DirectFuel: Final publishable summary report

Revised 18.3.2015

Grant Agreement number: 256808

Project acronym: DirectFuel

Project title: Direct biological conversion of solar energy to volatile hydrocarbon fuels by engineered cyanobacteria (DirectFuel)

Funding Scheme: Collaborative STREP project EC grant; THEME [ENERGY.2010.3.5-1]

Period covered: from 1.10.2010 to 30.09.2014

Scientific representative of the project's coordinator:
Professor Eva-Mari Aro (Period 3)

Organisation:
Molecular Plant Biology, Department of Biochemistry, University of Turku, FINLAND

Tel: +358 2 333 5931
Fax: +358 2 333 8075
E-mail: evaaro@utu.fi
Project website address: www.directfuel.fi

Executive summary

There is an urgent need to find novel solutions for the production of fuels and chemicals to replace the use of fossil oil and gas resources, challenging our globe with greenhouse gas emissions. An obvious solution to obtain these compounds in demand would be photosynthesis; the conversion of CO_2 with solar energy to plant biomass providing a sustainable way for renewable resources. However, cultivation of plants needs arable land, and should be reserved for the production of food, in ever increasing demand on our globe. An alternative is the photosynthetic micro-organisms, such as cyanobacteria, commonly also known as blue green algae, which you could grow in large reactors with likewise provision of only light and CO_2. The present project was addressing this option. In spite of the fact that these kind of biotechnological solutions are today hard to prove economic, compared to production processes based on fossil resources, it is incremental to commence the research and development now, to have the solutions ready for the future.

The objective of the DirectFuel project was to develop cyanobacteria that would be able to directly convert solar energy and CO_2 into useful energy carrier chemicals. The focus was on hydrocarbon end-products, which could serve as direct drop-in replacements for non-renewable petroleum-derived fuels currently in use. Propane, a volatile short-chain alkane naturally not produced by microbes, was selected as the ultimate target due to its favorable physicochemical properties and existing global market and infrastructure for transport and utilization.

The foundation of the DirectFuel project was on enzyme discovery and molecular level characterization, which from the beginning aimed at designing biosynthetic pathways for the production of propane in cyanobacterial hosts. The collective efforts culminated in the introduction of consecutive enzyme-catalysed steps responsible for converting the native carbon compound intermediates from photosynthesis towards the target product, propane. A major share of the research efforts was allocated to the study of the very last and still uncharacterized key enzymatic step, formation of the alkane product. This involved detailed *in vitro* enzymatic studies, structure-based enzyme modification, and extensive evaluation of the enzyme(s) in biosynthetic contexts *in vivo* in *E. coli*. In parallel, systematic studies of the cyanobacterial host were carried out in order to evaluate and understand the consequences of the introduced pathways on the primary metabolism, viability and growth. This was supported by extensive *in silico* modelling of the cyanobacterial metabolism, which was developed throughout the project to identify gene targets for engineering and to predict the feasibility of the production processes. Novel photobioreactor systems were developed for the production, evaluation and quantitation for volatile target products, serving as the first stage towards up-scaled industrial processes. At the same time, theoretical assessments on the feasibility of different production strategies were carried out, in order to determine the most prominent approaches for the future.

The DirectFuel consortium made a significant collective achievement in generating fundamental understanding, biotechnological tools and know-how towards the direct production of carbon-based fuels from CO_2 and water using photosynthetic cyanobacterial hosts. The microbial biosynthesis of propane was shown to be possible for the very first time on an *E. coli* model platform, and the metabolic bottlenecks of transferring the system into cyanobacteria were carefully evaluated one biosynthetic step at a time. Altogether, the substantial novel information and knowledge generated in the DirectFuel project will pave the way for future sustainable solutions to meet the global demands for renewable fuels and chemicals.

A summary description of project context and objectives

The DirectFuel project context

The complementation and subsequent replacement of fossil oil and gas with renewable production of chemicals and fuels is a major global challenge with consequences for greenhouse gas emissions, energy independence, food supply and finances. Solar energy and CO_2 will need to be used as the source of renewable substrate in order to achieve a long-term sustainable solution. Although plant photosynthesis is able to generate renewable energy, it has become increasingly clear that biomass production is not a sustainable solution since agricultural land should be reserved as far as possible for the production of food. Instead of terrestrial plants, it is theoretically possible to use liquid systems for cultivation of photosynthetic microorganisms that likewise convert solar to chemical energy, and this biotechnology does not need land for operation. Although this sounds ideal, our ability to use these photobiological catalysts for biotechnological production is at present limited. A particular constraint is the need to use microbial species that are sufficiently robust to dominate large-scale biotechnological environments. By genetically engineering suitable photobiological strains, however, we can expand their photobiocatalytic capabilities and make metabolic short cuts towards desired target products far before biomass accumulation.

In the present project, the emphasis was placed on low market value chemicals such as alkanes and alkenes that currently derive from fossil sources. The major bottleneck for enabling renewable production of such chemicals is the challenge of economically competing with traditional fossil fuel industries. In order to commercialize photobiotechnology it is therefore important to first create production systems that also generate higher value products in order to reach above the threshold of economic sustainability[1]. Once financially sustainable production of the combined products has been achieved, it will become possible to progressively increase the distribution of low to high market value products through continually improving the economic efficiency of the entire system. As a result of this process, the scale of demand and production are likely to rise, and only then is it likely that this new technology will actually have a measureable impact on bioeconomy.

The project objective

The objective with the DirectFuel project was to use the "solar biofuel" concept and develop new understanding and technologies that would enable future production of low market value fuel and chemicals using cyanobacteria as a photobiological catalyst and solar energy, water and CO_2 as substrates. As noted above, a major issue hindering commercial realization of such technology is likely to be the excessive cost of production relative to the retail value of the product. R&D projects that can identify technical solutions to enhance the economic sustainability are therefore important, if not essential in order to actually realize such a process. Looking into detail at the dominating proposed solutions at that time (2009), algal biodiesel and H_2, obvious gaps were noted.

For example, algal biodiesel production requires that the biocatalyst is harvested and destroyed in order to collect the precursor for further chemical conversion. Intuitively, this appears to be a wasteful method given that there is an implicit resource cost associated with the amplification of the catalyst in the first place, and it also precludes the continuous use of the photosynthetic cells as catalysts, in addition to loss of product associated with chemical processing. Several benefits can be anticipated with an alternative approach. Firstly, if the fuel can be produced in a ready-to-use state, greater overall efficiency can be expected from a reduced requirement for further processing. Secondly, if the fuel can be excreted from the organism in which it was synthesised, the catalyst can continue to be used, rather than having to be destroyed in order to access internalized product. We can also imagine other follow-on benefits from this, including a thermodynamic effect and enhanced tolerance (some fuels are indeed highly toxic), both as a consequence of product removal.

Some photobiological organisms already produce such a fuel. For example, H_2 is a ready-to-use fuel that readily separates from the production process. It suffers from other downstream issues, however, as it is highly challenging to store H_2 at a high energy density, an issue also shared with electricity. Nevertheless, the previous EU SolarH and SolarH$_2$ projects provided an interesting start point for a carbon-based solar biofuel project. In the DirectFuel project, something similar to H_2, in terms of utility and process separation, was sought, yet with greater compatibility with the existing infrastructure for utilization. *Propane* was proposed as a key target given that it straddles the interface between gas and liquid; Gaseous at atmospheric conditions, thereby allowing facile product-process separation, yet requiring only a modest amount of energy to liquefy in comparison to both H_2 and methane. A central issue was that no pathway for the biosynthesis of propane was yet known. Hence, a primary target in the project was to first create a metabolic pathway for propane biosynthesis using synthetic biology. In order to enable such pathways to be constructed, however, fundamental knowledge was first needed regarding the key reaction steps of the pathway, the conversion of aldehyde to alkane. At the time of application, no such enzyme had been identified. Fortuitously, just prior to the start of the project, the first such enzyme was described, a breakthrough discovery by the US company LS9 Inc[2]. Despite the unexpected availability of the "missing" key enzyme, it rapidly became obvious that substantial more research, both *in vitro* and in *E. coli,* was needed in order to both understand (WP1) and improve (WP2) this novel enzyme.

As engineered photobiotechnology was not yet an existing commercial process, as described above, it was also decided to in parallel with these above synthetic biology tasks to establish a comprehensive program to deliver fundamental understanding and tools for this novel biotechnology. The focus was placed on the most well-studied cyanobacterium known, *Synechocystis sp.* PCC 6803. As we planned to introduce a new metabolic pathway for propane biosynthesis (WP4), the availability of primary "solar fuel" substrates i.e. the fixed CO_2 and excited electrons was likewise addressed in WP4, and strong emphasis was placed on characterizing (WP5) and mapping out (WP3) the propane metabolism in *Synechocystis sp.* PCC 6803 cells. Finally, in order to understand how far our developed technology was from industrial commercialization we engaged two companies to engineer bioreactor systems for volatile products and to quantify the economic and environmental parameters of proposed complete production systems (WP6).

1. Stephens, E. et al. An economic and technical evaluation of microalgal biofuels. *Nat Biotechnol* 28, 126-128 (2010).
2. Schirmer, A., Rude, M., Li, X., Popova, E. & del Cardayre, S. Microbial biosynthesis of alkanes. *Science* 329, 559-562 (2010).

Main S&T results and foregrounds

Contents

The DirectFuel project context 3

The project objective 3

1. The concept of DirectFuel 6

1.1. The Consortium 6

1.2. Ultimate aims and objectives 6

1.3. Scope of the research 6

2. Pathway Enzyme Discovery (WP1) 6

2.1. Background and research objectives in WP1 6

2.2. Initial characterization of cADO reaction and kinetics 7

2.3. Mechanistic insights from radical clock substrates 7

2.4. Solvent isotope effects 8

2.5. cADO structural studies 8

3. Enzyme engineering (WP2) 9

3.1. Background and research objectives in WP2 9

3.2. Phenylacrylic acid decarboxylase/ferulic acid decarboxylase (PAD/FDC1) 9

3.3. Aldehyde deformylating oxygenase (cADO) 10

4. Computational Biology and Metabolic Modeling (WP3) 11

4.1. Background and research objectives in WP3 11

4.2. The metabolic network of *Synechocystis sp.* PCC 6803 12

4.3. Computational pathway analysis 12

5. Metabolic engineering (WP4) 13

5.1. Background and research objectives in WP4 13

5.2. Plasmid-based system for cyanobacterial ethylene production 13

5.3. Coenzyme A dependent pathway for propane production in *Synechocystis* 14

5.4. Enhanced production of short chain length fatty acids in cyanobacteria as precursors of alkane biosynthesis 15

5.5. Enhancement CO_2 fixation and reallocation of carbon 16

5.6. Tailoring the electron flow to target fuel molecules 16

5.7. Product toxicity to cyanobacterial hosts 17

5.8. Propane biosynthesis in *E. coli* 17

6. Bottleneck analysis (WP5) 19

6.1. Background and research objectives in WP5 19

6.2. *Synechocystis* sp. PCC6803 as a model organism in DirectFuel 19

6.3. Distribution of alkane biosynthesis genes in cyanobacteria and their expression 20

6.4. Transcriptomic and microarray technologies within DirectFuel 21

6.5. Establishment and application of a high efficiency transposon-mediated differential hybridisation approach 21

7. Process considerations and photobioreactor technology (WP6) 23

7.1. Background and research objectives in WP6 23

7.2. Life cycle analysis for the production of ethylene 23

7.3 Literature survey of photobioreactor designs 24

7.4. Construction of a photobioreactor optimized for autotrophic production of volatile hydrocarbon fuel 24

1. The concept of DirectFuel

1.1. The Consortium

The European Union Seventh Framework Programme (FP7-ENERGY-2010-1) project DirectFuel [grant agreement no. 256808] is an international collaborative project carried out over four years between 1.10.2010 - 30.09.2014. The DirectFuel consortium consists of seven universities/ research institutes and two industrial partners (Turun Yliopisto, Finland; Valtion Teknillinen Tutkimuskeskus/ VTT, Finland; Albert-Ludwigs-Universität Freiburg, Germany; Humboldt-Universität zu Berlin, Germany; The University of Manchester, United Kingdom; Regents of the University of Michigan, United States; Københavns Universitet, Denmark; Biochemtex S.p.a., Italy; Photon Systems Instruments/ PSI, Czech Republic).

