curr Opin Struct Biol*. 2012;22(4):406-412. doi:10.1016/j.sbi.2012.03.010

95. Lauchli R, Rabe KS, Kalbarczyk KZ, et al. High-Throughput Screening for Terpene-Synthase-Cyclization Activity and Directed Evolution of a Terpene Synthase. *Angew Chem Int Ed*. 2013;52(21):5571-5574. doi:10.1002/anie.201301362

96. Nielsen AZ, Ziersen B, Jensen K, et al. Redirecting Photosynthetic Reducing Power toward Bioactive Natural Product Synthesis. *ACS Synth Biol*. 2013;2(6):308-315. doi:10.1021/sb300128r

97. Falls KC, Williams AL, Bryksin AV, Matsumura I. Escherichia coli Deletion Mutants Illuminate Trade-Offs between Growth Rate and Flux through a Foreign Anabolic Pathway. Cirino PC, ed. *PLoS ONE*. 2014;9(2):e88159. doi:10.1371/journal.pone.0088159

7 Chapter 7
References

98. Burg JM, Cooper CB, Ye Z, Reed BR, Moreb EA, Lynch MD. Large-scale bioprocess competitiveness: the potential of dynamic metabolic control in two-stage fermentations. *Curr Opin Chem Eng*. 2016;14:121-136. doi:10.1016/j.coche.2016.09.008

99. Liu W, Zheng P, Yu F, Yang Q. A two-stage process for succinate production using genetically engineered Corynebacterium acetoacidophilum. *Process Biochem*. 2015;50(11):1692-1700. doi:10.1016/j.procbio.2015.07.017

100. Chubukov V, Sauer U. Environmental Dependence of Stationary-Phase Metabolism in Bacillus subtilis and Escherichia coli. Spormann AM, ed. *Appl Environ Microbiol*. 2014;80(9):2901-2909. doi:10.1128/AEM.00061-14

101. Chubukov V, Desmarais JJ, Wang G, et al. Engineering glucose metabolism of Escherichia coli under nitrogen starvation. *Npj Syst Biol Appl*. 2017;3(1):16035. doi:10.1038/npjsba.2016.35

102. Kasari M, Kasari V, Kärmas M, Jõers A. Decoupling Growth and Production by Removing the Origin of Replication from a Bacterial Chromosome. *ACS Synth Biol*. 2022;11(8):2610-2622. doi:10.1021/acssynbio.1c00618

103. Li S, Jendresen CB, Nielsen AT. Increasing production yield of tyrosine and mevalonate through inhibition of biomass formation. *Process Biochem*. 2016;51(12):1992-2000. doi:10.1016/j.procbio.2016.09.007

104. Baba T, Ara T, Hasegawa M, et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. 2006;2(1). doi:10.1038/msb4100050

105. Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*. 2013;152(5):1173-1183. doi:10.1016/j.cell.2013.02.022

106. Li S, Jendresen CB, Landberg J, et al. Genome-Wide CRISPRi-Based Identification of Targets for Decoupling Growth from Production. *ACS Synth Biol*. 2020;9(5):1030-1040. doi:10.1021/acssynbio.9b00143

107. Menzel R, Gellert M. Regulation of the genes for E. coli DNA gyrase: homeostatic control of DNA supercoiling. *Cell*. 1983;34(1):105-113.

108. Vijayan V, Zuzow R, O'Shea EK. Oscillations in supercoiling drive circadian gene expression in cyanobacteria. *Proc Natl Acad Sci*. 2009;106(52):22564-22568. doi:10.1073/pnas.0912673106

109. Ruhdal Jensen P, van der Weijden CC, Bogø Jensen L, Westerhoff HV, Snoep JL. Extensive regulation compromises the extent to which DNA gyrase controls DNA supercoiling and growth rate of *Escherichia coli*: Control by DNA gyrase. *Eur J Biochem*. 1999;266(3):865-877. doi:10.1046/j.1432-1327.1999.00921.x

110. Loeschcke A, Dienst D, Wewer V, et al. The photosynthetic bacteria Rhodobacter capsulatus and Synechocystis sp. PCC 6803 as new hosts for cyclic plant triterpene biosynthesis. *PLoS ONE*. 2017;12(12):e0189816. doi:10.1371/journal.pone.0189816

