# Kloniranje, ekspresija i karakterizacija fosforiboziltransferaza iz bakterije Escherichia coli

**Maršić, Tin**

**University of Zagreb**

**Faculty of Science**

**Department of Biology**

**Tin Maršić**

# CLONING, EXPRESSION AND CHARACTERIZATION OF *ESCHERICHIA COLI* PHOSPHORIBOZYLTRASFERASES

# MASTER'S THESIS

**Zagreb, 2019**

This master‘s thesis was written at the chair of Molecular Biotechnology of the Technical University of Dresden, Faculty of Biology, under the supervision of Dr. Christoph Loderer. Master's thesis was submitted to the University of Zagreb, Department of Biology in fulfillment of the requirements for the degree of Master in Molecular biology.

I would like to express my gratitude to my family, friends and mentors for their patience, kindness and expertise that made my education and this thesis possible.

# TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Biološki odsjek Diplomski rad

## KLONIRANJE, EKSPRESIJA I KARAKTERIZACIJA FOSFORIBOZILTRANSFERAZA IZ BAKTERIJE *ESCHERICHIA COLI*

Tin Maršić

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Nukleotidi su jedna od temeljnih klasa bioloških spojeva potrebnih za ispravnu funkciju, rast i replikaciju svih živih organizama.

Alternativni nukleotidi nalaze primjenu u medicini kao citostatici i antivirotici. Novonastajuće polje sintetičke biologije koristi alternativne nuleotide kao gradivne jedinice ksenonukleinskih kiselina (XNA).

Fosforiboziltransferaze (PRT) su skupina enzima koja sintetizira nukleotide u procesu zvanom salvage pathway. Zajedničko svojstvo ovih enzima je dodavanje dušičnih baza na ribozu, koja je prethodno aktivirana u oblik 5-fosforibozil 1-pirofosfat (PRPP).

Svrha ovoga rada bila je klonirati, eksprimirati i pročistiti pet fosforiboziltransferaza iz *E. coli* te utvrditi optimalne uvjete ekspresije, protokol za pročišćavanje, radnu temperaturu te optimalni pH kao i specifičnost za suptrate. Ovaj rad je zamišljen kao prvi korak u nastojanju da se proizvede umjetni biosintetski put ribonukleotida koji bi omogućio sintezu alternativnih nukleotida za upotrebu u znanstvene ili biotehnološke svrhe.

Pet fosforiboziltransferaza iz *E.coli* je klonirano, eksprimirano preko noći pri 37°C, pročišćeno putem His-Tag afinitetne kromatografije te dodatno putem gel filtracije. Kao optimalni radni pufer za sve enzime, pokazao se HEPES. pH pri kojem su enzimi HPRT i UPRT najaktivniji su 7.5 odnosno 8**,** a optimalne temperature su 60°C odnosno 50°C. Pokazalo se da su XPRT i HPRT sposobni procesuirati 6-merkaptopurin, međutim to nije bio slučaj sa APRT-om. Također, 5-fluorouracil se pokazao supstratom za UPRT, ali ne i za OPRT.

(46 stranica, 44 slike, 4 tablice, 27 literaturnih navoda, jezik izvornika: engleski

Rad je pohranjen u Središnjoj biološkoj knjižnici

Ključne riječi: *E. coli*, nukleotidi, alternativni nukleotidi, fosforiboziltransferaza, karakterizacija

Voditelj: Dr. sc. Christoph Loderer, doc.

Suvoditelj: Dr. sc. Nenad Malenica, doc.

Ocjenitelji: Dr. sc. Višnja Besendorfer, red. prof., dr. sc. Martina Šeruga Musić, izv. prof., dr. sc. Inga Marijanović, izv. prof.

Rad prihvaćen: 18.9.2019.

# BASIC DOCUMENTATION CARD

University of Zagreb

Faculty of Science

Department of Biology

Graduation Thesis

## CLONING, EXPRESSION AND CHARACTERIZATION OF *ESCHERICHIA COLI* PHOSPHORIBOZYLTRASFERASES

Tin Maršić

Rooseveltov trg 6, 10000 Zagreb, Croatia

Nucleotides are an essential class of biological compounds necessary for proper function, growth and replication of all living organisms.

Alternative nucleotides have an application in medicine as anticancer and antiviral drugs. Newly emerging field of synthetic biology uses alternative nucleotides as the building blocks of xenonucleic acids (XNA)

Phosphoribosyltransferases (PRTs) are a group of enzymes utilized by living systems to synthetise nucleotides via salvage pathway. The shared property by these enzymes is the addition of nucleobases to ribose, which is previously activated in the form of 5-phosphoribosyl 1-pyrophosphate (PRPP).

The purpose of this work was to clone, express and purify five *E. coli* phosphoribosyltransferases and determine their optimal expression conditions, establish a purification protocol, working temperature and optimal pH as well as supstrate specificity. This work is a first step proof of the concept as part of a wider effort to produce a biosynthetic ribonucleotide synthesis pathway that would potentially allow for the production of alternative nucleotides for use in scientific research and biotechnology.

Five *E.coli* phosphoribosyltransferases were cloned, expressed at 37°C overnight, purified by His-Tag affinity and size-exclusion chromatography. Optimal buffer for all enzymes was established to be HEPES, and pH optima for HPRT and UPRT are 7.5 and 8 respectively**,** while optimal temperatures are 60° and 50°C respectively; 6-mercaptopurine was found to be taken up by XPRT and HPRT but not by APRT, while 5-fluorouracil is processed by UPRT but not by OPRT.

(46 pages, 44 figures, 4 tables, 27 references, original in: English

Thesis deposited in the Central Biological Library

Key words: *E. coli*, nucleotides, alternative nucleotides, phosphoribozyltransferase, characterization

Supervisor: Dr. Christoph Loderer, Asst. Prof.

Cosupervisor: Dr. Nenad Malenica, Asst. Prof.

Reviewers: Dr. Višnja Besendorfer, Prof., Dr. Martina Šeruga Musić, Assoc. Prof., Dr. Inga Marijanović, Assoc. Prof.

Thesis accepted: 18.9.2019.

**TABLE OF CONTENTS**

# 1. INTRODUCTION

## 1.1. Nucleotide biosynthesis

Nucleotides are one of the key classes of biological compounds necessary for proper function, growth and replication of all living organisms. Their importance and ubiquity is apparent in the number and range of different biological processes in which they are involved. They are the building blocks of nucleic acids but are also utilized in many other cellular functions. For example, ATP, as well as GTP are involved in energy and phosphate group transfer, and sometimes serve as signal molecules. Nucleotide derivatives such as UDP-glucose are involved in glycogen biosynthesis.

There are two general strategies utilized by the cell to synthetize nucleotides. In *de novo* pathway, nucleotides are assembled from simpler precursors. Pyrimidine bases are synthetized *de novo* piece by piece and then added to PRPP (Ross 1981).

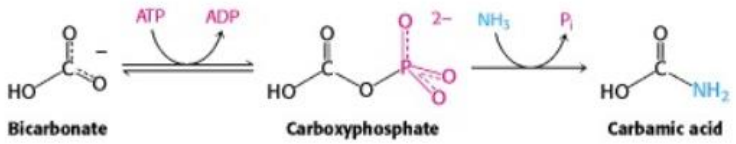



**Figure 1.1**. Carbamic acid synthesis by carbamoyl phosphate synthetase.

The first step involves synthesis of carbamic acid from bicarbonate by carbamoyl phosphate synthetase (CPS) in a two-step process that cleaves two molecules of ATP (Figure 1.1). In the first step carbamic acid is produced using $NH_3$ molecule derived from glutamine.

