View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by OTHES

DIPLOMARBEIT

Titel der Diplomarbeit

A Directed Evolution Approach to Engineering Oxygen Resistant Fe-Fe Hydrogenases

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer.nat.)

Verfasser:	Gerald Grandl
Matrikelnummer:	0407199
Studienrichtung:	A 490 - Molekularbiologie
Betreuer:	Prof. Karl Kuchler, Prof. Pamela Silver

Wien, im November 2010

Contents

I Abstract

This thesis presents an approach to design an artificial selection for hydrogenase function in E. coli, and use it to screen large mutant libraries if Fe-Fe hydrogenases, with the aim of isolating mutants with an increased half-life in aerobic environments.

The biological production of hydrogen by the direct connection of photosynthesis to hydrogenases is a promising option for renewable energy but requires oxygen-tolerant hydrogenases.

A directed evolution approach entails making large libraries of randomly mutated versions of an enzyme and then selecting variants that have an increased functionality for the desired parameter – in our case hydrogenase activity in increased O_2 partial pressures. Directed evolution is the most powerful tool available to protein engineering, provided a system to select for beneficial mutations can be applied.

To that end an E. coli strain has been developed at the Silver lab that is dependent upon an active hydrogenase for cysteine and methionine synthesis. An artificial electron transport pathway shuttles electrons from the hydrogenase to a transgenic Sulfite-reductase, allowing the reduction and incorporation of sulfite, greatly improving growth on selective media (lacking reduced sulfur).

A protocol to generate large plasmid libraries of mutated hydrogenases of different origins using error-prone PCR and screen them in the artificial selection strain at different O_2-levels was developed. Also knock-out strains to increase the selectivity of the screening process were generated and alternative approaches to improve the selectability of hydrogenases in our system as well as to increase the range of selectable oxygen levels were developed. Characterization of putative hits by growth

assays was followed by measuring activities of hydrogenases using both in vitro and in vivo assays in order to characterize promising mutants. A protocol for the anaerobic purification of hydrogenases was also applied.

Several mutants could be isolated from the growth assays, most of which were active but showed reduced activity compared to wild type. Their half-lives in O_2 were not improved either. This suggests that the system we made did not achieve the desired goal of linking growth of E. coli to hydrogenase activity and we were not able to isolate an O_2-resistant hydrogenase. However, as a first attempt to engineer a selection for hydrogenases and one of the first attempts at engineering Fe-Fe hydrogenase at all our approach uses innovative methods and our findings provide valuable insight into the challenges and obstacles to be overcome to ultimately be successful in engineering new function in hydrogenases, as well as biological systems at large.

I.2 Zusammenfassung

Diese Arbeit stellt einen Ansatz vor eine künstliche Selektion für FeFe-Hydrogenasen in E. Coli zu entwerfen, und damit große Bibliotheken mutierter Hydrogenasen auszusortieren, mit dem Ziel Mutanten mit einer erhöhten Halbwertszeit in aerober Umgebung zu isolieren.

Die biologische Herstellung von Wasserstoff durch Kopplung von Photosynthese und Hydrogenasen ist ein vielversprechender Ansatz, der allerdings Sauerstofftolerante Hydrogenasen voraus setzt. Für einen Ansatz mittels gerichteter Evolution müssen große Bibliotheken zufällig mutierter Varianten eines Enzyms hergestellt werden und von diesen dann solche auszuwählen, die eine erhöhte Funktionalität in dem

erwünschten Parameter aufweisen – im Fall diese Projektes eine erhöhte Aktivität von Hydrogenasen in erhöhten Sauerstoffpartialdrücken. Gesetzt den Fall, dass eine Selektion zum herausfiltern erwünschter Mutationen angewendet werden kann ist gerichtete Evolution die leistungsfähigste Methode für die Entwicklung neuer Proteine.

Zu diesem Zweck wurde im Silver -Lab ein E. Coli Stamm entwickelt, dessen Wachstum auf eine aktive Hydrogenase für die Herstellung von Cystein und Methionin angewiesen ist. Eine künstliche Elektronentransportkette transferiert Elektronen von der Hydrogenase zu einer transgenen Sulfit-Reduktase, wodurch Sulfit reduziert und somit verwendet werden kann und das Wachstum unter selektiven Bedingungen (fehlender reduzierter Schwefel) stark erhöht wird.

Ein Protokoll zur Herstellung großer Mutanten-Bibliotheken von verschiedenen Hydrogenasen mittels fehlerhafter-PCR und deren aussortieren in dem synthetischen Selektionsstamm bei verschiedenen Sauerstofflevels wurde entwickelt. Weiters wurden verschiedene knock-out-Stämme um die Selektivität des Sortierungsprozesses zu erhöhen erzeugt sowie alternative Ansätze die Selektivität zu erhöhen getestet.

Nach der Charakterisierung möglicher Treffer aus den Wachstumsuntersuchungen wurden die Aktivitäten der Mutanten mit in vivo und in vitro Ansätzen gemessen. Außerdem wurde ein Protokoll zur biochemischen Aufreinigung der Hydrogenasen angewendet.

Mehrere Mutanten wurden aus den Wachstumsuntersuchungen isoliert. Die meisten wiesen eine stark verminderte Aktivität im Vergleich zum Wildtyp auf. Auch ihre Halbwertszeiten in O_2-atmosphere waren nicht erhöht. Das legt den Schluss nahe, dass das Ziel das Wachstum von E. Coli streng an Hydrogenaseaktivität zu koppeln nicht

geglückt ist und wir daher nicht in der Lage waren, O_2-resistente Hydrogenasen zu isolieren. Dennoch stellt das Projekt, als ein erster Ansatz einer Selektion zur Entwicklung verbesserter Hydrogenasen und einer der ersten Versuche Hydrogenasen weiterzuentwickeln eine sehr innovative Methode dar und unsere Ergebnisse gewähren wertvolle Einsichten in die Herausforderungen und Hindernisse die es zu meistern gilt um letztendlich erfolgreich Hydrogenasen – und biologische Systeme im Allgemeinen – weiterzuentwickeln.

II Introduction

1. Perspectives on Biofuels

The quest to find efficient and sustainable ways to provide clean energy has been going on since the middle of the last century. In my interpretation the industrial revolution was possible because of innovation in engineering that allowed a shift from using bioenergy directly – in the form manpower or animal power – to using bioenergy indirectly – that is using bioenergy from the past – in the form of burning coal or biomass initially or fossil fuels later.

Population as well as energy consumption has been surging as a result of this for a while and alternatives to burning fossil fuels, although heavily sought after, have not been found so far. While the current production of crude oil is sufficient to fuel our needs and worldwide reserves can be expected to last for another few decades the real issue is sustainability and protecting the environment. Photovoltaic solutions seem possible and are heavily researched and debated but the fundamental downside of solar cells is the cost of making them – not only in monetary terms but also in terms of energy and the availability ions necessary for making semiconductors on a large scale. Research into biofuels offers an interesting and promising alternative to photovoltaics for harvesting the energy of the sun.

The biggest challenge in making biofuels lies in the conversion of energy into a fuel molecule and avoiding the production of non-utilizable biomass – the bulk of which is cellulose. Different types of biofuel molecules offer different advantages. Ethanol is fairly easy to make owing to a long history in adaptation of microorganisms to making it but

has a relatively low energy density and is generally made by using a primary source of biomass rich in glucose which greatly decreases the overall energy efficiency of the process. Some success has been made in engineering higher energy biofuels based on fatty acids but the approaches are not refined yet and suffer from the main problem of directing metabolic flux towards the biofuel molecules rather than non-utilizable biomass. (Savage et al., 2009)

Hydrogen is a very popular candidate for a biofuel but it is also a basic chemical for industrial purposes, as in the Haber-Bosch process for making ammonium for fertilizer (Lee et al., 2010). While hydrogen as a fuel is not very attractive from a strategic and logistic point of view – it is energy intensive to make and store, there is no infrastructure in place to distribute it, the attractiveness of hydrogen as a fuel stems from the fact that it has the greatest energy density per mass of all potential fuels and can be converted to electrical or kinetic energy at efficiencies between 50-80%, versus 20-37% via combustion of fossil fuels, and it produces only H_2O as a waste-product (Song, 2002). Hydrogen is however the most challenging candidate as a biofuel, the reasons being a difficulty to direct metabolic flux and the oxygen sensitivity of hydrogenases.

The unifying theme in biofuels is the need for innovation and novel approaches – while the concept of using biology to harvest solar energy as a fuel is very promising and is something that will have to be implemented in some way if we are to solve our energy problems we have to admit that molecular biology and bioengineering are still fairly young fields and biological systems are very complex and somewhat resistant to meddling – new approaches and deeper understanding are necessary.

2. Synthetic Biology

Synthetic biology can be best understood as a new paradigm to biological engineering. At heart synthetic biology aims to be bioengineering done in a more rational, model driven, predictive and standardized fashion than traditional bioengineering. (Endy, 2005). In addition synthetic biology wants to go beyond the current capacity of biology and bioengineering. Rather than just using and shuffling the existing reservoir of biological function one aims at completely redesigning parts of biology – be it networks and circuits or even protein function and metabolism.

Presently the borders to traditional bioengineering are vague – synthetic biology incorporates all the elements of traditional bioengineering such as recombinant DNA technology, optimizing pathways or reactions in the way metabolic engineering has done for a while, or various -omics approaches. However, originating from an area at the interface of biology and electrical- or computer engineering the self declared aim of many synthetic biologists is to ultimately allow the rational design of complete biological circuits or systems to perform specific tasks. To achieve that goal synthetic biology incorporates various elements in addition to traditional bioengineering.

First of all synthetic biology emphasizes modularity, trying to generate and use parts that can be functionally separated and recombined (Agapakis and Silver, 2010). Parts in synthetic biology can be genes as well as plasmids, promoters, terminators or regions coding for functional RNA – in short any DNA sequence that has a functional capacity for the engineer – and part of the modularity endeavor is to make parts with a

normalized and conserved behavior (this is of particular importance for promoters whose strength very often depends on their context) (Canton et al., 2008).

Secondly synthetic biology tries to incorporate mathematical and computational tools to accurately model biological systems. This can have applications also in the basic research area of systems biology, where synthetic biology can be used to test and improve a model. (Mukherji et al., 2009). The primary focus of synthetic biology is however, to get the modeling capacity (of modular circuits) to a level where the output data from the model can serve as a useful proxy for the in vivo characteristics of the actual designed system which can then be built.

Thirdly, although the concept of directed evolution has perhaps not been associated with synthetic biology from the beginning, directed evolution is likely to be a key element in synthetic biology both for the evolution of new, synthetic, proteins or functions as well as for the optimization (or in the best case scenario testing) of circuits of other products of synthetic biology (Haseltine and Arnold, 2007).

The last and perhaps most defining aspect of synthetic biology is the increasing ability to synthesize DNA. Both the length of fragments that can be synthesized as well as the accuracy at which they can be synthesized and assembled have increased tremendously over the last few years while the price or ease of production for synthetic DNA has gone down a lot (Gibson et al., 2010). This of course has profound implications, not only for synthetic biology but for the whole field of biological sciences. While I doubt that our understanding of biology will advance to the point where whole biological systems will be designed in silico, printed as DNA and then put inside an

empty cell to boot a functional cell that makes a biofuel or destroys cancer any time soon, the ability to synthesize large DNA molecules accurately will be very useful for all biologists – be it for the microbiologist that can synthesize genes or whole genomes of an organism that is hard to grow in lab conditions or the neuroscientist that can substitute a week of designing a plasmid and an order at his DNA-synthesis facility for a few weeks or even months of cloning.

In a way synthetic biology is the ultimate goal of biological science. Realizing it in the way described here entails a profound understanding of molecular biological processes and a very integrative and interdisciplinary mindset in the field of molecular biology, enabling big impact bioengineering and furthering our insights into the science – perhaps like physics and electrodynamics did with engineering in the early 20th century.

3. Directed Evolution

Directed evolution is the practice of engineering a protein with desired properties by subjecting it to successive rounds of followed by selection for the desired properties. Given that our current understanding of how precisely a biological structure causes a specific function is limited at best and we don't have a good understanding of how the primary sequence and the cellular context of a protein causes its structure it is very challenging to rationally design a protein with a desired function (Romero and Arnold, 2009). Consequently directed evolution is the most successful and in my mind also most promising protein engineering tool to date.

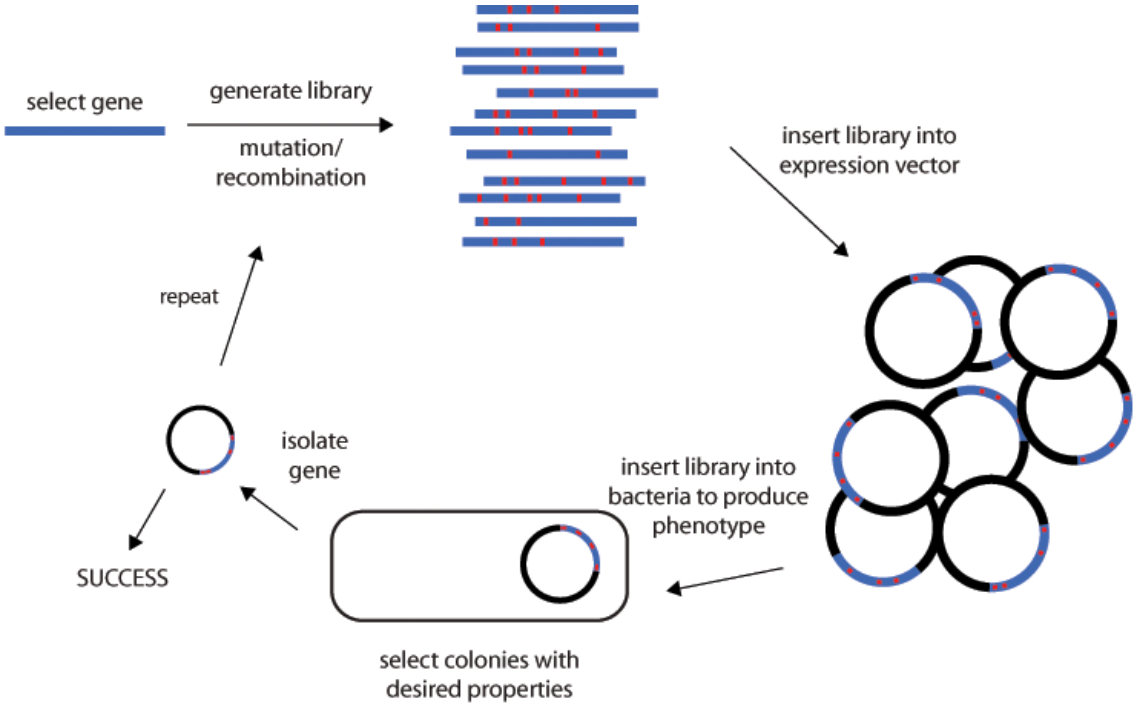

Figure. 1: Work flow of a directed evolution experiment.

Directed evolution has been used extensively in protein engineering. Successful applications include the generation of a recombinase evolved to remove the HIV gene from the host genome (Sarkar et al., 2007), evolving cytocrome P450 to efficiently hydroxylate propane (Fasan at al., 2007), increasing the biostability of lipase A by 40°C (Reetz at el., 2006) as well as the multi-parameter optimization of a metabolic pathway in E. Coli to produce lycopene, an industrially relevant isoprenoid through multiplex automated genome engineering (MAGE) – a clever automation of directed evolution (Wang et al., 2009).

A concept that is central to directed evolution is to think of the evolving protein as moving through a three dimensional fitness space, given by the protein structure and the relative fitness for a specific parameter. By evolving proteins in the lab directed evolution helps to elucidate and understand the way evolution works on a molecular

level because it shows how sequence or function space is sampled, how proteins can move in fitness space or the importance of stability and epistasis effects in the evolution of a specific function.

