A toolkit to enable hydrocarbon conversion in aqueous environments

Eva K. Brinkman1, Kira Schipper1, Nadine Bongaerts1, Mathias J. Voges2, Alessandro Abate2, S. Aljoscha Wahl1

1Department of Biotechnology, Delft University of Technology
2Delft Center for Systems and Control, Delft University of Technology

Correspondence to: S. Aljoscha Wahl at s.a.wahl@tudelft.nl

URL: http://www.jove.com/video/4182/
DOI: 10.3791/4182

Keywords: Oil remediation, alkane metabolism, alkane hydroxylase system, resting cell assay, prefoldin, Escherichia coli, synthetic biology, homologous interaction mapping, mathematical model, BioBrick, iGEM

Date Published: 6/26/2012


Abstract

Pollution of soil and aqueous environments by crude oil has been, and still is, an important environmental threat. Remediation of polluted areas has proven to be challenging 3. In this research, we propose a synthetic biology approach for sustainable solutions.

This work puts forward a toolkit that enables the conversion of alkanes by Escherichia coli and presents a proof of principle of its applicability. The toolkit consists of multiple standard interchange-able parts (BioBricks) 8 addressing the conversion of alkanes, regulation of gene expression and survival in toxic hydrocarbon-rich environments.

A three-step pathway for alkane degradation was implemented in E. coli to enable the conversion of medium- and long-chain alkanes to their respective alkanols, alkanals and ultimately alkanoic-acids. The latter were metabolized via the native β-oxidation pathway. To facilitate the oxidation of medium-chain alkanes (C5-C13) and cycloalkanes (C5-C6), four genes (alkB2, rubA3, rubA4 and rubB) of the alkane hydroxylase system from Gordonia sp. TF6 8, 21 were transformed into E. coli. For the conversion of long-chain alkanes (C15-C36), the ladA gene from Geobacillus thermodenitrificans was implemented 7, 15, 18, 22. For the required further steps of the degradation process (from the alcohol, to the aldehyde and finally to the alkanoic acid), ADH and ALDH (alcohol and aldehyde dehydrogenase, also originating from G. thermodenitrificans) were introduced 10, 11. The activity was measured by resting cell assays. For each oxidative step, enzyme activity was observed.

To optimize the process efficiency, the expression was only induced under low glucose conditions: a substrate-regulated promoter, pCaiF, was used 6. pCaiF is present in E. coli K12 and regulates the expression of the genes involved in the degradation of non-glucose carbon sources 6. Low levels of glucose indirectly activate this promoter 14. Thus, using the pCaiF promoter prevents unnecessary pathway activation during cell growth on glucose (e.g. to prepare large amounts of cells). At the end of the batch (glucose limitation), the alkane-degrading genes were expressed without the addition of an inducer.

The last part of the toolkit - targeting survival - was implemented using solvent tolerance genes, PhPFDα and β, both from Pyrococcus horikoshii OT3 17. Organic solvents can induce cell stress and decreased survivability by negatively affecting protein folding 12. As chaperones, PhPFDα and β improve the protein folding process e.g. under the presence of alkanes. The expression of these genes led to an improved hydrocarbon tolerance shown by an increased growth rate (up to 50%) in the presences of 10% n-hexane in the culture medium were observed. The function as chaperones was also suggested by a software tool that was designed to find putative interaction partners in silico between the recombinant PhPFDα and β proteins and the endogenous proteins of the receiving host E. coli.

Summarizing, the results indicate that the toolkit enables E. coli to convert and tolerate hydrocarbons in aqueous environments. As such, it represents an initial step towards a sustainable solution for oil-remediation using a synthetic biology approach.

Introduction

Oil pollution is among the most serious causes of environmental contamination, and greatly affects ecosystems, businesses and communities 3. Solutions are therefore required to battle the continuous oil pollution originating from the oil sands tailing waters in Alberta, Canada. During the process of oil extraction from oil sands, bitumen, a semi-solid oxidized form of oil, is removed using thermal recovery techniques that consume about 3.1 barrels of water per single barrel of oil 1. Oil contaminated process water, mainly originating from a local river, is stored in tailing ponds after bitumen extraction. A more effective recycling of process water in order to reduce the need for freshwater uptake is needed. To facilitate the bitumen extraction and to ensure that downstream sites meet water quality guidelines for the protection of aquatic ecosystems, process water treatments are rapidly evolving 3.

To treat pollution of organic compounds, bioremediation technologies employing microorganisms are presently encouraged 1. Alkanes are the most abundant family of hydrocarbons in crude oil, containing 5 to 40 carbon atoms per molecule 7, 21. Many bacteria are known to degrade...
Within this iGEM project several enzymes from different organisms were expressed and characterized, and made available via the BioBrick standard and Registry of Standard Biological Parts.