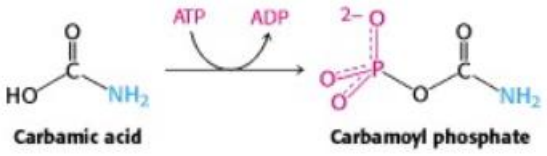



**Figure 1.2.** Carbamoyl phosphate synthesis by carbamoyl phosphate synthetase.

The second step catalyzed by the same enzyme yields carbamoyl phosphate (Figure 1.2)

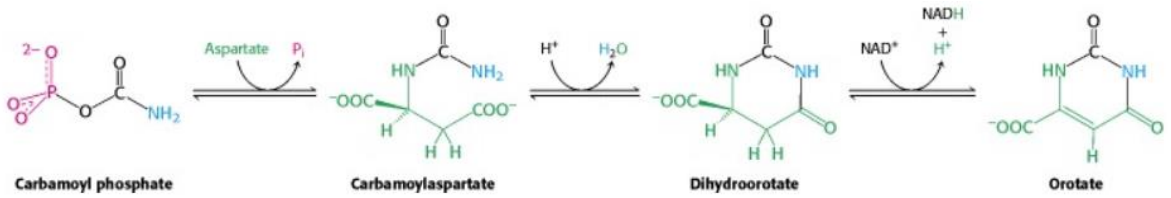



**Figure 1.3.** Orotate synthesis by CAD protein complex.

Carbamoyl phosphate is converted into orotate in a series of reactions that are catalyzed by a single protein complex in mammals called CAD (carbamoyl phosphate synthetase, aspartate transcarbamoylase and dihydroorotase) (Figure 1.3). Orotate is then added to PRPP to yield orotidylate by the enzyme pyrimidine phosphoribosyltransferase, which is homologous to other similar enzymes that catalyze the addition of other groups of pyrimidine bases to PRPP. Uridylate is formed from orotidylate by decarboxylation mediated by the enzyme orotidylate decarboxylase. CTP is formed by replacing the carboxyl group by amino group of UTP by the enzyme cytidine triphosphate synthetase.

Purine nucleotides are synthetized on PRPP scaffold in a succession of steps, which involve phosphorylation followed by the replacement of the latter with ammonia, or amine group acting as a nucleophile (Figure 1.4)

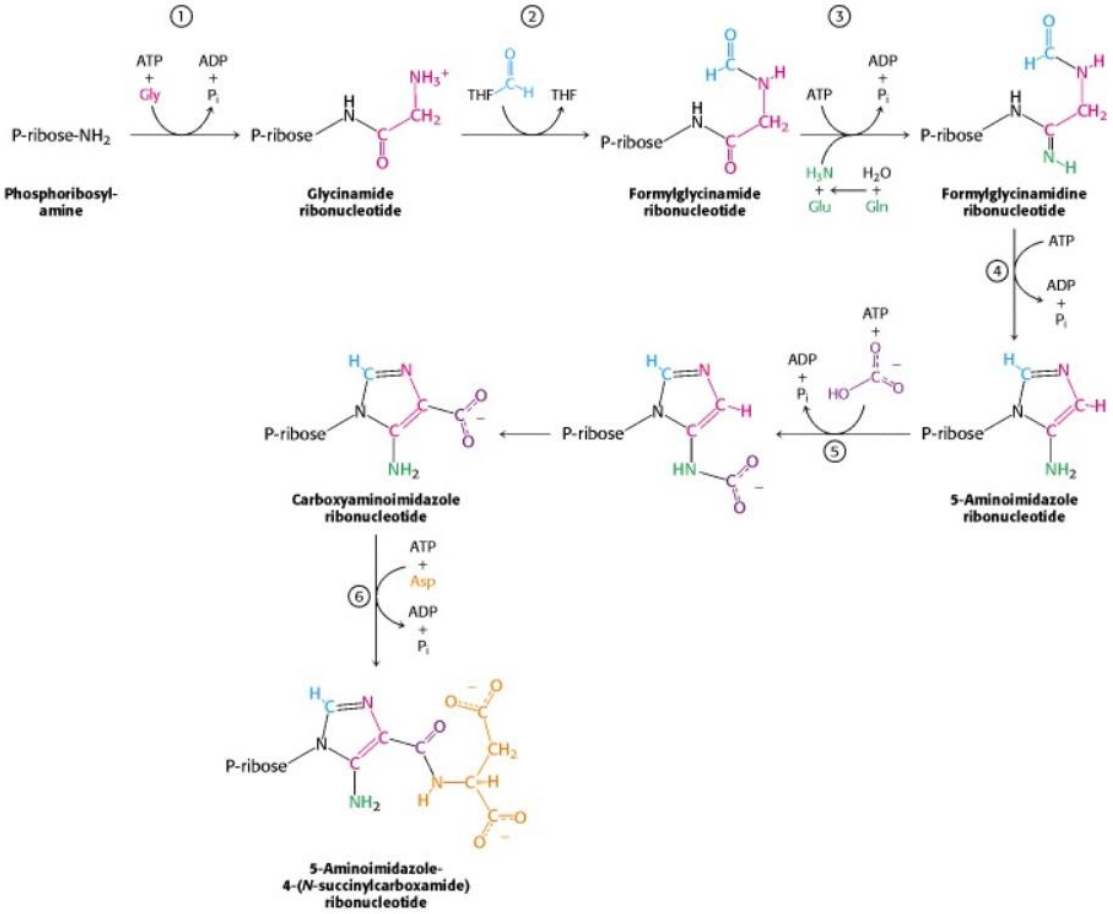



**Figure 1.4.** Partial scheme of purine nucleotide synthesis.

A particularly relevant aspect of nucleotide metabolism for our purposes are salvage pathways. Salvage pathways are a useful alternative for generating nucleotides which are more energy conserving and allow for the bases taken by digestion or generated from the breakdown of DNA or RNA to be converted into their respective nucleotide analogues. Salvage enzymes adenine phosphoribosyltransferase and hypoxanthine-guanine phosphoribosyltransferase catalyze the formation of AMP and GMP respectively (Figure 1.5).

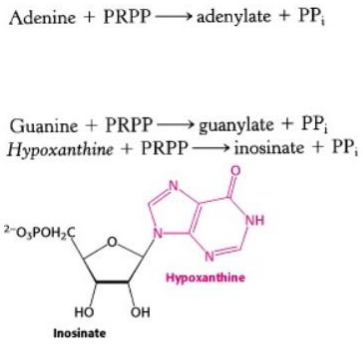



**Figure 1.5.** Overview of reactions catalyzed by APRT and HGPRT.

Analogous reactions take place for pyrimidine nucleotide synthesis. Adenosine deaminase and a purine nucleoside phosphorylase catalyze the conversion of adenine to hypoxanthine, which is a substrate of hypoxanthine phosphoribosyltransferase (HPRT). The product, IMP, is a precursor for AMP and GMP. In most bacteria and nearly all eukaryotes, HPRTs also catalyze the salvage of guanine and, in some cases, xanthine. It is interesting to note that many parasitic organisms are unable to synthesize purines via *de novo* pathways and rely completely on the enzymes in salvage pathways for the synthesis of purine nucleotides thus making parasitic enzymes interesting targets for antiparasitic drugs (Craig & Eakin 2000).