1.2. Ultimate aims and objectives

The objective of the project was to develop photosynthetic microorganisms (cyanobacteria) that catalyze direct conversion of solar energy and carbon dioxide to engine-ready fuels. The main goal was to design and evaluate platforms which allow microbial biosynthesis of engine-ready transport fuels which would be easily extractable, and with minimal modication, compatible with the current infrastructures for storage and utilization. The main focus of the project was on volatile short-chain hydrocarbons such as ethylene and propane.

1.3. Scope of the research

The research in the DirectFuel project was divided between the consortium partners in Work Packages which spanned from fundamental research towards evaluation and practical implementation. The core of the research centered on the enzymology of dedicated microbial alkane pathways, and included molecular-level kinetic characterization and structure-based enzyme modification. This was linked with the design and assembly of artificial biosynthetic pathways and metabolic engineering to produce the desired end-metabolites in microbial hosts. The pathways were evaluated using computational biology, involving the design of metabolic models which help to identify and eliminate limiting steps and to find targets for further modification. Process evaluation and design of photobioreactor technology were crucial for advancing towards up-scaled production systems. The key results and findings of the DirectFuel project are shortly described in the following sections (2 – 7) in context with the associated Work Packages.

2. Pathway Enzyme Discovery (WP1)

2.1. Background and research objectives in WP1

The main thrust of WP1 (led by Regents of the University of Michigan, United States) has been the discovery and characterization of new enzymes that may be useful for the production of short chain alkanes such as propane. The focus of the studies has been on characterizing a newly-discovered enzyme

involved in alkane biosynthesis, *cyanobacterial aldehyde deformylating oxygenase* (cADO; a.k.a. *aldehyde decarbonylase* cAD) which catalyzes the deformylation of long-chain aldehydes derived from fatty acids to alkanes. In the context of the DirectFuel project, the objective of our investigations has been to understand the kinetics and mechanism of this enzyme in sufficient detail to allow rational engineering of substrate range (WP2) and incorporation of the enzyme into photosynthetic organisms (WP4) to facilitate photosynthesis-driven alkane production.

2.2. Initial characterization of cADO reaction and kinetics

The C-1 bi-product of the cADO reaction was predicted to be either CO or CO_2 based on work on animal and plant decarbonylases. We showed that, in fact, the aldehyde carbon was converted to formate and provided the first evidence that the reaction is iron dependent. We also established that the aldehyde hydrogen is retained in formate and that the proton in the alkane derives from solvent (findings independently demonstrated by other groups).

We have delineated the kinetics and substrate range of cADO, including the requirements for an external reducing system. Most non-heme iron oxygenases require an auxiliary reducing system to reduce the Fe(III) resting state to Fe(II) and initiate the catalytic cycle. The physiological reducing system for cADO is unknown but presumed to be ferredoxin. We demonstrated that ferredoxins, including cyanobacterial ferredoxins, while supporting the reaction, give extremely slow rates of turn-over. We investigated a variety of chemical reducing systems and found the most effective to be PMS (as mediator) combined with NADH (as reductant): this system increased turn-over by over 100-fold from initially reported values. Nevertheless, the maximal turn-over rates achieved under steady state conditions are only ~ 1 min^{-1}; despite much effort by various labs no substantive improvements in the kinetics have been achieved. We investigated the substrate range of cADO and showed it was active on straight-chain aldehydes ranging from 5 to 18 carbons. In collaboration with group UMA we re-engineered the substrate-binding channel to increase the enzyme's selectivity for short-chain aldehydes. Through 2 mutations, V41Y and A134F, we were able to increase selectivity for aldehydes shorter than 6 carbons by 3 – 6-fold while selecting against C-14 and longer aldehydes by 5 – 10-fold over wild-type.

2.3. Mechanistic insights from radical clock substrates

We have obtained valuable insights into the C1-C2 bond scission step from the reaction of an octadecanal analog that incorporates a strategically placed cyclopropyl group between C3 and C4 that can act as a "radical clock". When reacted with cADO, this substrate produced only 1-octadecene as the rearranged product, providing evidence for a radical mechanism for C-C bond scission. From the known rate constant for ring opening of cyclopropylcarbinyl radicals, the lifetime of the radical generated during C-C scission was estimated to be ≥ 10 ns. Unexpectedly, we identified an alternate pathway in which the cyclopropyl aldehyde acts as a mechanism-based irreversible inhibitor of cADO. LC-ES-MS of inactivated cAD demonstrated the formation of a covalent adduct between deformylated aldehyde and cADO. Proteomic analysis identified the site of modification as a Phe residue lining the substrate-binding channel. Thus a plausible mechanism for inactivation is that the alkyl radical resulting from ring-opening reacts with the Phe side-chain to covalently modify the substrate-binding site and inactivate the enzyme. Studies are in progress to evaluate the effects of mutating this residue on the inactivation pathway.

Further mechanistic insights followed the discovery that cADO catalyzes the deformylation of α-oxyranyl aldehydes to produce the corresponding oxiranes. Thus by performing the reaction in D_2O we could determine the facial selectivity of proton addition by ^{1}H-NMR spectroscopy. We found that the proton is delivered with equal probability to either face of the oxirane ring, indicating the formation of an oxiranyl radical intermediate that is free to rotate during the reaction. Unexpectedly, the enzyme also catalyzes a side reaction in which oxiranyl-aldehydes undergo tandem deformylation to produce alkanes two carbons shorter. This involves the rearrangement of the intermediate oxiranyl radical to an φ-carbonyl radical, leading to an aldehyde one carbon shorter that is deformylated in a second step. The rate constants for oxiranyl radical rearrangements are known, so these compounds could be used as "slow" radical clocks allowing the lifetime for the radical on the enzyme to be estimated as 10^{-4} s.

2.4. Solvent isotope effects

To gain insight into the proton donor to the alkane, we have measured solvent isotope effects on the incorporation of deuterium into the product, which suggest that metal-bound water (or hydroxide) is the immediate proton donor. The solvent kinetic isotope effect on V_{max} is 1.0 +/- 0.05, i.e. the *rate* of reaction is independent of the mole fraction, ? $_{(D2O)}$. However, careful measurements of the deuterium content of the product (determined by MS) show that protium transfer is favored over deuterium by a factor of 2.18 +/- 0.02 independent of ? $_{(D2O)}$. This is unlikely to be a kinetic isotope effect as other evidence points to C-C bond scission being irreversible. This observation is best explained as arising from a reactant state equilibrium isotope effect in which the proton is donated from a site with a low fractionation factor, α = 0.43, resulting in protium being enriched over deuterium. The most likely candidate for such a low fractionating site is a water molecule bound to iron, which has been identified in the crystal structure.

2.5. cADO structural studies

Previously structures of cADO with fatty acids bound in the active site had been reported. However we have recently obtained structures of cADO with a long-chain aldehyde bound. These structures provide insight into how both the substrate and the metal cofactor gain access to the active site, which is deeply buried. To accomplish this we synthesized an undecanal analog that incorporates 3 ethylene glycol units at its terminus; this makes the compound far more water soluble (the compound displays similar activity to octadecanal), allowing it to be included in the crystallization buffer. Most interestingly, the protein crystallized as a mixture of two structures that we interpret as metal-free and metal-bound structures. The structures provide important insights into the pathway by which the required metal ions enter the active site; they also identify a new substrate binding channel and a pathway for water and oxygen to enter/leave the active site.

References for WP1

Insights into Substrate and Metal Binding from the Crystal Structure of Cyanobacterial Aldehyde Deformylating Oxygenase with Substrate Bound Benjamin C. Buer, Bishwajit Paul, Debasis Das, Jeanne A. Stuckey, and E. Neil G. Marsh (2014) *ACS Chem. Biol.*, Articles ASAP

Solvent Isotope Effects on Alkane Formation by Cyanobacterial Aldehyde Deformylating Oxygenase and Their Mechanistic Implications Matthew W. Waugh and E. Neil G. Marsh (2014) *Biochemistry*, *53*, pp 5537–5543

Mechanistic Insights from Reaction of α-Oxiranyl-Aldehydes with Cyanobacterial Aldehyde Deformylating Oxygenase Debasis Das, Benjamin Ellington, Bishwajit Paul, and E. Neil G. Marsh (2014) *ACS Chem. Biol.*, *9* , pp 570–577

Aldehyde Decarbonylases: Enigmatic Enzymes of Hydrocarbon Biosynthesis E. Neil G. Marsh and Matthew W. Waugh (2013) *ACS Catal.*, *3* , pp 2515–2521

Probing the Mechanism of Cyanobacterial Aldehyde Decarbonylase Using a Cyclopropyl Aldehyde Bishwajit Paul, Debasis Das, Benjamin Ellington, and E. Neil G. Marsh (2014) *J. Am. Chem. Soc.*, *135* , pp 5234–5237

Aldehyde-forming fatty acyl-CoA reductase from cyanobacteria: expression, purification and characterization of the recombinant enzyme. Fengmin Lin, Debasis Das, Xiaoxia Lin, E. Neil G. Marsh (2013) FEBS J.,280 pp 4773-81

Production of propane and other short-chain alkanes by structure-based engineering of ligand specificity in aldehyde-deformylating oxygenase. Khara B, Menon N, Levy C, Mansell D, Das D, Marsh ENG, Leys D, Scrutton NS. (2013) ChemBioChem. 14 pp 1204-8

Alkane Formation by Non-Heme Iron-Dependent Cyanobacterial Aldehyde Decarbonylase: Investigation of Kinetics and Requirement for an External Electron Donor Bekir E. Eser, Debasis Das, Jaehong Han, Patrik R. Jones, and E. Neil G. Marsh (2011) *Biochemistry*, *50* pp 10743–10750

Decarbonylation of aldehydes by cyanobacterial aldehyde decarbonylase: a new reaction of diiron enzymes. Das D, Eser BE, Han J, Sciore A, Marsh ENG. (2013) Angew Chem 50 pp 7148-52

3. Enzyme engineering (WP2)

3.1. Background and research objectives in WP2

The overall objective of the DirectFuel project was to develop a novel photobiological system that catalyzes direct conversion of solar energy and carbon dioxide to engine-ready fuel such as propane. To achieve this goal, the aim in Work Package 2 (led by The University of Manchester, United Kingdom) was to redefine the specificity of enzymes involved in alkane/alkene generation to enable their use in metabolic engineering in Work Package 3. Towards this objective, the research focused on two different enzymatic systems (i) phenylacrylic acid decarboxylase/ferulic acid decarboxylase (PAD/FDC1) and (ii) aldehyde deformylating oxygenase (cADO) system in an attempt to engineer and improve the enzyme catalytic efficiencies with target substrates to enable efficient processing in *de novo* engineered biosynthetic pathway.

