Abstract:

Mathematical modeling of the synthesis of recombinant hydrogenase in E. coli as a tool for increasing its functional yield

Co-supervisor 1: Prof. Dr. Peter Neubauer, Technische Universität Berlin, Institute of Biotechnology, Chair of Bioprocess Engineering, Straße des 17. Juni 135, D-10623 Berlin, Germany, Germany

Co-supervisor 2: Dr. Oliver Lenz, Technische Universität Berlin, Institute of Chemistry, Straße des 17. Juni 135, D-10623 Berlin, Germany, Germany

Co-supervisor 3: Dr.-Ing. Nicolas Cruz Bournazou, Technische Universität Berlin, KIWI-biolab at the Chair of Bioprocess Engineering, Straße des 17. Juni 135, D-10623 Berlin, Germany, Germany

Hydrogenases are complex metal-containing biocatalysts that are used in hydrogen production and in hydrogen-based biocatalytic reactions, e.g. for nicotinamide cofactor recycling. However, they are difficult to produce in large scale. While the efficient *E. coli*-based production of the relatively simple regulatory hydrogenase from the bacterium *Cupriavidus necator* could be recently solved in an UniSysCat project, large-scale heterologous synthesis of more complex, biotechnologically relevant hydrogenases remains even more challenging.

Within this project we will develop a mechanistic mathematical model for hydrogenase production in *E. coli* which is adaptable to different hydrogenases and involves aside from the modules which describe the synthesis of the enzyme components, also the cell growth and the environmental factors, such as the oxygen and metal dependency. This model serves as the basis for parametrisation of the specific expression clones and the consecutive bioprocess development in an automated high throughput laboratory. The final aim is the significantly faster and straight forward development of processes for the production of hydrogenases with a complex subunit composition.

Extended description version of the project:

Mathematical modeling of the synthesis of recombinant hydrogenase in E. coli as a tool for increasing its functional yield

Co-supervisor 1: Prof. Dr. Peter Neubauer, Technische Universität Berlin, Institute of Biotechnology, Chair of Bioprocess Engineering, Straße des 17. Juni 135, D-10623 Berlin, Germany, Germany
Co-supervisor 2: Dr. Oliver Lenz, Technische Universität Berlin, Institute of Chemistry, Straße des 17. Juni 135, D-10623 Berlin, Germany, Germany
Co-supervisor 3: Dr.-Ing. Nicolas Cruz Bournazou, Technische Universität Berlin, KIWI-biolab at the Chair of Bioprocess Engineering, Straße des 17. Juni 135, D-10623 Berlin, Germany, Germany

1. Overall goal of the project

The energy revolution, i.e., the transition from fossil fuels to alternative, CO_2-neutral energy sources, is currently one of the greatest and, at the same time, the most unavoidable social tasks in the fight against global warming. Thus, the search for and utilization of so called "green" alternative energy sources is an important task. A promising solution is the use of biohydrogen, which can be generated by specific bacterial biocatalysts called hydrogenases. However, currently the production of hydrogenases is limited to low yields, as they are complex proteins consisting of more functional subunits, containing metal cofactors, and their synthesis crucially depends on the environmental redox situation. While the production of the regulatory hydrogenase (RH) from *C. necator* as a model enzyme could be recently markedly improved by a step-by-step laborious one-factor-by-time strategy [1-3] so that enough protein can be provided for functional studies, the process is not yet optimal and robust enough for efficient production for practical applications.
Especially in the past years, the methods of model-based bioprocess optimization in connection with the development of automated laboratory systems have become an efficient tool for the development and optimization of complex bioprocesses. By combining digital cell and process models, it is possible to run *in silico* optimization scenarios and design dynamic experiments that lead to an optimum robust bioprocess significantly faster than standard trial and error approaches.