## 1.2. Phosphoribosyltransferases

The focus of this research is the characterization of phosphoribosyltransferases (PRTs), which are a group of enzymes utilized by living systems to synthetise nucleotides via salvage pathways. The common property of these enzymes is that they add activated nucleobases to ribose, which is previously activated in the form of 5-phosphoribosyl 1-pyrophosphate (PRPP). The reaction is ordered and sequential with PRPP binding first followed by the purine base (Figure 1.6). After catalysis, pyrophosphate ($PP_i$) is released before the nucleotide (Craig & Eakin 2000). The illustration below shows the general reaction scheme for purine bases (Scism et al. 2007):

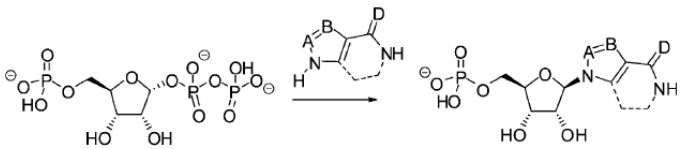


**Figure 1.6.** Scheme showing a general reaction catalized by phosphoribosyltransferases.

*E.coli* PRPPs included in this research were hypoxanthine phosphoribosyltransferase, xanthine phosphoribosyltransferase, adenine phosphoribosyltransferase, uracil phosphoribosyltransferase and orotate phosphoribosyltransferase (Scism et al. 2007).

Purine salvage enzymes can be divided into three monophyletic sets, each of which catalyzes mechanistically similar reactions: adenine-, xanthine-, hypoxanthine and guanine-phosphoribosyltransferases. These enzymes share considerable homology among themselves, as well as among nucleoside phosphorylases such as adenine deaminase, adenosine deaminase, adenosine monophophate deaminase, guanine reductase and inosine monophosphate dehydrogenase. These homologies suggest that substrate specificity is the result of gene duplication, and that the purine nucleotide salvage pathways evolved by a process that probably took place before the divergence of the three cell domains, namely Bacteria, Archaea, and Eucarya (Becerra & Lazcano 1998).

### 1.2.1. Adenine phosphoribosyltransferase

Adenine phosphoribozyltransferase catalyzes the formation of AMP from adenine and PRPP using either $Mg^{2+}$ or $Mn^{2+}$ as a cofactor. It is a homodimer with two catalytic sites. Enzyme activity is inhibited by both pyrophosphate and AMP as well as dAMP, ATP, dATP, ADP and dADP and unaffected by cyclic phosphate derivatives. Literature data suggest that guanine, hypoxanthine, xanthine, aminoimidazolecarboxyamide and 6-mercaptopurine are not substrates for APRT (Hochstadt-Ozer 1972). In humans, APRT defects are not lethal, however they do cause 2,8-dihydroxyadenine urolithiasis that manifests in a rare form of kidney stones. The purine recycling process is particularly important for tumor cells which points to APRT and other enzymes of the purine salvage pathway as potential targets for chemotherapeutic treatment.

The crystal structure of adenine phosphoribozyltransferase from *Leishmania donovani* was determined by Phillips et al. in complex with adenine, AMP with sulphate and citrate ions mimicking phosphate moieties (Figures 1.7 and 1.8). The structure was found to be similar to that of other phosphoribosyltransferases, but the adenine binding domain shows significant differences. The active site is composed of residues from both subunits which indicates that dimerization is necessary for activity (Phillips et al. 1999).

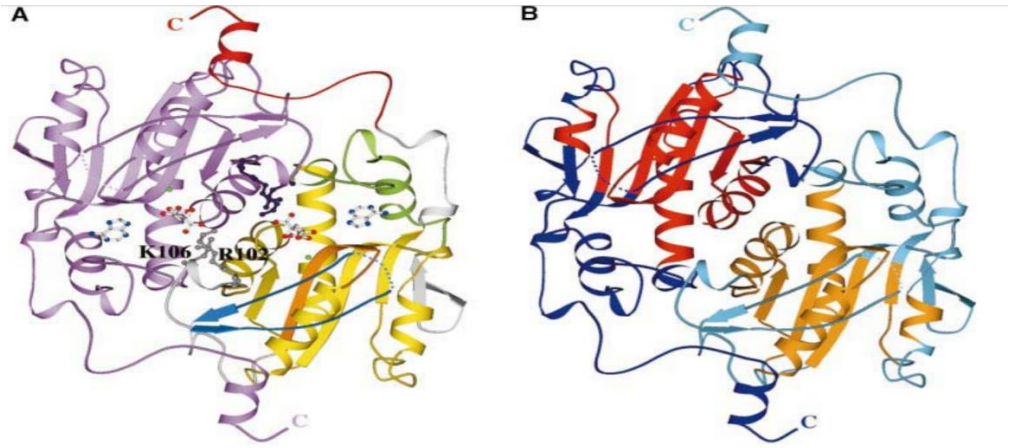



**Figure 1.7.** Crystal structure of adenine phosphoribozyltransferase (APRT) homodimer from *Leishmania donovani*. One subunit is colored in purple, $Mg^{2+}$ is colored in green and sulphate and citrate are colored in gray.