3.2. Phenylacrylic acid decarboxylase/ferulic acid decarboxylase (PAD/FDC1)

For the decarboxylation system, we have selected PAD1/FDC1 from *Saccharomyces cerevisiae* as it naturally decarboxylates phenylpropanoic acid to the corresponding vinyl product. However, the wild-type enzyme of PAD1/FDC1 does not have detectable activity with short chain linear carboxylic acids (e.g. crotonic acid, required for propene production). As the target substrate C4 crotonic acid is not naturally accepted by PAD1/FDC1, we have performed a directed evolution approach on the PAD1/FDC1 decarboxylation system in an attempt to alter the substrate specificity to enable production of short chain alkenes.

The PAD1/FDC1 genes do not have significant homology with other enzyme families for which structural information is available. This lack of inferred structural similarity with any known structurally determined

proteins makes rational design by directed mutagenesis a difficult task. Consequently, a random mutagenesis strategy was employed in which mutations were introduced throughout the whole FDC1 gene via error prone PCR. Mutation libraries were assayed using an agar plate format supplemented with pH indicators (phenol red), thus allowing identification of active decarboxylase variants from colony screens by formation of a red coloured halo caused by an increase in pH attributed to local alkalinification. Screening results showed that the substrate specificity for the wild-type PAD1/FDC1 system is relatively stringent; PAD1/FDC1 solely accepted substrates with a six-carbon chain length and two double-conjugated bonds. In principle, one could alter the substrate specificity of this enzyme using a random mutagenesis approach. However, screening of all possible variants is a time-intensive method, and given the exceptionally large sequence space involved in screening it became apparent that this is not an effective route forward. For that reason, we focused our attention on a second enzymatic system, cADO from *Prochlorococcus marinus* strain MIT9313, which has remained the primary focus of the DirectFuel project. The major difference between the two classes of enzyme (PAD1/FDC versus cADO) is that cADO catalyses the conversion of aldehydes to alkanes, whereas PAD1/FDC1 converts carboxylic acid to alkenes, indicating that cADO could be a useful catalyst for the generation of short chain alkanes.

3.3. Aldehyde deformylating oxygenase (cADO)

A variety of aldehyde decarbonylases derived from a wide range of organisms such as green alga *B. braunii*, *P. marinus* MIT9313, *N. punctiformes* PCC73102, *Synechococcus sp.* RS9917, and *Synechocystis sp.* PCC6803 have been explored and characterised in this DirectFuel project (Work package 1). Among all these decarbonylases, several studies confirmed that cADO from *P. marinus* was an attractive candidate, which was found to be capable of catalysing the conversion of variety aldehydes at the C-1′ position into corresponding alkanes with the release of formate in the fatty acid biosynthetic pathway of cyanobacteria. However, kinetic results from UNIMAN showed that the catalytic turnover for the conversion of short chain aldehyde, such as butanal to propane by cADO, is very low.

We have employed two strategies in an attempt to overcome the low turnover number and restricted specificity range for cADO. In Work Package 2 we were able to determine the nature of the active site ligand in cADO and also improve the structural analysis of the protein by X-ray crystallography. The availability of structural information then enabled us to use structural-based design/directed mutagenesis targeted at residues in the active site channel to alter the substrate specificity toward that of short chain aldehydes. Based on available X-ray crystallographic structure of cADO we designed and isolated two variant forms of cADO (Val41Tyr and Ala134Phe) that were shown to have improved specificity towards short chain aldehyde substrates. In particular, we have found that Ala134Phe showed approximately a four-fold improved catalytic efficiency with the target butyraldehyde substrate together with ten-fold increase in specificity for octadecanal substrate. The discovery of the Ala134Phe variant was an important milestone for the DirectFuel project. For the first time, we have found a variant cADO, which showed a high degree of substrate specificity and increased propane production compared to WT cADO activity. Other members of the DirectFuel consortium for use in cyanobacteria have since used this variant in metabolic engineering programmes. We have also used this variant form of cADO in metabolic engineering studies with *E. coli* in which propane generation has been demonstrated both in cell-based biotransformation reactions and also following the construction of a de novo butanol production pathway which was intercepted by inclusion of the cADO variant to convert butanal to propane.

In order to identify second and third generation cADO variants with improved activity it has been necessary to develop ultra-high throughput screens that enable us to screen large mutant libraries in which residues outside the immediate active site have been targeted. With this in mind, we have utilised a directed evolution strategy to construct a non-targeted cADO randomised library via error prone PCR method. To enable the screening of novel variant cADO activity present in whole cells, in Work Package 2 we have focused on the development of an ultra-high throughput GFP based activity assay that is suitable for short chain alkane screening in an aqueous system. During period 3 of the grant period, we have completed characterisation and optimisation of the *AlkS/palkB* biosensor reporter system, which is specifically designed for endogenous alkane detection from T7 driven cADO-producing strains. In this study, our work on the AlkS/*palkB* biosensor has developed a robust screening method that can be used for the detection of novel cADO variants, which will be beneficial to all consortium partners. In summary, future work will include continued enzyme evolution of cADO to enhance product yield (through improvements in specificity and turnover number), mainly toward short chain alkanes production.

4. Computational Biology and Metabolic Modeling (WP3)

4.1. Background and research objectives in WP3

Strain improvement is increasingly assisted by computational approaches to identify suitable modifications of host metabolism. While computational approaches are already well-developed for a small number of industrial heterotrophic microorganisms, the application of such methods to phototrophic metabolism is still in its infancy (Steuer et al., 2012). Currently, the most suitable computational approaches to investigate microbial metabolism are methods based on constrained-optimization, such as flux-balance analysis (FBA). Importantly, such methods do not require extensive knowledge of the kinetic parameters of biochemical reactions, but rely on knowledge of the reaction stoichiometries only – and are therefore readily applicable to many host organisms of biotechnological interest for which kinetic data are scarce. Constrained-based optimization builds on large-scale stoichiometric reconstructions of metabolism. A large-scale metabolic reconstruction seeks to provide a comprehensive compendium of all biochemical interconversions of small molecules (metabolites) taking place within a cell or organism. Starting point for a reconstruction is usually the annotated genome sequence. Based on an initial set of annotated enzymes, the completeness of synthesis routes for all known cellular constituents can be systematically tested and, if necessary, additional reactions can be identified and added to the reconstruction. Once a draft reconstruction is established, individual synthesis routes of cellular components must be experimentally tested and verified, thereby iteratively improving the accuracy of the reconstruction. Metabolic reconstructions provide a highly useful knowledge base for further computational interrogation of metabolism, including carbon and energy balances of biofuel synthesis and the identification of suitable knock-out strategies.

To support experimental analysis and to enable future cyanobacterial biotechnology, our objectives within DirectFuel WP3 (led by Humboldt-Universität zu Berlin, Germany) were to provide a state-of-the art metabolic reconstruction of the model cyanobacterium *Synechocystis* sp. PCC 6803, to iteratively improve the reconstruction over the course of the project using data generated within as well as outside the consortium, and to make use of the metabolic reconstruction for computational pathway analysis,

bottleneck analysis, and identification of synthesis routes for precursors of biofuel, in particular the bioproducts ethylene and propane.

4.2. The metabolic network of *Synechocystis sp*. PCC 6803

A main result developed within WP3 is the improvement of a high-quality manually curated reconstruction of the metabolic network of the cyanobacterium *Synechocystis sp.* PCC 6803. The reconstruction accounts for all primary metabolic pathways, transport reactions, light absorption, photophosphorylation, photosynthetic NADP+ reduction, biomass formation, and cellular maintenance. The model describes uptake of several macro- and micronutrients, such as bicarbonate, sulphur, nitrate, phosphate and inorganic ions. Details of the reconstruction process, as well as the current reconstruction are described in Knoop et al. (2013). Network reconstruction relied on input and expertise of all partners within the project, making it a unifying platform to represent and summarize state-of-the-art knowledge on cyanobacterial metabolism. To ascertain unlcear reaction steps, the presence of selected reactions were investigated experimentally within the project. For example, the alleged presence of a glyoxylate shunt could be refuted (Knoop et al., 2013). The network currently accounts for 677 genes, giving rise to 759 metabolic reactions among 601 metabolites. The model is encoded in SBML format. To provide a user-friendly output, a large-scale network visualization is provided and available on DirectFuel homepage (www.directfuel.fi).

4.3. Computational pathway analysis

Throughout the project the reconstructed network was employed to detail stoichiometric properties and energetic constraints of cyanobacterial biofuel production. Our main objectives were to conduct computational pathway analysis and identification of precursor synthesis routes, as well as bottleneck analysis and identification of knock-out strategies.

Specifically, the large-scale stoichiometric reconstruction of *Synechocystis* sp. PCC 6803 was used to evaluate synthesis routes and metabolic efficiencies of various products, including the alkene ethylene, the alkane propane, as well as butanol, octadecanoic acid, heptadecane, octadecanol, pentadecane. Thermodynamic properties of the pathways were compiled and considered in the analysis. As key results, we obtained that maximal stoichiometric and energetic yield does not depend appreciably on pathway length, even though in practice longer pathways are expected to be subject to competing side reactions (Kämäräinen et al., 2012). A notable exception is the synthesis of ethylene, whose stoichiometric and energetic yield is significantly below that of other considered target products.

Subsequently, the large-scale metabolic reconstruction of *Synechocystis* sp. PCC 6803 was used to scrutinize the metabolic map and predict potential bottleneck and suitable knock-out strategies that can be expected to enhance or stabilize the synthesis of desired products or precursors thereof. Several methods were employed, including flux transition analysis (FTA): The objective of FTA is to gradually shift cyanobacterial *in-silico* metabolism from a wildtype growth-only phenotype to a production phenotype with ethylene as the main product. During this transition all changes in metabolic fluxes are monitored. Changes in reaction usage are classified according to their predicted (differential) fluxes in growth-only versus production-only phenotypes. Of particular interest are those reactions that increase their flux in the production phenotype but are not part of the core synthesis pathway. These are likely bottlenecks for fuel synthesis. Based on this strategy, a number of potential modification targets could be identified. A

promising strategy, for example, is to induce a shift in ATP/NADPH ratio. The synthesis of most target compounds uses an ATP/NADPH ratio significantly below that of biomass synthesis. Therefore, elimination of excess ATP production is suggested as a suitable option to push flux towards product synthesis (Erdrich et al., 2014). Other identified modifications concern the turnover of cofactors and by-products of fuel production. It is noted that not all computationally identified strategies can be readily implemented into the host metabolism. However, their identification provides highly valuable suggestions for further improvements and thereby can provide guidelines for experimental strain improvement beyond trial and error.