7 Chapter 7
References

111. Pattanaik B, Englund E, Nolte N, Lindberg P. Introduction of a green algal squalene synthase enhances squalene accumulation in a strain of Synechocystis sp. PCC 6803. *Metab Eng Commun*. 2020;10:e00125. doi:10.1016/j.mec.2020.e00125

112. Dietsch M, Behle A, Westhoff P, Axmann IM. Metabolic engineering of Synechocystis sp. PCC 6803 for the photoproduction of the sesquiterpene valencene. *Metab Eng Commun*. 2021;13:e00178. doi:10.1016/j.mec.2021.e00178

113. Matsudaira A, Hoshino Y, Uesaka K, Takatani N, Omata T, Usuda Y. Production of glutamate and stereospecific flavors, (S)-linalool and (+)-valencene, by Synechocystis sp. PCC6803. *J Biosci Bioeng*. 2020;130(5):464-470. doi:10.1016/j.jbiosc.2020.06.013

114. Choi SY, Lee HJ, Choi J, et al. Photosynthetic conversion of CO2 to farnesyl diphosphate-derived phytochemicals (amorpha-4,11-diene and squalene) by engineered cyanobacteria. *Biotechnol Biofuels*. 2016;9:202. doi:10.1186/s13068-016-0617-8

115. Xu W, Yao J, Liu L, et al. Improving squalene production by enhancing the NADPH/NADP+ ratio, modifying the isoprenoid-feeding module and blocking the menaquinone pathway in Escherichia coli. *Biotechnol Biofuels*. 2019;12(1):68. doi:10.1186/s13068-019-1415-x

116. Troost K, Loeschcke A, Hilgers F, et al. Engineered Rhodobacter capsulatus as a Phototrophic Platform Organism for the Synthesis of Plant Sesquiterpenoids. *Front Microbiol*. 2019;10:1998. doi:10.3389/fmicb.2019.01998

117. Frohwitter J, Heider SAE, Peters-Wendisch P, Beekwilder J, Wendisch VF. Production of the sesquiterpene (+)-valencene by metabolically engineered Corynebacterium glutamicum. *J Biotechnol*. 2014;191:205-213. doi:10.1016/j.jbiotec.2014.05.032

118. Chen H, Zhu C, Zhu M, et al. High production of valencene in Saccharomyces cerevisiae through metabolic engineering. *Microb Cell Factories*. 2019;18(1):195. doi:10.1186/s12934-019-1246-2

119. Arnesen JA, Kildegaard KR, Cernuda Pastor M, Jayachandran S, Kristensen M, Borodina I. Yarrowia lipolytica Strains Engineered for the Production of Terpenoids. *Front Bioeng Biotechnol*. 2020;8:945. doi:10.3389/fbioe.2020.00945

120. Lee J, Hilgers F, Loeschke A, Jaeger KE, Feldbrügge M. Ustilago maydis Serves as a Novel Production Host for the Synthesis of Plant and Fungal Sesquiterpenoids. *Front Microbiol*. 2020;11:1655. doi:10.3389/fmicb.2020.01655

121. Mills LA, Moreno-Cabezuelo JÁ, Włodarczyk A, et al. Development of a Biotechnology Platform for the Fast-Growing Cyanobacterium Synechococcus sp. PCC 11901. *Biomolecules*. 2022;12(7):872. doi:10.3390/biom12070872

122. Wendt KE, Walker P, Sengupta A, Ungerer J, Pakrasi HB. Engineering Natural Competence into the Fast-Growing Cyanobacterium Synechococcus elongatus Strain UTEX 2973. *Appl Environ Microbiol*. 2022;88(1):e0188221. doi:10.1128/AEM.01882-21

7 Chapter 7
References

123. Cha Y, Li W, Wu T, et al. Probing the Synergistic Ratio of P450/CPR To Improve (+)-Nootkatone Production in Saccharomyces cerevisiae. *J Agric Food Chem*. 2022;70(3):815-825. doi:10.1021/acs.jafc.1c07035

124. Durante-Rodríguez G, de Lorenzo V, Nikel PI. A Post-translational Metabolic Switch Enables Complete Decoupling of Bacterial Growth from Biopolymer Production in Engineered Escherichia coli. *ACS Synth Biol*. 2018;7(11):2686-2697. doi:10.1021/acssynbio.8b00345