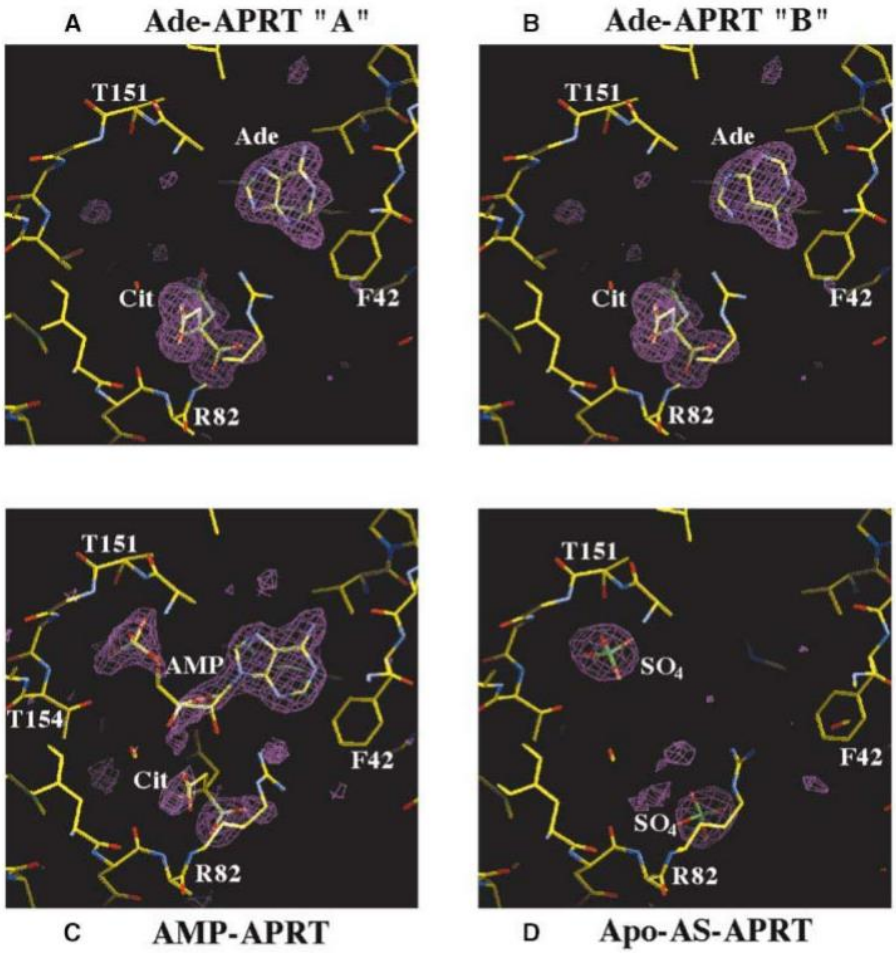



**Figure 1.8.** Crystal structure of APRT from *Leishmania donovani* with focus on active site. (**A**) Positions of adenine and citrate within the active site (**B**) Positions of adenine and citrate within the same active site in a different orientation (**C**) Position of AMP and citrate within the active site (**D**) Positions of sulphate moieties within the active site

### 1.2.2. Xanthine-guanine phosphoribosyltransferase

In *E.coli* an enzyme called xanthine-guanine phosphorybosyltransferase (XGPRT) is responsible for replenishing XMP and GMP via the salvage pathway. The enzyme is membrane-bound and acts as purine transporter, which simultaneously converts its substrates to nucleotides. Two different XGPRTs were found in *E. coli*, the first preferentially uses hypoxhantine and guanine as substrates and the second uses hypoxanthine, xanthine and guanine as substrates (Deo et al. 1985). An analogous mammalian enzym called hypoxanthine phosphoribosyltransferase (HPRT) was characterized and it was determined that it catalyzes the conversion of hypoxanthine and guanine to IMP and GMP respectively, but the difference is that the bacterial enzyme can convert xanthine to XMP far more efficiently than its mammalian analogue (Mulligan & Berg 2006). The structure of XPRT was determined to be a tetramer both in solution and in the crystal form, and contains four active catalytic sites (Figures 1.9 and 1.10). The presence of $Mg^{2+}$ is necessary as it stabilizes PRPP within the active site, and it probably departs with pyrophosphate after nucleotide formation (Vos et al. 1998).

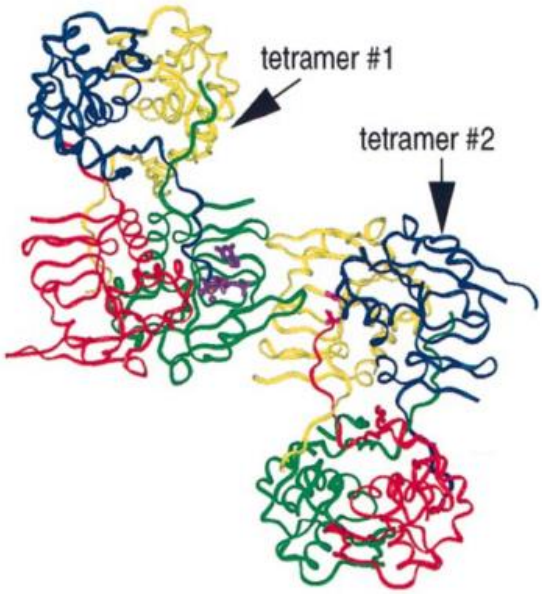



**Figure 1.9.** Crystal structure of two XPRT tetramera from *E. coli* rotated in two different positions. Supstrate (guanine) is shown in purple.

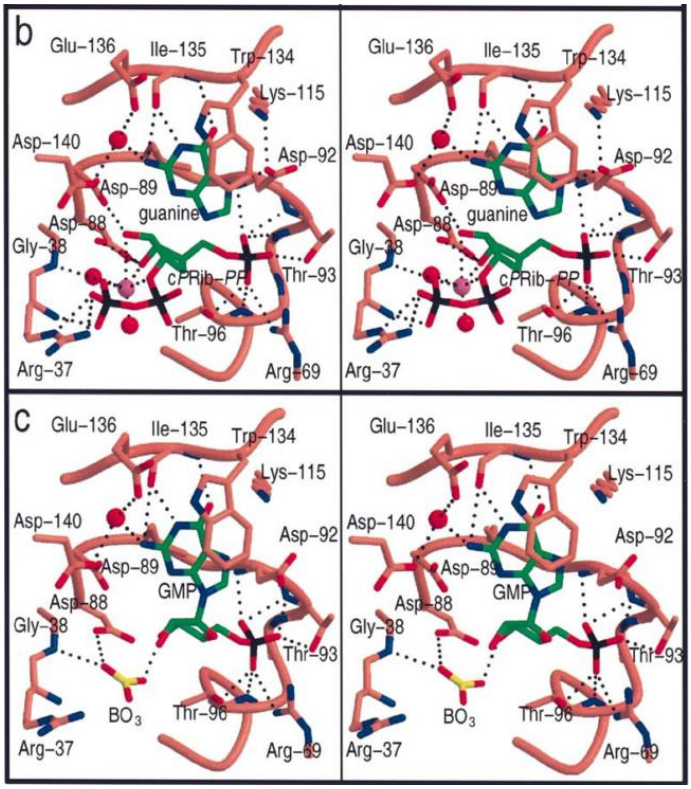



**Figure 1.10**. Crystal structure of XPRT from *E. coli* active site. Interactions between supstrates (guanine and PRPP), product (GMP) and surrounding amino acids moieties are shown.

### 1.2.3. Hypoxanthine-guanine phosphoribosyltransferase

HGPRT acts primarily on hypoxanthine, while its affinity for guanine is rather low. The enzyme shows no activity towards xanthine. The human HPRT can salvage xanthine, but at low levels. Xanthine is normally converted to uric acid for excretion. HPRTs likely descended from a common ancestral *hpt* gene of prokaryotes, and substrate specificity can be significantly modified by single amino acid substitution (Craig & Eakin 2000). *E. coli* protein is built from four identical subunits (Figures 1.11 and 1.12). Two $Mg^{2+}$ cations are bound per subunit (Keough et al. 2002). *E. coli* HPRT is distinct from other known 6-oxopurine PRTases in that it preferes hypoxanthine as substrate over both xanthine and guanine. Its substrate specificity is due to the modes of binding of the bases. Unlike mammals and most parasites, *E. coli* cells express two 6-oxopurine phosphoribozyltransferases, with different specificities for hypoxanthine, guanine, and xanthine. Salvage enzymes generally allow for a more energy efficient synthesis of purine nucleoside monophosphates compared with the *de novo* pathway (Ullman & Carter 1997). The kinetic analysis suggests that *E. coli* HGPRT is mainly responsible for the synthesis of IMP and that XGPRT primarily salvages guanine and xanthine. In *E. coli* HPRT, the carbonyl oxygen of Asp163 likely forms a hydrogen bond with the 2-exocyclic nitrogen of guanine (in the HPRT-guanine-PRib-PP-$Mg^{2+}$ complex) (Keough et al. 2002). Work by Subbayya et al. suggests that the N-terminal residues of the *P. falciparum* are not part of the active site pocket. However in human HGPRT the N-terminus is a part of the binding site for GMP and PRPP (Subbayya et al. 2000). It is interesting to note that mutations in HPRT cause a rare genetic disease in humans called Lesch-Nyhan disease. The afflicted have specific neurological and behavioral symptoms such as self-injury by biting, aggression, spitting and the use of foul language. Other symptoms include mild mental retardation, the accumulation of uric acid which leads to gout and sandy sludge or stones in the urinary system (Jinnah 2009)