References for WP3

Knoop H, Gründel M, Zilliges Y, Lehmann R, Hoffmann S, Lockau W, Steuer R. (2013) Flux balance analysis of cyanobacterial metabolism: the metabolic network of Synechocystis sp. PCC 6803. PLoS Comput Biol. 2013;9(6):e1003081.

Steuer R, Knoop H, Machné R. (2012) Modelling cyanobacteria: from metabolism to integrative models of phototrophic growth. J Exp Bot. 2012 Mar;63(6):2259-74.

Erdrich P, Knoop H, Steuer R, Klamt S. (2014) Cyanobacterial biofuels: new insights and strain design strategies revealed by computational modeling. Microb Cell Fact. 2014 Sep 19;13(1):128.

Kämäräinen J, Knoop H, Stanford NJ, Guerrero F, Akhtar MK, Aro EM, Steuer R, Jones PR. (2012) Physiological tolerance and stoichiometric potential of cyanobacteria for hydrocarbon fuel production. J Biotechnol. 2012 Nov 30;162(1):67-74.

5. Metabolic engineering (WP4)

5.1. Background and research objectives in WP4

The overall objective of WP4 (led by Valtion Teknillinen Tutkimuskeskus/ VTT, Finland) was to engineer cyanobacteria (*Synechocystis sp* PCC8803 and *Synechococcus* sp. PCC 7002) and generate associated tools in order develop strains for the production of desired metabolites, with the focus on volatile hydrocarbons ethylene and propane. The work centred on the design and construction of various production platforms to assess the production of metabolic intermediates and end products in regards to yields and metabolic limitations. Research involved the development and analysis of cyanobacterial ethylene production systems, introduction of pathways for propionyl/butyryl-CoA biosynthesis, modification of FA biosynthesis for enhanced production of short-chain FA precursors, assembly and evaluation of first-generation propane pathway in *E. coli*, and investigation/modification of cyanobacterial photosynthetic processes for enhanced direction of the fuel primary substrates, the solar-energy-excited electrons and atmospheric CO_2, to target biofuels or their precursors.

5.2. Plasmid-based system for cyanobacterial ethylene production

In WP4 we generated and evaluated a plasmid-based cyanobacterial platform for the production of ethylene (C_2H_4) (Guerrero et al 2012), which then served as the general model system for the study of

volatile hydrocarbons in the DirectFuel project. The overall strategy is based on ethylene forming enzyme (*efe*) originating from a prokaryotic plant pathogen *Pseudomonas syringae*. The enzyme catalyses the conversion of 2-oxoglutarate, a citric acid cycle intermediate, into ethylene, which then diffuses out of the cell and as a gas separates spontaneously into headspace of the culture. Several different cyanobacterial ethylene production systems have been previously described in literature, yet there are still unanswered questions regarding the genetic stability of different *efe* constructs and the most optimal over-expression strategies.

Alternative promoters and variations of the *efe* gene, with and without an affinity tag, were compared in an autonomously replicating plasmid-based system in *Synechocystis sp* PCC6803. While most promoters were either constitutively active (P_{trc}) or resulted in low ethylene production, a variation of a common P_{lac} promoter allowed high and regulatable expression dependent on inducer concentration.

Designed constructs were functional, and allowed prolonged production of ethylene in *Synechocystis sp* PCC6803. No indications of metabolic stress caused by protein over-expression, depletion of cellular primary precursors (2-oxoglutarate) or formation of ethylene was observed. Instead, throughout the consecutive batch cultures the cells appeared healthy, retaining a blue-green phenotype characteristic to the wild-type strain. The levels of ethylene produced were typically in the range of 100-200nl/ ml culture per hour (normalized to cell density), which correspond to the amounts reported earlier.

Although the production yields still remain clearly below commercial viability, we provide evidence that the cyanobacterial strain *Synechocystis sp* PCC6803 is a potential host for biotechnological production of ethylene in regard to metabolic effects of *efe* over-expression, and expression system stability. In addition, the work resulted in the generation of a tunable *E. coli* / *Synechocystis sp* PCC6803 shuttle vector system, which can be used for protein over-expression as an alternative to chromosomal integration constructs. The next step in improving ethylene production would be to maximize the metabolic flux through the TCA cycle towards the production of 2-oxoglutarate intermediate, which may currently be the most significant limitation of cyanobacterial ethylene production platforms. The best performing system has further been evaluated in photobioreactor cultivations as part of WP5.

5.3. Coenzyme A dependent pathway for propane production in *Synechocystis*

The mutated *P. marinus* ADO enzyme has been shown to convert butanal to propane (Khara et al. 2013). Another fuel molecule, butanol, is also produced from butanal. Several microbes have been engineered for butanol production and numerous biosynthetic pathway options for butanal biosynthesis are known (Lan and Liao 2013). A modified coenzyme A (CoA) dependent pathway was selected and *Synechocystis* was first engineered to produce the propane pathway intermediate butyryl-CoA. The functionality of the pathway was confirmed by measuring the activity of pathway enzymes and by analysing intracellular metabolites, which demonstrated that butyryl-CoA was present in the engineered strain but not in the parent strain.

Butyryl-CoA can be converted to butanal directly by an acylating aldehyde dehydrogenase (ALDH) or alternatively, via a butyric acid intermediate. The first option was to co-express an acylating aldehyde dehydrogenase (ALDH) and the ADO variant A134F (Khara et al. 2013) in the butyryl-CoA producing *Synechocystis* strain. Production of butanol, butanal and propane into gas phase of the cyanobacteria cultures was assessed. The introduction of the ALDH resulted in enhanced production of butanal and

butanol as could be expected, but introduction of both ALDH and ADO together did not, which suggests that butanal was converted to another compound. However, propane was not detected in the cultivation experiments. This suggests that the ADO enzyme was not able to compete for butanal with the alcohol dehydrogenase type of activity present in cyanobacteria, or that the amount of propane possibly produced was too low to be detected with the analytical methods used.

5.4. Enhanced production of short chain length fatty acids in cyanobacteria as precursors of alkane biosynthesis

Besides propane, other short chain alkanes and fatty alcohols are suitable as fuel products since they can be used directly as fuels without further chemical modifications. Because they are also volatile, they can be separated from the liquid cultivation medium and harvested. Alkanes can be produced by redirecting a fraction of fatty acid biosynthetic pathway towards production of the desired compounds by introducing novel enzyme activities into the production host. As a first step towards production of short chain alkanes, we modified cyanobacteria to produce increased amounts on short chain length fatty acids, i.e. unsaturated octanoic and decanoic fatty acids, consisting of 8 or 10 carbon molecules, respectively. This was achieved by expressing an acyl-ACP thioesterase that is specific for C8 and C10 substrates. The C8 and C10 free fatty acids were found inside the recombinant cells as well as excreted into the culture medium, while the unmodified host cell did not produce these fatty acids. The production level of the C8 and C10 fatty acids was modified by adjusting the copy number of the acyl-ACP thioesterase encoding gene, by altering the non-translated region controlling expression, and by inactivating an endogenous metabolic pathway that competes with short chain fatty acid production for the same metabolic precursors. Increase in gene copy number and elimination of a competing pathway each resulted in significant increase in the excreted short chain fatty acids.

Two alternative parallel strategies were chosen for further conversion of the C8 and C10 fatty acids to alkanes. One strategy takes advantage of the carboxylic acid reductase (CAR) enzyme (Akhtar et al. 2013) that converts free fatty acids to the corresponding aldehydes in one step. The other strategy relies on two consecutive enzymatic steps that convert the fatty acids first to fatty acyl-CoA esters and second to aldehydes. Finally, conversion of the aldehydes to short chain alkanes depends on the ADO enzyme. Each of the two strategies for the conversion of C8 and C10 fatty acids to the corresponding aldehydes and further to alcohols was demonstrated to function in *E. coli*: externally added C8 and C10 fatty acids were converted to C8 and C10 fatty alcohols mainly, probably via an aldehyde intermediate. The presence of the C8 and C10 aldehydes was also demonstrated in the strains expressing the acyl-ACP thioesterase and CAR. Native *E. coli* enzymes are responsible for the conversion of aldehydes to alcohols. The functionality of the CAR enzyme was also demonstrated in cyanobacteria cultivations. In a co-culture of two transformant strains, one strain produced short chain fatty acids into the medium and the other strain internalized the fatty acids and converted them, probably via C8 and C10 aldehyde intermediates, to C8 and C10 alcohols that were excreted into the medium. Thus, it can be concluded that the tested pathway components can be used for fatty aldehyde biosynthesis in cyanobacteria. However, the endogenous cyanobacterial ADO (Schirmer et al. 2010) activity was not sufficient to enable detectable production of alkanes. Overexpression of ADO activity together with the rest of the pathway has not been achieved yet but our hypothesis is that optimization of the ADO step will allow short chain alkane production.

5.5. Enhancement CO_2 fixation and reallocation of carbon

As a first approach we overexpressed the Rubisco operon in the cyanobacterium *Synechococcus* sp. PCC 7002 through exchange of the promoter sequence, however this did not enhance the photosynthetic and growth capacities of the cyanobacterium. In the second approach, we overexpressed sedoheptulose/fructose bisphosphatase (SBP/FBPase), also involved in the Calvin cycle, and observed that both growth and the photosynthetic capacity of the mutant strain were considerably higher than in the wild type. Subsequently, the overexpression of SBP/FBPase and the expression of the ethylene forming enzyme (EFE) were performed in the same strain and tested if this results in enhancement of photobiological production of ethylene. Preliminary results showed nearly 10-fold increase in ethylene production as compared to the control strain wherein EFE was present but the overexpression cassette SBP/FBPase was not.

We showed that the *pmgA* mutant of *Synechocystis* sp. PCC 6803 accumulates large amount of glycogen. Genome-wide microarray analysis in collaboration with ALU-FR has identified a set of genes differentially regulated in the *pmgA* mutant, of which a non-coding RNA, *ncr0700*, was found considerably reduced in the abundance in the *pmgA* mutant. Deletion and overexpression mutant of *ncr0700* was generated in the wild-type and the *pmgA* mutant backgrounds. The *ncr0700* deletion mutant was found to be sensitive to glucose when grown photomixotrophically, a phenotype also seen in the pmgA mutant. These results suggest that ncr0700 is a new genetic target for altering the carbon allocation in the cyanobacterium and together with PmgA are potentially useful targets for strain improvement towards sustainable biofuel production.