125. Koch M, Berendzen KW, Forchhammer AK. On the Role and Production of Polyhydroxybutyrate (PHB) in the Cyanobacterium Synechocystis sp. PCC 6803. *Life Basel Switz*. 2020;10(4):47. doi:10.3390/life10040047

126. Tian R, Liu Y, Cao Y, et al. Titrating bacterial growth and chemical biosynthesis for efficient N-acetylglucosamine and N-acetylneuraminic acid bioproduction. *Nat Commun*. 2020;11(1):5078. doi:10.1038/s41467-020-18960-1

8 Chapter 8
Acknowledgements

Ich möchte mich bei Prof. Dr. Ilka Axmann für das in mich geschenkte Vertrauen und für die schöne Zeit bei ihr in der Arbeitsgruppe bedanken. Ilka, ich habe dich als eine faire Vorgesetzte kennenlernen dürfen, die mir in meiner wissenschaftlichen Ausbildung viele Freiheiten eingeräumt hat, jedoch immer mit gutem Rat zur Seite stand. Danke dafür. Bei Prof. Dr. Michael Feldbrügge bedanke ich mich dafür, dass er mich als Co-Supervisor unterstützt hat.

Ich bedanke mich explizit bei Rainer Machné für die spannende Zusammenarbeit und die daraus resultieren Ergebnisse. Rainer, dein Fachwissen sowie deine Vernarrtheit den Coil zu hacken, haben die Qualität meiner Arbeit maßgeblich gesteigert.

Während meiner Zeit als Promotionsstudent habe ich viel Unterstützung erhalten. Dabei möchte ich meinen Dank besonders an Dennis Dienst, der mich als Masterstudent betreute, mir aber darüber hinaus auch in späterem Verlauf nie eine meiner Fragen offenließ, aussprechen. Weiterhin möchte ich meinen besonderen Dank an Alice Pawlowski, Thomas Drepper, Sabine Metzger, Vera Wever, Philipp Westhoff, Oliver Klaus, Paul Hudson, Jungho Lee, Arsenij Kokorin und Oliver Klaus aussprechen.

Ich hatte das Vergnügen während meiner Zeit als Promotionsstudent Bachelor und Masterstudierende bei ihrer Abschlussarbeit zu betreuen. Ich bedanke mich bei Monika, Sebastian, Daniel, Kathi, Kiku und Andreas. Ich hoffe das ich euch einiges mitgeben konnte, mir habt ihr auf jeden Fall viel zurückgegeben.

Die Zeit wäre ohne Nic, Lutz, Nik, Kai und Kai und vielen anderen aus dem SynMibi Labor nicht die gleiche gewesen. Vielmehr war es eine wunderschöne Zeit mit euch zusammen!

Ich bedanke mich zum Schluss bei meiner Familie, im Besonderem bei meiner Mutter und Oma, die doch regelmäßig den Status meiner Arbeit abgefragt haben.

Ein ganz besonderer Dank geht an meine liebe Freundin, die Anna. Du warst mir eine gute Partnerin in der Wissenschaft und bist mir eine im Leben.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass die eingereichte Dissertation selbständig und ohne unzulässige fremde Hilfe unter Beachtung der ''Ordnung über die Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf'' verfasst worden ist. Weiter versichere ich andere als die in der Dissertation angegebene Literatur nicht benutzt zu haben und dass ich alle ganz oder annähernd übernommenen Textstellen sowie verwendete Grafiken, Tabellen und Auswertungsprogramme kenntlich gemacht habe. Außerdem versichere ich, dass die vorgelegte elektronische mit der schriftlichen Version der Dissertation übereinstimmt und die Abhandlung in dieser oder ähnlicher Form noch nicht anderweitig als Promotionsleistung vorgelegt und bewertet wurde.

Declaration of academic honesty

I hereby declare an oath that the submitted dissertation has been written independently and without undue assistance, considering the ''Rules and Principles for Safeguarding and Good Scientific Practice at HHU''. Further, I assure that I did not use other references than the ones specified in the dissertation and that I have highlighted complete or partially adopted texts as well as graphics, tables, and analysis programs used in this work. Furthermore, I assure that the submitted electronic version is in accordance with the written version of the dissertation and that the dissertation in this or similar form has not yet been presented and evaluated as a doctoral thesis.

_____