The well-studied alkane hydroxylase system of *Gordonia* sp. TF6 facilitates the initial oxidation step of C_{7-13} alkanes along with that of C_{2-3} cycloalkanes using a minimum of four components: alkB2 (alkane 1-monoxygenase), rubA3, rubA4 (two rubredoxins) and RubB (rubredoxin reductase) \(^6,21\). Oxidation of long-chain alkanes (ranging from C_{15} up to C_{28}) is reported to be performed by ladA, a flavoprotein alkane monoxygenase from *Geobacillus thermolithotrophicus NG-80-2* \(^7,15,18,22\). LadA forms a catalytic complex with flavin mononucleotide (FMN) that utilizes atomic oxygen for oxidation. This results in the conversion of alkanes into the corresponding primary alkanol. The alcohols are further oxidized by alcohol and aldehyde dehydrogenases to fatty acids, which readily enter the \(\beta\)-oxidation pathway \(^7,21\). A zinc-independent alcohol dehydrogenase from the thermophilic bacterium *Geobacillus thermoleovorans* B23 oxidizes medium-chain alkanols into their respective alkanals, using NAD\(^+\) as a cofactor \(^1^2\). Aldehyde dehydrogenase from the same bacterium is able to catalyze the NAD\(^+\)-dependent final step in the medium-chain oxidation \(^1^1\).

In order to reduce induction costs and to maintain optimal proliferation of the bacterial system, a promoter inducing expression at low glucose concentrations was used. The promoter pCaIF from *E. coli* was used to regulate the expression of the hydrocarbon degradation pathway components. This promoter was regulated by cAMP-Crp levels, which in turn depend on glucose levels \(^6\). At high extracellular glucose concentrations in the environment the cellular cAMP (cyclic Adenosine Mononucleotide Phosphate) level was low through the inhibition of adenylcyclase as a side effect of PTS mediated glucose transport. Conversely, during limitation (low glucose concentrations) the cAMP level increased and Crp bound to cAMP forming the complex, cAMP-Crp, which bound pCaIF and activated transcription of the downstream components \(^6,14\).

Wildtype *E. coli* can only tolerate moderate concentrations of hydrocarbons. To complete the toolkit, tolerance to hydrocarbons had to be addressed. Several organic solvent-tolerant bacteria are known to survive in water-solvent two-phase systems \(^1^2\). Molecular components known to increase tolerance are chaperones that facilitate the correct folding of proteins. The prefoldin system from *Pyrococcus horikoshii* OT3, consisting of the proteins phPFD\(\alpha\) and phPFD\(\beta\), was shown to increase hydrocarbon-tolerance \(^1^7\).

The alkane conversion toolkit was constructed following the BioBrick principle, which is documented at the Registry of Standard Biological Parts \(^8\). BioBricks are plasmids containing a specific functional insert that is flanked by 4 predefined restriction sites. The BioBrick inserts can be extended flexibly, allowing the construction of biological systems with new functions.

### Protocol

#### 1. BioBrick assembly

1. BioBricks from the Registry of Standard Biological parts are provided by iGEM headquarters. To construct a new BioBrick from existing BioBricks, digest the donor BioBrick (up to 1,0 \(\mu\)g) with the enzymes EcoRI and SpeI for positioning the donor part downstream of the acceptor part. Digest with XbaI and PstI for positioning the donor part upstream of the acceptor part. Add a third appropriate restriction enzyme that cuts in the backbone of the donor. Perform the digestions in a total volume of 20-25 mL with the appropriate buffer, according to the supplier (final concentration 1X). Use 5 units/\(\mu\)g DNA for the restriction enzymes.
2. Perform the transformation reaction with circa half of the ligation mix.
3. Incubate the digestion for (at least) one hour at 37 °C. Inactivate the restriction endonucleases by heat, incubation at 80 °C for 10 minutes and centrifuge shortly.
4. Perform transformation reaction with the remaining ligation mix.
5. Confirm the BioBrick to be correct with sequencing.
6. To construct a new BioBrick from synthesized DNA, modify the genes for synthesis for optimal expression in *E. coli* by means of the JCat website tool (http://www.jcat.de/). Make further modifications in accordance with the requirements of the BioBrick standard. This standardization implies that every BioBrick is composed of a DNA sequence of interest preceded by a prefix and followed by a suffix. The prefix and suffix are sequences that contain predefined restriction sites, that are absent in the remaining plasmid sequence. These standard restriction sites make it possible to interchange and extend the BioBrick inserts flexibly \(^9\). The prefix and suffix have the following sequences:
2. Alkane conversion resting cell assay, *in vivo*

This assay was performed based on the method described by Fujii *et al.* (2004)

1. Culture *E. coli* cells expressing the Alkane Hydroxylase system (**BBa_K398014**) and cells carrying an empty vector (**BBa_J13002**) in 5 mL LB medium with appropriate antibiotics over night.
2. Transfer 500 μL of the overnight culture into 50 mL of fresh LB (with antibiotic) and incubate until the cell turbidity reached an OD (optical density) of 0.3 at 600 nm.
3. Centrifuge for 10 minutes at 4,000 rpm and resuspend the pellet in 5 mL 0.1 M phosphate buffer (pH 7.4).
4. Centrifuge again for 10 minutes at 4,000 rpm and resuspend the pellet in 5 mL 0.1 M phosphate buffer now containing E2 salts and 0.66% v/v glycerol (nitrogen-deficient medium).
5. Measure the cell turbidity (OD$_{600}$).
6. Prepare cell-mixture aliquots of 6 mL in 25 mL closed-cap glass flasks and no-cell controls (E2 salts + 0.66% v/v glycerol).
7. Add 100 nmoles of alkane to each flask.
8. Incubate the mixtures at 37 °C for 24 hours (shorten when higher rates are obtained).
9. Measure the OD$_{600}$ after incubation.
10. Extract the hydrocarbons in the culture media by ethyl acetate and determine hydrocarbon concentration in cell culture by gas chromatography (see protocol 3).
11. Calculate the degradation per unit of biomass by dividing the total amount of alkane converted by the total biomass present in each flask (convert OD$_{600}$ to dry weight). Dividing the result by the experimental duration yields the degradation activity per unit biomass per time unit. The average is taken from three individual runs.