5.6. Tailoring the electron flow to target fuel molecules

Production of any "biological solar fuel", requires a safe and efficient channelling of light energy into photosynthetic reaction centers for conversion into chemical energy and subsequent efficient electron flow to target biofuel molecules without production of reactive oxygen species. In WP4, we evaluated auxiliary electron transfer pathways in *Synechocystis sp.* PCC 6803 using various flavodiiron protein (*flv1, flv2, flv3, flv4*) and ferredoxin (*fad7*) mutants and tested their performance under different growth conditions. In general, the presence of strong electron sinks most often eliminated the need for auxiliary electron transfer pathways in scavenging the production of reactive oxygen species, indicating that the "loss of electrons" can be avoided by constructing the system such that the target biofuel production provides a strong sink for electrons.

It was shown that the Flv1/Flv3 heterodimer can in certain conditions transfer from photosystem (PS) I more than 60 % of electrons, originally derived from water splitting PS II, back to atmospheric oxygen to form water again. Moreover, even in high CO_2 growth conditions up to 20% of electrons were directed to molecular oxygen via Flv1/Flv3 heterodimer. Fad7 had a more general regulatory role in redirecting electrons from PSI to different pathways under varying environmental conditions. Another pair of flavodiiron proteins in *Synechocystis sp.* PCC 6803, the Flv2/Flv4 heterodimer, together with the Sll0218 protein (all three proteins encoded by the same operon) was shown to accept electrons from PS II and

being particularly important for PS II photoprotection under ambient CO_2 conditions and/or high light, and in this task having an intimate collaboration with the phycobilisome antenna system. Extensive knowledge obtained from various auxiliary electron transfer pathways allows us to construct tailored *Synechocystis sp.* PCC 6803 host cells with maximal electron transfer efficiency to target biofuels.

5.7. Product toxicity to cyanobacterial hosts

As part of WP4 *Synechocystis sp* PCC6803 and *Synechococcus elongatus* PCC7942 were evaluated in regards to product toxicity by supplementation of selected chemicals (Kämäräinen et al 2012). The chemicals representing the potential biosynthetic end-products of biotechnological applications included a selection of alcohols, aldehydes, fatty acids and alkanes of different chain-lengths. These were roughly divided into categories C2-C4 (volatile-liquid), C6-C7 (liquid) and C11-C12 (liquid-solid), supplemented into the batch cultures of *Synechocystis sp* PCC6803 and *Synechococcus elongatus* PCC7942, and evaluated for the effect on growth.

Based on the findings it seems that *Synechocystis sp.* PCC6803 is slightly more tolerant to some of the supplemented chemicals than *Synechococcus elongatus* PCC7942, particularly butyrate and undecane. In context with the project goals the results may have several implications: (i) Alkanes appear to be less toxic than the corresponding aldehydes, alcohols or fatty acids, making them a feasible target for biotechnological applications. (ii) *Synechocystis sp PCC6803* is more tolerant to supplemented butyrate, which may suggest that it would be a more suitable host for propane biosynthesis using an ADO/ CAR pathway, which proceeds via free butyrate intermediate. Furthermore, (iii) the apparent high toxicity of aldehydes, although expected, highlights the importance of very efficient coupling between the successive biosynthetic steps catalyzed by CAR (fatty acid > aldehyde) and ADO (aldehyde > n-1 alkane).

From the perspective of the DirectFuel project, alkanes appear to be optimal product candidates due to relatively low toxicity effects in comparison to other alternatives such as alcohols, aldehydes or fatty acids. In addition, comparison of the alternative host strains in several different contexts has not revealed significant differences in performance, which would provide grounds to select *Synechococcus elongatus* PCC7942 over *Synechocystis sp* PCC6803 – the strain used as a default in most of the experimental set-ups.

5.8. Propane biosynthesis in *E. coli*

Major part of the WP4 focused on studying the *cyanobacterial aldehyde defomylating oxygenase* (ADO) enzyme in different biosynthetic contexts in *E. coli*. The aim was to construct an ADO-based heterologous pathway for the production of short-chain hydrocarbons, and to evaluate the performance, restrictions and possibilities of the system in regards to subsequent transfer into photosynthetic cyanobacterial hosts. The study was a direct continuum to the molecular structural studies and functional characterization of ADO carried out in WP1 and WP2.

The work resulted in the reconstitution of a synthetic oxygen-tolerant alkane pathway producing propane and heptane in *E. coli* (Kallio et al 2014). The core of the system consisted of an *ACP-thioesterase* responsible for releasing the C4 buryrate fatty acid precursor from the endogenous fatty acid biosynthesis

machinery, *carboxylic acid reductase* responsible for conversion to butyraldehyde, and ADO (*P. marinus*) catalyzing the last step to form propane. Although previously ADO has been shown to be able to use C4 substrate and produce propane *in vitro* (Khara et al 2013), this was a first time the production of propane has been accomplished as part of a biological system *in vivo*. The productivity of the *in vivo* pathway remained very low even under the best conditions (~30mg/L), and continuous production could not be maintained in the closed-vial batch system. However, the pathway demonstrated the proof-of-concept of microbial biosynthesis of propane, providing new valuable information on the possibilities and potential bottlenecks.

As part of the main findings, we were able to show that the pathway flux in *E. coli* was limited by (i) the availability of butyrate precursor, (ii) unoptimal kinetic properties of the pathway enzymes (CAR and ADO) towards the C4 substrate, (iii) the availability of reducing equivalents for ADO, and (iv) competition of the endogenous aldehyde reductases for the butyraldehyde precursor.

Taken together, the *in vitro* and *in vivo* characterization of ADO has significantly paved the way towards engineering the systems in photosynthetic hosts, as part of the ultimate goals of the DirectFuel project. The studies have revealed significant constraints which have to be overcome before reaching the final goal of generating efficient autotrophic production systems. One of the main conclusions of the work regarding the use of ADO in synthetic alkane pathways concerns the supply of electrons for the deformylation reaction. The ADO reaction requires an input of four electrons per catalytic cycle, and our experiments demonstrated the significance of a functional electron relay to ADO, a concept previously only studied in vitro: The *E. coli* hydrocarbon pathways reported before have relied on the endogenous capacity of the host to reduce ADO, whereas we co-expressed cyanobacterial ferredoxin (PetF) and NADPH:ferredoxin oxidorecuctase (Fpr) as part of the pathway. The indispensable role of both redox components highlighted the need for efficient coupling of the host metabolism with the electron relay to ADO, and even if this is expected to be less of an issue in autotrophic hosts like cyanobacteria, the possible limitation for high-efficiency production systems must be taken into account in the pathway design. Perhaps most importantly, comparison of pathway variations producing propane versus heptane suggested that the kinetic properties of ADO were inferior towards C4 substrate in comparison to the C7 substrate. Thus, like earlier suggested *in vitro*, it appears that the metabolic flux of the propane pathway is at least in part compromised by low affinity and low conversion efficiency of ADO towards butyraldehyde. Consequently, modified ADO variants or enzyme homologs with higher specificity towards short-chain precursors would be required to improve productivity in any biological host.

Overall, the study demonstrated that ADO can be used in a heterologous biosynthetic context for the production of short-chain alkanes such as propane and heptane *in vivo*. The study serves as the first case of proof-of-concept for the biosynthesis of short-chain hydrocarbons, and addresses some of the primary limitations of the system. The results presented here serve as a solid starting point for the subsequent design and transfer of the second-generation pathways into photosynthetic hosts.

References for WP4

Akhtar MK, Turner NJ, Jones PR (2013) Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. Proc Natl Acad Sci U S A 110: 87-92

Khara B, Menon N, Levy C, Mansell D, Das D, Marsh EN, Leys D, Scrutton NS (2013) Production of propane and other short-chain alkanes by structure-based engineering of ligand specificity in aldehyde-deformylating oxygenase. Chembiochem 14(10): 1204-1208

Lan EI, Liao JC (2013) Microbial synthesis of n-butanol, isobutanol, and other higher alcohols from diverse resources. Bioresour Technol 135:339–349

Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB (2010) Microbial biosynthesis of alkanes. Science 329:559-562

Kallio, P., Pásztor, A., Thiel, K., Akhtar, M. K. & Jones, P. R. (2014) An engineered pathway for the biosynthesis of renewable propane. Nat Commun 5(4731): 1-8

Guerrero, F., Carbonell, V., Cossu, M., Correddu, D. and Jones. P. R. (2012) Ethylene synthesis and regulated expression of recombinant protein in Synechocystis sp. PCC 6803. PLoS ONE 7(11):e50470.

Kämäräinen, J., Knoop, H., Stanford, N.J., Guerrero, F., Akhtar, M.K., Aro, E-M., Steuer, R., Jones, P.R. (2012) Physiological tolerance and stoichiometric potential of cyanobacteria for hydrocarbon fuel production. Journal of Biotechnology 162 (1), 66-73.

6. Bottleneck analysis (WP5)

6.1. Background and research objectives in WP5

The objectives of WP5 (led by Albert-Ludwigs-Universität Freiburg, Germany) were to identify possible bottlenecks in order to optimize host metabolism to favour the introduced pathways in WP4. Goals included evaluating the effects of cultivation conditions on the host physiology, metabolism and yield of products by a combination of targeted metabolite, lipid, and non-targeted transcriptome measurements, followed by an evaluation against the refined metabolic model.

6.2. *Synechocystis* sp. PCC6803 as a model organism in DirectFuel

The cyanobacterial strain *Synechocystis* sp. PCC6803 was chosen as a model organism in the DirectFuel project. Currently, there are more than 5,000 publications listed in PubMedCentral when searching for 'Synechocystis', part of it due to the fact that this organism was the third ever for which a total genome sequence was determined (Kaneko et al., 1996), more than 15 years ago. Because bacteria are notoriously prone to constantly accumulate mutations over time, we (Partner ALU-FR) subjected *Synechocystis* sp. PCC6803 strain "PCC-M" to a resequencing analysis. This resequencing analysis (Trautmann et al., 2012) revealed indeed several SNPs potentially of interest for the scope of the DirectFuel project. A transition in gene *slr1609* leads to a L608S (L548S) substitution in the long-chain acyl-CoA-synthetase Slr1609 that has been found crucial for fatty acid activation and the biosynthesis of alkanes (Gao et al., 2012). Interestingly, an unrelated SNP exists at position 488923 within the *slr1609* coding sequence in *Synechocystis* sp. PCC6803 strain 'YF', leading to a G546L (G486L) substitution (Aoki et al., 2012). Furthermore, it should be noted that we could re-annotate the *slr1609* reading frame as being 60 codons shorter (636 instead of 696 amino acids), compared to the original annotation (Kaneko et al., 1996). The shorter Slr1609 protein of 636 amino acids is also consistent with the start site of transcription at position mapped by us at position 487352, located 115 nt upstream of the revised start codon (Trautmann et al., 2012).