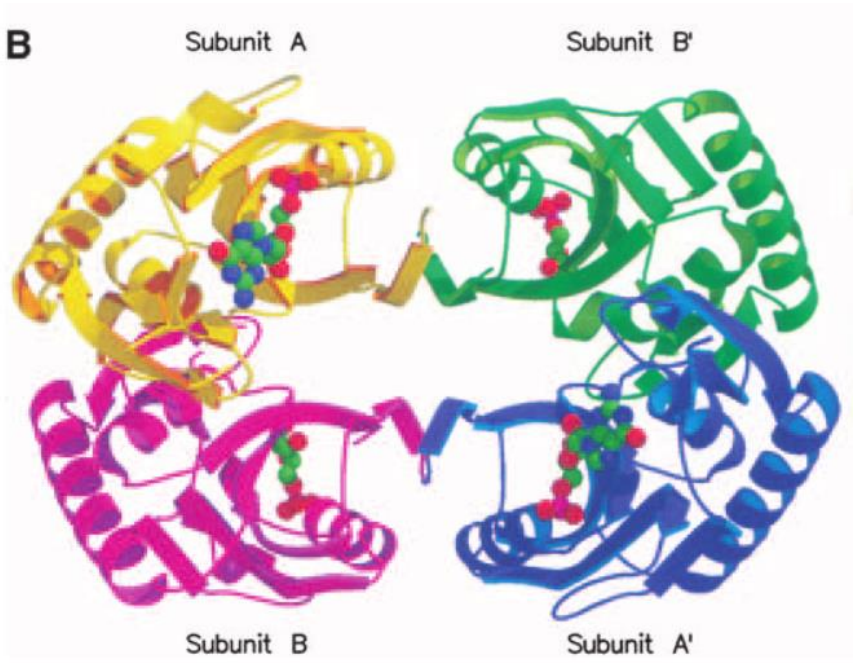



**Figure 1.11.** Crystal structure of HPRT tetramer from *E. coli*. Subunits A contain active sites that bind PRPP and hypoxanthine or guanine.

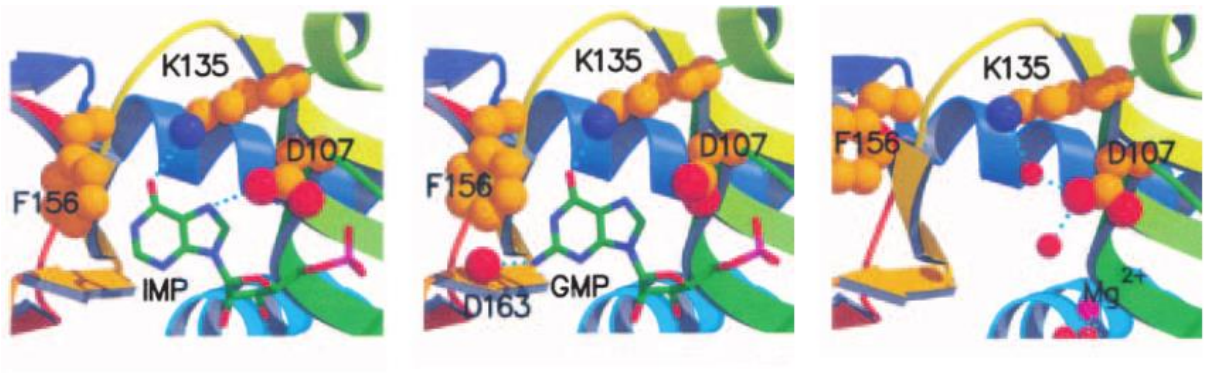



**Figure 1.12.** Crystal structure of *E. coli* HPRT active site of subunit A

### 1.2.4. Uracil phosphoribosyltransferase

Uracil phosphoribosyltransferase (UPRT) catalyzes the transfer of a ribosyl phosphate group from α-D-5- phosphoribosyl-1-pyrophosphate to the N1 nitrogen of uracil. It has three similar subunits and molecular weight of 75 000. Uracil phosphoribosyltransferase from *E. coli* is activated by GTP and inhibited by uridine nucleotides. The specificity of UPRT towards other uracil analogues was determined by Rassmusen et al. 6-azauracil showed 97% activity, 5-fluorouracil showed 216% activity, orotate, cytosine, thymine and hypoxanthine showed less than 1% activity where uracil uptake as a substrate represents 100% activity (Rasmussen et al. 1986). Studies carried out on purified UPRT from the *E.coli* (Rasmussen et al. 1986) and *Acholeplasma laidlawii* (McIvor et al. 1983) indicated that they exist as a homotrimer and homodimer respectively. Studies on UPRT purified from the protozoan *Crithidia luciliae* (Asai et al. 1990) and the yeast *Saccharomyces cerevisiae* (Natalini et al. 1979) revealed that these proteins exist as a homodimer and a heteroligomer, respectively (Figures 1.13 and 1.14). Importantly, pyrimidines containing the substituents larger than fluorine at position 5, such as the methyl group of thymine, are not utilized as substrates by the *Toxoplasma gondii* UPRT (Iltzsch and Tankersley 1994). In contrast, 5-fluorouracil is bound by UPRT and converted to the nucleotide level by this enzyme. Thus, this pyrimidine functions as a subversive substrate (Schumacher et al. 1998).