3. Alkane conversion enzyme assay, *in vitro*

This assay was performed essentially according to the method described by Li *et al.* (2008).

1. Culture *E. coli* cells expressing the *ladA* gene (**BBa_K398017**) and cells carrying an empty vector (**BBa_J13002**) in 50 mL LB medium with appropriate antibiotics over night.
2. Transfer 500 μL of the overnight culture into 50 mL of fresh LB (with antibiotic) and incubate until the cell turbidity reached an optical density of 0.6 at 600 nm.
3. Centrifuge 10 minutes at 4,000 rpm (4 °C) and resuspend the pellet in 5 mL of 50 mM Tris buffer.
4. Sonicate (Cell disrupter, LA Biosystems) the cells at 40% duty cycle with an output control of 4, keep the solution on ice for the entire duration.
5. Centrifuge the resulting mixture for 5 minutes at 4,000 rpm at 4 °C to remove the cell debris. Transfer the supernatant to a fresh vial.
6. Determine the total protein concentration of the cell extracts by Bradford assay. Note: Use glass vials to prevent the increase of background and/or loss of protein.
7. Prepare a 100 mL mixture containing 0.1% v/v alkane and 50 mM Tris-HCl buffer.
8. Heat the mixture at 100 °C for 5 minutes. Note: Perform this step only for medium-long chain alkanes with high boiling point. (e.g. C16: 287°C).
9. To achieve an optimal solubility of the alkane, sonicate for 1 minute while still warm until a homogenous, viscous mixture is obtained.
10. Add 1mM of NADH, 1mM FMN, 1mM MgSO$_4$ and 0.01 v/v Triton X-100.
11. Prepare 6 mL aliquots in 25 mL closed-cap flasks.
12. Add adequate amounts of cell extract (depends on Bradford assays, final concentration of 5 mg protein/L). Prepare a no-protein control.
13. Incubate at 60°C (for optimal enzymatic activity) for 24 hours.
14. Extract the hydrocarbons in the culture media by ethyl acetate and determine hydrocarbon concentration in cell culture by gas chromatography (see protocol 3).
15. Calculate the degradation per unit of biomass by dividing the total amount of alkane converted by total protein added to each flask (by Bradford calibration curve). Further dividing by experiment duration yields the degradation activity per unit of cellular protein per time unit. The average is taken from three individual runs.

4. Ethyl acetate hydrocarbon extraction and concentration measurements

1. Alkanes are extracted from the aqueous solution by adding 2.5 mL ethyl acetate (apolar solvant) to 6 mL experimental solution. An internal standard is added to the solvent at a concentration of 0.1 % (v/v). The standard (e.g. cyclo-decan) varied depending on the expected range of the peaks of interest.
2. Optional: Add Triton-X100 to the aqueous mixture and centrifuge the samples for 10 minutes at 4,000 rpm in order to get a proper bi-phasic system.
3. Vortex the mixture for 5 seconds (1,500 rpm) and incubate at room temperature until the two phases separate.
4. Remove a maximal amount of the organic layer (top) and dry the solvent using anhydrous magnesium sulphate.
5. Remove MgSO$_4$ by filtration (0.2 micron) and transfer the filtrate into gas chromatograph vials for measurements, or store at -20 °C.
6. Determine the concentration by gas chromatography using a CP-SIL 5CB column (length = 5m). Inject 10 μL of sample in split mode (1:10, 230°C). Set the column gas flow to 1 mL/min (Helium). The following oven temperature program is used:

<table>
<thead>
<tr>
<th>Rate</th>
<th>Temperature [°C]</th>
<th>Time [min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>7.5</td>
</tr>
<tr>
<td>50</td>
<td>90</td>
<td>1.0</td>
</tr>
<tr>
<td>50</td>
<td>110</td>
<td>2.0</td>
</tr>
<tr>
<td>50</td>
<td>130</td>
<td>2.0</td>
</tr>
<tr>
<td>50</td>
<td>145</td>
<td>2.0</td>
</tr>
<tr>
<td>50</td>
<td>160</td>
<td>2.0</td>
</tr>
<tr>
<td>50</td>
<td>170</td>
<td>2.0</td>
</tr>
<tr>
<td>50</td>
<td>185</td>
<td>2.0</td>
</tr>
<tr>
<td>50</td>
<td>210</td>
<td>2.0</td>
</tr>
<tr>
<td>50</td>
<td>250</td>
<td>2.0</td>
</tr>
<tr>
<td>50</td>
<td>320</td>
<td>2.0</td>
</tr>
</tbody>
</table>

7. Integrate peaks and correct the concentrations with respect to the internal standard.

5. Alcohol/aldehyde dehydrogenase activity assay

This assay was performed essentially according to the method described by Kato et al. (2010)