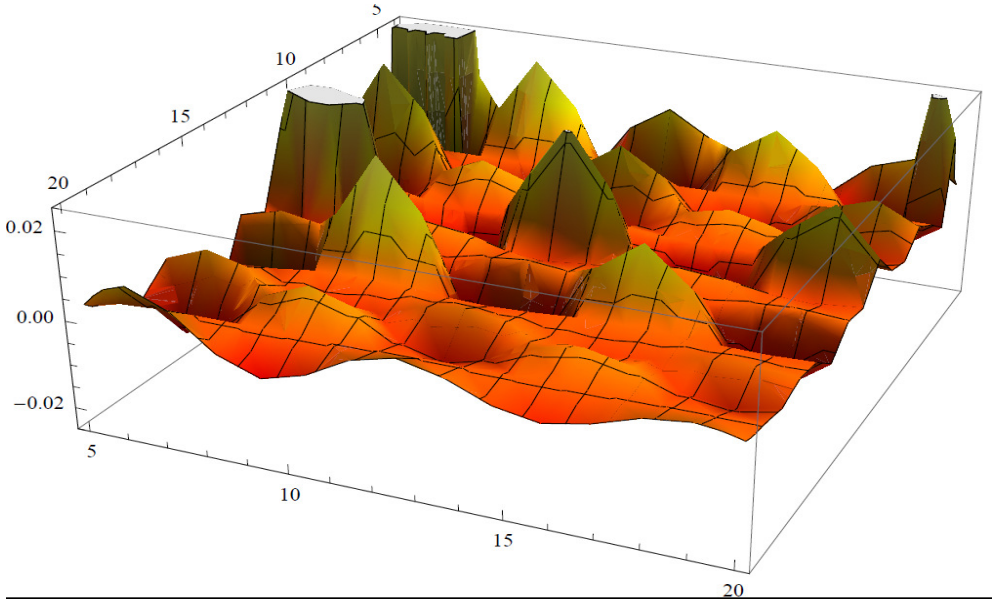

Figure. 2: Graphic rendering of a random Fitness landscape.

For example, data derived from directed evolution studies strongly corroborates the central role that neutral drift plays in evolution (Gupta and Tawfik, 2008). A protein that is able to drift neutrally is able to explore a large sequence and structure space while still maintaining functionality and thus has a higher chance of acquiring mutations that may lead to an additional or changed function. This is especially important if one considers that the majority of mutations that elicit a structural change are likely to be deleterious and that furthermore in order to acquire a radically new function the structure of a protein very likely has to change significantly. Thus, to use the concept of the fitness space as a space of various different functions, rather than climbing one particular fitness mountain, neutral drift enables a protein to sample a large number of different fitness mountains because a partial loss of function does not present a

problem. It is obvious that any meaningful strategy to evolve novel function – be it in the lab or in the biosphere – will increase its likelihood for success by several orders of magnitude if it incorporates neutral drift.

At the level of directed evolution neutral drift ideally means to select for correct protein folding without selecting for a specific function – generating a large, high quality library that does not contain misfolded protein– which is the bulk of large scale mutagenesis protocols.

While generating large libraries of mutated proteins can be trying well established protocols exist for randomly mutagenizing as well as recombining DNA. By far the most challenging aspect of any directed evolution experiment in the lab is how to select or screen for the desired property. Clearly the best way is to do a selection – which means that only the bacteria with the desired property can grow – but this is only possible in rare cases. Alternatively high throughput screening techniques such as FACS (Yang and Withers, 2009) or simple colorimetric screens (see Wang et al., 2009) can be used to find mutants with the desired properties and manual low-throughput screens (see Fasan at al., 2007) have also been successful.

4. Hydrogenases and Bioenergy

4.1 Hydrogenases

Hydrogenases are enzymes that catalyze the reversible reaction of $2e^- + 2H^+ \rightarrow H_2$. They play a crucial role in anaerobic metabolism, being used as an electron acceptor and oxidizing protons in the absence of molecular oxygen as an electron acceptor.

Perhaps as a consequence of this, nearly all known hydrogenases, with the exception of a few NiFe-hydrogenases, are extremely sensitive to oxygen, being irreversibly damaged by it.

Hydrogenases are classified through the ion composition of their active sites. There are nickel-iron [NiFe]-, iron-iron [FeFe]-, mono-iron [Fe]- and a few nickel-iron-selenium [NiFeSe]-hydrogenases (which are of potential interest because some of them may be oxygen tolerant). Because of their rather complex chemistry hydrogenases generally require intricate, multi-step maturation involving multiple protein- and non-protein co-factors in order to assembly their active sites.

[FeFe]-hydrogenases are typically small (around 48kD), often have relatively few maturation factors, which, along with their high catalytic rates and the fact that a lot of them bind ferredoxin as an electron carrier, makes them the most promising hydrogenases to engineer.

4.2 Photosynthetic Hydrogen

The bulk of reduction power used in biological systems comes from H_2O through photosynthesis. Therefore, considering hydrogen production through biotechnology it seems a very promising strategy to use this reduction power (and energy) directly at the photosystem, where it is created, rather than much later through fermentation of Glucose or other high energy carbohydrates. The most straight forward way to do so would be to simply add a hydrogenase that accepts electrons from ferredoxin – the universal electron carrier that links the photosystem to the rest of metabolism – to the

system. (Brenner et al. 2006). From a thermodynamic point of view algal or bacterial FeFe-hydrogenases would be ideal but they are irreversibly damaged even by small quantities of O_2.

The concept has been shown to be functional and make H_2 but proofs to be completely unfeasible because of the deleterious effect of the produced O_2.

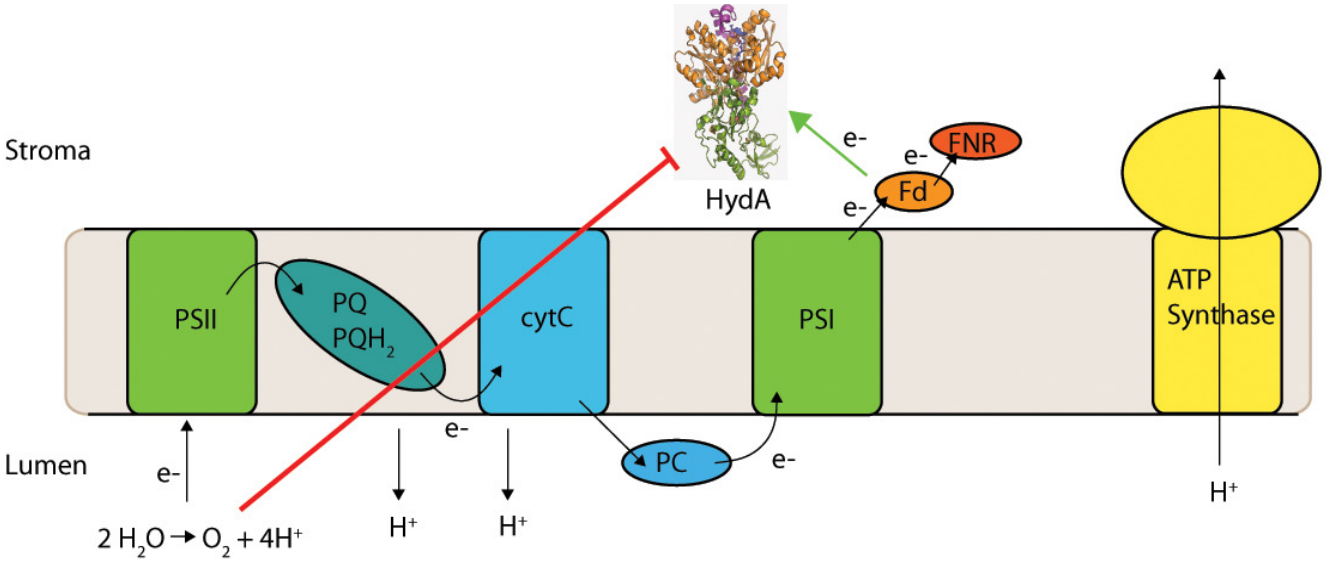

Fig. 3: Schematic of a possible Hydrogen generation pathway through photosynthesis: Electrons from H_2O reduce ferredoxin which can reduce an orthogonal FeFe-hydrogenase to generate H_2. The approach is unfeasible because all known FeFe-hydrogenases are extremely oxygen-sensitive.

4.3 [FeFe]-hydrogenases

[FeFe]-hydrogenases play a crucial role in microbial metabolism. While they catalyze the reversible reaction of reducing two protons to H_2 they are typically committed to run in the direction of H_2 evolution in their cellular environment. They are mostly monomeric, in the cytoplasm, with a molecular weight ranging from 45 to 130kDa (Tard and Pickett 2009). This, along with the fact that they have very high catalytic rates for H_2 evolution makes them ideal candidates for bioengineering purposes.

The exact mechanism of hydrogen evolution by [FeFe]-hydrogenases, or their inactivation by oxygen is not well understood but over the last years tremendous advances in their research have been made. The first structure was solved for *clostridium pasteurianum* HydA (Peters et al., 1998). The structure, resembling a mushroom shape, consists of 4 subunits; the largest one termed the catalytic subunit making the head and the three smaller subunits making up the stem of the structure. The active site, called H-cluster, contains six Fe atoms arranged as a [4Fe-4S] cubical iron-sulfur cluster linked to a [2Fe2S]-complex by a single cysteinyl S. The cubical [4Fe-4S]-cluster is coordinated to the enzyme through four cysteins, one of them linking it to the [2Fe2S]-complex. In addition to this bridging cysteinyl S the [2Fe2S]-complex is completed by three CO and two CN molecules – biologically unique non-protein ligands – as well as a dithiolate ligand.

Within the hydrogenase the H-cluster is located in the center of the catalytic subunit. The stem contains three accessory cubical [4Fe-4S]-clusters and a ferredoxin binding domain. The route of electrons through the enzyme is from the ferredoxin via the three accessory [4Fe-4S]-clusters to the H-cluster. The site of H2 evolution is believed to be the distal Fe atom, in the [2Fe-2S]-cluster of the active site, which gets charged with the two electrons. A putative proton shuttle consisting of polar amino acid residues and water molecules ends at the active site close to the distal iron. There is a hydrophobic gas channel at the top of the enzyme connecting the active site to the cytoplasm.

Recently a lot of headway has been made in understanding the maturation of the active site. While the [4Fe-4S]-clusters containing cystein residues can be synthesized by the

general Fe-S maturation host-machinery the [2Fe-2S]-cluster containing CO and CN as co-factors has to be matured by specialized enzymes. In most [FeFe]-hydrogenases the synthesis of this cluster requires only three maturation factors, HydE, HydF and HydG. Structural analysis of hydrogenases in an HydEFG knockout background suggest that the assembly of the active site is stepwise, assembly of the [2Fe-2S]-cluster following the prior assembly of the active site [4Fe-4S] cluster at the base of a transient assembly channel that disappears once mature conformation is reached (Mulder et al. 2010). Figure 4 depicts the Fe-Fe hydrogenase active site and whole structure.

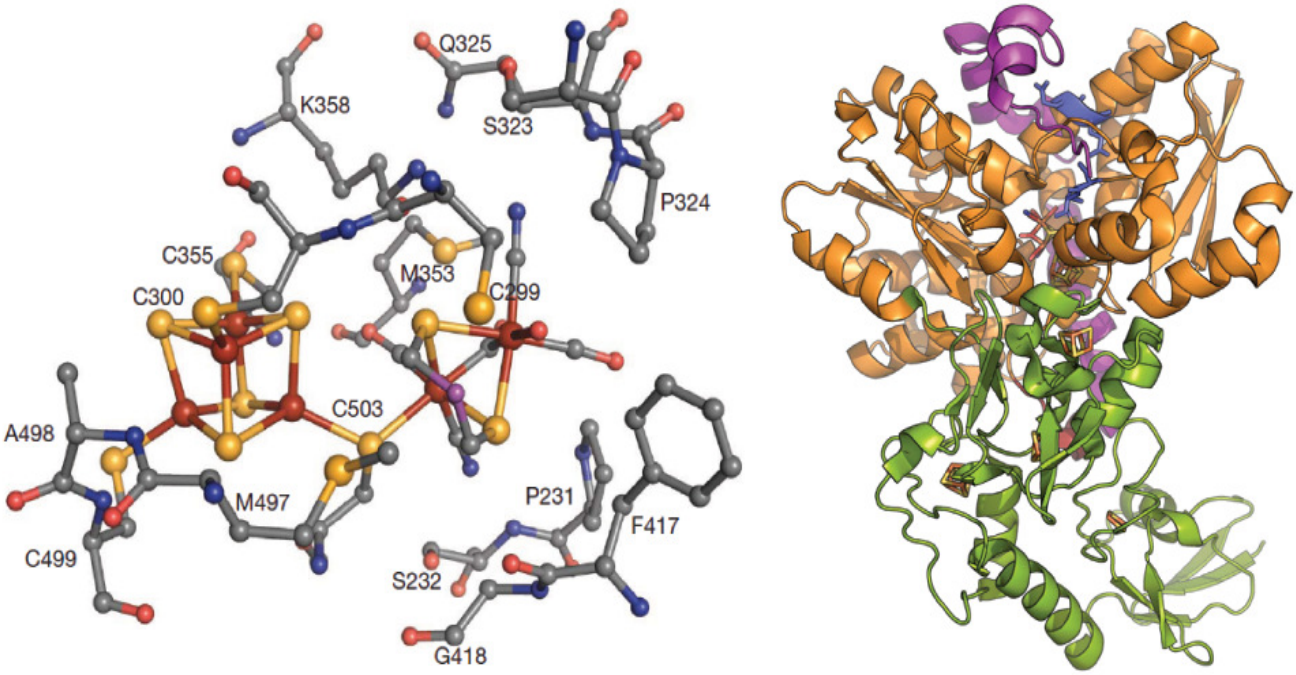

Fig. 4: Left side: H-cluster with the cubical [4Fe-4S]-cluster coordinated by the four cysteins and the [2Fe-2S]-cluster coordinated with the CO, CN and dithiolate groups. Black: Carbon, Red: Iron, Yellow: Sulfur, Blue: Nitrogen, Light Red: Oxygen, Magenta: Central Atom of dithiolate group (Mulder et al. 2010). Right side: Structure of the complete CpI HydA highlighting the positions of the three iron-sulfur clusters as well the H-cluster. The catalytic subunit is orange; the stem structure with the ferredoxin-binding domain at the left is green. Residues is blue are the putative proton shuttle (image made in PyMol, courtesy of Buz Barstow). A putative gas channel is located at the top of the enzyme left of the proton shuttle, connecting the active site to the cytoplasm.

The mechanism of degradation of the active site by O_2 is poorly understood. Interestingly X-ray absorption spectroscopy experiments suggest that the [2Fe-2S]-cluster remains intact while the [4Fe-4S]-cluster gets destroyed. A current model is that oxygen binds to the distal Fe of the [2Fe-2S]-cluster. Consequently it gets reduced by the active site and forms reactive oxygen species (ROS). The ROS can then destroy the proximal [4Fe-4S]-cluster either by migrating the short distance or by remaining bound at the distal Fe and oxidizing the cluster via through electron transfer.

5. Aim of the project

From the previous chapters it becomes clear that engineering FeFe-hydrogenases to become oxygen resistant is a very challenging goal, but would have enormous impact both from a scientific and economic standpoint. From what we know about the active site we can assume that a metal-cluster containing iron and sulfur in this way, due to its simple electrochemical properties, will always react with oxygen in some way. Therefore we presume that engineering oxygen resistance in FeFe-hydrogenases can only be achieved by somehow protecting the active site from O_2 – either by sterically preventing O_2 from entering the active site, or perhaps by somehow scavenging the O_2 intramolecularly before it can reach the active site.

From a protein engineering standpoint it seems clear that any attempt at rationally designing changes in FeFe-hydrogenases are prone to failure because of our underlying ignorance of the precise involved biophysics and protein maturation. Therefore we decided to try a directed evolution approach to engineering

hydrogenases. In terms of high throughput the best strategy for directed evolution, particularly of hydrogenases, where a rapid screening technique for functionality does not come to mind, is a selection, where desired traits of the enzyme can be selected for at the level of growth or survival of the host. The fundamental difficulty with this is linking growth of the host to functional hydrogenases.

An artificial electron transport pathway was developed at the Silver lab to achieve that goal. The idea is to introduce a reaction that is necessary for growth and depends on the presence of a functional hydrogenase. H_2 is not a molecule that any commonly grown lab-strains depend upon. It does however play a – potentially underestimated – role as short term electron sink in multiple anaerobic pathways. Furthermore FeFe hydrogenases naturally interact with various ferredoxins and so it seems a promising approach to link the growth linked reaction to the hydrogenase via ferredoxin.