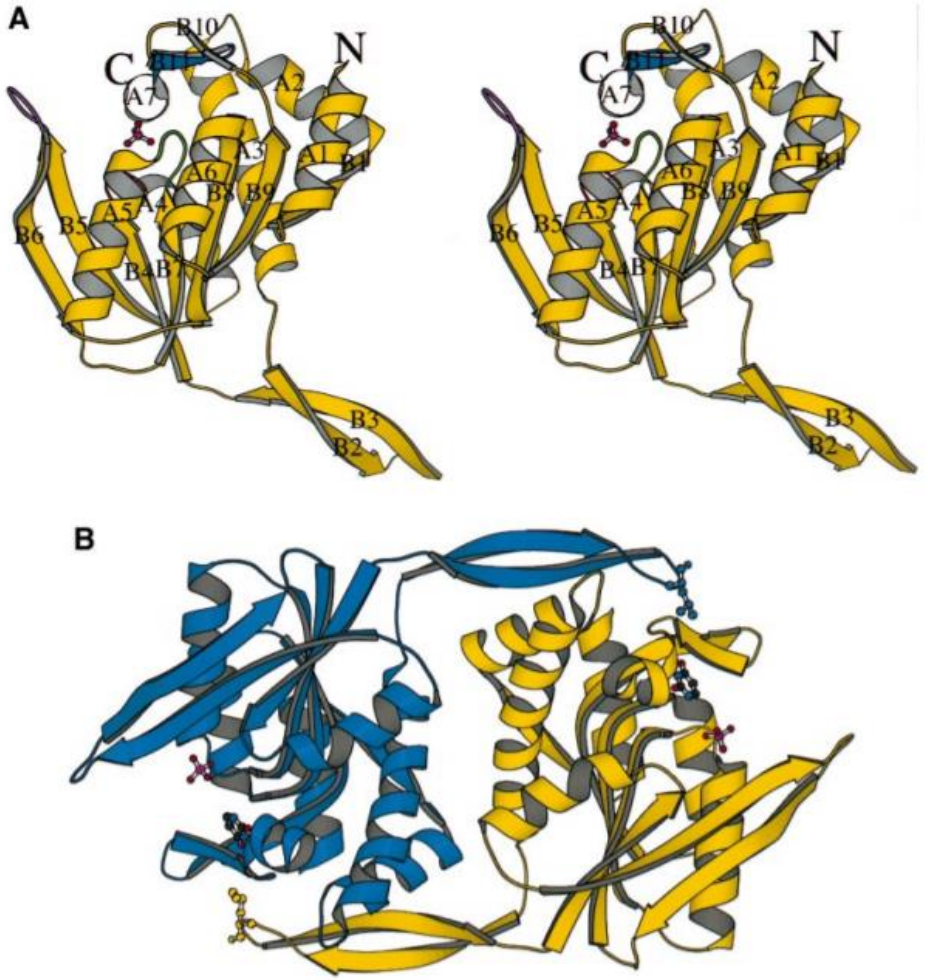



**Figure 1.13.** (A) Crystal structure of *Toxoplasma gondii* UPRT monomer (B) Crystal structure of Toxoplasma gondii UPRT dimer.

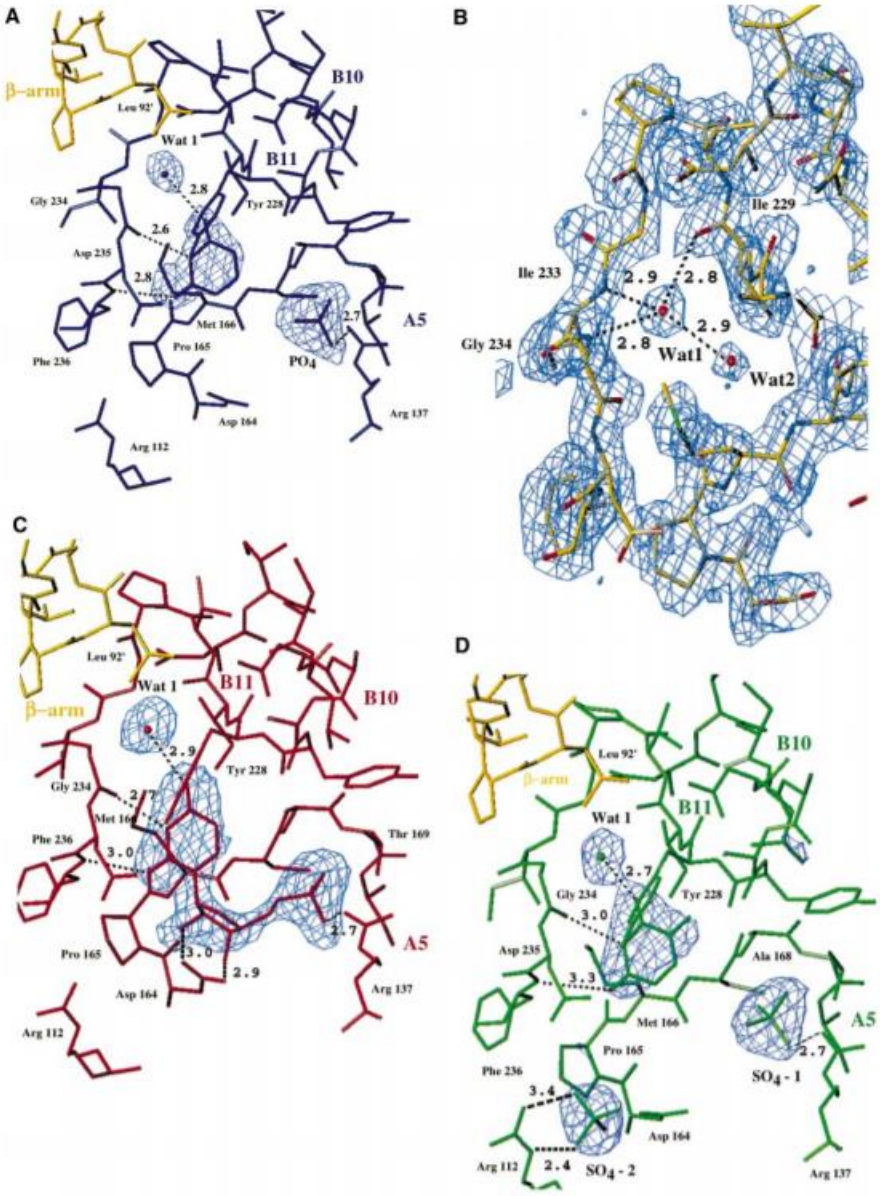



**Figure 1.14.** Crystal structure of *T. gondii* UPRT active site. Interactions between various supstrates and surrounding amino acids are shown.

### 1.2.5. Orotate phosphoribosyltransferase

Orotate phosphoribosyltransferase or OPRT utilizes PRPP and orotate to form pyrophosphate and orotidine 5′-monophosphate (OMP) in the presence of divalent cations. Crystalline *E. coli* OPRT is a homodimer (Fig. 1.15). A study conducted by Henriksen et al. determined that a closed conformation of the flexible loop also conserved in *S. typhimurium* is involved in OPRT catalysis at the pyrophosphate binding site. Structures of OPRT co-crystallized with orotate and sulfate, or with OMP, $Mg^{2+}$, and sulfate indicate the existence of two orotate/OMP binding sites (Henriksen et al. 1996). OPRT in humans is the main enzyme involved in phosphoribozylation of 5-fluorouracil (5-FU), which leads to therapeutic response and inhibition of tumor growth. Study by Ochiai et al. concluded that lower OPRT activity is associated with poor survival in

patients with colorectal cancer and recommend OPRT assay prior to 5-FU administration (Ochiai et al. 2006).