2. State of the art

Hydrogenases are abundant metalloenzymes that catalyze the reversible conversion of dihydrogen into protons and electrons. They can be found in many bacteria, archaea or unicellular eukaryotes. According to the metal composition of their active site, they are classified as [Fe]-, [FeFe]- or [NiFe]-hydrogenases. In addition to the metal ion-containing catalytic center, hydrogenases contain multiple Fe-S clusters that serve as electron relay. A unique feature of some [NiFe]-hydrogenases is their O_2-tolerance that makes these enzymes particularly interesting as this facilitates their biotechnological application (e.g. for biohydrogen production, in fuel cells or for cofactor regeneration). Unfortunately, the low production yields in their native hosts result in high costs for strain cultivation and enzyme purification, thus limiting their investigation and application. To circumvent this problem the heterologous production in a robust and genetically tractable host (e.g. *E. coli*) is a promising strategy. Moreover, the difficult maturation process required for the incorporation of the catalytic metal center and the multi-subunit architecture of the enzymes make them difficult-to-produce proteins. To tackle this, until now several approaches have been made to improve the heterologous production of [NiFe]-hydrogenases [4]. Using the regulatory hydrogenase (RH) from *C. necator* as a model enzyme, we already investigated relevant process parameters (e.g. overproduction strain, inducer concentration, temperature, production time) [1]. Furthermore, implementing a fed-batch like production process further improved the RH yield in shake flask scales several 100-fold compared to RH yields obtained from *C. necator* [1]. In initial attempts, only an inactive, cofactor-free RH was produced [1, 2]. In a further step, an enzyme with similar activity to the native hydrogenase was isolated from *E. coli* by carefully adapting the culture conditions and co-production of specific maturation genes [3]. Currently this recombinant regulatory hydrogenase can be produced as a functional enzyme with a yield of about 130 mg/L.

3. Specific aims and how they may be reached

So far, the heterologous hydrogenase production was achieved in shake flasks with a scale up to 500 ml. While other projects currently aim at the scaling to bioreactors, a more strategic optimization of complex bioprocesses for difficult products can be achieved by a combination of model-based experimental designs, laboratory automation, and model-based continuous redesign of the experiments. With the parametrized model for the product producing cell, an *in silico* simulation and optimization of different process scenarios becomes possible, which finally replaces many laborious and non-optimal wet-lab experiments.

An important basis for this is (i) the design of the digital model for the cell factory and the product synthesis pathway and (ii) the computer-based design of experiments which allow a parametrization of the cellular parameters. While the tools for doing this have been implemented in different projects at the KIWI-biolab at the Chair of Bioprocess Engineering, the principles so far have been applied only to Fab antibody fragments, which are considered as difficult-to-produce molecules but much less difficult to produce than the complex hydrogenases.

Therefore, it is the specific aim of this project to apply the available computation tools together with the high throughput laboratory robots to design a digital twin of the hydrogenase process and to parameterize it. This mathematical model will then be applied to design process scenarios for optimal production of the NAD^+-reducing [NiFe]-hydrogenase.

WP1: Development of a cellular mathematical model for hydrogenase production in *Escherichia coli*

Based on cellular mechanistic models describing growth and product formation for recombinant proteins [5], a model is developed that contains the different factors for the production of functional hydrogenase. The focus is on the implementation of measurable factors that can be experimentally determined and thus used for its parameterization. Thereby it is obvious that the currently available model must be extended with the different metabolic pathways which lead to active hydrogenase (synthesis of the protein subunits, the Fe-S clusters, and the $[NiFe(CN)_2(CO)]$ cofactor). Furthermore, also the environmental conditions (aerobic/anaerobic) must be properly considered in the model.

WP2: Design of dynamic experiments for parameter identification of the cellular model towards a digital twin

Dynamic fed-batch experiments are performed in the robot station to allow the determination of the model parameters. As the key parameters of the cell are dependent on the different phases of the process, a moving horizon estimator [6] must be applied, which describes the cellular parameters dependent on the external environmental conditions. This allows for adaptive fitting of the model, even if process conditions change suddenly. Experiments are run under different conditions, so that the model is calibrated and can be used under these different scenarios. Additionally, an extended Kalman filter is used to estimate the states even with low measurement frequency to enable better process understanding [7]. With the knowledge gained in each experiment, a new experiment is designed followed by principles of model-based design of experiments, sometimes referred to as optimal experimental design. The objective is to improve the parameter accuracy by running the experiment under conditions which offer a higher sensitivity. The parameters which are obtained from fitting the model are also determined with their inaccuracies using Bayesian approaches so that the model can be used for probabilistic *in silico* simulations yielding not only the mean output, but also their uncertainty [6].