Sulfite reductases are enzymes that catalyze the reduction of SO_3^{2-} (sulfite) to HS^- (hydrogen sulfide ion) or H_2S (hydrogen sulfide), depending on the pH, with different enzymes using NADPH, ferredoxins or flavins as electron donors. They play an important role in sulfur metabolism, both in bacteria or in sulfur assimilation in plants providing reduced sulfur that can enter metabolism.

E. coli does not naturally express a sulfite reductase able to reduce SO_3^{2-} and can therefore not grow on media that does not contain a source of reduced sulfur. A *Zea mays* (corn) sulfite reductase that uses ferredoxin as an electron donor was chosen to create a synthetic electrochemical pathway to reduce sulfur and allow growth on media with sulfite as the only sulfur source. The pathway uses electrons provided by H2 via a

FeFe-hydrogenase that are transferred to sulfite reductase via a *Spinachia oleracia* (spinach) ferredoxin. Figure 5 is a schematic of the pathway.

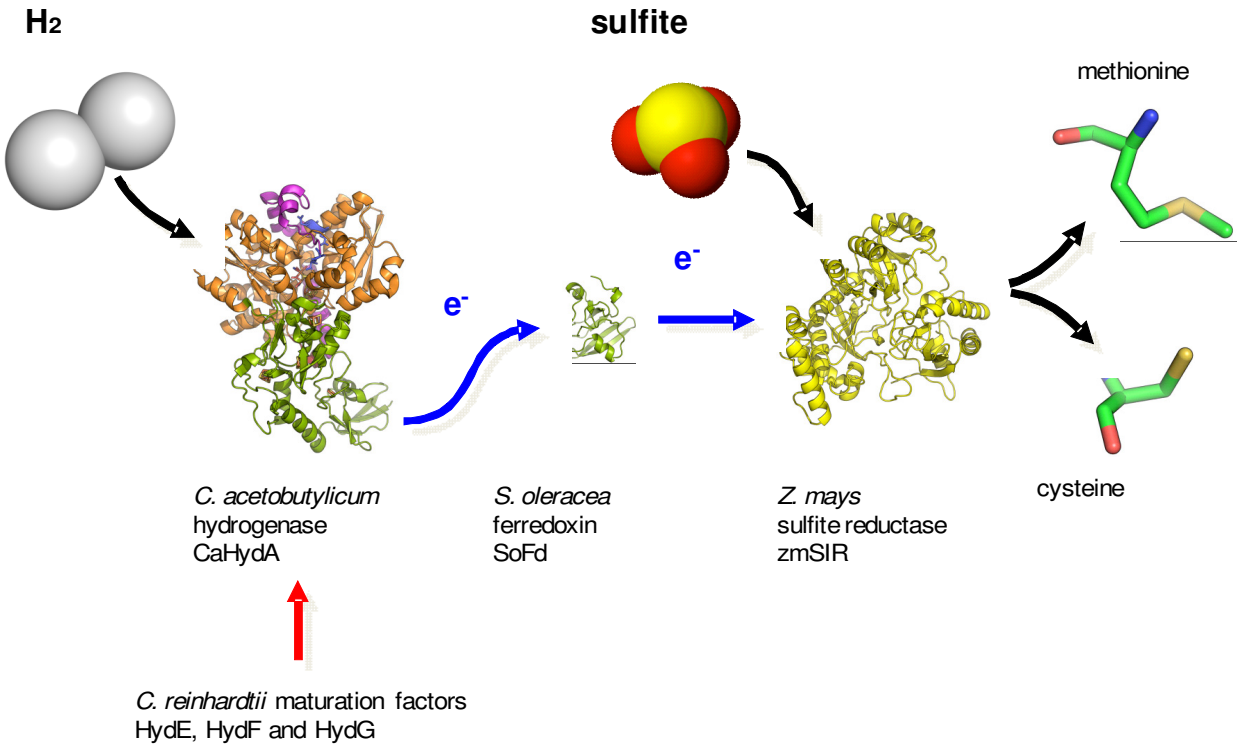

Fig. 5: Illustration of the synthetic sulfite reduction pathway in E. coli. Electrons flow from H2 to a corn SIR via a chlostridial hydrogenase and a spinach ferredoxin, reducing SO_3^{2-}, thereby providing reduced sulfur, allowing growth on media lacking an adequate sulfur source. Graphic courtesy of Buz Barstow.

This pathway, although it is completely orthogonal in E. coli, allows growth on media with only SO_3^{2-} as a sulfur source in anaerobic conditions. Indeed the initial problem with the pathway was not a lack of electrons entering the pathway and reducing sulfur but rather a lot of background interactions providing electrons to the pathway from different sources. As a result growth was possible with and without the transgenic hydrogenase. This did not come as a huge surprise, since E. coli possesses several hydrogenases on its own, as well as various other enzymes that might interact with the pathway. These hydrogenases as well as other potential interaction partners, based on structural or functional homology, where knocked out and produced an E. Coli strain that did indeed

show a markedly increasing growth rate in the presence of a transgenic hydrogenase and hydrogen as depicted in figure 6.

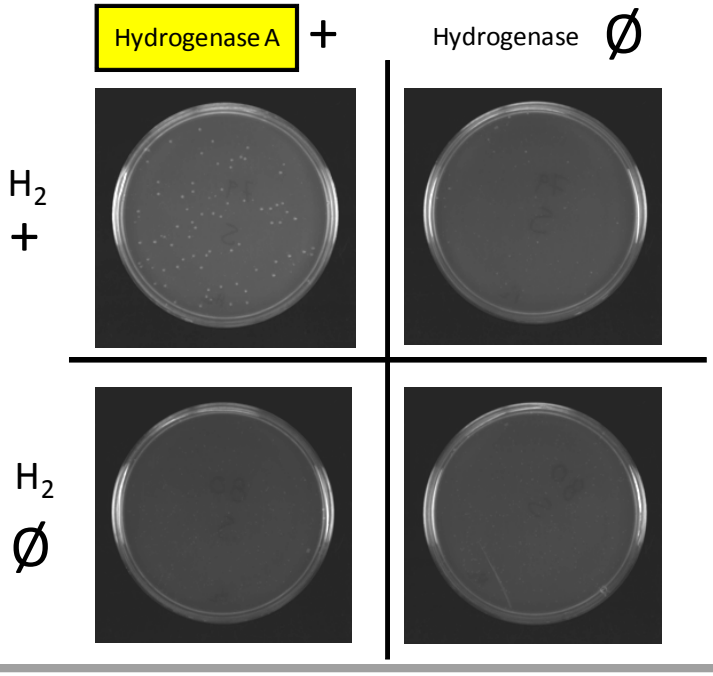

Fig. 6: Growth of the knock-out strain ko32 at three days under anaerobic conditions with and without H2 and hydA. Image and data courtesy of Edwin Wintermute.

It should be noted that this pathway does not constitute a true selection where survival and growth is strictly dependent on the presence of the transgenic hydrogenase. As seen in the photos there is some level of background growth on all the plates, with very small colonies after three days growth in anaerobic conditions. It is also clear that the cells containing hydA and hydrogen grow much better.

The aims of the project were to try to further isolate this pathway for use as a synthetic hydrogenase selection, to establish protocols to make large mutant libraries of hydrogenases to test if the pathways can be used to select for active hydrogenases at various partial pressures of O2 and to analyze and characterize mutant hydrogenase that come out of the screen.

III Materials and Methods

1. Bacterial Strains and Culture conditions

1.1 Strains

Bacterial strains used were DH5a as a host strain for cloning and BL21 for protein expression as well as BL21 derived knock out strains for the selection. DH5a and BL21 $T1^R$ (resistant to phage T1) were generated by P1-transduction from the Keio collection. Table 1 shows a list of all used strains.

Table 1: List of used bacterial strains. For BL21 derived knock-out strains only the knock-outs are listed in the genotype.

Strain	Genotype	Source
DH5α	F^-, φ80d*lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *deoR*, *ΔphoA8*, *glnV44*(AS), *recA*1, *endA*1, *hsdR*17(rk^-, mk^+), *phoA*, *supE*44, λ^-, *thi*-1, *gyrA*96, *relA*1, $T1^R$	Invitrogen
BL21	F–, *omp*T, *hsd*S_B(r_B-, m_B-), *dcm*, *gal*, rne131, λ(DE3)	Invitrogen
BL21 $T1^R$	F–, *omp*T, *hsd*S_B(r_B-, m_B-), *dcm*, *gal*, rne131, λ(DE3), $T1^R$	this study
BL21 derived		
ko22	ΔhycE, ΔhybC, ΔhyaB	E. Wintermute
k32	ΔcysI, Δfpr, ΔydbK, Δhcr, ΔyeaX, ΔhcaD, ΔfrdB, ΔhycE, ΔhyaB, ΔhybC, hyfG::kan	E. Wintermute
ko32	ΔcysI, Δfpr, ΔydbK, Δhcr, ΔyeaX, ΔhcaD, ΔfrdB, ΔhycE, ΔhyaB, ΔhybC, ΔhyfG	this study
k33	ΔcysI, Δfpr, ΔydbK, Δhcr, ΔyeaX, ΔhcaD, ΔfrdB, ΔhycE, ΔhyaB, ΔhybC, ΔhyfG, fre::kan	this study
ko33	ΔcysI, Δfpr, ΔydbK, Δhcr, ΔyeaX, ΔhcaD, ΔfrdB, ΔhycE, ΔhyaB, ΔhybC, ΔhyfG, Δfre	this study
Jc104	k32 with plasmids D108 and E113, Kan^R, Amp^R, Chl^R	E. Wintermute

Selection strain

Several knock-outs were performed by E. Wintermute and C. Agapakis in order to insulate the synthetic SIR-pathway prior to starting the screening. Genes were knocked out based on structural homology or shared binding motifs with the SIR. Knock out strains were generated sequentially by P1-transduction from the Keio collection and subsequent flipping out of the Kan^R-cassette. The nomenclature is k(number) for strains containing the Kan^R-cassette in the last gene and ko(number) for strains with the flipped out cassette. Table 2 lists the genes knocked out and their function. k32 with the plasmids a4 and E113 (carrying hydEFG and ZmSIR), termed Jc104, was the selection strain used during most of the experiments.

Table 2: List of knocked-out genes

Knock-out	Gene	Keio strain	Gene Function
ko32	Δcysl	JW2733	sulfite reductase, beta subunit, NAD(P)-binding
	Δfpr	JW3895	ferredoxin-NADP reductase
	ΔydbK	JW1372	fused predicted Fe-S subunit of pyruvate-flavodoxin-oxidoreductase
	Δhcr	JW5117	NADH oxidoreductase hcr
	ΔyeaX	JW1792	predicted oxidoreductase
	ΔhcaD	JW2526	phenylpropionate dioxygenase, ferredoxin reductase subunit
	ΔfrdB	JW4114	fumarate reductase, Fe-S subunit
	ΔhycE	JW2691	formate hydrogenlyase, subunit 5
	ΔhyaB	JW0955	hydrogenase 1
	ΔhybC	JW2962	hydrogenase 2, large subunit
	ΔhyfG	JW2472	hydrogenase 4, subunit
k33	fre::kan	JW3820	flavin reductase

1.2 Media

List of the used media. All media were prepared in miliQ H_2O and autoclaved for 45 min prior to use. For plates 1.5% (w/v) agar was added. LB was used for all standard purposes, SOC was used for recovery

LB (per L)

Tryptone	10 g
Yeast Extract	5 g
NaCl	10 g

SOC (per L)

Tryptone	20 g
Yeast Extract	5 g
5M NaCl	2 ml
1M KCl	2.5 ml
1M MgCl2	10 ml
1M MgSO4	10 ml
20% (w/v) Glucose	20 ml

Selection Media 1L

Selection media consists of a standard M9 minimal media with IPTG, ferric citrate and sulfur dropout powder

5x M9 salts	200ml
1L: $Na_2HPO_4 . 7H_2O$	64 g
KH_2PO_4	15 g
NaCl	2.5 g
NH_4Cl	5 g
1M $CaCl_2$	0.1mL
1M $MgSO_4$	10 mL
Ferric citrate	25 mg
Sulfur-dropout Powder	0.5g
20% Glucose	100mL
1M IPTG	1mL

Sulfur dropout powder

A standard rich amino acid mix without cysteine and methionine.

Amino Acids
Alanine - 2.0 g
Arginine - 2.0 g
Asparagine - 2.0 g
Aspartic acid - 2.0 g
Cysteine - 0.0 g
Glutamic acid - 2.0 g
Glutamine - 2.0 g
Glycine - 2.0 g
Histidine - 2.0 g
Isoleucine - 2.0 g
Leucine - 10.0 g
Lysine - 2.0 g
Methionine - 0.0 g
Phenylalanine - 2.0 g
Proline - 2.0 g
Serine - 2.0 g
Threonine - 2.0 g
Tryptophan - 2.0 g
Tyrosine - 2.0 g
Valine - 2.0 g
Nucleotide Bases
Adenine - 0.5 g
Uracil - 2.0 g
Vitamins
p-Aminobenzoic acid - 0.2 g
Inositol - 2.0 g

Buffered LB for Hydrogenase Assay and Half-Life Measurements (1L)

Tryptone	5 g
Yeast Extract	2.5 g
Sodium chloride	5 g
Monopotassium phosphate	1.2 g
Dipotassium phosphate	7.2 g
20% Glucose	2.5 mL
Ferric citrate	12.5 mg
IPTG Stock 1M	1 mL
Baker's Antifoam B	50 µL

Terrific Broth for Protein Purification (per L)

Tryptone	12 g
Yeast Extract	24 g
Glycerol	4 mL
Phosphate buffer	100 ml

prepare separately: 0.17M KH2PO4 and 0.72M K2HPO4 in milliqH_2O, autoclave

Antibiotics

Antibiotics were used for selection for up to 4 plasmids at a time. Antibiotics were added after autoclaving and cooling the media down in the following concentrations: Ampicillin (Amp) 100 µg/ml, Kanamycin (Kan) 50 µg/ml, Spectinomycin (Sm) 50 µg/ml, Chloramphenicol (Cm) 25 µg/ml.

1.3 Culture Conditions

Cells were grown at 37℃. Liquid cultures were grown at 37℃ on a shaker with 250rpm to provide aeration. For anaerobic growth cells were backdiluted, sparged for at least 10min with N_2, Argon or 10%H_2 in N_2, capped with air tight rubber seals (suba-seal cap, Sigma-Aldrich, St. Louis, MO, USA) and grown as usual. Cells were stored either at 4℃ or at -80℃ in 50% glycerol.

1.4 Gas

Gas and gas mixtures where purchased from Airgas Inc. Gas mixtures used were H_2, N_2 and Argon (pure), as well as 10%O_2 in N2 10%O_2 in H_2 and 10% H_2 in N_2.

1.5 Plasmids

Below all the plasmids used in this study are listed. The majority of the work was done in vectors of the Duet-series that allow co-expression from two different multiple cloning sites (MCS) under an IPTG inducible T7-promoter. In accordance with the BioBrick standard used at the Silver lab the vectors were mostly used with modified MCS that accept Silver-biobricks.

Table 3: List of plasmids used

Name	Based on	Description	Marker	Source
pACY-Duet	-	2 MCS under IPTG inducible T7 promoter	Cm^R	Novagen
pET-Duet	-	2 MCS under IPTG inducible T7 promoter	Amp^R	Novagen
pCDF-Duet	-	2 MCS under IPTG inducible T7 promoter	Sm^R	Novagen
pCOLA-Duet	-	2 MCS under IPTG inducible T7 promoter	Kan^R	Novagen
v23	pACY	Biobricked Duet vector	Cm^R	E. Wintermute
v24	pET	Biobricked Duet vector	Amp^R	E. Wintermute
v25	pCDF	Biobricked Duet vector	Sm^R	E. Wintermute

v26	pCOLA	Biobricked Duet vector	Kan^R	E. Wintermute
D106	v25	Carrying CrHydA	Sm^R	E. Wintermute
D108	v25	Carrying CaHydA	Sm^R	E. Wintermute
D109	v25	Carrying CsHydA	Sm^R	E. Wintermute
a4		Carrying *C. ralstonia* maturation factors HydEFG	CmR	C. Agapakis
E113		PFOR		
P1	V25	Carrying n-term His-tagged CaHydA	Sm^R	this study
P2	V24	Carrying c-term His-tagged CaHydA	Amp^R	this study
705-FLP	pSC101	Carrying the FLP-recombinase; T-sensitive ori and promoter	Cm^R	Silver lab
V0120		Silver lab cloning vector, MCS without promoter	Amp^R	Silver lab
DFS656		Carrying *Zymomonas Mobilis* ADH	Kan^R	D. Savage

1.6 Oligonucleotides

All oligonucleotides were purchased from Integrated DNA Technologies (IDT) Inc., Coralville, Iowa. Below is a list of oligonucleotides used in this study.