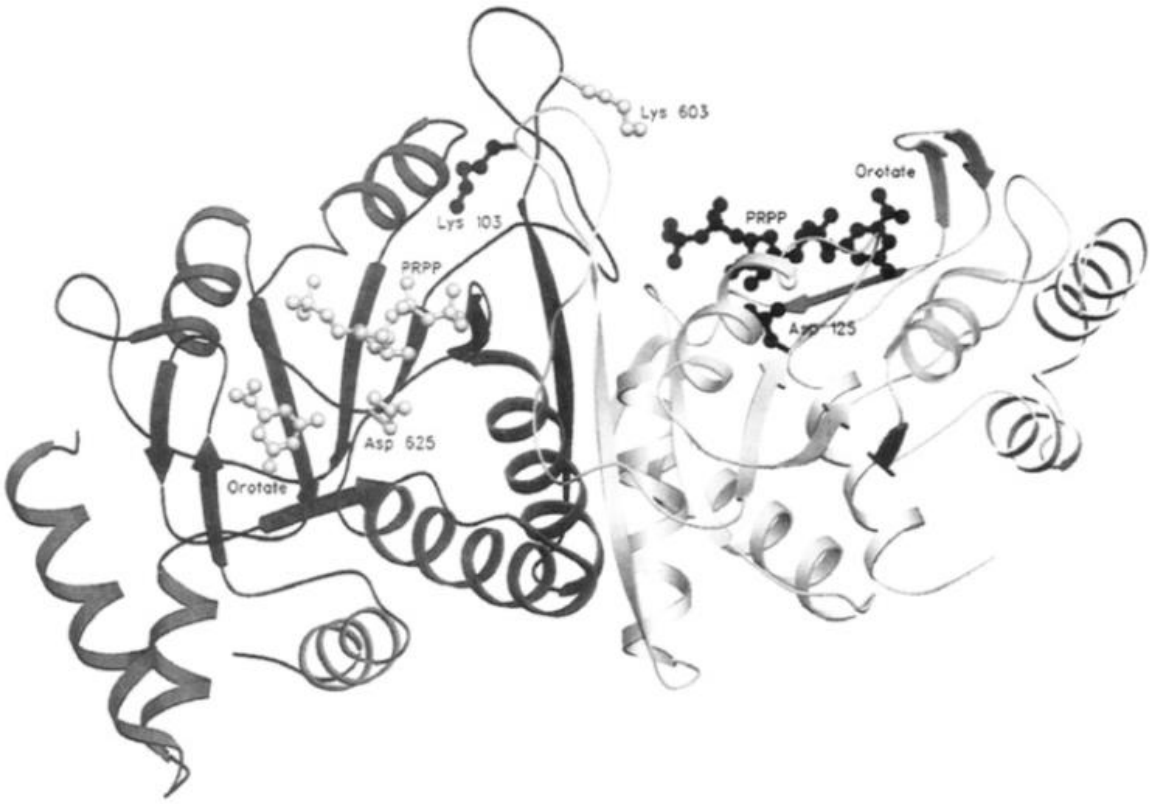



**Figure 1.15.** Crystal structure of OPRT from *E. coli* and positions of active sites on two subunits

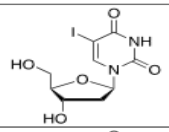
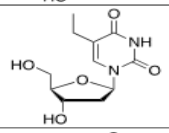
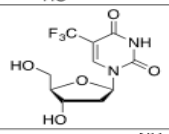
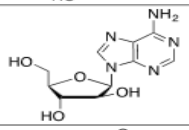
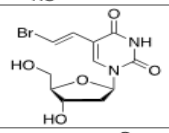
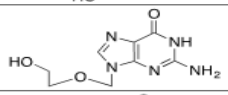
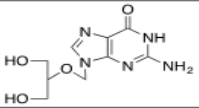
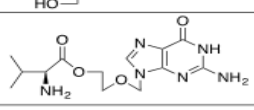
## 1.3. Potential applications of phosphoribosyltransferases

Nucleotide analogues have a wide range of applications in clinical practice. Nucleoside analogues and nucleobases are a family of drugs, which include cytotoxic compounds, antiviral agents, and immunosuppressive molecules. They were one of the first chemical agents used in cancer treatment. The anticancer drugs include derivatives of natural purine and pyrimidine nucleosides and nucleobases. The most widely used representatives of purine analogues cladribine and fludarabine have been used for the treatment of low-grade blood malignancies. Pyrimidine analogues such as cytarabine have their application in acute leukemia treatment and gemcitabine can be used to treat various solid tumors. The mechanism of action of these compounds is well understood. Most of these agents enter the cells via specialized transporter proteins. Inside the cells, they are phosphorylated by intracellular enzymes to yield triphosphate derivatives. Active derivatives of nucleoside analogues act as cytostatics by incorporation into DNA and RNA or by interfering with various enzymes involved in the synthesis of nucleic

acids, such as DNA polymerases and ribonucleotide reductase. These actions result in the inhibition of DNA synthesis and apoptotic cell death (Galmarini et al. 2002).

Nucleotide analogues have been traditionally used as antiviral drugs since 1959, starting with a compound called iodoxuridine which was originally synthetized as an anticancer drug. Over the years many similar compounds have been developed for addressing various viral diseases. Classical examples include vidarabine and brivudine**,** which are used for the treatment of herpes viruses and act as DNA polymerase inhibitors. Acyclic nucleosides as acyclovir, ganciclovir, valacyclovir, cidofovir, valganciclovir, penciclovir and famcyclovir have been used for the treatment of herpes virus infections. A class of 2`,3`-dideoxynucleosides (ddNs) are known to be the most effective therapeutic agents against HIV. They owe their efficacy to the absence of 3`-hydroxyl which makes elongation of DNA sequence impossible. Examples include zidovudine, didanosine, zalcitabine, stazudine and abacavir. L-nucleosides as lamivudine, emtricitabine in addition to acyclic nucleoside prodrugs such as tenofovir disoproxil fumarate, tenofovir alafenamide inhibit the viral cycle of retroviruses such as HIV and HBV by acting as reverse transcriptase inhibitors. Ribavirin was developed in 1986 to treat hepatitis C and viral hemorrhagic fever. The newest addition to the arsenal, sofosbuvir, has recently been approved by FDA as a novel therapy for hepatitis C (HCV) as a nonstructural polymerase inhibitor (Mahmoud et al. 2018).

**Table 1.1.** Examples of frequently used nucleotide analogues in treatment of viral infections

| Structure | Compound | Virus | Company |
|---|---|---|---|
|  | Idoxuridine, IDU, Herplex | HSV, VZV | GSK, 80s, no FDA label |
|  | Edoxudine, EDU, Aedurid | HSV | Upjohn, 1969 |
|  | Trifluridine, TFT, Viroptic | HSV | King Pharmaceutical 1980 |
|  | Vidarabine, Ara-A, Vira-A | HSV, VZV | Parkedale pharmaceuticals, no FDA label |
|  | Brivudine, BVDU, Helpin | HSV, VZV | Berlin Chemie, 80s, no FDA label |
|  | Acyclovir, ACV, Aciclovir | HSV, VZV | GSK, 1982 |
|  | Ganciclovir, GCV, Cytovene | CMV | Hoffmann-La Roche, 1989 |
|  | Valaciclovir, VACV, V altrex | HSV, VZV, CMV | GSK, 1996 |

Mechanistic studies and enzyme structures could be used for development of inhibitors of PRTs, that would act as chemical blockers of their activity in the forseeable future (Wang 1984). Such inhibitors would represent a novel group of drugs for the treatment of diseases caused by protozoan parasites (Craig & Eakin 2000). In the recent years, a prospect of designing chemical inhibitors of parasitic phosphoribozyltransferases began to emerge (Somoza et al. 1998). An example is the atempt to develop an inhibitor of *Plasmodium falciparum* HG(X)PRT, exploiting the fact that many parasites completely depend on salvage pathways to convert purine bases from the host to nucleotides needed for DNA and RNA synthesis (Eakin et al. 1997). An approach to developing novel antimalarial drugs atempted use *P. falciparum* HG(X)PRT to convert artificial purine base analogs to nucleotides toxic to the parasite. This strategy requires that these compounds be good substrates for the parasite enzyme but poor substrates for the human analogue (Keough et al. 2006).