1. Culture E. coli cells expressing the ADH (BBa_K398018) or ALDH (BBa_K398030) gene and cells carrying an empty vector (BBa_J13002) in 50 mL LB medium with appropriate antibiotics overnight.
2. Transfer 500 μL of the overnight culture into 50 mL of fresh LB (with antibiotic) and incubate until the cell turbidity reached an optical density of 0.6 at 600 nm.
3. Centrifuge 10 minutes at 4,000 rpm at 4 °C and resuspend the pellet in 5 mL of 50 mM Tris buffer.
4. Sonicate the cell solution at 40% duty cycle with an output control of 4 (keep the solution on ice during the sonication).
5. Centrifuge the resulting mixture for 5 minutes at 4,000 rpm at 4 °C to remove cell debris.
6. Determine the total protein concentration of the cell extracts by Bradford assay (Note: use a glass vial to prevent protein binding).
7. Load a 96 well plate with 180 μL 57 mM glycine buffer containing NAD (final concentration 1 mM, pH 9.5) in each well.
8. Add 5 μL of the alcohol (aldehyde) to be tested to the wells. Note: Heat long chain alcohols before the start of the assay to have a liquid. For each alcohol (aldehyde) a control without substrate should be added (negative control). In addition, prepare a blank containing a mixture of buffer and the substrate without cell extract.
9. Preheat the plate of the plate reader (Tecan Magellan v7.0) for 15 minutes at 37 °C to allow equilibration of the system.
10. Add adequate amounts of cell extract (depends on Bradford assays, final concentration of 5 mg protein/L). Prepare a no-protein control.
11. Measure the NADH production using a spectrophotometer at a wavelength of 340 nm every 2-3 minutes for 1 hour at 37 °C.
12. Calculate the NADH production rate from the slope of the OD (340 nm). Take into account the light path length and the extinction coefficient of NADH of 6220 M⁻¹ cm⁻¹. Divide the NADH production rate by the total amount of protein added to express the activity of the dehydrogenase reaction in the cell extract (U/mg whole cell protein). Calculate the mean and the standard deviation from three independent runs.

6. pCaiF characterization

1. Culture E. coli cells expressing the pCaiF-GFP construct (BBa_K398331) and cells carrying the promoter alone (BBa_K398326) overnight in 5 mL LB medium the appropriate antibiotics overnight.
2. Inoculate 5 mL M9 containing 10 g/L glucose and antibiotics with 50 μL of the overnight culture and grow overnight.
3. Subculture 50 μL of the overnight outgrowth into 5 mL of fresh M9 with10 g/L and incubate until the cell turbidity reached an optical density of 0.2 at 600 nm.
4. Load a 96 well plate with 100 μL of fresh M9 medium containing the desired amount of carbon source for testing in each well. Perform triplicate experiments for statistical evaluation and add the respective negative controls (wild-type E. coli K12).
5. Add 5 μL of the overnight culture to the medium containing wells.
6. Measure the growth curve (OD₆₀₀) and GFP fluorescence (485 nm excitation and 520 emission) using a plate reader every 10 minutes for 18 hours at 37 °C with constant shaking.
7. Calculate the growth rate and the specific GFP content from the respective measurements and compare to the control. Calculate the mean and the standard deviation of at least three independent experiments.

7. Tolerance assay

1. Culture E. coli expressing PhPFDα and β gene (BBa_K398406) overnight in 5 mL LB medium and the appropriate antibiotics. E. coli expressing ladA gene (BBa_K398017) is used as negative control.
2. Subculture 10 μL of the overnight outgrowth into fresh 5 mL of LB (with antibiotic) and incubate until the cell turbidity reached an optical density of 0.3 to 0.4 at 600 nm.
3. Dilute the cultures with fresh medium until an OD₆₀₀ of 0.1 is reached.
4. Load a 96 well plate with 180 μL M9 medium containing the appropriate antibiotics and the proper final concentration of the toxic compound (e.g. 0, 4, 8, 10% of n-hexane) in triplicate. Because alkane-water mixtures could lead to two-phase systems it is essential to have appropriate controls on the plate (e.g. different strains and the respective blank experiments).

5. Add 20 μL of the culture (control) into the wells.

6. Measure the biomass concentration (OD₅₆₀) using a plate reader every 10 minutes for 24 hours at 37 °C with constant shaking.

7. Calculate the growth rates from the respective measurements for the different agent concentrations and compare to the negative control. Calculate the mean and the standard deviation of at least three independent runs.

8. Homolog Interaction Mapping

1. The application HIM was developed to perform protein queries on a PostgreSQL server running the STRING database (free for academic use) \(^{20}\). The homologue interaction mapping application identifies interacting proteins in the original host organism using the STRING database. The sequences of the respective interacting genes are used in a BLAST search to find homologous genes in the target organism. Cytoscape \(^{4}\) is used to visualize the result of the mapping. The HIM software tool can be downloaded at: https://github.com/jcnossen/InteractionHomologMapping.

2. To perform a mapping, (1) enter the BioBrick ID and the application will automatically download the part sequence data from the Registry of Standard Biological Parts \(^{9}\), or (2) enter (paste) the sequence data in the application.

3. Use the STRING Database website to find the STRING protein ID for the entered amino acid sequence.

4. A protein with high homology is determined using BLAST. Subsequently the application lists each known interacting protein in the source organism and searches for homologs in the host organism (e.g. E. coli).