Table 4: List of oligonucleotides used.

Name	Description	SEQUENCE
JW.376	error prone PCR CrHydA F	GCTGCACCAGCCGCAGAAGCTCCTTT
JW.377	error prone PCR CrHydA R	GCGACTCTCAATGTTATGCCGCCCC
JW.382	error prone PCR CaHydA F	GGAGATATACCATGGGCGCGGCCGCATCTAGAATG
JW.383	error prone PCR CaHydA R	CGATTATGCGGCCGTGTACAATACGATTACTTTCTGTTCGAC
JW.380	error prone PCR CsHydA F	GGAGATATACCATGGGCATAAACATAGTAATTGATG
JW.381	error prone PCR CsHydA R	TTATGCGGCCGCTTATTTTTCAAACTGCGG

JW.395	fre KO locus F	GGCTGGGTTCAAAAATGGGGCTGG
JW.396	fre KO locus R	CGCCTTATCCGGCCTACGGTTCGG
JW.254	Keio Kan^R R	CGGCCACAGTCGATGAATCC
JW.261	T7 locus F	TGGTCTTCGTGGCATAAGGA
JW.262	T7locus R	GTAAACACGACCGCGCCAGT
JW.384	CsHydA.seq.1	ATGGGCATAAACATAGTAATTGATG
JW.385	CsHydA.seq.2	CCTTGTCCAAGAAGAGAAAATTGCG
JW.392	CsHydA.seq.9	CTTTCAATAAACTCTGTTGCTTCTTCC
JW.389	CsHydA.seq.6	GCAAACTTAGCTGAATTCATGAATAGCGG
JW.390	CsHydA.seq.7	CTCCACGCGCTCGGCGCGCC
JW.384	CsHydA.seq.1	ATGGGCATAAACATAGTAATTGATG
JW.374	CsHydA colony F	TTAATAAGGAGATATACCATGGGC
JW.375	CsHydA colony R	TTACTTTCTGTTCGACTTAAGTTAAC
JW.353	CaHydA.seq.F2	TCAAAATGTGTACTATGCGG
JW.354	CaHydA.seq.F3	AAGCTACTGAACTTTTAGGC
JW.355	CaHydA.seq.F4	GGTGCTGGAGCTATCTTTGG
JW.356	CaHydA.seq.R1	GGTAAAAATGGTTTTGAAGCTTTTGC
GG.37	CaHydA n-His TEV F	AACCATGGCCCATCACCATCACCATCACGAAAAC CTGTATTTCCAGGGAGGCGCGGCCGCATCTAGA ATGGG
GG.20	CaHydA n-His R	GCGGCCGTGTACAATACGATTACTTTCTGTTCGA CTTAAGTT
GG.59	CaHydA c-His F	TTCGCCATGGAAAACAATAATCTTAAATGGCAAT GAAGT
GG.58	CaHydA c-His R	CCACTAGTCTAGTGATGGTGATGGTGATGTTCAT GTTTTGAAACATTTTTATCTTTTGT

2. Molecular Biology Techniques

2.1 DNA purification & Plasmid Prep

DNA was purified using the plasmid purification kit from QIAGEN according to the manufacturers' instructions. For purification of plasmids the QIAGEN plasmid miniprep or fastmidiprep kits were used. For column purification of DNA after PCR or digestion the QIAquick PCR purification kit was used. For gel purification the QIAquick Gel Extraction Kit was used.

2.2 DNA Gel Electrophoresis

DNA gel electrophoresis was performed on agarose gels containing 0.8 to 1.5% (g/v) agarose. Gels were prepared in TAE-buffer (40mM Tris, 1mM EDTA, 0.35% (v/v) acetic acid). 10 µg/ml ethidium bromide was added to the gels to visualize the bands in UV light. Gels were run at a voltage of 90-150V in TAE-buffer.

2.3 PCR

Various polymerases were used for PCR. Standard PCRs for cloning were done using the Finnyzmes Phusion High-Fidelity PCR Kit (NEB). Whole plasmid PCRs were done either with Phusion or with Platinum Pfx from Invitrogen. Colony PCRs were performed using a Taq Polymerase Master Mix from QIAgen. For colony PCR template preparation colonies were picked into 100µl ddH_2O and boiled for 5 min. Error prone PCR was performed using the Mutazyme II Polymerase mix from Genemorph.

2.4 Restriction and Ligation

For DNA-digestion restriction enzymes were from the Fermentas FastDigest series were used according to instructions – digestion times varied from 5min to overnight. Dephosphorylation was done using FastAP from Fermentas or the Roche DNA dephosphorylation and ligation Kit. For ligations T4 DNA ligase from NEB or the Rapid DNA dephosphorylation and ligation Kit from Roche were used. Typically ligations were done using a 1:3 ratio of vector to insert.

2.5 Transformation of E. coli

Chemically competent cells were prepared with the CaCl2 method. Cells were grown to an OD of around 0.5, harvested by centrifugation and incubated with ice cold 100mM $CaCl_2$ o/n at 4℃ and stored 100mM $CaCl_2$ + 15% glycerol at -80℃.

Cells were transformed by adding around 100ng of plasmid DNA, heat shocking for 30sec, placing the cells on ice for 2min, recovering in SOC with agitation and plating on selective media. Recovery times varied from a few minutes (Amp) to 1 hour (Cm). For multiple transformations up to 100ng per plasmid were used and recovery times were extended to up to 3 hours.

2.6 Sequencing

Sequencing was performed by Genewiz Inc. Sequencing samples were prepared according to manufacturers' instructions. For Sequence analysis software from the DNAStar Lasergene package was used.

2.7 Knock outs in E. Coli

E. Coli knock out strains were generated using the P1-transduction method to insert knockout-cassettes from strains from the Keio collection. The Keio collection is an open repository of deletion strains of all nonessential genes in *E. coli.* Each deleted gene is replaced with a kanamycin resistance cassette that can be excised by FLP recombination. The strain background is E. coli K-12 BW25113 (Baba et al., 2006). The strains from the Keio collection are available via the Coli Genetic Stock Center (CGSC) at Yale University. The P1-transduction uses the bacteriophage P1 to randomly package bacterial chromosomal DNA and transfer it to a new strain upon infection. P1 requires Ca^{2+} to infect cells and can therefore easily be removed from a culture by using a Ca^{2+} chelator.

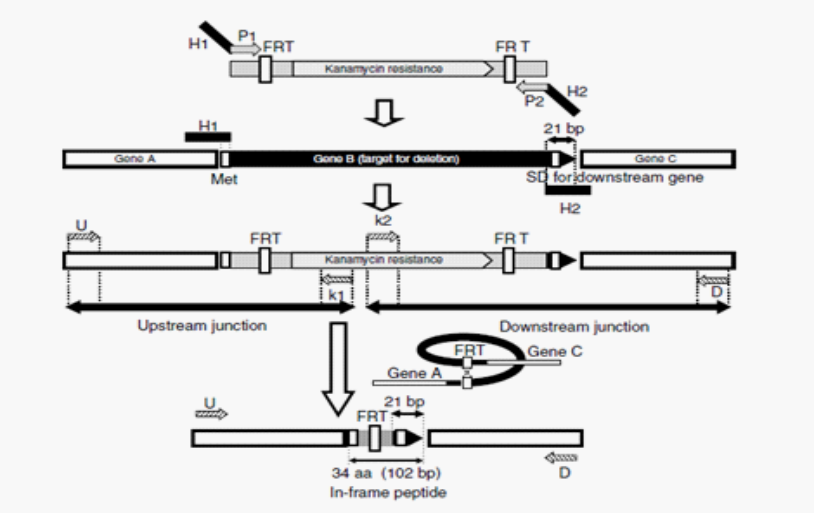

Figure 7: Cartoon outlining the strategy used to generated knock-outs in the keio collection. (Baba et al, 2008)

P1 – transduction

Lysate preparation: lyophilized donor strains from the CGSC were grown at 37℃ to saturation, backdiluted 1:1000 in LB 5 mM CaCl2 0.2% glucose and grown for 1h at 37℃. 100µl of P1-lysate was added and cells grown for ~3hours until complete lysis.

Several drops of chloroform were added and vortexed, the phage cleaned by centrifugation at 14000rpm in a table centrifuge and the supernatant left in the hood for 1h to evaporate the chloroform and stored at 4°C. (Note: the phage cannot be stored at temperatures below 0°C because it is not stable).

Transduction: Recipient strain was grown, cells harvested by centrifugation and resuspended in LB 100 mM MgSO4 5 mM CaCl2. Four reactions: A. 100 µL undiluted P1 lysate + 100 µL recipient cells B. 100 µL 1:10 diluted P1 lysate + 100 µL recipient cells C. 100 µL LB + 100 µL recipient cells D. 100 µL undiluted P1 lysate + 100 µL LB (transfections and controls) were set up and incubated at 37°C for 30min to 1h. 200 µL 1 M Na-Citrate (pH 5.5) and 1 mL LB were added and cells recovered at 37°C for 1h. Cells were pelleted, resuspended in 100 µL LB 100 mM Na-Citrate (pH 5.5) and plated on selective media. Colonies were picked and presence of the Kan^R marker confirmed by PCR.

Flipping out the Kan^R cassette: positive colonies were used to make competent cells transfected with 705-FLP plasmid and grown at 30°C for ~48h on LB/Cm. Cells were restreaked on LB without Cm and grown at 40°C to flip out the cassette and lose the plasmid. Flip-out of the Kan^R was confirmed by colony PCR.

2.8 Library Generation

Generation of large plasmid libraries of mutated hydrogenases in E. Coli requires three separate steps: randomizing the hydA gene by error prone PCR, cloning the randomized gene to an expression vector and transfecting the plasmid into the selection strain. Initially the Genemorph EZ Clone Domain Mutagenesis Kit (Stratagene) was used for library generation but due to very inefficient steps two and three of this method

a new protocol incorporating standard restriction and ligation cloning as well as electroporation of the selection strain was developed. Figure 7 outlines and compares the two approaches.

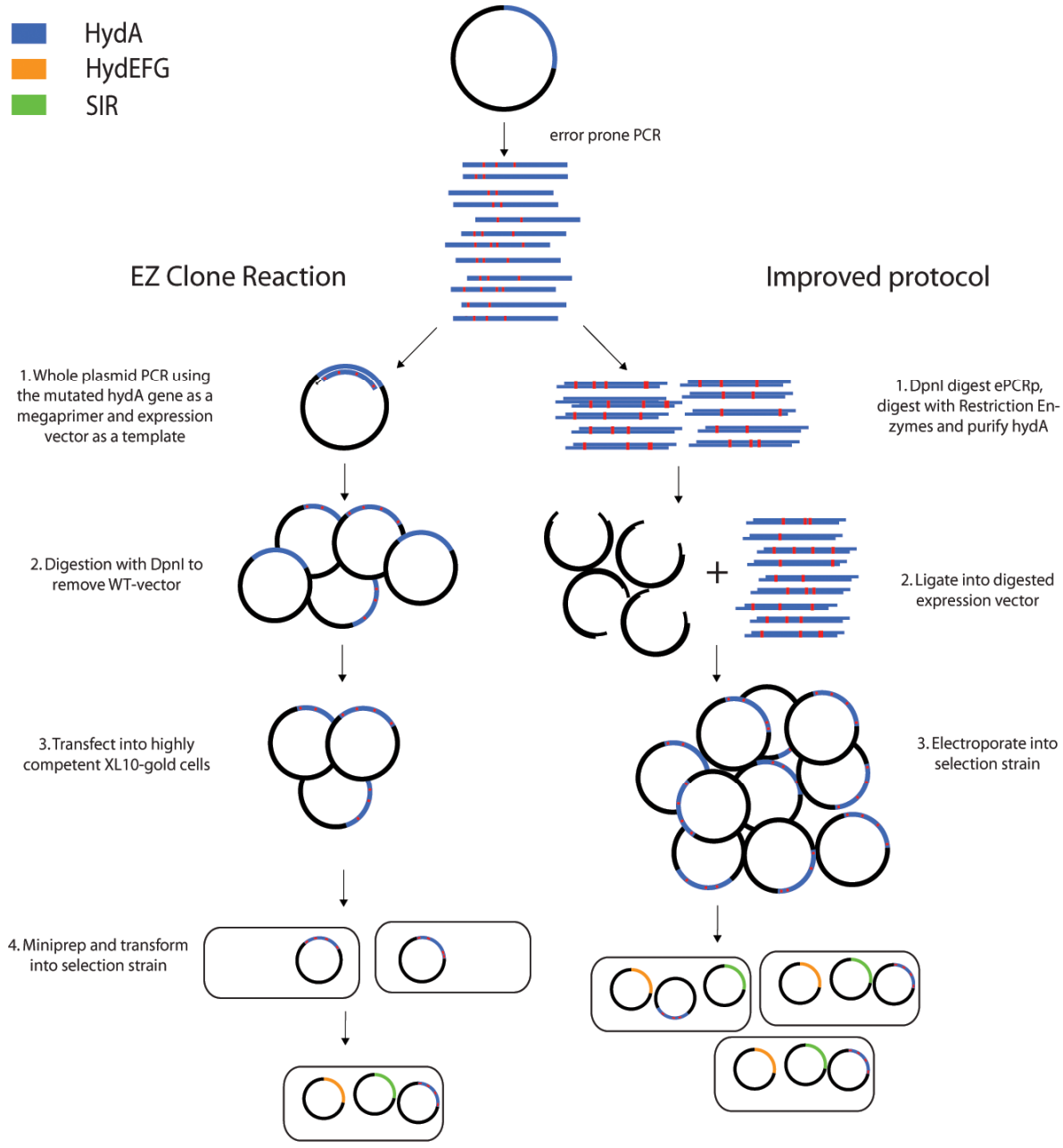

Figure 8: Comparing the EZClone method with the developed improved protocol to make mutant libraries.

The limiting steps in the EZClone reaction were the whole plasmid PCR, which could not be optimized to yield high amounts of correct vector and was hardly scalable and the detour of transfecting commercial XL10-Gold cells to generate a highly redundant library of miniprepped plasmid to be used for transfection into the selection strain. In the improved protocol all steps could be optimized and were easily scalable.

Libraries were initially created from three different templates: *Clostridium acetobutylicum*, *Clostridium saccharobutylicum*, and *Clostridium reinhardtii* *hydA, but after development of the optimized protocol efforts were focused on the former two.* Error prone PCR was performed using the Genemorph Mutazyme II (Stratagene), a mix of two different error prone DNA-polymerases giving a relatively unbiased mix of random mutants. Mutation rates could be tuned by varying PCR-conditions. Two different mutation rates were used. Mutation rates were confirmed by sequencing 5 random complete hydA reading frames and adjacent sites from the library. Table 5 shows the mutations rates used and parameters modified.

Table 5: Tuning the mutation rate of the error prone PCR. Table shows error rates used and parameters changed.

Mutation Frequency	**Amount Template Used**	**Replication Cycles**
4.5 – 9 mutations/kb	200 ng	35
0 – 5 mutations/kb	500 ng	25

PCR-products were purified and digested with NcoI and NotI (CaHydA) or NcoI and SpeI (CsHydA) o/n. Expression vector was digested with appropriate enzymes and alkaline phosphatase o/n. Linearized vector and sticky end mutant hydAs were ligated with a ratio of 1/3 using T4-DNA ligase (NEB). Ligation products were analyzed by gel electrophoresis and electroporated into the selection strain.

Electrocompetent cells were generated as follows: a saturated o/n culture of the preinduced selection strain k32 carrying plasmids for hydEFG and SIR, was backdiluted 1/50 and grown in LB/Amp+Cm+100mM IPTG at 37°C with shaking to an OD_{600} of around 0.5. Cells were harvested by centrifugation at 4°C (20 min, 2500 rpm) and kept at 4°C from then on. Cells were washed 1 x with ice-cold deionized H_2O and three times with ice-cold deionized H_2O 10% glycerol by pelleting (25 min of centrifugation at 4000rpm) and placed on ice until use. For reasons of increasing electrocompetence cells were prepared fresh for making libraries.