WP3: Production of the soluble NAD^+-reducing [NiFe]-hydrogenase (SH) from *C. necator* in *E. coli*

The model will be used for process design, this is, to find the best industrial production strategy based on model-based simulations of possible scenarios. The most promising of them are then validated in test fermentation runs with the focus on the final yield of active hydrogenase and robustness, i.e. reproducibility of the process also under consideration of its scaling from the minibioreactor system to a standard benchtop bioreactor. Using the high-throughput cultivation platform in the Neubauer lab, it is possible to run several cultivations in parallel. This enables to perform rapid screening experiments, especially when supported by the model-based framework

developed in **WP2**. To optimize the whole process, the findings of the simulation need to be confirmed and optimized using the experimental settings. Hence, cultivations under different process conditions are performed, which are optimized using model-predictive control (MPC). MPC uses the fitted model from WP2 to iteratively optimize the feed rate in the remaining cultivation, considering several constraints [8]. One possible beneficial constraint would be to run the process under reducing, but not anaerobic conditions, which is a difficult task to be performed when using standard approaches. MPC enables to find an optimal feeding regime, even when only limited information about the strain is present under the used cultivation conditions. By this, the process can be iteratively improved until optimal conditions to produce the biotechnologically relevant NAD^+-reducing hydrogenase are found.

4. References:

1. Fan Q, Caserta G, Lorent C, Lenz O, Neubauer P, Gimpel M. Optimization of Culture Conditions for Oxygen-Tolerant Regulatory [NiFe]-Hydrogenase Production from *Ralstonia eutropha* H16 in *Escherichia coli*. Microorganisms. 2021 May 31;9(6):1195. doi: 10.3390/microorganisms9061195.
2. Fan Q, Neubauer P, Gimpel M. Production of soluble regulatory hydrogenase from *Ralstonia eutropha* in *Escherichia coli* using a fed-batch-based autoinduction system. Microb Cell Fact. 2021 Oct 18;20(1):201. doi: 10.1186/s12934-021-01690-4.
3. Fan Q, Caserta G, Lorent C, Zebger I, Neubauer P, Lenz O, Gimpel M. High-Yield Production of Catalytically Active Regulatory [NiFe]-Hydrogenase From *Cupriavidus necator* in *Escherichia coli*. Front Microbiol. 2022 Apr 29;13:894375. doi: 10.3389/fmicb.2022.894375.
4. Fan Q, Neubauer P, Lenz O, Gimpel M. Heterologous Hydrogenase Overproduction Systems for Biotechnology-An Overview. Int J Mol Sci. 2020 Aug 16;21(16):5890. doi: 10.3390/ijms21165890.
5. Anane E, García ÁC, Haby B, Hans S, Krausch N, Krewinkel M, Hauptmann P, Neubauer P, Cruz Bournazou MN. A model-based framework for parallel scale-down fed-batch cultivations in mini-bioreactors for accelerated phenotyping. Biotechnol Bioeng. 2019 Nov;116(11):2906-2918. doi: 10.1002/bit.27116. Epub 2019 Jul 30.
6. Krausch N, Barz T, Sawatzki A, Gruber M, Kamel S, Neubauer P, Cruz Bournazou MN. Monte Carlo Simulations for the Analysis of Non-linear Parameter Confidence Intervals in Optimal Experimental Design. Front Bioeng Biotechnol. 2019 May 24;7:122. doi: 10.3389/fbioe.2019.00122.
7. Kemmer A, Fischer N, Wilms T, Cai L, Groß S, King R, Neubauer P, Cruz Bornazou MN. Nonlinear state estimation as tool for online monitoring and adaptive feed in high-throughput cultivations. A*uthorea.* 2022 August 09, doi: 10.22541/au.166005030.05586352/v1.
8. Krausch N, Kim JW, Barz T, Lucia S, Groß S, Huber MC, Schiller SM, Neubauer P, Cruz Bournazou MN. High-throughput screening of optimal process conditions using model predictive control. Biotechnol Bioeng. 2022 Sep 15. doi: 10.1002/bit.28236. Epub ahead of print