5. Export resulting putative interaction list to text or Cytoscape.

9. Representative Results

Alkane conversion

The activity of the three oxidation steps from the alkane to the respective fatty acid was evaluated using resting cell assays and enzyme activity measurements. The results are presented following the pathway reactions (1) alkane hydroxylase, (2) alcohol dehydrogenase and (3) aldehyde dehydrogenase.

For the first step, different plasmids were constructed for medium and long-chain alkanes. The plasmid BBa_K398014 contains the four genes of the alkane hydroxylase (AH) system for medium chain alkane oxidation: alkB2, rubA3, rubA4 and rubB under the control of a single constitutive promoter (BBa_J23100). As negative control, E. coli K12 with the plasmid BBa_J13002 was used. The activity was measured by an enzyme activity assay using n-octane as short-chain hydrocarbon source. The consumption of alkane, respectively formation of the corresponding alkanol was analyzed using gas chromatography. Based on hydrocarbon consumption, the specific enzymatic activity of the system was determined. For cell extracts from E. coli K12 carrying BBa_K398014 an enzymatic activity of 4.49×10⁻⁸ U/mg was obtained (1-octanol production from octane). For the negative control an activity of 0.12×10⁻⁸ U/mg was found (Fig. 2B).

The ladA gene from Geobacillus thermodenitrificans was synthesized and cloned downstream of the constitutive promoter BBa_J23100 and expressed in E. coli (BBa_K398027). The enzyme activities of cells transformed with recombinant ladA protein or cells containing the vector control were determined in vitro by a resting cell assay using cell extracts and hexadecane as long-chain hydrocarbon substrate. GC-analysis indicated an enzyme activity of 3.33×10⁻³ U/mg of the E. coli carrying ladA while the activity of the negative control was 0.55×10⁻³ U/mg (Fig. 2B).

The following two oxidation steps were performed using NAD dependent alcohol and aldehyde dehydrogenases (ADH and ALDH genes, BioBricks BBa_K398018, BBa_K398030 respectively). The degradation activity was measured using a resting cell assay with E. coli overexpressing ADH, E. coli overexpressing ALDH and E. coli carrying the vector control (BBa_J13002). For the ADH system, octanol-1 and dodecanol-1 were used as substrates, for the ALDH system octanal and dodecanal. No difference was observed between E. coli transformed with ADH, ALDH plasmids and the control (data not shown). We speculate that the alcohols and aldehydes did not cross the cell membrane in vivo and were therefore not metabolized. The activity of the respective enzyme was also measured in cell extracts (in vitro). Since the alcohol dehydrogenases use NAD as co-factor, the dehydrogenase activity can be determined from the reduction of NAD to NADH \(^{11}\).

For cell extracts of the E. coli control strain, the activity for dodecanol-1 was measured at 0.58×10⁻³ U/mg. The recombinant strain, expressing ADH, had an activity of 1.76×10⁻³ U/mg, which is slightly higher (2-fold) than the wild type activity (Fig. 2B). The expression of ALDH increased the dodecanal dehydrogenase activity of E. coli cell extracts 3-fold compared to the control strain, activities of respectively 1.75×10⁻² U/mg and 0.52×10⁻² U/mg (Fig. 2B) were obtained. This suggests that the recombinant strains E. coli carrying ALDH can functionally convert octanal and dodecanal.

Sensing

The aim of the sensing module is to have an optimal use of cellular resources, e.g. only produce the alkane degradation pathway enzymes when required. To characterize and monitor the activity of the pCaIF promoter, GFP was cloned downstream the pCaIF promoter and the fluorescence signal of GFP in time was measured. E. coli cells with pCaIF-GFP and a control construct containing only the promoter were monitored at different glucose concentrations (2 g/L, 5 g/L and 10 g/L) and substrate mixtures containing lauric acid. At lower glucose concentration, the substrate was depleted more rapidly leading to a decreased growth rate, respectively no growth, after about 8 hours. At the same time, an increase in the GFP signal was observed (Fig. 3B). Strains grown in high glucose medium did not produce GFP, because nitrogen limitation is reached before carbon limitation (data not shown). These results showed that pCaIF promoter is activated at low glucose levels. The result
suggests that the promoter can be used to enable catabolic shifts from glucose to new degradation pathways. To further analyze the pCaiF promoter a mathematical model was derived using the following hypotheses.

1. **cAMP levels are coupled to the carbon influx:**

   It was assumed that the concept of flux sensing, as shown recently by Kotte et al.\(^2\), can be applied to 'measure' the flux for the two carbon degradation pathways respectively glucose and fatty acid degradation. Kremling et al.\(^3\) has shown that scrY that is also dependent on the transcription factor Crp has a hyperbolic Hill relation (with n=6) to the EIIA\(^{Crr}\) ratio. Assuming that the that EIIA\(^{Crr}\) ratio is related to the cAMP levels, we used following relation between cAMP and the uptake flux \(q_{glc}\):

   \[
   \text{cAMP}_{glc}^{\text{cAMP}} = 1 - \frac{q_{glc}^3}{K_{glc,cAMP}^3 + q_{glc}^3}
   \]