Electroporation Cells were aliquoted (25µl per reaction) and pipetted into electroporation cuvettes (Biorad) on ice. 3µl of ligation product were added and cells were electroporated using a Biorad Electroporator. Immediately 1ml of SOC with 100mM IPTG was added and cells were recovered for 1h at 37°C. After recovery cells were washed 2x with 2ml of PBS and plated on selective media in 150µl of deionized H_2O.

2.9 Screening/Selection and growth Assay

Plated libraries were grown at 37°C for 72 hours at 37°C. Cells were grown in anaerobic chambers (add product) at 10% O_2 in H_2. Desiccant pillows were added to reduce condensation of H_2O. To adjust the atmosphere chambers were evacuated to a pressure of ½ atmosphere (10% O_2) and filled up with H_2 to one atmosphere.

Colonies from clean plates were restreaked on selective media and grown at 10% O_2 and air to check for contamination. Colonies that only grew at 10% O_2 and not in air where restreaked twice more at 10% O_2.

Isolation of mutant-plasmids: colonies that passed the growth assay were grown up in 5 mL LB/Sm 100mM IPTG (removing the selection pressure for the plasmids carrying SIR and hydEFG) and restreaked twice on LB/Sm. HydA- carrying plasmids were purified, and retransformed into the selection strain. Colonies were again restreaked on selective media and grown in 10% O_2 and air. Plasmids that rescued growth selectively at 10% O_2 were considered a hit of the selection.

3. Biochemistry Techniques

3.1 SDS-PAGE

SDS-PAGE was performed using NuPAGE Novex Bis-Tris gradient gels with a bisacrylamide gradient from 4-12%. Samples were diluted in 4x NuPAGE LDS loading buffer with 10% β-mercaptoethanol and put on 100°C for 5min. PAGE gels were run in NuPAGE Tris MOPS buffer at 120 to 140V for 100-120min. For bacterial whole cell extracts cells were lysed, debris was pelleted and the lysate (supernatant) run on the gel. Protease inhibitors (Sigma) were added as necessary. As a standard the Biorad Precision Plus Dual Color ladder was used. Protein gels were washed with H_2O and stained with SimplyBlue SafeStain (Invitrogen) or used for a Western Blot.

3.2 Western Blot

Proteins were transferred to a NuPAGE Millipore Immobilon-p membrane with NuPAGE transfer buffer with 10% methanol. Membranes were activated with methanol for 30 seconds and washed 3 times in transfer buffer. Transfer was done using a semi-dry blot apparatus for 45min at 140mA. Membranes were blocked in 5% dry milk in PBS-T (PBS

with 0.1% Tween) for 1h or o/n. Binding of 1:2000 HRP-conjugated anti-His antibody (Santa Cruz Biotechnology) was done in blocking buffer and membranes were washed 3 times for 10 min in PBST. To activate HRP 500µl of each chemiluminescence luminal reagent (Perkin Elmer, enhanced luminal) were added and left for 5-10minutes. To detect the signal Hyblot Cl Autoradiography Film (Denville Scientific) was exposed to the membrane and developed.

3.3 Protein Expression

All proteins were expressed from Novagen Duet expression vectors of their IPTG-inducible T7-promoters. Expression was induced by adding IPTG to a final concentration of 100µM. In order to get functional HydA the *Clostridium ralstonia* maturation factors hydEFG were co-expressed on a plasmid. The correct folding of the used hydrogenases was confirmed by measuring hydrogenase activity. The selection strain as well as strains for the hydrogenase assays were kept at continuous expression since the expression load would kill of a fraction of the cells after IPTG induction.

3.4 Protein Purification

3.4.1 Aerobic

CaHydA was tagged n-terminally and c-terminally with a 6x HIS tag. Expression was induced with IPTG from the pET-Duet vector in BL 21 (DE3) cells carrying the a4 plasmid (HydA maturation factors HydEFG). Purification was performed under aerobic and anaerobic conditions using Ni Sepharose Fastflow resin (GE Healthcare).

Buffers: all buffers had a pH8.0

Lysis buffer (1 L)

50 mM NaH2PO4	6.90 g NaH2PO4·H2O (MW 137.99 g/mol)
300 mM NaCl	17.54 g NaCl (MW 58.44 g/mol)
10 mM imidazole	0.68 g imidazole (MW 68.08 g/mol)

Wash buffer (1 L)

50 mM NaH2PO4	6.90 g NaH2PO4·H2O (MW 137.99 g/mol)
300 mM NaCl	17.54 g NaCl (MW 58.44 g/mol)
20 mM imidazole	1.36 g imidazole (MW 68.08 g/mol)

Elution buffer (1 liter):

50 mM NaH2PO4	6.90 g NaH2PO4·H2O (MW 137.99 g/mol)
300 mM NaCl	17.54 g NaCl (MW 58.44 g/mol)
250 mM imidazole	17.00 g imidazole (MW 68.08 g/mol)

Purification was done at 4°C; protease inhibitors (Sigma) were added to the lysis buffer to prevent protein degradation. 500ml of culture were grown in terrific broth, pelleted and resuspended in lysis buffer. Lysis was achieved by incubating for one hour with lysozyme (1mg/ml) and sonication (6 x 10sec with 10sec cooling). Lysate was loaded on an equilibrated column (1-2ml bed volume) of Ni Fastflow resin and washed twice (5 volumes and two volumes) by gravity flow. Protein was eluted with elution buffer and fractions of 500µl of eluates were stored separately.

All fractions were stored for analysis by SDS-PAGE and Western Blot. Protein content was measured by Bradford Assay (Bio-Rad, Hercules, CA).

Fractions containing hydrogenase were transferred to PBS using Sephadex Gel Filtration desalting columns (GE Healthcare) according to manufacturers' instructions.

3.4.2 Anaerobic

In terms of resin and protocol anaerobic protein purification was performed in a similar fashion as aerobic protein purification. All steps except centrifugation where performed inside an Mbraun labmaster glovebox (Stratham, NH) in an anaerobic N_2-athmosphere at 10-12°C. Anaerobic lysis was achieved by 1 hour of incubation with lysozyme (1 mg/mL) inside the anaerobic chamber and then using a bead-beater (BioSpec) with 0.5mm glass beads (Cole-Parmer) for 2 minutes.

Buffers: All buffers were degassed by sparging with Argon for 15 min and leaving them uncapped in the anaerobic chamber for at least 12 hours. Dithiothreitol and Na^+-dithionite were added to all buffers at 2mM to keep a reductive milieu and protect the hydA active site.

3.5 Hydrogenase Assay

Hydrogenase activity was measured in whole cell extracts. Cells were lysed with Bacterial Protein Extraction Reagens (B-PER, Thermo Scientific, Rockford, IL, USA); hydA was reduced via methyl-viologen (SIGMA) with sodium dithionite (Fisher Scientific) as a reducing agent. O/n cultures in buffered LB of cells carrying HydA and HydEFG plasmids were backdiluted 1/25, aliquots of 20ml sparged with N_2 for 15 min, capped and grown to saturation at 37°C with agitation (~24 h). 1ml of assay buffer (prepared fresh) was added anaerobically with a syringe and cells lysed o/n at 37°C with agitation. H_2 concentration in headspace was assayed by gas chromatography.

Assay buffer (20mL)

1M Tris-B-PER II pH 7.5

Methyl viologen 0.2 g

Na^+-dithionite 2 g

Tris-B-PER was sparged continuously with N_2 or Argon for at least 15 min, MV and Na^+-dithionite was added anaerobically under continuous sparging and the bottle capped.

3.5.1 Hydrogenase Assay for purified protein

Purified protein was stored in PBS at 10℃ inside anaerobic chamber (assay was done within 1 day after purification). Methyl viologen and Na^+-dithionite solutions where prepared inside the anaerobic chamber in anaerobic TrisCl. 1ml of each was added to 8ml of anaerobic TrisCl containing 100µl of the purified hydA-fraction in PBS.

Buffers (100 mL): at pH8.0

500mM dithionite TrisCl 8.7 g

20mM methyl viologen TrisCl 0.514320 g

3.6 Half Life Measurements

Measuring the half life in aerobic conditions of hydrogenases was done indirectly by measuring hydA activities after different times of exposition to O_2. 20ml samples of cells were grown to saturation anaerobically in buffered LB as previously described. Cells were lysed by anaerobically adding lysis buffer under sparging and occasional shaking for 30min. Six triplicates of lysates were exposed to 0-25 minutes (increments of 5) of sparging with 10% O_2 followed by 10min of sparging with pure N_2 to remove residual N_2.

1ml of assay buffer was added anaerobically with a syringe to each sample and H_2 evolved at 37℃ with agitation. After 2 hours the reaction was stopped by injection of 1ml methanol. H_2 in headspace was measured by gas chromatography.

Lysis buffer (20mL)

1M Tris-B-PER II pH 7.5

Baker antifoam 200µl

Na^+-dithionite 50 mg

Assay buffer (20mL)

1M Tris pH 7.5

Methyl viologen 0.2 g

Na^+-dithionite 2.5 g

3.7 Gas Chromatography

Gas chromatography was performed using a Shimadzu GC-14A, with a thermal conductivity detector (TCD) and N_2 as a carrier gas. 1ml of headspace of the assay samples was loaded on the GC for analysis with an air-tight syringe. H_2 and O_2 peaks were recorded and peak areas integrated digitally.

4. Homology modeling

Homology models of *C. acetobutylicum* HydA1 and mutants were made using the SWISS-MODEL server with the *Clostridium pasteurianum* HydA X-ray structure as a template.

IV Results

1. Insulating the pathway

1.1. Knocking out of the fre gene

As seen from the preliminary data from E. Wintermute shown in the introduction (Fig.6) the orthogonal pathway in E. coli used to link growth to Hydrogenase function does not represent a true selection. Knocking out the endogenous sulfite reduction capacity, endogenous hydrogenases as well as genes implicated in ferredoxin interaction increased the growth linkage to a high level (knocked-out genes listed in Table 1 of the Materials and Methods section). The hypothesis is that the sequential removal of interaction partners for the pathway gradually reduced the amount of non-hydrogenase derived electrons entering the pathway. One approach in the project was therefore to try to further insulate the pathway by additional knock-outs of potential interaction partners. The first additional gene to be knocked out was fre (KEGG entry JW3820), coding a flavin reductase that potentially could bind a plant type ferredoxin as used in the selection. The gene was knocked out using P1-transduction and the Kan^R-cassette was flipped out as described on the Materials & Methods section. To test for improved selectability growth of the Δfre-strain (k33) was compared to that of the previous knock-out strain (k32) under selective conditions with the CaHydA. Figure 8 shows an agarose gel confirming the knock-out of fre as well as the different strains grown under selective conditions and in air.

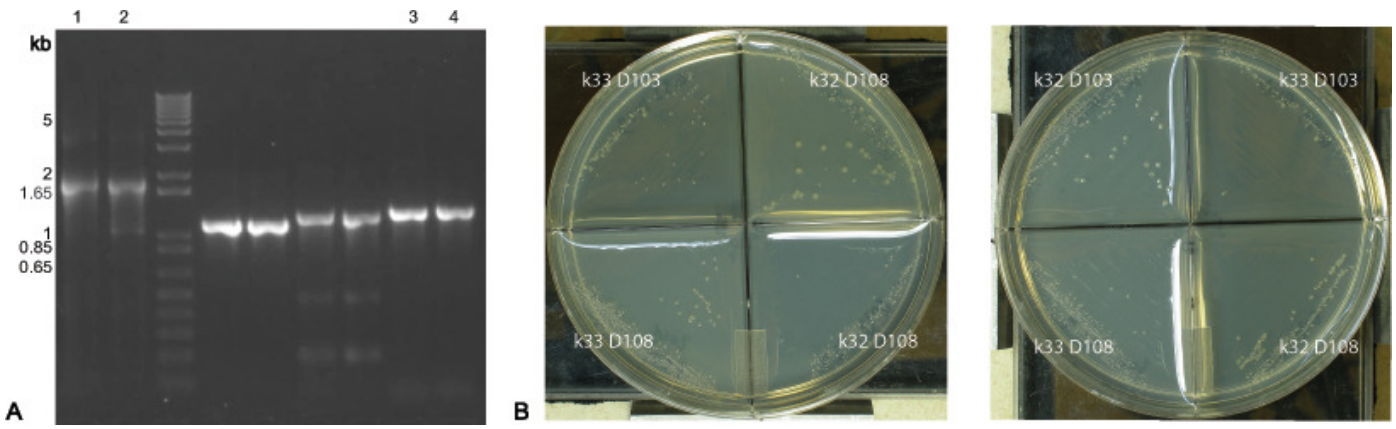

Figure 9: **A**: Agarose gel of colony PCR products, each lane represents one colony. Lanes 1 and 2 show the product with primers specific to a few hundred kb up and downstream of the fre locus, lanes 3 and 4 show the product with the fre-locus forward primer and the primer k2 (Baba et al. 2008) a reverse primer in the middle of the Kan-cassette. Expected bands are at around 2kb for the locus and 1kb for the KanR-product. **B**: Plates of the knock out strains grown on selective media at 0% O_2 and in air. FNR is a positive control, k33 the fre-knock-out strain and k32 the previous knock-out strain.

At zero percent O_2 the selection strains k32 and k33 don't show a significant difference in growth (amount or size of colonies), with and without H_2. This suggests that the fre gene did not contribute to the observed background.

Interestingly, flipping out of the Kan^R-cassette failed several times; yielding no viable colonies, while flipping out the cassette from different loci, such as tonA (ΔtonA cells are resistant to T1-phage, the knock-out was generated to deal with problems of T1-phage contaminations in some lab areas) was possible. This suggests that in the case of k33 carrying the Kan^R-cassette flanked by FRT-loci in the fre-locus there was some interaction between one of the FRT-sites and a distant FRT-site – the background strain already contains 11 knock-outs, resulting in a lethal deletion. However since the removal of the fre-gene did not seem to have an effect on selectability of the SIR-pathway different mutants could be generated starting from ko32. However, the efficiency of flipping out the cassette was already low in the case of k32, so the

likelihood for problems of this nature to arise would seem to increase with each additional FRT-locus added by knocking out a further gene. This work was done in parallel to setting up a library screening protocol and testing the selection strain and as a consequence of the lack of effect of the fre-knock out the mother strain k32 was used for this.

1.2. Using NADH-ADH interaction as an artificial electron sink

An interesting observation from working with the selection strains and libraries of hydrogenases was that there was a strong inverse correlation between growth and oxygen levels, not only in the presence of H_2, where it is obviously due to the inactivation of the hydrogenase, but also for the background growth observed when growing the strains without hydrogen. If plated at O_2-levels between 5% and 10% background growth could rarely be observed. This circumstance might be interpreted by there being more reducing power or electrons to enter the pathway present in the hypooxygenic E. coli metabolism or reflect a difference in expression patterns of multiple electrochemical enzymes in E. coli.

The principle of metabolism is the transfer of high energy electrons to a lower energy level and using part of that energy difference. The key element is disposing of low energy electrons, or rather using redox pairs with a large enough ΔG to run the reactions. The best and most prevalent terminal electron acceptor in nature is O_2. For anaerobic microbial metabolism there are two ways to run metabolism in the absence of oxygen, anaerobic respiration, which runs an electron transport chain but uses a different terminal electron acceptor than O_2 and is mostly used by obligate anaerobes or

fermentative processes that run glycolysis or related metabolic pathways and recycle NAD^+, by using the excess reductive power of generated NADH to generated fermentation products such as lactate or ethanol. E. coli uses various fermentation pathways to grow anaerobically. Given the strong need for E. coli to get rid of reduction equivalents in the form of NADH when grown anaerobically one hypothesis was that providing an artificial electron sink to E. coli to dispose of electrons might reduce part of the background electron flow into the SIR-pathway. This is of course a very long shot, since to truly isolate the pathway it would be necessary to eliminate all possible interactions between E. Coli and the *S.o. ferredoxin*, and E. coli would provide electron sinks (fermentation pathways) on its own when grown anaerobically but given that the proteins of the SIR-pathway were highly overexpressed, we tested what effect the overexpression of a different electron sink, zymomonas mobilis alcohol dehydrogenase (ADH) might have.

