Abstract:

Structural studies on pyrimidine nucleoside phosphorylases to improve kinetics for sugar-modified nucleosides

Co-supervisor 1: Prof. Dr. Maria Andrea Mroginski, UniSysCat, TU-Berlin

Co-supervisor 2: Prof. Dr. Peter Neubauer, UniSysCat, BIMoS, TU-Berlin

Pyrimidine nucleoside phosphorylases (PyNPs) catalyze the reversible phosphorolysis of pyrimidine nucleosides in the presence of phosphate. Research on this class of enzymes is highly interesting due to the following reasons: I. PyNPs play a key role in the purine salvage pathway. II. PyNPs are valuable catalysts for the synthesis of nucleosides and their analogues, which form a major class of small-molecule pharmaceuticals: 50% of anti-viral drugs and 25% of anti-cancer drugs are nucleoside analogs. In the proposed project structure function relationships of PyNPs will be investigated in detail by means of a multidisciplinary approach combining bioengineering, computational modelling and structural biology. While molecular modelling will be performed in the group of Prof. Mroginski, cloning, expression and biochemical characterization of mutant enzymes will be done in the group of Prof. Neubauer. In addition, collaborations with Prof. Müller (machine learning), Prof. Hildebrandt (spectroscopy) and Prof. Bange (x-ray crystallography, Univ. Marburg) are planned.

Extended description version of the project:

Structural studies on pyrimidine nucleoside phosphorylases to improve kinetics for sugar-modified nucleosides

Co-supervisor 1: Prof. Dr. Maria Andrea Mroginski, TU-Berlin (UniSysCat)
Co-supervisor 2: Prof. Dr. Peter Neubauer, TU-Berlin (UniSysCat, BIMoS)

1. Overall goal of the project

This project aims to study the structure function relationship of pyrimidine nucleoside phosphorylases (PyNPs) in detail. PyNPs have not yet been studied intensively although they are highly valuable biocatalyst for the synthesis of nucleoside analogues with pharmaceutical and biotechnological importance. Although PyNPs share a high degree of sequence similarity and conservation of active site residues, differences in kinetic properties are observed. Since initial studies revealed that PyNPs undergo large conformational changes during catalysis, an improved understanding of PyNPs is not possible simply on crystal structure analysis. In this project, flexible docking and molecular dynamics simulations of the enzyme's flexibility will be taken into account for in-depth studies on PyNPs. In the Neubauer group interesting modified PyNPs will be heterologously expressed in *E. coli* and analyzed with respect to kinetic parameters.

2. State of the art

PyNPs catalyze the reversible phosphorolysis of pyrimidine in the presence of phosphate. PyNPs are valuable catalysts in the synthesis of nucleosides and their analogues. Nucleoside analogs are important small molecule drugs to treat cancer and viral infections, but all of these drugs are currently only synthesized by laborious and inefficient chemical syntheses. Based on their structural characteristics nucleoside phosphorylases (NPs) are classified into two main groups. The NP-I family consists of Purine NPs (PNPs). PNPs have been widely studied as the severe symptoms of PNP deficiency suggested possible chemotherapeutic applications of potent inhibitors of this enzyme[1]. The NP-II family includes thymidine phosphorylases (TP) and the wider spectrum PyNPs[2,3].
PyNPs have been mainly isolated from thermostable organisms such as *Bacillus stearothermophilus*, *Geobacillus thermoglucosidasius* and *Thermus thermophilus*. Thermostable NPs are interesting as they withstand harsh reaction conditions such as high pH or organic solvents[4]. Also, they are expected to have a wider substrate spectrum than their mesophilic counterparts. Indeed, compounds with modifications at the C2′ position of the nucleoside are accepted as substrates by PyNP. Arabinosides as well as 2′-deoxy-2′-fluororibosides and 2′-deoxy-2′-fluoroarabinosides are substrates for some PyNPs, predominantly for those isolated from thermophilic organisms such as *T. thermophilus*.
In depth understanding of the structure-function relationship of proteins in general requires precise knowledge of the structural and dynamical properties of the entire system and of the electronic properties of the corresponding active site and potential substrates. This information can be obtained with the help of computation and state-of-art molecular modelling techniques. Commonly used in modelling studies are molecular docking techniques which aim to predict the best binding mode of a ligand to a protein, by sampling conformations/orientations of the ligands in the active site pocket. Since, most docking algorithms consider rigid or semi-rigid entities where the protein atoms are held