   Similarly, a function was defined for cAMP levels in dependence of the lauric acid uptake:

   \[
   \text{cAMP}_{lauric}^{\text{cAMP}} = 1 - \frac{q_{lauric}^2}{K_{lauric,cAMP}^2 + q_{lauric}^2}
   \]

   To combine the two levels, a multiplication was applied: in the presence of only one high flux, the level of cAMP will be close to zero, while in a situation with two low fluxes, the cAMP level will be high:

   \[
   \text{cAMP} = \text{cAMP}_{lauric}^{\text{cAMP}} \times \text{cAMP}_{glc}^{\text{cAMP}}
   \]

2. **Growth Stoichiometry: Yield, maximal uptake rates:**

   The stoichiometric coefficients were taken from Lin et al., with a slight adaptation to enable reaching the observed biomass concentrations\(^4\):

   \[
   Y_{X/glcp} = 0.6 \frac{g_{CDW}}{g_{glc}}
   \]

   Since for lauric acid no explicit value was found, it was assumed to be comparable to glucose:

   \[
   Y_{X/lauric} = 0.5 \frac{g_{CDW}}{g_{lauric}}
   \]

   The maximal uptake rate on glucose was determined by Lin et al.\(^5\) as:

   \[
   q_{glc} = \frac{q_{glc,max} g_{glc}}{K_{glc} + g_{glc}}, \quad q_{glc,max} = 1.3 \frac{g_{CDW}}{g_{lauric}} h, \quad K_{glc} = 0.035 \frac{g}{L}
   \]

   For lauric acid, a hyperbolic uptake relation was assumed and the pathway activity was directly modeled assuming a 'pathway enzyme' \(E_{lauric}\) with a specific activity \(k_{E_{lauric}}\):

   \[
   q_{lauric} = \frac{q_{lauric,max} \frac{g_{lauric}}{K_{lauric} + g_{lauric}}}, \quad K_{lauric} = 0.15, \quad q_{lauric,max} = k_{E_{lauric}} E_{lauric}, \quad k_{E_{lauric}} = 1000 \frac{U}{\mu mol}
   \]

   The growth rate depends on the uptake rate of both substrates:

   \[
   \mu = Y_{X/glcp} (q_{glc} - m) + Y_{X/lauric} q_{lauric}
   \]

   Since in some experiments nitrogen limitation is reached before carbon limitation, growth was additionally assumed to depend on the nitrogen availability:
\[ \mu = \mu N_{lim}, \quad N_{lim} = 1 - \frac{X^2}{X_{0.5max}^2 + X^2}, \quad X_{0.5max} = 0.12 \frac{g_{CDW}}{g_{CDW}} \]

Notice that this represents a simplification of the real mechanism. In the experiments saturation was observed at OD=0.45, which corresponds to about 0.13 gCDW/L. The balances for biomass and substrates then read:

\[ \frac{dX}{dt} = \mu X - k_{lys} X \]
\[ \frac{d\text{glc}}{dt} = q_{glc} X \]
\[ \frac{dlau}{dt} = q_{lau} X \]

3. **Enzyme expression:**

\[ \beta = \beta_{max} \frac{cAMP^{1.5}}{K_{E,cAMP} + cAMP^{1.5}}, \quad K_{E,cAMP} = 0.9, \quad \beta_{max} = 50 \frac{\mu\text{mol}}{g\text{CDW h}} \]

The expression of GFP under the control of pCaiF depends on the intracellular cAMP levels with a hyperbolic relationship (n=1.5, the concentration of Crp was assumed to be constant and not included in the model) as:

The parameters were chosen based on estimates found in Alon et al. \(^2\). With the dilution by growth and estimated protein degradation (\(k_{deg}\)), the balance for the enzyme read:

\[ \frac{dE_{lauric}}{dt} = \beta - (\mu + k_{deg})E_{lauric}, \quad k_{deg} = 0.03 \text{ l/h} \]

4. **Measurement Model:**

The model state variables biomass concentration \(X\) and GFP levels (intracellular) were measured using OD and fluorescence (GFP):

\[ OD = \frac{X}{0.37*0.33} \]
\[ GFP = r_{GFP} \frac{GFP_{intra}}{X} \]

The factor for OD/biomass was taken from Debreczeny and Davies \(^5\). Because the wells were filled with 100 μL a correction for the shorter light pass was used. The response factor for GFP was estimated from parameter optimization.

Data from several experiments was used to parameterize the model described above (parameter estimation problem). The reproduction of the experimental results is shown in Figure 3. Please note that with the same parameters different experimental conditions (10 g/L glucose, 2g/L glucose, no lauric acid) can be reproduced (Fig. 3B). From the model parameters it is concluded that the response of the pCaiF has a 'sharp' response especially to the glucose uptake rate, the PTS system activity. Given that the alkane degradation might reduce the cAMP levels, an auto-regulation is obtained. The enzyme is then produced in amounts required to sustain growth, which is also favorable for the purpose of alkane degradation in as much active cells as possible. More simulation results can be found at the following URL:


**Hydrocarbon tolerance**

To increase hydrocarbon tolerance, genes from various organisms were cloned into a plasmid and expressed in *E. coli*. The selection of genes was accompanied by an *in silico* approach to identify possible unwanted protein-interactions in the host organism. The novel software HIM was applied to search for potential homolog interaction partners of the cloned proteins in the *E. coli* host. We blasted the solvent tolerance protein, prefoldin from *P. horikoshii* in HIM (Fig. 4A), which resulted in a list of interaction partners (Fig. 4B). The proteins depicted in green are *E. coli* homologs of *P. horikoshii* proteins that interact with prefoldin (based on interactions described in the STRING database). The proteins depicted in
red are interacting proteins in *P. horikoshii* that do not have a homolog counterpart in *E. coli*. An interaction with proteins for RNA processing was found in the host and in the target organism, hinting that recombinant prefoldin interacts similarly (Fig. 4B).