E. coli expressing the complete SIR-pathway as well as the *Z.m. ADH* were plated at a very low dilution on selective media and grown anaerobically with and without H_2 for 5 days to test a possible effect on non-hydrogenase linked growth. Figure 10 shows plates of cells grown with and without H_2.

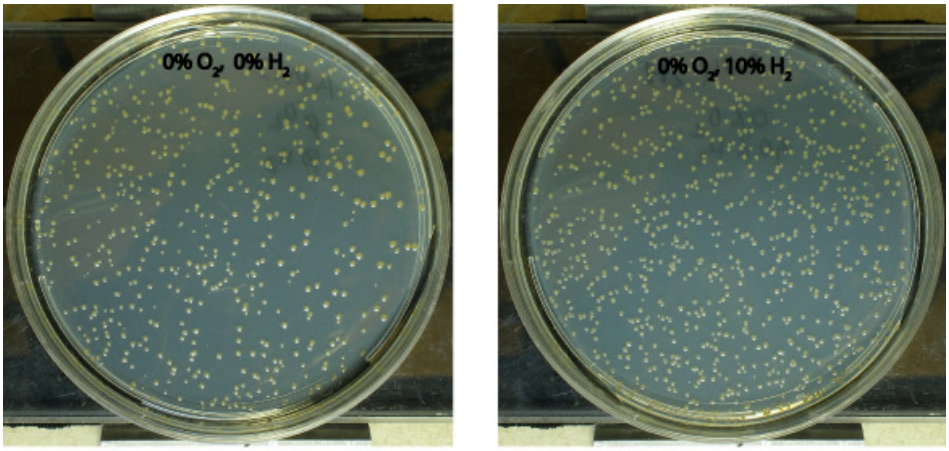

Figure 10: k32 selection strain expressing the zmSIR, soFd, CaHydA and ZmADH grown with and without H_2 at 0% O_2.

As seen on the plates no strong growth difference was observed between the strains expressing the ADH with or without H_2 at 0% O_2, while it seems that the plate grown with H_2 has more colonies growing. This shows that the artificial expression of ADH as an electron sink did not resolve the issue of background electron flow to the pathway.

2. Quantifying Growth and Selectability of the Selection Strain

Given the initial data presented in the introduction that showed a growth advantage in the selection strain k32 expressing the SIR-pathway conferred by the hydrogenase as well as the lack of success in further isolating this pathway we sought to quantify the extent of the growth advantage conferred by the hydrogenase to assess the selectability of hydrogenases using this system. To analyze this, strains were grown at various levels of O_2 and H_2 for 72 hours, expressing different hydrogenases and spinach-ferredoxin. Cells were grown up in liquid culture, diluted, counted and plated so that between 50 to 100 cells would be plated per plate in triplicates. To measure growth, colony size was measured by image segmentation and calculation of the area using a

Matlab script. Figure 11 shows the growth under various conditions expressed as average colony size vs. the oxygen concentration.

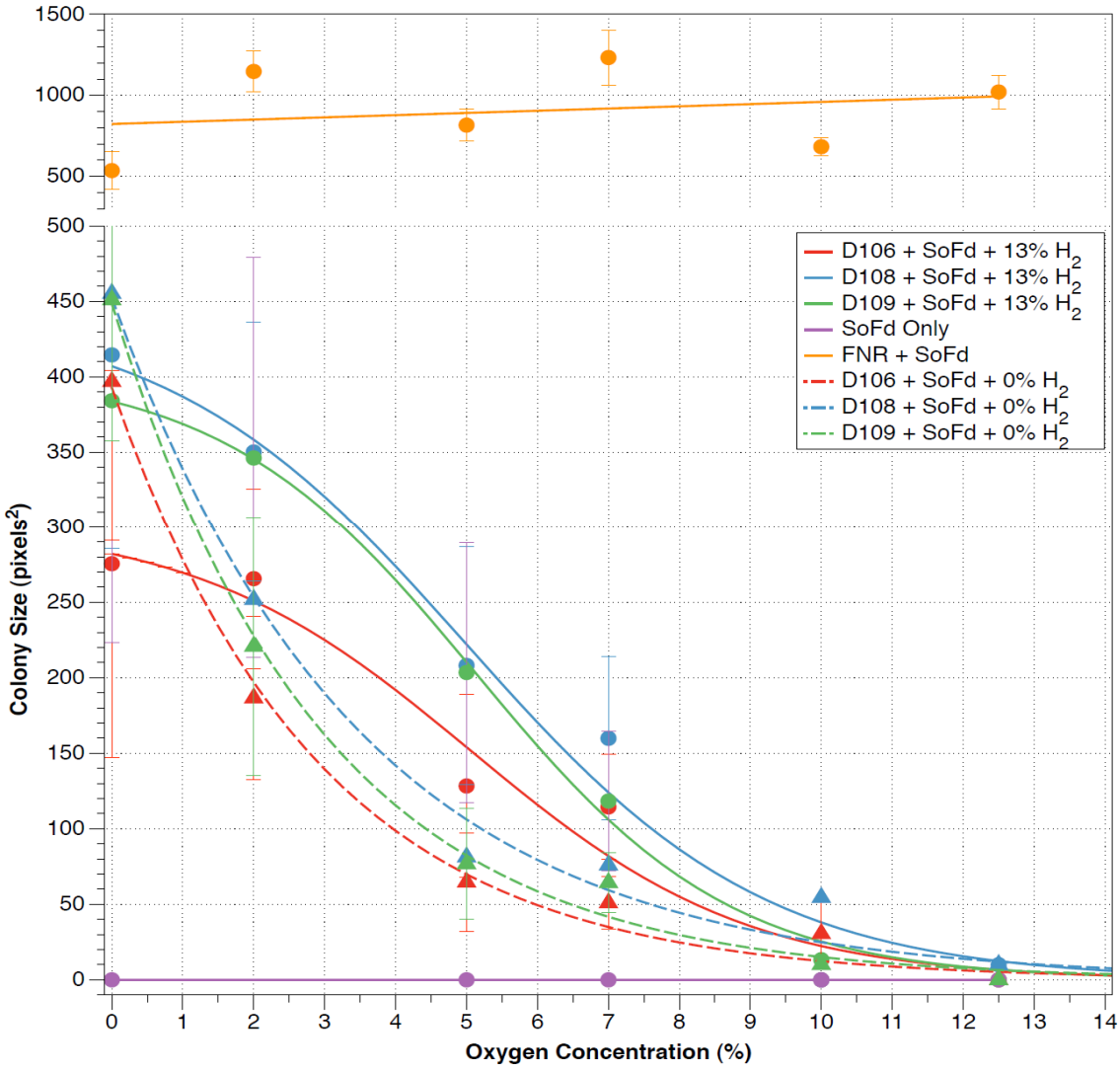

Figure 11: Growth of the ko32 selection strain with various hydrogenases; red: D106, CrHydA, blue: D108, CaHydA, green: D109, CsHydA and spinach ferredoxin grown with and without H_2 at various O_2-concentrations. Magenta represents the strain grown without hydrogenase and orange is the strain expressing a ferredoxin-NADH-oxidoreductase FNR, short-circuiting the pathway and serving as a negative control. Thanks to Buz Barstow for sharing his data and Bruno Afonso for helping with image segmentation and analysis.

It can be seen that the hydrogenase does indeed give a growth advantage to the strain grown under selective conditions. The magenta line not expressing a hydrogenase shows no discernible growth – meaning none of the 50 to 100 cells plated grew to a

colony. It should be noted however, that at 0% O_2 if planted at high density growth without a hydrogenase was also occasionally observed, although to a much lesser extent than with a hydrogenase. In Figure 10 it is apparent that the growth does not strictly depend on the electron flow from H_2 via the hydrogenase to ferredoxin, some non-specific interaction allow flow of background interactions into the pathway. Interestingly this background interaction also inversely correlates with O_2 concentration, reflecting either a role of iron-sulfur clusters of the hydrogenase at the ferredoxin binding site or an overall change in E. coli expression profile and redox-chemistry shifting from anaerobic to aerobic metabolism.

3. Generating and Screening Libraries

3.1. Optimizing a protocol to generate and screen large mutant libraries

As outlined in chapter 2.8 of the Materials and Methods section of this thesis (in particular figure 8) two approaches were used to generate and screen libraries. The limiting steps in the process were transferring the randomized DNA molecules to the expression vector and transfecting the expression vector library into the selection strain with high efficiency. The initially used EZClone reaction used a whole plasmid PCR for the first step and transfection into commercial highly competent WT- E. coli and subsequent miniprepping to generate a transfectable plasmid library for the second step. For the genes and plasmids used in this study it was not possible to optimize the whole plasmid PCR, by varying conditions, polymerases and priming temperatures using gradient PCR to a point where the expression vector, having a length of around 5.7 kilobases, could be seen as a nice band on a gel, although after optimization some

weak band was visible. Figure 12 shows the method used (whole plasmid PCR vs. ligation), an agarose gel of the expression vector carrying the mutated hydA and a control plate of a dilution of a library in E. Coli plated on LB to measure library size.

The data shown illustrates the success in optimizing a protocol to screen large mutant libraries. Using a conventional restriction and ligation strategy and electroporating the selection strain hugely increased the library size that was possible to be screened per plate.

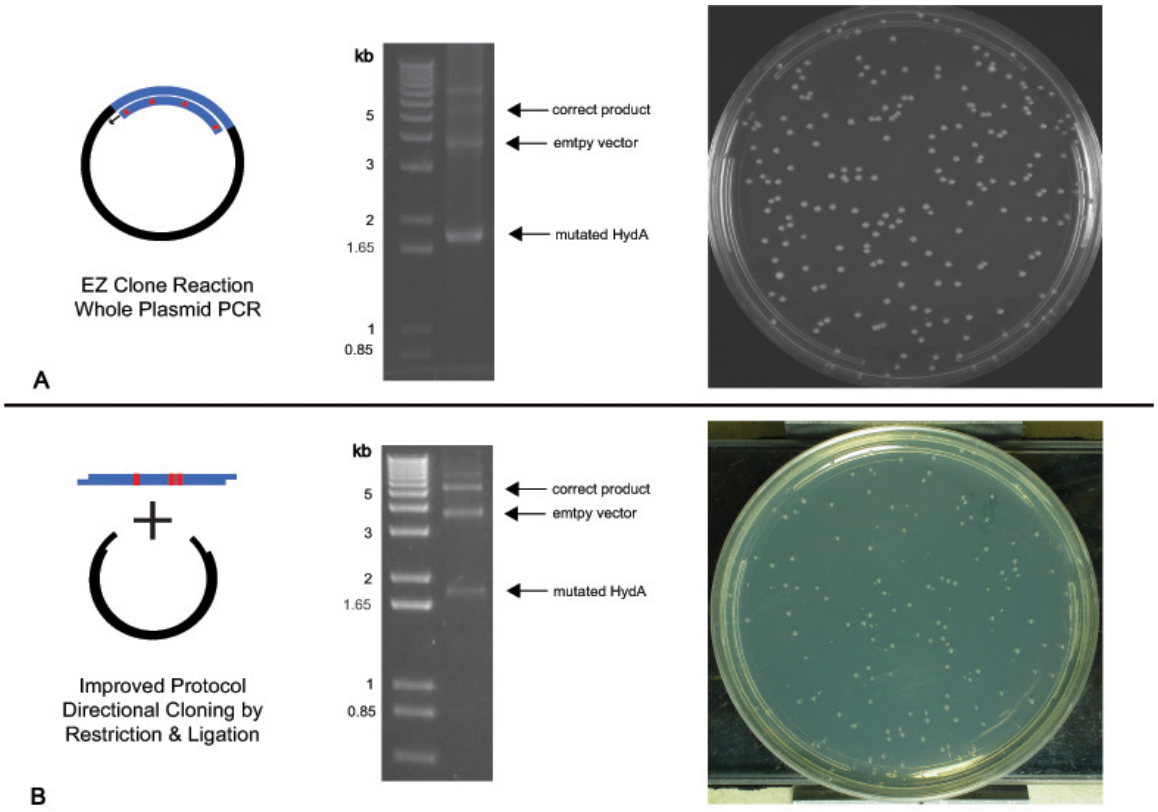

Figure 12: **A**: Cartoon, gel and LB-plate of a 1/20 dilution of a library using the EZClone reaction. Note that the plate is a library in the highly competent XL10 strain, not the selection strain, meaning that the miniprepped library from this plasmid library will be redundant and will have to be oversampled by at least a factor of 5 if screened in the selection strain. **B**: Shows a cartoon, gel and LB-plate of a 1/10000 dilution of a library in the k32 selection strain. Correct product is at ~5.7kb.

The efficiency of the various methods was further quantified, shown in figure 13. Note that in figure 13 the increase of library size going from commercial electrocompetent cells (Lucigen) cells is not due to an increased electroporation efficiency of the k32

selection strain. However, to get libraries in the area of 10^7 colonies per plate, up to 200ng were used per electroporation (reducing efficiency of the reaction but increasing total number of transformants). Also note that the libraries in the Lucigen cells and XL10 gold cells would be highly redundant if miniprepped and used in the selection strain.

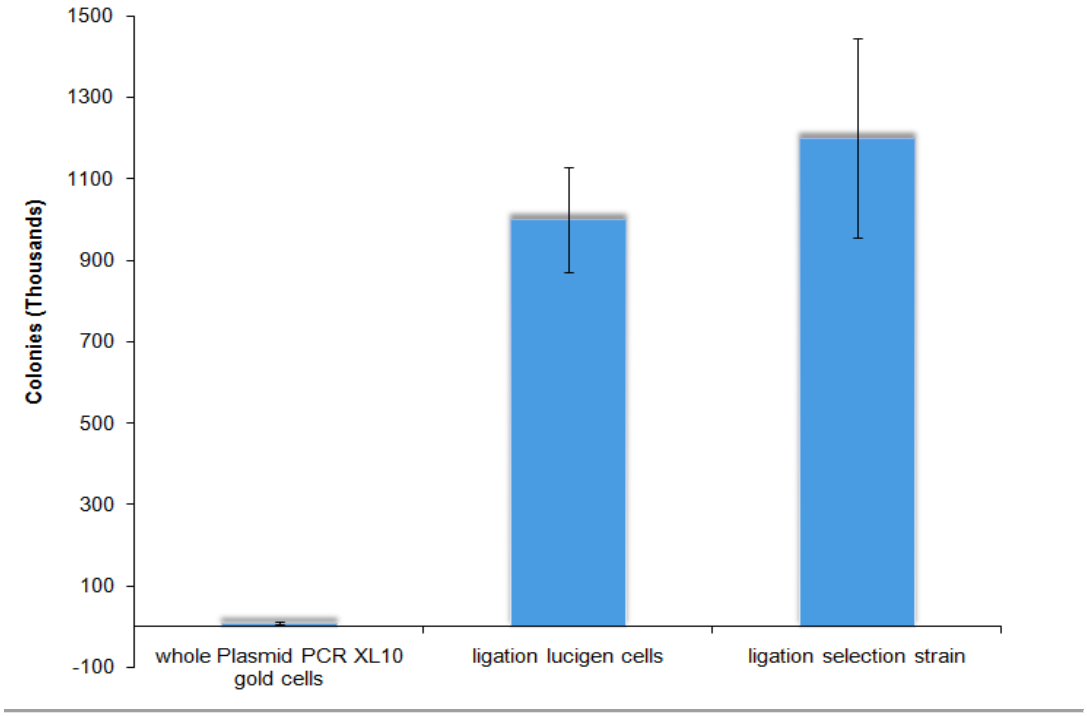

Figure 13: Quantification of library sizes using the different methods.