6.3. Distribution of alkane biosynthesis genes in cyanobacteria and their expression

In cyanobacteria, alkanes are synthesized from a fatty acyl-ACP by two enzymes (Schirmer et al., 2010), *acyl-acyl carrier protein reductase* (AAR) and *aldehyde deformylating oxygenase* (ADO). Nothing has been known thus far about the physiological function of alkane synthesis or the transcriptional organization of their genes. However, such information could be very meaningful to boost the production of biofuels through this pathway. We screened 181 cyanobacterial genomes for orthologs of AAR and ADO. Orthologs of genes for both enzymes were identified in 90% (163/181) of the genomes, which underlines their importance. Notably, all the other genomes possess an alternative pathway, which relies on a polyketide-based chemistry to produce 1-alkenes (Coates et al., 2014), reinforcing the putative physiological relevance of alkane/alkene synthesis in cyanobacteria further (Klähn et al., 2014).

We determined the primary transcriptomes of the closely related *Synechocystis* strains PCC 6803 and PCC 6714 by the dRNAseq approach (Mitschke et al., 2011a; Kopf et al., 2014). Beside the identification of all transcriptional start sites (TSS) and active promoters, these data allowed the definition of transcriptional units thereby enabling us to distinguish between monocistronic and polycistronic transcripts (operons). Meanwhile, the primary transcriptomes of many other cyanobacteria, including *Anabaena* PCC 7120, *Nodularia spumigena* CCY9414, *Prochlorococcus* MIT9313 & MED4 and *Trichodesmium erythraeum* IMS101 became also available (Mitschke et al., 2011b; Pfreundt et al., 2014; Voigt et al., 2014; Voß et al., 2013). Thus, we performed a comprehensive analysis of the genomic arrangement of genes encoding the enzymes ADO and AAR that are responsible for alkane synthesis combined with the investigation of the transcriptional organization of these genes (Klähn et al., 2014). Despite of the conserved clustering of *ado* and *aar* the data revealed an independent transcription of both genes in *Synechocystis* 6803 (Klähn et al., 2014). The presence of specific promoter elements upstream of both genes was verified by reporter gene fusion experiments.

Thus, despite their close functional connection and co-appearance in most genomes, our data demonstrate that in all tested strains the *ado* and *aar* genes are transcribed from own, specific promoters. Thus, these two genes can be expressed independently from each other, in turn enabling their different or even divergent regulation. Due to the presence of separate regulatory elements, the dominance of monocistronic mRNAs and additionally the fact that the two genes are split in some species we speculate that physiological situations might exist for which an independent transcription of *ado* and *aar* could be advantageous. Additionally, in some strains *ado* is transcribed from two separate TSS, which increases the transcriptional complexity further.

A highly interesting aspect of the observation that *ado* and *aar* genes are transcribed from separate and distinct promoters in cyanobacteria are the possible metabolic implications. Their monocistronic layout allows the regulatory autonomy of these two genes. Thus, they may signify also other, unknown, pathway(s) branching off from aldehyde synthesis that would require their separate and non-stoichiometric expression. Such pathways might, e.g. lead to free fatty acid (from aldehyde oxidation) or even fatty alcohols (from aldehyde reduction) for which separate regulation of these two genes would be required. These results point to the possibility that other, still unknown, aldehyde-derived pathways might remain to be discovered in cyanobacteria.

6.4. Transcriptomic and microarray technologies within DirectFuel

The development and use of transcriptomic and microarray technologies has been another key activity in the Work Package. We have characterizied appropriate mutants with the whole system transcriptomic analysis in order to identify bottlenecks of carbon fixation and new molecular targets for redirecting carbons from growth towards products. The *pmgA* mutant was generated as a strain with enhanced carbohydrate accumulation. We identified the non-coding RNA ncr0700 that is likely acting downstream of PmgA and is involved in a regulatory cascade controlling the glucose tolerance and utilization in *Synechocystis* sp. PCC 6803.

A high density microarray covering the entire genome of *Synechocystis* sp. PCC 6803 was designed for the IMDH approach used in WP3 and WP5. For the design of this microarray, the sequence of the chromosome and all plasmids was fragmented *in silico* into 52,000 fragments of 75 bp length each. Afterwards, an oligonucleotide probe matching all necessary physicochemical properties was designed for both strands of all fragments yielding 104,000 individual probes. Finally, these probes were manufactured in a 105K Agilent custom microarray format and utilized for the high resolution mapping of IMDH-mediated insertions.

6.5. Establishment and application of a high efficiency transposon-mediated differential hybridisation approach

As a primary objective in WP5 we established an experimental protocol similar to a method known as 'Transposon-mediated Differential Hybridization' (TMDH) which combines high-throughput transposon mutagenesis and microarray hybridization in order to characterize simultaneously all mutation events in a pool of mutants analysis. TMDH is a combined approach that includes high-throughput transposon mutagenesis and microarray hybridization to map simultaneously all transposon insertions in a pool of mutants (e.g., Chaudhuri et al. 2009). TMDH presents the advantage to generate a large number of mutants which can all be screened in a single step by microarray hybridization. TMDH allows functional assignments to be made based on whether genes are essential (= no transposon insertion) for survival under particular conditions or not. Conditionally essential genes are essential at a defined condition. Moreover, TMDH can be used to identify unknown genes for metabolic processes that are essential at standard growth conditions but not when a particular metabolite is supplied and taken up by the cells. Thus, this approach (because of technical improvements here renamed to IMDH for "Insertion-Mediated Differential Hybridization") allows functional assignments to be made based on whether genes are essential for survival under particular conditions or not. In this context, the tool was demonstrated as being helpful to validate the computational model of primary metabolism in *Synechocystis* sp. PCC 6803. First, we generated a strain with a disruption of ORF *sll1354* encoding the RecJ exonuclease, causing an increase of the transformation efficiency by two orders of magnitude (Kufryk et al., 2006). Using the Δ*recJ* strain as platform, the mutant library was generated by *in vitro* cleavage of genomic DNA with diverse restriction enzymes, followed by ligation with a Km^R cassette containing outward directed T7 promoters, which subsequently were transformed into *Synechocystis* PCC 6803 where recombination into genomic DNA took place. From an initial number of about 1.7 million kanamycin resistant colonies several different experiments started in order to select conditionally essential genes. By using microarray hybridization the library was characterized in detail. As we used an approach independent of transposons the protocol is now referred

to as "Insertion-Mediated Differential Hybridization (IMDH)", a technical modification of the original TMDH protocol. The microarray hybridization was performed as described earlier (Georg et al., 2009) but for the data analysis a dedicated bioinformatics pipeline had to be developed. Focusing on protein-coding genes on the chromosome, 2,839 of 3,317 protein-coding genes were found disrupted by the insertion of a kanamycin resistance cassette. To evaluate if the disruption of a gene might be (dis-)advantageous under a particular condition, microarrays were performed after long-term cultivation of the mutant library under standard conditions as well as in presence of 3% NaCl and 25 µM methionine, respectively. Assuming negative selection for mutations of essential genes negative fold changes would be expected comparing the library at day 0 and later stages even if cells are grown under standard conditions. Indeed, we identified 65 genes for which a negative $\log_2$-fold change < -3.0 was observed, indicating these mutations are mainly negative for growth and/or survival. Overall, for 435 genes a mutant was found in the initial library which was absent after 175 days. We concluded that the IMDH technique validated the computational model of primary metabolism in *Synechocystis* sp. PCC 6803 and led to the identification of genes encoding metabolic activities and transport functions that were functionally unknown thus far. In context with the systemic analysis of mutants in regulatory genes these techniques allows the identification of regulatory and metabolic bottlenecks and contributes to an improved model of the regulatory-metabolic network.

References for WP5

Aoki, R., Takeda, T., Omata, T., Ihara, K. and Fujita, Y. (2012) MarR-type transcriptional regulator ChlR activates expression of tetrapyrrole biosynthesis genes in response to low-oxygen conditions in cyanobacteria. *J. Biol. Chem.* 287: 13500–13507.

Chaudhuri R.R., Allen A.G., Owen P.J., Shalom G., Stone K., Harrison M., Burgis T.A, Lockyer M., Garcia-Lara J., Foster S.J., et al. (2009) Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics* 10: 291.

Coates, R. C., Podell, S., Korobeynikov, A., Lapidus, A., Pevzner, P., Sherman, D. H., Allen, E. E., Gerwick, L., and Gerwick, W. H. (2014) Characterization of cyanobacterial hydrocarbon composition and distribution of biosynthetic pathways. *PLoS ONE* 9: e85140.

Gao, Q., Wang, W., Zhao, H. and Lu, X. (2012) Effects of fatty acid activation on photosynthetic production of fatty acid-based biofuels in *Synechocystis* sp. PCC6803. *Biotechnol. Biofuels* 5: 17.

Georg J., Voss B., Scholz I., Mitschke J., Wilde A., Hess W.R. (2009) Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. *Mol. Sys. Biol.* 5*:* 305.

Kaneko, T., Sato, S., Kotani, H., 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 (supplement). *DNA Res.* 3: 185–209.

*Klähn S., Baumgartner D., Pfreundt U., Voigt K., Steglich C., Hess W. R. (2014) Alkane biosynthesis genes in cyanobacteria and their transcriptional organization in *Synechocystis* sp. PCC 6803. *Front. Bioeng. Biotechnol.: Synth. Biol.* 2: 24.

Kopf, M., Klähn, S., Scholz, I., Matthiessen, J.K.F., Hess, W.R., and Voss, B. (2014) Comparative analysis of the primary transcriptome of *Synechocystis* sp. PCC 6803. *DNA Research* 21: 527-539.

Kufryk G. I., Sachet M., Schmetterer G., Vermaas W.F.J. (2002) Transformation of the cyanobacterium *Synechocystis* sp. PCC 6803 as a tool for genetic mapping: optimization of efficiency. *FEMS Microbiol. Lett.* 206: 215-219.

Mitschke, J., Georg, J., Scholz, I., Sharma, C.M., Dienst, D., Bantscheff, J., et al. (2011a) An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. *Proc. Natl. Acad. Sci. U.S.A.* 108: 2124–2129.

Mitschke, J., Vioque, A., Haas, F., Hess, W.R., and Muro-Pastor, A.M. (2011b) Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. *Proc. Natl. Acad. Sci. U.S.A.* 108: 20130–20135.

Pfreundt, U., Kopf, M., Belkin, N., Berman-Frank, I., and Hess, W.R. (2014) The primary transcriptome of the major marine diazotroph *Trichodesmium erythraeum* IMS101. *Scientific Reports (Nature Publishing Group)* 4*: 6187.*

Schirmer, A., Rude, M. A., Li, X., Popova, E., and del Cardayre, S. B. (2010) Microbial biosynthesis of alkanes. *Science* 329, 559–562.

*Trautmann, D., Voss, B., Wilde, A., Al-Babili, S., and Hess, W.R. (2012) Microevolution in cyanobacteria: re-sequencing a motile substrain of *Synechocystis* sp. PCC 6803. *DNA Res.* 19: 435–448.

Voigt, K., Sharma, C.M., Mitschke, J., Lambrecht, S.J., Voß, B., Hess, W.R., and Steglich, C. (2014) Comparative transcriptomics in two environmentally relevant cyanobacteria reveals unexpected transcriptome diversity. *ISME J;* doi: 10.1038/ismej.2014.57 [Epub ahead of print].