[1] Bzowska et al. 2000. Purine nucleoside phosphorylases: Properties, functions, and clinical aspects. *Pharmacol Therapeutics* 88(3), 349–425.
[2] Pugmire & Ealick. 2002. Structural analyses reveal two distinct families of nucleoside phosphorylases, *Biochem J* 361, 1–25.
[3] Yehia *et al.* 2017. Substrate spectra of nucleoside phosphorylases and their potential in the production of pharmaceutically active compounds. *Curr Pharm Design*, 23(45), 6913–6935.
[4] Kamel *et al.* 2020. Thermophilic nucleoside phosphorylases: Their properties, characteristics and applications. *BBA - Proteins Proteomics*. Elsevier, 1868(2).

fixed to their crystallographic positions, they are not applicable to flexible proteins, such as PyNPs. Here, flexible docking strategies combining classical docking algorithms with molecular dynamics simulations are the best option for quantifying substrate-protein interactions[5]. Furthermore, significant improvement of the docking technique is achieved by using hybrid quantum mechanics/molecular mechanics- based scoring functions. Such a flexible-docking-QM/MM approach has been already successfully applied to investigate substrate specificity and potent inhibitors of PyNP from *B. subtilis*[6].

3. Specific aims and how they may be reached:

To study structure function relationships for PyNPs, initial experiments will focus on 2´-modified nucleosides. The effect of a fluorine substituent in the 2´-position of ribo- and arbinonucleosides will be investigated. A fluorine atom at the C2´-arabino position of nucleosides or nucleotides is known to exert a wide variety of physicochemical effects like a strengthening of the glycosidic bond. In earlier studies we could show that the specific activity is decreasing if substrates with 2´-modifications are used. Interestingly, kinetic parameters strongly differ between PyNPs of different origins. These differences are not attributed to changes in active side residues. Given the fact that the catalytic process in the NP-II family is coupled to large conformational changes of the protein matrix, this project will extend the characterization of PyNPs by investigating the effect of protein dynamics on their catalytic activity. Therefore a multidisciplinary approach combining, molecular biology/bioengineering, computational modelling and structural biology will be iteratively applied to investigate different substrate specificity of two selected thermostable PyNPs.

WP 1: Gene synthesis, cloning and expression of mutant PyNP. Single and multi-position mutants will be cloned in a suitable expression strain after gene synthesis. If an increasing number of mutant PyNP variants will be studied, the high-throughput facility of the Neubauer group will be applied. It allows for an automated cloning in 96- up to 384-well plate format.

WP 2: Kinetic characterization of the mutant PyNP. After the successful expression of the mutant PyNP variants, kinetic parameters (specific activity, K_m, K_{cat}) will be determined using 2´-sugar modified nucleosides. To determine kinetic parameters a recently developed photometric assay[7] (Kaspar *et al.*, 2019) in combination with HPLC analysis will be applied.

WP 3: Dynamic computational modelling. Realistic models of the Enzyme-Substrate Complex will be generated using state-of-art flexible docking algorithms. Existing crystal structures of PyNP and mutants, which have been solved in collaboration with G. Bange will be used as template; otherwise structural models will be generated by homology modelling. Dynamic properties of the wild type and mutated enzymes in absence and in presence of specific substrates will be further investigated by classical molecular dynamics simulations, while interaction energies between the active site and the ligand will be accurately predicted by quantum mechanics using hybrid QM/MM approaches.

Initial protein crystallization experiments were performed in cooperation with **Prof. Bange (Marburg)**. Further studies are carried out either with him or, due to closer geographical & scientific vicinity, with **Prof. Dobbek** or **Dr. Scheerer**. Spectroscopic experiments will be done in collaboration with **Prof. Hildebrandt** and **Prof.Esser**. Development and implementation of machine-learning based quantum chemical estimation of scoring functions in docking studies will be performed in collaboration with **Prof. Claus-Robert Müller (TU-Berlin).**

---

[5] Salmaso & Moro 2018. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process : An Overview. Front. Pharmacol. 9, 1–16.

[6] Balaev *et al.* 2016. Substrate specificity of pyrimidine nucleoside phosphorylases of NP-II family probed by X-ray crystallography and molecular modeling. *Cryst Rep.* 61(5), 830–841.

[7] Kaspar *et al.* 2019. A UV / Vis Spectroscopy-based assay for monitoring of transformations between nucleosides and nucleobases. *Meth Prot.* 2(60), 1–13.