3.2. Library quality

The advantage of using a whole plasmid PCR based method for transferring randomly mutated genes to an expression vector for use in a screen or selection is that no empty or WT plasmid should be contained in the reaction because of a step of digestion with DpnI (digesting DNA Dam1 methylated DNA from minipreps). Using a more traditional approach of directed cloning with restriction enzyme digestion and subsequent ligation carries the risk of contaminating a library with recircularized vector or plasmids carrying

multiple insertions on the gene (however there is no risk of contamination with WT gene since only mutated DNA is used). While we always dephosphorylated the vector before ligation with alkaline phosphatase we would always get some level of religation in the controls using just vector. To test the composition of the libraries 10 colonies of titration plates of libraries on LB were picked at random, plasmids isolated, digested and run on a gel. Levels of correctly inserted product routinely were between 70% and 100%, giving a sufficient library quality. To test mutations rates 5-10 plasmids per library were sequenced over the range of the gene. Mutation rates were in the range expected from the protocol (data not shown) and no WT-plasmid was detected. Figure 14 shows two agarose gels with digested plasmids from 10 randomly picked colonies of two different libraries, illustrating the composition of libraries in regard to empty vector/correct vector.

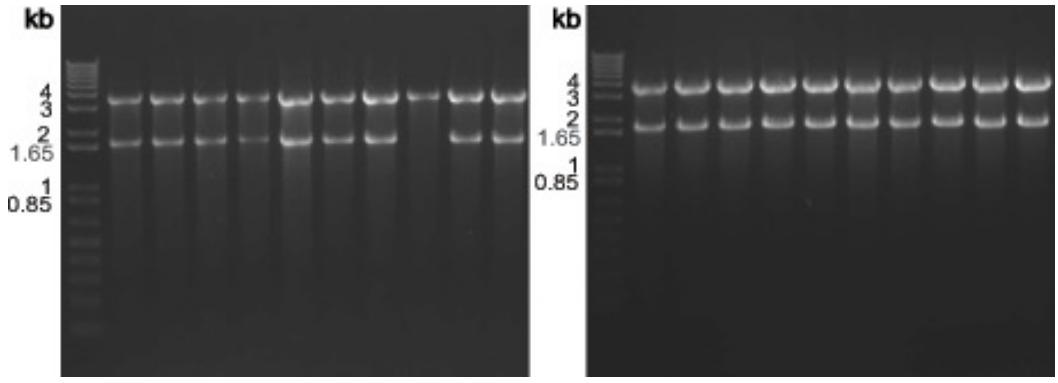

Figure 14: Two gels showing 10 randomly selected plasmids from libraries digested with NcoI/NotI (CaHydA) and NcoI/SpeI (CsHydA). The vector has a size of 3.5 kb, the hydrogenase has 1.8kb.

4. Hits of the Growth Assay – Mutations

Mutant hydrogenases isolated from colonies coming out of the growth assay that rescued growth when retransformed into the selection strain were isolated, and tested further. Table 6 shows a list of all the mutations (on a peptide level) of the selected

mutants. Only one hit was derived of a CsHydA, several more were isolated from CaHydA.

Table 6: List of peptide sequence mutations of selected hydrogenases.

Mutant	Parent	No. Mutations	Mutations
mutant 1	CsHydA	15, 1 stop	I12T, I20V, A26T, C46Y, L193Y, V203I, L244I, Y246F, I353L, A378T, F435I, A475T, S543T, T547S, D555X
mutant 2	CaHydA	5	I6T, D39E, R59T, E460D, N464D
mutant 3	CaHydA	3	F58Y, E488D, K676I
mutant 4	CaHydA	3	D273E, N487I, N519K
mutant 5	CaHydA	1	P547S
mutant 6	CaHydA	2, 1 stop	M50R, T116A, I166X

5. Mutant HydA Activities

Mutant hydrogenases that were selected from the growth assay were tested for activities. The first hit of the selection, mut1 was initially tested in an in vivo system, developed in the Silver lab by Agapakis et al. (2010), using a *Desulovibrio africanus* pyruvate ferredoxin oxidoreductase (PFOR) to shuttle electrons from pyruvate to a clostridial ferredoxin which reduces the hydrogenase. This approach was chosen to test in vivo functionality and functional interaction with ferredoxin. Activity was measured by measuring H_2 after H_2 evolution over night. Mut1 showed an activity of approximately 5% of the wild-type CaHydA (data not shown). To confirm this and account for potential problems of hydA interactions in the in vivo system, activity was also measured in vitr, in whole cell lysates reducing the hydrogenase with methyl viologen, using dithionite as an electron source. All measurements, both in vivo and in vitro were done anaerobically.

For the later hits of the selection initial in vivo measurements were omitted and in vitro measurements were performed right away. Figure 15 lists activities of selected mutant hydrogenases.

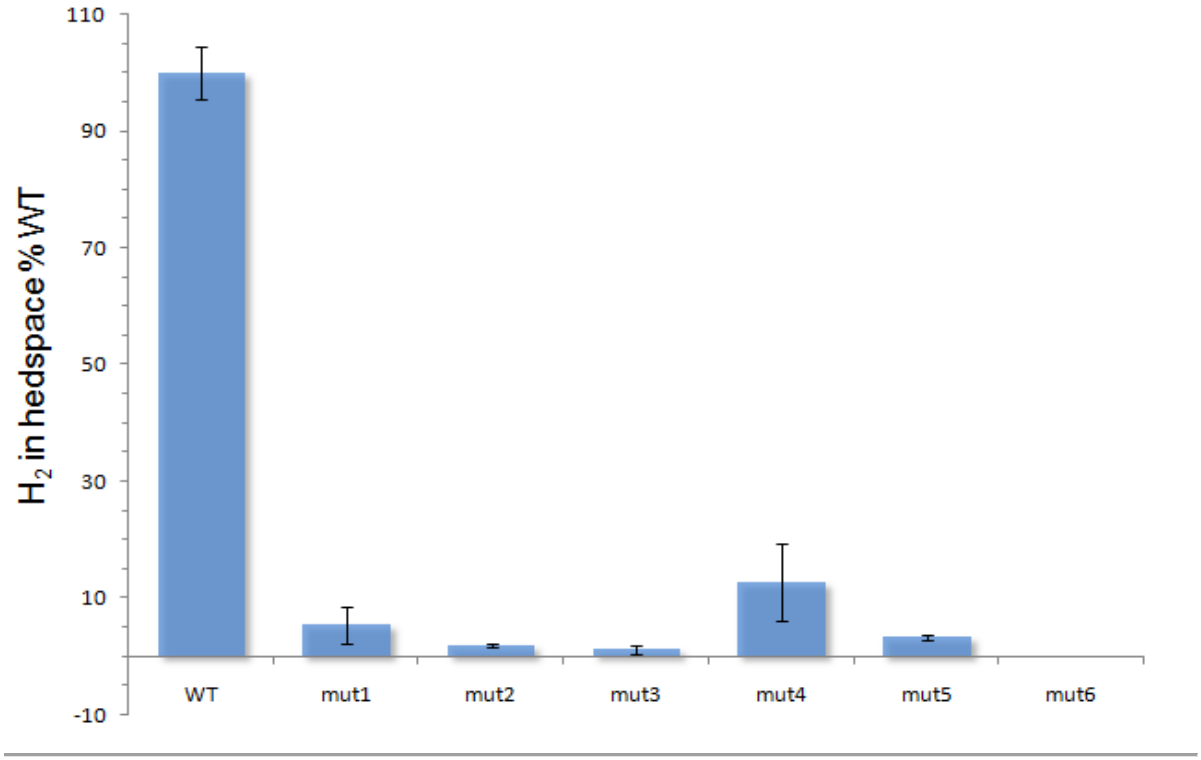

Figure 15: Activities of mutant hydrogenases measured in whole cell lysates, activities are represented as the amount of H_2 in the headspace relative to the WT-CaHydA.

The Data shows clearly that none of the selected hydrogenases retained a very high activity and one of the isolated mutants, mutant 6 had no activity at all. Some loss of activity is to be expected, since random mutagenesis was performed on the hydrogenases but considering that the growth assay should have ideally been a selection for functional hydA active sites one would have expected a much lower loss in activity for selected mutants. On the other hand, a majority of the selected mutants show at least some activity, suggesting that perhaps some level of active site integrity was indeed under selection.

6. Mutant Hydrogenase Half-Lives in Oxygen rich Atmosphere

Since the objective of the project was to engineer hydrogenases with an increased oxygen tolerance the primary measure of success was not to measure overall activity or rates but to measure activity or half-lives in an oxygen rich atmosphere. Initial tests were done with all mutants, as well as wild type enzymes as a control to measure activities of hydAs grown in a 10% O_2 atmosphere. No activity of any of the mutants or WT hydrogenases was detected at 10% O_2. However, other than being completely resistant effects could also be explained by hydrogenases having an increased half-life in O_2, providing electrons to the system for a longer time before being inactivated. To test this hypothesis half-lives were measured. We developed a protocol to measure half-lives in whole cell extracts by exposing hydrogenases to 10% O_2 for increasing times and measuring H_2 evolution of these extracts after reduction by methyl viologen with dithionite as an electron donor. Unlike electrochemical experiments performed on purified enzymes on a cathode – for which we were not equipped – this method has the disadvantage that measurements become rather noisy if the activities of the hydrogenase are low. Using reductive power in large excess we were able to get measurements for the wild type C.a. and S.a. hydrogenases as well as for the mutant 4 – showing the highest activity of the selected mutants. Figure 16 shows the measured half-lives.

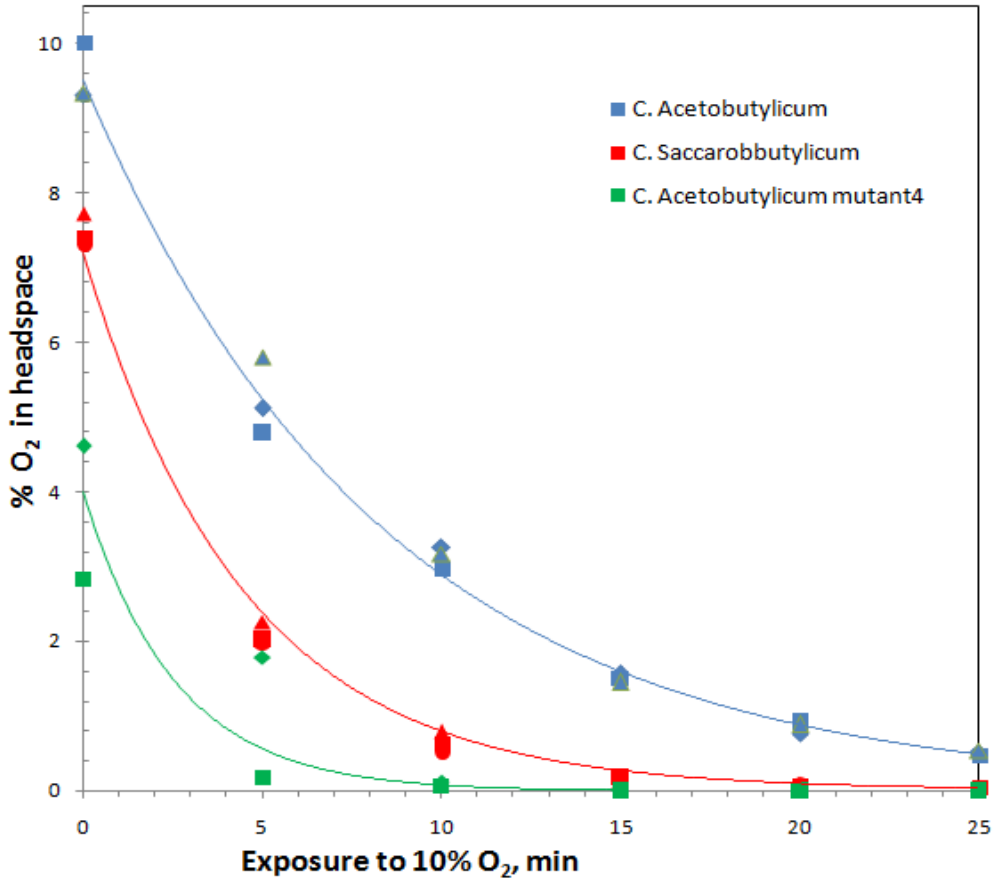

Figure 16: Measured half-lives of the *Clostridium acetobutylicum* (blue) and *Clostridium saccharobutylicum* (red) hydrogenases as well as the hydrogenase mutant 4 (green).

Half-lives of around 6 minutes for the CaHydA and around 3 minutes for the CsHydA were measured – this is in accordance with the values reported in the literature. For the mutant 4 – a mutant derived from a Ca-hydrogenase the measured half-life is at around 2-3 minutes, meaning a strong decrease in half-life compared to the wild-type.

This is unfortunate, meaning that whatever trait of the hydrogenases rescued growth at 10% O_2 it was not their activity of converting H_2 to electrons and protons. It suggests that rather than selecting for a hydrogenase with an increased half-life we selected mutants that showed an increased ability to short-circuit the system or shuttle background electrons into the pathways.

7. Purification of Clostridium Acetobutylicum HydA

In order to be able to assay mutant hydrogenases in more detail by doing electrochemistry experiments or analyze structures the Ca HydA was also purified. It was tagged c- and n-terminally with a HIS-tag and purified over a Ni-sepharose column. While n-terminally tagged hydrogenase was not soluble and went to inclusion bodies c-terminally tagged hydrogenase could be purified without problems at high purity and quantity.

Purified enzyme was tested for activity in an in vitro assay using methyl viologen. No activity could be measured, however there was no oxygen detected in the sample. According to the literature problems can arise if purifying metalloproteins with a HIS-tag because the Ni-resin can remove necessary ions from the protein and destroy active sites. Ca-hydrogenase has been functionally purified using a Strep-tag under strictly anaerobic conditions in the past. Most likely using a HIS-tag for purification approach is not feasible for Ca-hydrogenase purification because the Fe-S clusters are not stable under the conditions used or in interaction with the resin.

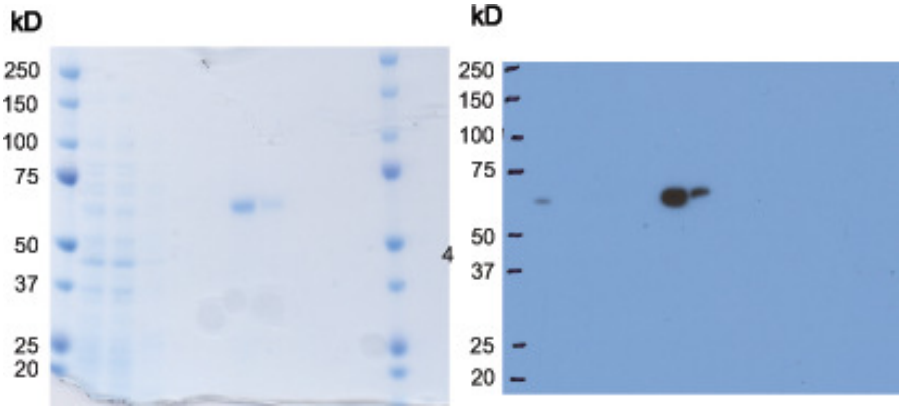

Figure 17: Stained SDS-Page gel and Western-blot of purified Ca hydrogenase. The bands are the flow through, two washes and the elutions of a HIS-tag purification. The molecular weight of the CaHydA is approximately 66 kilodalton.

V Discussion

1. Failure to Isolate O_2-tolerant Hydrogenases

The fact that all the hydrogenases isolated using the selection had both a reduced activity under anaerobic conditions as well as a decreased half-life in O_2-rich atmosphere so far as it could be measured and no residual activity at 10%-O_2 clearly means that the approach to evolve improved oxygen tolerance failed.

In principle getting a slight decrease of function is nothing unusual in a first step of a directed evolution experiment – although decrease of function will typically not occur in the parameter under selection – simply because in order to achieve a fundamental change in properties of an enzyme structural change is necessary which will obviously affect an enzymes function. In general, directed evolution experiments are done iteratively and good results can only be expected once several beneficial mutations come together to achieve a new desired function. However, the data presented in this study shows that the artificial pathway we tried to use as a selection function did not allow to select for hydrogenase function, precluding the meaningfulness of iterations.

On the other hand, the designed pathway did have an effect on redox-chemistry and was linked to hydrogenase in some way. In the aftermath of the project, when it became apparent that the hits from the selection were not promising, my collaborators tested the various knock-outs leading to the selection strain for their effect on the insulation of the pathway by growing them in liquid culture with and without the ferredoxin-NADH-oxidoreductase as a positive control. The data clearly shows a strong quantitative effect on growth caused by insulating the pathway (figure 18).