We constructed a prefoldin BioBrick containing the phPFDα and phPFDβ genes. This solvent tolerance cluster (BBa_K398406) was expressed in *E. coli* K12. As a control, *E. coli* expressing BBa_J13002-ladA was used. Experimentally, the prefoldin expressing cells were challenged by growth experiments in the presence of different amounts of *n*-hexane. The BBa_K398406 BioBrick clearly improved the growth rate under high *n*-hexane conditions. At 10% *n*-hexane, the *E. coli* strain expressing prefoldin showed a 50% growth rate increase compared to the control (Fig. 5). Since prefoldin is functional in *E. coli*, one could argue that this might be a result of interacting with homologue proteins identified *in silico*, though no proof is available yet (Fig. 4B).

**Figure 1.** The BioBrick assembly method for combining two BioBrick inserts. The donor BioBrick is digested with either EcoRI and SpeI for positioning the donor part downstream of the acceptor part (A) or with XbaI and PstI for positioning the donor part upstream of the acceptor part (B). A third appropriate restriction enzyme is added that cuts the backbone of the donor to improve the success rate of the subsequent ligation reaction. The acceptor BioBrick is digested with either EcoRI and XbaI (A) or SpeI and PstI (B). The BioBrick parts are ligated together forming a BioBrick containing both inserts are again flanked by the 4 standard restriction sites.
A three-step pathway for alkane oxidation was implemented in *E. coli*. (A) Flow chart of the implemented genes. Four genes (*alkB2*, *rubA3*, *rubA4* and *rubB*) of the alkane hydroxylase (AH) system of *Gordonia* sp. TF6 were used to facilitate the oxidation of the medium chain alkanes to their respective alkanols. The *ladA* gene from *Geobacillus thermodenitrificans* was implemented for the conversion of long-chain alkanes. To improve the degradation process for the alcohols or aldehydes, respectively, *ADH* and *ALDH* genes (alcohol and aldehyde dehydrogenase) were expressed in *E. coli*. The resulting alkanoic-acids were metabolized via the host’s β-oxidation pathway. (B) Enzyme activities of alkane conversion proteins for medium- and long-chain alkanes. The biotransformation of *n*-octane was monitored using resting cells assays in *E. coli* K12 carrying the BBa_K398014 plasmid, the BBa_K398017 plasmid, the BBa_K398018 plasmid, the BBa_K398030 plasmid or control plasmid BBa_J13002. The enzyme activities of *ADH* and *ALDH* were tested in cell extracts *in vitro* by monitoring the reduction of NAD to NADH. Error bars represent the SEM of three independent experiments. Click here to view larger figure.
Figure 3. Mechanism and characterization experiment of the pCaiF promoter. (A) Mechanism of the regulation of the pCaiF promoter by cAMP-Crp levels. When the glucose concentration in the environment is high, cAMP levels are low. During depletion the cAMP level increases and so does the concentration of the complex cAMP-Crp. This complex will bind and activate the pCaiF promoter. (B) The biomass concentration (OD) and the fluorescence signal of E. coli K12 carrying pCaiF-GFP with 2g/L glucose. Click here to view larger figure.
Figure 4. Homolog Interaction Mapping. (A) Schematic view of the HIM application. HIM generates a list of putative interactions of proteins in the host organism. Protein sequences can be entered in HIM and the software searches for the corresponding proteins in the STRING database. The application lists each known interacting protein in the source organism and searches for homologs in the host organism, resulting in a map of putative interactions partners for the cloned gene in the host organism. (B) HIM of the prefoldin protein from *P. horikoshii*. The proteins depicted in green are *E. coli* homologs of a *P. horikoshii* protein that interacts with prefoldin, according to the STRING database. The proteins depicted in red are interacting proteins in *P. horikoshii* that do not have a homolog in *E. coli*.

Figure 5. Solvent tolerance. The growth rate (OD$_{600}$) was measured of *E. coli* strains with a periodicity of 10 minutes. *E. coli* K12 expressing the solvent tolerance cluster (BBa_K398406) was tested by different amounts of *n*-hexane (0, 4, 8, 10%). As a control, *E. coli* expressing BBa_J13002-ladA was used. Error bars represent the SEM of three independent experiments.
The BioBrick principle is used to construct a chassis for the degradation of alkanes and a proof of principle for the single components of the toolkit was obtained. Several assays are proposed to measure the in vivo and in vitro activity of alkane degrading pathway enzymes. The presented work successfully demonstrates a number of methods that can be used to determine enzyme activities and expression in the host organism E. coli after implementation of suitable BioBricks. Furthermore, it is shown that the BioBrick principle can be used to design an organism that expresses proteins necessary for the degradation of alkanes, provides a regulatory system that produces these enzymes only when necessary and augments the tolerance of the host organism in the presence of hydrocarbons. Once further developed, this chassis could be used for the biological degradation of residual oil, e.g. in oil sands tailing waters, or for the treatment of wastewater from the oil industry.