Voß, B., Bolhuis, H., Fewer, D.P., Kopf, M., Möke, F., Haas, F., et al. (2013) Insights into the physiology and ecology of the brackish-water-adapted cyanobacterium *Nodularia spumigena* CCY9414 based on a genome-transcriptome analysis. *PLoS ONE* 8: e60224.

7. Process considerations and photobioreactor technology (WP6)

7.1. Background and research objectives in WP6

The objectives of WP6 (led by Biochemtex S.p.a., Italy) were (i) to make a comprehensive evaluation of the microbiological fuel production process and (ii) to design and construct a photobioreactor for commercial production of volatile hydrocarbons fuels. Towards these goals, a computational Life Cycle Analysis was performed to evaluate the production process form an environmental viewpoint, in order to identify weaknesses and strengths which would allow further modification for optimal output. In parallel, an extensive literature and patent survey was carried to out to optimize photobiorector technology for maximal performance and minimal cost. The second part focused on the design, development and evaluation of a bench photobioreactor (PBR), equipped with sensors, regulatory modules and detectors suitable for continuous automated production of volatile hydrocarbons.

7.2. Life cycle analysis for the production of ethylene

In order to assess the photobiological production systems studied in the project from an environmental viewpoint, we carried out a Life Cycle Analysis (LCA) for the production of ethylene through photochemical fixation of carbon dioxide with cyanobacterium *Synechocystis* sp PCC6803. The preliminary analysis made based on literature and patent survey data was to estimate from an environmental perspective the most critical points in the process. The most significant positive contribution in the process derives, as expected, from CO_2 uptake, while the most relevant negative impact is the electricity consumption. The results served as a starting point for further evaluation, and were adapted to a pilot-scale model involving experimental data obtained from DirectFuel consortium partners.

The pilot scale LCA took into consideration the direct impact of the material requirements and the energy input/ output of the process, while indirect factors such as plant construction or product storage were ignored. The model was based on a continuous production strategy, with various parameters derived from data/ approximations from affiliated DirectFuel partners, and compared the use of two different light sources (sun and artificial lamps) and two different cultivation media (sea water and fresh water supplemented with necessary nutrients). The environmental competitiveness of the biological production strategy was evaluated by comparison with a petrochemical ethylene production system (steam cracking) of corresponding capacity.

The results implicate that the biological production system could provide a positive environmental impact in comparison to the common petrochemical production route, but more comprehensive evaluation is required for precise estimation. In the current model, electricity consumption is the most relevant contributor to the outcome. Using supplemented sea water for the medium clearly decreases the negative environmental impact by conserving fresh water for other uses, thus reinforcing the view that production plants should be located in brackish water bodies or coastal areas where they would not compete for arable land and fresh water.

7.3 Literature survey of photobioreactor designs

An extensive literature and patent survey was carried out to compare different photobioreactor designs and materials for optimizing the process performance versus associated costs. Various types of existing reactor systems with different geometries and process parameters such as light requirement, mixing, CO_2 transfer systems, O_2 removal were investigated. The comparison clearly indicates that none of the candidates is clearly superior over the others, and the most optimal choice is case-dependent with multiple overlapping considerations. Potential bioreactor designs which could meet the demands for a large-scale photobiological production strategy could be *horizontal tubular reactor* or *flat plate reactor* (closed systems) or an open *raceway pond with a paddle wheel* (open system).

7.4. Construction of a photobioreactor optimized for autotrophic production of volatile hydrocarbon fuel

One of the main objectives in WP6 was a construction of functional photobioreactor optimized for autotrophic production of volatile hydrocarbon fuel – ethylene (Photon Systems Instruments/ PSI, Czech Republic). The laboratory photobioreactor was upgraded with several regulatory modules facilitating optimized long-term continual cultivation. These modules allow maintaining constant cultivation conditions based on monitoring of parameters such as irradiance, pH, optical density, fluorescence, O_2, CO_2. For the purpose of volatile hydrocarbon determination the photobioreactor has been coupled either with membrane inlet mass spectrometry (MIMS) or with an optical sensor based on infra-red absorption (NDIR). Acquired data allows the quantification of ethylene yields per light and carbon units.

Two photobioreactors were constructed and used for performing of experiments focused on optimizing conditions for biomass production of target organism *Synechocystis* sp. PCC 6803. We subjected four sub-

strains ("Hagemann", "Kaplan", "Nixon", "Shestakov") of *Synechocystis* sp. PCC 6803 to a comprehensive testing using methodological foundation established previously. The motivation was a selection of the strain most suitable for further optimization, genetic modification and up-scaling. The strains were evaluated regarding the following criteria: 1) maximum growth rate and biomass production, 2) sustainable long-term cultivation, 3) growth reproducibility under given conditions (no tendency to population collapse), and 4) resistance to stress factors inevitably coupled with cultivation or harvesting (e.g. shear stress). The sub-strain "Kaplan" was identified as the most suitable candidate because of its robustness and stability in growth. Other tested sub-strains ("Hagemann", "Shestakov", and "Nixon") did not give required growth rates or biomass production or showed unexpected instability of population and collapsed in various phases of growth. Based on matrix analysis, a number of environmental factors were examined and the full set of optimal conditions for biomass production in developed flat-panel photobioreactor were identified for *Synechocystis* sp. PCC 6803 ("Kaplan"). The illumination optimum ranges between 200-300 μmol(photons).$m^{-2}.s^{-1}$; the spectral quality: red light with peak around 635 nm with additional at least 25 μmol(photons).$m^{-2}.s^{-1}$ of white or blue light; optimal temperature is 30°C ± 3 °C; CO^2 enrichment about 5000 ppm, pH: 8,0 ± 0,5.

The next step was to determine the optimal conditions for target metabolite production, ethylene in our case evolved via genetically modified "Kaplan" *Synechocystis* sp. PCC 6803 pDF-trc-EFE (Gerrero et al 2012) expressing gene for ethylene forming enzyme (EFE). PSI ran experiments that helped understand possible limitations and led to several conclusions which could be employed for the potential design of large-scale photobioreactor for volatile fuel production. Higher ethylene productivity was reported during the phase of active cyanobacterial growth, particularly in exponential phase that could correlate simply with higher metabolic activity and potentially higher concentration of ethylene precursors (e.g. 2-oxoglutarate) in the metabolic pools. One of the limiting factors for ethylene production seems to be nutrients availability or perhaps some imbalance among nutrients of BG11 medium. Further optimization of the medium to obtain the optimal ethylene yield is necessary. The light was not a limiting factor, the same ethylene production was detected already at low illumination conditions such as 50 – 100 μmol photons. m^{-2}. s^{-1} as well as under higher irradiance. The major advantage of low light would be lower cultivation cost and suboptimal photosynthesis with low oxygen evolution. Temperature optima for biomass growth and ethylene production are similar. The potential design of large-scale photobioreactor should comprise some thermoregulation system to avoid distinct temperature fluctuations that might be more important for culture maintenance than for ethylene production alone. Important is a large headspace enabling an adequate gas exchange between liquid and gaseous phases to prevent O_2 accumulation in the algal solution. Moreover, it seems the ratio of headspace/biomass volume influences the ethylene synthesis through unexplained factors as well. This experimental data serves as a background for possible large-scale photobioreactor design that is evaluated by life cycle assessment.

PSI has developed an ultimate analytical method for online determination of gases or dissolved gases concentration characterized by high sensitivity, real-time response and low detection limit in photobioreactor environment based on membrane inlet mass spectrometry (MIMS). As an alternative a low cost detection of ethylene concentration in gaseous phase by nondispersive infrared sensor (NDIR) has been designed.

To control the cultivation of target cyanobacterial strains and the ethylene production a new photobioreactor software was developed to enhance viewing, reporting, searching and control capabilities

of the cultivation. The new software is flexible and adjustable according to customers' requirements. It allows remote online control of multiple cultivation units including all external modules as gas mixing, additional sensors, medium supply and pumps. Moreover, the user may add other supplementary sensors considering individual experiment specifics. The implementation of fine pH control resulted in very precise pH adjustment and monitoring during the cultivation. Currently it is possible to record the actual medium consumption per time unit or as a cumulative parameter. Due to the script support the end user can modulate the experiment conditions in more sophisticated way. In addition, the new software enables not only to control the system and to collect and view the data but also to filter and search the database. All mentioned features contribute the full online and automatic cultivation control and monitoring.

References for WP6

Guerrero, F., Carbonell, V., Cossu, M., Correddu, D. and Jones. P. R. (2012) Ethylene synthesis and regulated expression of recombinant protein in Synechocystis sp. PCC 6803. PLoS ONE 7(11):e50470.

The potential impact

Important socio-economic and societal implications

One of the most urgent needs for our society is to combat the climate change and replace the fossil fuels as soon as possible with sustainable and renewable energy options. It is of pivotal importance to guarantee the basic requirements of life for future generations despite ever increasing human population and extinction of fossil fuel resources. The DirectFuel project is intimately related to this pertinent global need for the change from fossil fuel based economies towards sustainable bioeconomy. In essence, the project is in line with the EU Bioeconomy Strategy strongly encouraging scientists to seriously invest in development of novel innovations towards lean CO_2-neutral fuel production systems.

Life on Earth is based on photosynthesis, a process that harnesses solar energy and converts it to chemical energy to support all forms of life. Based on this platform, scientists have taken a global "solar fuel" action to exploit the principles of natural photosynthesis in multidisciplinary efforts to find replacements for nonrenewable fossil fuels. As the premise, the sun is an excessive and practically endless source of energy, but the fundamental challenge is how to collect and store this energy. Significant investments have been devoted to the development of photovoltaic (PV) and solar cell technologies, but the restriction to these strategies is that electricity as such does not suffice all the needs; Majority of the energy in the current infrastructure is needed in the form of fuels, and particularly the constraints related to the storage and transport of electricity, it is pertinent to develop sustainable fuel production systems. In parallel, biomass based fuel production is very inefficient, with harnessed solar energy to biofuel photon conversion efficiency (PCE) well below 0.5%, which is clearly below the target threshold level. Even more importantly, biomass to biofuel conversion competes for land necessary for global food production, which is a central consideration in the development of any sustainable long-term solution.

The solar fuel concept. It is evident that new scientific breakthroughs are needed to replace the fossil fuels, serving as the primary motivation for the photosynthesis researcher community to start considering alternative methods for sustainable and clean fuel production. In 2008 the European Science Foundation (ESF) published a science policy briefing "Harnessing solar energy for the production of clean fuel" (www.esf.org) calling for an action plan towards the development of novel strategies for sustainable energy solutions. The coordinator of the DirectFuel project (Aro) was one of the signing researchers of this first European initiative for solar fuels. The message was that multidisciplinary and "out-of-the-box" research is pertinently needed to gain understanding on photosynthetic solar energy conversion mechanisms and on metabolic pathways that convert the captured energy to suitable chemicals that can function as a fuel or as a precursor for chemical conversion to a suitable fuel. On one hand, the goal was to facilitate the construction of completely artificial manmade devices for collection and conversion of solar energy to a fuel. In parallel, a photobiological approach aimed at using living microorganisms as photocatalysts to efficiently convert solar energy, water and CO_2 to solar fuels. It was calculated that for the development of any sustainable and potentially competitive fuel production system, the solar-to-fuel photon conversion efficiency should be at least 7% but over 10 % efficiency should be a general goal.