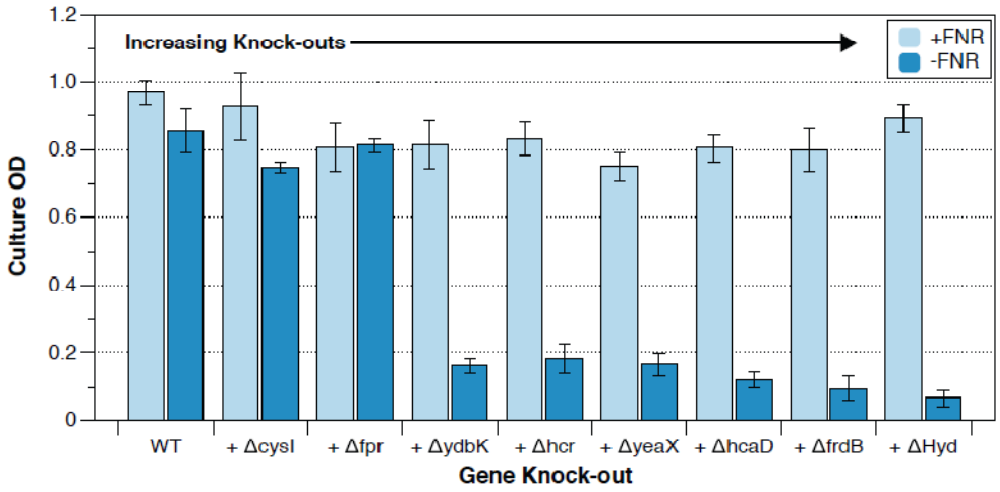

Figure 18: Liquid culture ODs of colonies grown with and without ferredoxin-NADH-oxidoreductase. The effect of insulating the artificial sulfur reducing pathway through the knock-outs is illustrated by the low levels of growth in the cultures not expressing the FNR. Figure courtesy of Buz Barstow.

Adding the hydrogenase to this circuit did increase the level of growth by supplementing electrons to the pathway. However we also saw that growth at a low level also was possible without the active site – illustrated by growth in the absence of hydrogen.

It is possible that the screening assay with the background growth was simply not stringent enough to select and that we were fooled by random events of growth popping through the screening but that seems unlikely.

A more plausible and interesting hypothesis is that some of the iron sulfur clusters of the used bacterial hydrogenases interacted with the plant-ferredoxin directing electrons to the system from an unknown source. In fact the ferredoxin-binding subunit of the FeFe hydrogenase, containing the bottom three accessory Fe2S2-clusters is believed to have evolved from the fusion of a ferredoxin with a hydrogenase (now the catalytic subunit). Interestingly in all the selected mutants the bottom part of this ferredoxin binding site including the first two Fe2S2-clusters seems to be intact (according to homology models) – even for the completely inactive mutant 6 that is truncated after the 166th amino acid (it is basically the ferredoxin binding subunit of the hydrogenase) as well as

for the mutant 1 from a library with a very high error rate carrying 15 mutations. It is plausible that rather than selecting for mutants that were able to oxidize H_2 at high O_2 levels by protecting the active site the fact that some background growth was possible meant that in fact structures facilitating this background growth were in fact selected for – although I have to admit that it is beyond my understanding of hydrogenase structural biology, theoretical as well as redox chemistry how hydrogenases could function as such conductors and what could be possible electron donors for such an interaction. There is an intriguing possibility that actually the *Clostridium ralstonia* maturation factors used in our system could be implicated in these reactions. But since this is based on recent data presented to me by E. Wintermute observing a significantly lower level in background growth when using *Clostridium acetobutylicum* maturation factors it is purely speculative at this point.

2. Bioenergy and Hydrogenases

In recent years research interest in hydrogenases – particularly microbial hydrogenases has risen significantly. From a bioenergy perspective the importance of finding or engineering hydrogenases with a high catalytic rate that might be used in the production of hydrogen that are resistant to O_2 is recognized as a clear goal and is pursued intensely. Approaches that have been taken toward that aim are frequently rational protein design efforts like changing residues in the gas-channel (King et al., 2009).

Another study that has been published recently took a directed evolution approach (Stapleton et al., 2010). They used a fully automated complex costly in vitro microtiter

plate based cell free screening system allowing medium throughput (10^3-10^4 mutants per day) to isolate CrHydA mutants with an up to 4 fold increase in O_2 tolerance.

Given these moderate successes and also the fact that in nature there is a wide range of FeFe-hydrogenases with different half-lives in O_2 and catalytic rates and there is no apparent trade-off between the two at all (which might be imaginable and detrimental to this engineering goal) evolving an improved FeFe-hydrogenase is definitely an attainable goal if the problem of how to select beneficial mutations or dealing with the vast numbers of mutants necessary to get there can be worked out.

On the other hand from a strict bioenergy, rather than biotechnology standpoint hydrogenases might be somewhat marginalized in the near future given recent research success in other areas. Recently a pathway in cyanobacteria was found that can synthesize C13-C17 alkenes (which is basically diesel) from fatty acid metabolism (Schirmer et al., 2010), which opens huge possibilities of making potentially carbon-neutral bio-diesel that is actually diesel. This would be compatible to current industrial standards and have a huge advantage over the use of H_2 as a fuel where a whole new logistics would have to be set up in order to use it on a large scale.

Another promising direction is research in microbial fuel cells, that use waste-product organic acids combining microbial redox-chemistry with fuel cell engineering and allowing the production of electricity assisted by microorganisms using the Gibbs energy of organic waste molecules or the production of hydrogen or other chemicals assisted by electrical energy from the grid or another electrical power source (Logan BE, 2009; Logan BE, 2009, Girguis PR et al., 2010).

3. Conclusions and Outlook

This work presents in interesting novel and rather out of the ordinary approach to using directed evolution for the engineering of hydrogenases. An artificial metabolic redox pathway consisting of plant and bacterial enzymes was assembled in E. coli, partially isolated and tested. Protocols to generate and screen large mutant libraries at different mutation rates were developed and implemented and mutants as well as wild-type hydrogenases were tested in various ways to analyze their activity and tolerance to molecular oxygen. While the goal of the project to make oxygen tolerant FeFe-hydrogenases was unfortunately not achieved the results from this project point to several fundamental issues pertinent to synthetic biology, systems biology and biological engineering and also leads to new questions regarding the chemistry and biophysics of hydrogenases. On the one hand we saw that it was possible to combine various enzymes of various species to generate somewhat functional pathways in E. Coli, both on the side of the selection as well as in using a PFOR-ferredoxin pathway for in vivo H_2 production and on the other hand it was striking to encounter the tremendous difficulty we had isolating this orthogonal redox pathway inside E. coli well. For a systems or synthetic biologist it is good news to see that creative and complex ideas like this can be implemented and that we can use some lever modularity in a way reminiscent of mechanical engineering, but as so often in biology the complexity of biological systems and the difficulty in pinning down or isolating particular processes within a living system are still very challenging. We can see that in order to unlock some of the promises of synthetic biology we should not only be able to engineer promoters or circuits but more importantly we have to understand the behavior of biological circuits and interactions in a larger context and more complex biological systems. A complete

systems level understanding of biology still is far away but to get there will not only be challenging but also very intriguing and rewarding in terms of insights and potential applications.

VI Acknowledgements

I especially thank Pam Silver for letting me join her lab and providing a wonderful, open, energetic and stimulating lab environment, Jake Wintermute for letting me join his project and collaborating, Buz Barstow for collaborating, sharing data and helpful discussions, Karl Kuchler for agreeing to advise my thesis and the Marshall Plan Foundation for funding. I also thank Christina Agapakis, Danny Ducat, Dave Savage, Bruno Afonso, Patrick Boyle, Michael Moore and Jeff Way for helpful discussions and much valued advice and the rest of the Silver lab for helpful discussions and not making fun of my hairstyle and choice of clothing too often.

Furthermore I thank Bernd Wolfstädter, Henrike Niederholtmayer, Renan Escalante-Chong for helpful discussions and much valued advice and Walter Bonacci for helpful discussions, much valued advice and sharing late night dinners at the lab, Dave Savage and Zeev Waks for good basketball games and the members of Harvard Capoeira and the Harvard Self Defense Club for being good sports and a lot of fun practicing martial arts.

VII References:

Savage, David F., Jeffrey Way, and Pamela A. Silver. “Defossiling Fuel: How Synthetic Biology Can Transform Biofuel Production.” *ACS Chemical Biology* 3, no. 1 (January 1, 2008): 13-16.

Lee, Hyung-Sool, Wim F.J. Vermaas, and Bruce E. Rittmann. “Biological hydrogen production: prospects and challenges.” *Trends in Biotechnology* 28, no. 5 (May 2010): 262-271.

Song, Chunshan. “Fuel processing for low-temperature and high-temperature fuel cells: Challenges, and opportunities for sustainable development in the 21st century.” *Catalysis Today* 77, no. 1-2 (December 1, 2002): 17-49.

Endy, Drew. “Foundations for engineering biology.” *Nature* 438, no. 7067 (November 24, 2005): 449-453.

Agapakis, Christina M., and Pamela A. Silver. “Synthetic biology: exploring and exploiting genetic modularity through the design of novel biological networks.” *Molecular BioSystems* 5, no. 7 (2009): 704.

Canton, Barry, Anna Labno, and Drew Endy. “Refinement and standardization of synthetic biological parts and devices.” *Nature Biotechnology* 26, no. 7 (July 2008): 787-793.

Mukherji, Shankar, and Alexander van Oudenaarden. “Synthetic biology: understanding biological design from synthetic circuits.” *Nat Rev Genet* 10, no. 12 (December 2009): 859-871.

Haseltine, Eric L., and Frances H. Arnold. “Synthetic Gene Circuits: Design with Directed Evolution.” *Annual Review of Biophysics and Biomolecular Structure* 36, no. 1 (6, 2007): 1-19.

Gibson, Daniel G., John I. Glass, Carole Lartigue, Vladimir N. Noskov, Ray-Yuan Chuang, Mikkel A. Algire, Gwynedd A. Benders, et al. “Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome.” *Science* 329, no. 5987 (July 2, 2010): 52-56.

Romero, Philip A, and Frances H Arnold. “Exploring protein fitness landscapes by directed evolution.” *Nature Reviews. Molecular Cell Biology* 10, no. 12 (December 2009): 866-876.

Sarkar, Indrani, Ilona Hauber, Joachim Hauber, and Frank Buchholz. “HIV-1 proviral DNA excision using an evolved recombinase.” *Science (New York, N.Y.)* 316, no. 5833 (June 29, 2007): 1912-1915.

Fasan, R., M. Chen, N. Crook, and F. Arnold. “Engineered Alkane-Hydroxylating Cytochrome P450BM3 Exhibiting Nativelike Catalytic Properties.” *Angewandte Chemie International Edition* 46, no. 44 (2007): 8414-8418.

Reetz, M. T., J. D. Carballeira, and A. Vogel. “Iterative Saturation Mutagenesis on the Basis of B Factors as a Strategy for Increasing Protein Thermostability.” *Angewandte Chemie International Edition* 45, no. 46 (2006): 7745-7751.

Wang, Harris H., Farren J. Isaacs, Peter A. Carr, Zachary Z. Sun, George Xu, Craig R. Forest, and George M. Church. “Programming cells by multiplex genome engineering and accelerated evolution.” *Nature* 460, no. 7257 (2009): 894-898.

Gupta, Rinkoo D, and Dan S Tawfik. “Directed enzyme evolution via small and effective neutral drift libraries.” *Nat Meth* 5, no. 11 (November 2008): 939-942.

Yang, G., and S. G. Withers. “Ultrahigh-Throughput FACS-Based Screening for Directed Enzyme Evolution.” *ChemBioChem* 10, no. 17 (2009): 2704-2715.

Brenner, M., Bildsten, L., Dyson, F., Fortson, N., Garwin, R., Grober, R., et al. (2006). Engineering Microorganisms for Energy Production. U. S. Department of Energy, Washington, DC, Report JSR-05-300.

Tard, Cédric, and Christopher J. Pickett. “Structural and Functional Analogues of the Active Sites of the [Fe]-, [NiFe]-, and [FeFe]-Hydrogenases†.” *Chemical Reviews* 109, no. 6 (June 10, 2009): 2245-2274.

Peters, J W, W N Lanzilotta, B J Lemon, and L C Seefeldt. “X-ray crystal structure of the Fe-only hydrogenase (CpI) from Clostridium pasteurianum to 1.8 angstrom resolution.” *Science (New York, N.Y.)* 282, no. 5395 (December 4, 1998): 1853-1858.

Mulder, David W., Eric S. Boyd, Ranjana Sarma, Rachel K. Lange, James A. Endrizzi, Joan B. Broderick, and John W. Peters. “Stepwise [lsqb]FeFe[rsqb]-hydrogenase H-cluster assembly revealed in the structure of HydA[Dgr]EFG.” *Nature* 465, no. 7295 (May 13, 2010): 248-251.

Baba, Tomoya, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A Datsenko, Masaru Tomita, Barry L Wanner, and Hirotada Mori. “Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection.” *Mol Syst Biol* 2 (February 21, 2006).

Agapakis, Christina, Daniel Ducat, Patrick Boyle, Edwin Wintermute, Jeffrey Way, and Pamela Silver. “Insulation of a synthetic hydrogen metabolism circuit in bacteria.” *Journal of Biological Engineering* 4, no. 1 (2010): 3.

King, Paul (Golden, CO, US), Ghirardi, Maria L. (Lakewood, CO, US), Seibert, Michael (Lakewood, CO, US), “Oxygen-resistant hydrogenases and methods for designing and making same", Patent No. 7501270, March 2009

Stapleton, James A., and James R. Swartz. “A Cell-Free Microtiter Plate Screen for Improved [FeFe] Hydrogenases.” *PLoS ONE* 5, no. 5 (May 10, 2010): e10554

Logan, Bruce E. “Scaling up microbial fuel cells and other bioelectrochemical systems.” *Applied Microbiology and Biotechnology* 85, no. 6 (12, 2009): 1665-1671.

Logan, Bruce E. “Exoelectrogenic bacteria that power microbial fuel cells.” *Nat Rev Micro* 7, no. 5 (May 2009): 375-381.

Girguis, Peter R, Mark E Nielsen, and Israel Figueroa. “Harnessing energy from marine productivity using bioelectrochemical systems.” *Current Opinion in Biotechnology* 21, no. 3 (June 2010): 252-258.

VIII Appendix: Curriculim Vitae

Gerald Hermann Grandl

gerald.grandl@gmail.com

Weissgasse 31, 1170 Wien, Austria

T: +43 1 4858398

C: +43 680 2092426

Education

10/2009 – 07/2010	Department of Systems Biology, Harvard Medical School Master's Thesis
10/2004 – Present	Molecular Biology, University of Vienna
10/2005 – 10/2006	Technical Physics, Vienna University of Technology Part-time pursuit of classes in maths and physics
1999 – 2003	Sir-Karl-Popper-School, Vienna

Experience

04/2009 – 07/2010	Master's Thesis Research, Harvard Medical School, Boston, Department of Systems Biology, Professor Pamela A. Silver *"Creation of Oxygen resistant Fe-Fe Hydrogenases by Directed Evolution"*
3/2008 – 5/2008	Visiting Student and Research Assistant, Centro Nacional de Biotecnología, Madrid, Dr. Mariano Esteban *"Cloning and Purification of Vaccinia-C10L using a pET27-vector His-Tag system"*
7/2007 – 8/2007	Visiting Student and Research Assistant, Insitute of Molecular Pathology, Vienna, Dr. Barry Dickson *"Promoter characterization and data analysis for an RNAi based screen for behavioral phenotypes in courtship-song in Drosophila"*
6/2006	Visiting Student and Research Assistant, Max F. Perutz Laboratories, Vienna, Prof. Gerhard Wiche *"The Role of Plectin in the assembly of Sarcomeric and Costameric Structures"*

Awards and Honors

Austrian Marshall Plan Scholarship awarded for diploma thesis project, 2009

University Entrance Examination (High School Examination Certificate) – Magna Cum Laude

Conference Presentations

Poster presentation, SynBERC Meeting 2009, Boston MA

"Engineering Oxygen Resistance in Hydrogenases through Directed Evolution"

Skills

Molecular biology and biochemistry lab techniques

Matlab and Mathematica modeling and data processing

German (native), English, Spanish (fluent), French (advanced), Chinese (beginner)