With regards to the alkane degradation, assays were performed for the characterization of the enzymes involved in the first step of alkane degradation (AH-system and ladA). It was demonstrated that an E. coli strain carrying the AH-system of Gordonia sp. TF6 was able to convert octane in vivo. This finding is in agreement with the results of Fujii et. al. 8, where biotransformation of n-alkanes using E. coli expressing the minimal component genes of the AH-system was achieved. The conversion activity was measured by a comparative study (wildtype vs. mutant) after a given time. For a full characterization, additional studies have to be performed on the activity of the AH-system over time and the kinetics of the system.

For the conversion of long-chain alkanes (C_{15+}-C_{36}), the ladA gene from Geobacillus thermodenitrificans was implemented. The enzyme activity was tested in vitro using hexadecane as long-chain hydrocarbon source. This modified E. coli strain showed an increased enzyme activity compared to the control strain, demonstrating the recombinant ladA protein enables the conversion of hexadecane by E. coli. In addition to the property of converting alkanes to alkanols, E. coli strains were developed with an improved degradation process for the alcohols and aldehydes by the implementation of ADH and ALDH genes respectively. The enzyme activity in cell extracts was tested in vitro by measuring the production of NADH from NAD +. The E. coli strain expressing ADH showed a two-fold increase in alcohol dehydrogenase activity compared to the wild type activity. Compared to the control strain, the expression of ALDH protein increased the dodecanal dehydrogenase activity in E. coli cell extracts three-fold. Since these assays clearly demonstrated increased enzyme activity in vitro, it can be speculated that the transformed E. coli strains have elevated levels of activity in vivo as well.

To explore the possibility of host-induced expression (for instance: non dependence on induction chemicals and low burden during growth), the degradation pathway was expressed under the control of the pCaIF promoter. This promoter activates gene expression when glucose levels are low, e.g. at the end of a batch growth phase. As a proof of principle, this promoter was tested via the production of GFP under difference glucose regimes. It was demonstrated that the pCaIF-GFP transformed E. coli produced more GFP during stationary (glucose-limited) phase than in the exponential phase (high glucose levels). In the presence of a secondary carbon source (e.g. lauric acid), the GFP production rate decreased again due to the catabolism of the secondary carbon-source.

The experiment showed that this promoter can be effectively used to enable catabolic shifts from glucose to new degradation pathways. This enables to rapidly grow a vast amount of pathways in rich medium before the degradation pathway is activated. We propose to use the pCaIF promoter upstream the AlkS transcriptional regulator. In Pseudomonas putida, AlkS recognizes C_{12}-C_{10} n-alkanes as effectors. Subsequently high levels of AlkS-alkane complex activate the PalkB promoter resulting in the expression of the alkBFGHJL operon encoding proteins involved in the conversion of n-alkanes to fatty acids 18.

Sensing and conversion of hydrocarbons is not enough for the cell to be viable. At increasing hydrocarbon levels in the environment, the growth of wild type E. coli is inhibited through the presence of hydrocarbons in the membrane and in the cell, which can cause protein misfolding. P. horikoshii prefoldin is a chaperone aiding the correct folding of protein in the presence of alkanes 17. Using HIM, it was predicted that this protein interacts with homolog E. coli chaperone proteins, suggesting that overexpression of prefoldin in E. coli could improve solvent tolerance. With the BioBrick BBa_K398406 (consisting of phPFDα and phPFDβ genes), the survivability of E. coli in the presence of alkanes was improved up to 50%. HIM is a suitable tool to select potential functional homologs.

Considering the toolbox as a whole, it can be concluded that it provides a first step in the construction of a chassis for bioremediation of hydrocarbons in aqueous environments. This represents a small, autonomous, self-replicating and relatively inexpensive method. The results were mainly acquired through enzyme assays and shake flash cultures and thus need to be validated on a larger scale. Further research on this system should focus on the coupling of the different enzymes, which have been shown here to work individually in order to generate the envisioned system for hydrocarbon degradation. Furthermore, additional of alternative pathways for the degradation of other pollutants could as well expand the use of this chassis beyond the bioremediation of hydrocarbons alone to a broader scope of pollutants, and possibly the implementation of product pathways could even utilize this system for the production of valuable chemicals. Our results emphasize the suitability of the BioBrick assembly strategy in synthetic biology to construct new organisms that can deal with oil pollution, and may additionally be useful to develop a wide variety of other applications.

Disclosures

No conflicts of interest declared.

Acknowledgements

The experiments performed in this video-article were developed for the international Genetically Engineered Machine competition 7. The authors would like to thank iGEM team members Luke Bergwerff, Pieter T.M. van Boheemen, Jelmer Cnossen, Hugo F. Cueto Rojas and Ramon van der Valk for the assistance in the research. We thank Han de Winde, Stefan de Kok and Esengül Yıldırım for helpful discussions and hosting this research. This work was supported by the TU Delft University Department of Biotechnology, The Delft Bioinformatics lab, TU Delft Department.
References


20. STRING 9.0 Known and Predicted Protein-Protein Interactions. http://string-db.org/newstring_cgi/show_download_page.pl