DirectFuel project in the solar fuel context. The vision behind the *solar fuel concept* is to produce desired chemical compounds using solar energy and water as the source of electrons. The very first solar fuel networks in EU were the projects SolarH and SolarH2 (coordinator prof. Stenbjörn Styring), which served as a background framework for the future projects such as DirectFuel. In addition to different alternative end-products, there are various alternative perspectives to the solar fuel research. These include, for example, *artificial photosynthesis*, which intends to develop synthetic devices that produce target products from solar energy and water. The DirectFuel project approached the concept from a biological perspective, and

aimed at using photosynthetic microorganisms, cyanobacteria, as biotechnological fuel factories. Specifically, the initiative was to generate cyanobacterial strains which use solar energy water and CO_2 to produce carbon-based solar fuels which could complement corresponding nonrenewable petroleum-derived products. DirectFuel was set off as a multidisciplinary project which focused on enzyme discovery, pathway construction and metabolic engineering towards the production of volatile hydrocarbon fuels such as propane in cyanobacterial hosts. Despite the ultimate goals of successful pathway engineering coupled to process upscale and optimization, it was realized that reaching the photon conversion efficiency close to the limit for commercialization (10%) was not realistic at this point.

Impact of the DirectFuel project. Despite numerous significant scientific findings and advances in the DirectFuel project, the achievements remained at the level of fundamental research and did not result in patents or applicable production platforms. However, the project has laid out an extensive foundation of detailed information, tools and know-how which is expected to pave the way in the research aiming at developing renewable production systems of carbon-based solar fuels. It is too premature, however, to predict the development and future course of the research field, and to evaluate the ultimate impact of the project. Estimating a realistic timeframe for reaching industrially relevant breakthroughs is difficult and largely subjective; Future progress will largely depend on both international policy-making, and multi-disciplinary collaboration between researchers working in fundamental-applied research and different fields of industry. Even though some scientists are optimistic and expect achievements in this front in the next five-year period typical estimations of the timespan for reaching commercially feasible strategies range from 10 years to 20 years depending on the context. These estimations coincide with the views of Professor Roel Bovenberg, a representative of Biotechnology DSM (a life science and material science company headquartered in Netherlands), who was presenting recent industrial developments in the field of synthetic biology and bioengineering in ChemBio EXPO in Helsinki Finland 18.3.2015. Based on his first-hand experience of developing industrial applications, new biotechnological strategies typically require a "…5-10 years timespan of building up scientific understanding and realizing it at practical scale…" – even in the case of more simple systems.

Altogether, current developments in the field of *synthetic biology* have opened new possibilities to transform the petroleum based chemistry towards renewable chemistry. In this progress, DirectFuel has been one of the pioreering projects using synthetic biology approaches to assemble novel metabolic pathways in the context of cyanobacterial cell factories. There is a strong current trend of synthetic biology rapidly expanding to industrial biotechnology, specifically in the USA but also in EU. Even in Finland, an industry survey was made for a specific synthetic biology project (including two DirectFuel partners) and the results were surprisingly positive and encouraging; There is definitely interest in combining the concepts of synthetic biology and microbial solar fuel production. It is possible that in this development the impact of DirectFuel on next-generation biofuel projects, which eventually approach production scale-up and commercialization, is much bigger than we can anticipate today. In the overall process, it is pivotal that the science is mature enough for industry to take over towards commercial implementation and exploitation.

Main dissemination activities and exploitation of results

EU level science policy. Active work towards societal realization of the importance of the solar fuel concept has continued throughout the DirectFuel project period. Activity forum has changed from ESF to the European Academies of Sciences Advisory Council (EASAC), which functions to provide EU Parliament and Commission with the best possible science-based knowledge for political decision making. The coordinator of the DirectFuel project (Aro) has been working in EASAC (first in the Biosciences Steering Committee and then in the Council) for four years promoting the different concepts associated with direct solar fuels. A

scientific report examining the future breakthrough energy solutions in EU is currently under preparation. In this context, a workshop on "EASAC Breakthrough in Renewable Energy Technologies" was held in Stockholm Sweden in 2013, with Aro as one of the four organisers. In addition, associated members have participated in numerous international meetings with specific focus on industrial and science policy aspects of the DirectFuel project. For example, Aro has given invited talks in "Molecular Science for Solar Fuel" (Sigtuna, Sweden, November 2009), in "2nd Symposium: Cyanobacteria - a new source for biofuels" (Cyano Biofuels GmbH, Berlin, Germany, May 2010) , in International Meeting "Photosynthesis Research for Sustainability" (Baku Azerbaijan, July 2011), in "5th Umeå Renewable Energy Meeting" (Umeå, Sweden, February 2013), in "International Workshop Photosynthesis: from Science to Industry". (Noordwijkerhout, Netherlands, October 2012), and in "Annual meeting BioSolar Cells (Ede, Netherlands, June 2013) - in all occasions introducing the DirectFuel concept.

Global level science policy. Direct solar fuel activities have been developed all over the continents during the past five years. The first approach to bring the different initiatives together was organized in Australia "Towards Global Artificial Photosynthesis" congress, Lord Howe Island, Australia, August 2011. Aro was an invited lecturer and explained the direct solar fuel approaches taken at the University of Turku, (DirectFuel and SolarH2). The following global congress was organised by the Royal Society "Do we need a global project on artificial photosynthesis (solar fuels and chemicals)?" July 8-10, 2014 in the Royal Society at Chicheley Hall, UK. The most prominent scientists in the field of solar fuels participated in the congress together with a number of different companies and funding bodies. One of the key questions was whether both the artificial and biological systems should be included in the same program; As the outcome, it was agreed that both "manmade" and biological "DirectFuel-type" systems should absolutely be developed under the same umbrella. In addition, continuation of the Global congresses and work for fund raising were agreed on, and the follow-up meeting with stronger integration of industry and other stakeholders was set on the agenda.

Exploitation of results. As the founding idea for the DirectFuel project, there is a general industrial interest towards photosynthetic microbial systems which could be used for the biotechnological production of renewable fuels such as short chain alkanes directly from solar energy, water and atmospheric CO_2. Thus the possibilities to exploit the scientific results, by means of patenting and interactions with industrial partners within and outside of the consortium, have been systematically discussed and evaluated throughout the project. In parallel, the patent databases and literature have been continuously followed in order to stay on track of inventions and new ideas outside the consortium.

The DirectFuel project has successfully combined multidisciplinary research and know-how to generate new platforms for solar fuel production, and evaluated the prospects to advance from fundamental research towards up-scaled optimized systems. However, despite the novel findings and cumulative achievements in the project, specific targets for patenting or biotechnological applications for industrial use have not emerged. The results still implicate that the DirectFuel strategies could be developed further, and that the findings could provide directions for the design of next-generation pathways. For example, molecular mechanistic studies and evaluation of the pathways *in vivo* have revealed that one of the central biotechnological bottlenecks would be the low specificity of the catalytic enzyme components towards the short-chain precursors in propane biosynthesis. Consequently, in parallel to further structure-based

enzyme engineering to obtain more efficient substrate-product conversion, future applications based on ADO are more likely to focus on heptane or longer chain-length hydrocarbons. As another general guideline for future designs, due to relatively broad chain-length specificity of most of the enzyme components involved, production of a mixture of hydrocarbons of different chain lengths (like typically found in jet fuels) would likely significantly enhance the efficiency of the biotechnological process in comparison to any single-product system.

As a conclusion, even if the development of direct applications fell short, the DirectFuel results serve as the foundation for the design and assembly of next-generation production platforms. The generated tools, methodology and know-how can be used for averting the identified metabolic constraints of the biological production systems, or as a basis for entirely new approaches.

- Scientific publications: By the end of the DirectFuel project period 24 scientific papers had been published or accepted for publication. Many of the papers were published in high impact factor multidisciplinary journals, indicating that the results from the DirectFuel project are of high scientific quality and importance. When writing this report, many manuscripts are in the phase of submission, finalization or in preparation. Moreover, conspicuous amount of research work is still needed to complete some parts of research initiated during the DirectFuel project. We anticipate the publication of about 10 more manuscripts that will acknowledge funding from DirectFuel.

- Presentation of DirectFuel results in international congresses: DirectFuel partners have extensively presented their scientific results as plenary or invited speakers, in addition to scientific posters, in at least 20 international congresses and symposia. The different forums have covered various areas of chemistry, biochemistry, biotechnology, bioengineering, biophysics, bioinformatics, computational sciences, molecular biology, microbiology, photosynthesis and plant biology, reflecting the multidisciplinary nature of the DirectFuel project. The 1st International Solar Fuels Congress will be organized in Uppsala, Sweden, April 26 –May1, 2015; Aro will participating as an invited speaker and will take the opportunity to introduce the DirectFuel concept and the overall final project results.

- Press releases and TV presentation: As an example of high quality research in DirectFuel, The Nature Communication publication "An engineered pathway for the biosynthesis of renewable propane" was announced in several press releases in the United Kingdom (Imperial College London) and in Finland (University of Turku and Suomen Tietotoimisto/ STT). In addition, this has been covered in TV (Jones, London) and on radio (Kallio, Turku).

- General societal dissemination: Dozens of general articles in newspapers, magazines and online as well as interviews in radio and newspapers - that either exclusively or partially included the DirectFuel concepts - have been released

- Special Lectures for high school students on DirectFuel concept have been delivered in London (Jones), Manchester (Scrutton) and Turku (Aro).

- Special international lecture for university students in the context of the 18th SPPS PhD Student Conference as an invited keynote lecturer on applied aspects of photosynthesis in production of solar fuels. Introduction of the solar fuel and DirectFuel concepts and results were presented by Aro.

- DirectFuel webpages (www.directfuel.fi) have been regularly updated throughout the entire project period, providing latest information, for example, on the different consortium events and scientific publications.

- A DirectFuel Leaflet was prepared, printed and delivered to introduce the overall research concept and aims of the DirectFuel project to scientists as well as the general public.

- BioFuel Workshop: DirectFuel workshop was organized in May 2014 in Turku, Finland. The workshop was designed to bring together some of the front-line cyanobacterial bioenergy researchers to discuss the current status of different focal areas in the field, particularly emphasizing the improvement of the photon conversion efficiency in production of novel carbon-based direct fuels, like propane in the Directfuel project. The workshop centered around lectures given by six international speakers from outside the DirectFuel consortium. The event started off with an introductory lecture on DirectFuel by the coordinator, which gave an overview to the biofuel research within the consortium. In addition, the research carried out by individual consortium partners was described in more detail in scientific posters and a separate poster session. DirectFuel was also advertised in the form of printed leaflets.

Project website address & relevant contact details

Project website address: www.directfuel.fi

Contact:
Professor Eva-Mari Aro (Coordinator)
Molecular Plant Biology
Department of Biochemistry
University of Turku
FI-20014 TURKU
Finland

Tel: +358 2 333 5931
Fax: +358 2 333 8075
E-mail: evaaro@utu.fi