Another, newly emerging use of alternative nucleotides can be found in the field of synthetic biology. One of the goals of synthetic biologists is to try to produce viable biological organisms that do not occur in nature. The aim is to design an alternative genetic material different from DNA and RNA, namely XNA (xenonucleic acid). XNA would theoretically consist of a variety of structural and chemical alterations which would make this novel information-storing biopolymer incompatible with their natural biological counterparts. The incompatibility with the natural world offers an opportunity to create a "genetic firewall" that renders the exchange of genetic information with the natural world impossible, which means it could be implemented as the ultimate biosafety measure (Schmidt 2010).

Recently a breakthrough in creating a semisynthetic organism was reported by Lavergne et al. The team managed to grow *E. coli* in the presence of unnatural nucleoside triphosphates dNaMTP and d5SICSTP as well as provide the means to import them via expression of a plasmid-borne nucleoside triphosphate transporter. These newly created semisynthetic organisms were also able to replicate their DNA containing a single dNaM-d5SICS unnatural base pair. The results were not entirely satisfying because *E.coli* cells engineered in this way grew slowly and were prone to losing information stored in unnatural base pairs. For that reason the semisynthetic organisms were additionally modified to transport and use a more chemically optimized unnatural base pair, and to ensure that the organism eliminates any DNA that might have lost the unnatural bases by using Cas9 system. The optimized semisynthetic *E. coli* grew faster, constitutively imported unnatural triphosphates, and was able to indefinitely retain multiple unnatural base pairs resulting in a form of life that can stably store genetic information using a six-letter, three-base-pair alphabet (Lavergne et al. 2017).

Production of different nucleotides, their analogues and pharmaceuticals still relies on chemical synthesis, which makes them rather expensive in many cases. Attempts are in progress to utilize the family of phosphoribosyltransferases involved in salvage pathways of nucleotide synthesis from various organisms in order to make them more readily available for medical uses and scientific research. Recently, an approach was reported in developing a biocatalytic system for the synthesis of nucleotide analogues. It involves one-step biocatalytic process for the synthesis of several nucleotide analogues from commercially available starting materials by utilizing *E. coli* whole-cell extract as the biocatalyst and a purification procedure using anion-exchange chromatography. The goal of the study was to overcome potential practical problems such as the specificity of the enzymatic systems involved, difficulty in obtaining purified enzymes, and

the separation of polar reaction products from polar starting materials and byproducts. The study identified a specific mutant of *E. coli* hypoxanthine phosphoribosyltransferase (HPRT), named 8B3PRT, produced by error-prone PCR library of the *hpt* gene in a protein expression host with improved in vivo transformation of triazole carboxamide to ribavirin monophosphate. The newly obtained mutant was also examined with respect to its ability to utilize a structurally diverse range of commercially available purine and purine base analogues. Purine phosphoribosyltransferase mutant 8B3PRT demonstrated enhanced activity and relaxed specificity in processing a wide variety of nucleoside base analogues. In the case of ribavirin monophosphate formation from triazole carboxamide, the enzyme showed 8-fold improvement of turnover versus wild type HPRT and 17-fold improvement for thioguanine (Scism et al. 2007).

### 1.4. Objectives

The purpose of this thesis was to clone, express and purify five *E.coli* phosphoribosyltransferases and determine their optimal expression conditions, establish a purification protocol, working temperature and pH as well as substrate specificity. This work is a first step proof of concept as part of a wider effort to produce a biosynthetic ribonucleotide synthesis pathway that would potentially allow for the production of alternative nucleotides for use in scientific research and biotechnology. Additional work is needed to genetically engineer enzymes for more efficient and more specific uptake of alternative nucleotides.

## 2. MATERIALS AND METHODS

### 2.1. Chemicals

1. Gibson assembly

5X ISO buffer:

dGTP

dATP

dTTP

dCTP

1M DTT

PEG-8000

10mM NAD

Master mix:

5X ISO buffer

T5 exonuclease

Phusion polymerase

2. Bacterial cultivation and protein expression

Agar

LB medium

Kanamycin

IPTG

Glycerol

3. Electrophoresis

Agarose

10X TAE buffer

Acrylamide

SDS

APS

TEMED

SDS buffer

Isorpropanol

Staining solution

4. Other chemicals used

Thymine

Orotic acid

Hypoxanthine

Xanthine

Cytosine

Guanine

Uracil

Adenine

NaCl

HEPES

TRIS-hydrochloride

MES

Potassium dihydrogen phosphate

Potassium hydrogen phosphate

Potassium phosphate

TBAH

Methanol

PMSF

DNase I

Lyzozyme

Imidazole

$MgCl_2$

PRPP

**2.2. Equipment**

Thermomixers and heaters:

Eppendorf Thermomixer Compact thermomixer

Eppendorf Thermomixer Comfor thermomixer

PEQLAB table heater

Micro centrifuge and shaker:

Sprout Micro centrifuge MZ001-S

NeoLab D-6012 shaker

PCR thermocycler:

Eppendorf Mastercycler gradient

Photometer:

Eppendorf BioPhotometer

Centrifuges:

Sigma 1-15K

Eppendorf mini Spin plus

Sigma 3-18K

Sigma 3K30

Size exclusion and His-tag purification:

GE Amersham Biosciences AKTA FPLC System

pH meter:

InoLab pH Level 1

Magnetic stirrers and heaters:

Heidolph MR 3001

Heidolph MR 3001

Scales:

Acculab ATILON

KERN 770

Sartorius TE214S

HPLC:

Knauer D-141463

Sonicator:

Sartorius LABSONIC®M

Laminar:

HERA safe 12 1/PE

Incubators:

Sanyo incubator MIR-153

Infors Minitron

Pippetes

Eppendorf Research Plus

## 2.3. Methods and experimental procedures

### 2.3.1. Primer design

DNA sequences for designated proteins adenine phosphoribosyltransferase, orotate phosphoribosyltransferase, xanthine phosphoribosyltransferase, hypoxanthine phosphoribosyltransferase, uracil phosphoribosyltransferase (labeled APT, OPT, XPT, HPT, UPT respectively) were found using UniProt online database and transferred to Serial Cloner. Gene sequences were checked for restriction sites, copied and pasted into pET28b(+) vector sequence between NdeI and BamHI restriction sites with the N-terminal His-tag. Gibson assembly primers were designed using NEBuilder program, and checked for self-annealing and hairpin formation using Oligo Calc internet website. The primers were order from Sigma Aldrich.

### 2.3.2. *E.coli* genomic DNA extraction

For the isolation of *E. coli* DNA 20 mL of LB medium were inoculated with BL21 (DE3) from a glycerol stock and incubated at 37°C at 220 rpm over night. The genomic DNA was isolated by taking two times 2mL from the overnight culture and using the RTP Bacteria DNA Mini Kit according to the following protocol. Bacterial culture was spun down at 11 000 x g for 3 minutes and the supernatant was carefully removed. This step was repeated after adding the remaining 2 mL aliquot. 400 µL of resuspension Buffer R was added to the pellet by pipetting up and down. The resuspended sample was transferred into the extraction tube and vortexed shortly. The sample was incubated in a thermomixer at 65°C for 10 minutes, and then at 95°C for 10 minutes. 400 µL of Binding Buffer B6 was added to the sample followed by short vortexing. The sample was loaded onto the RTA Spin Filter Set and incubated for 1 minute at room temperature and then centrifuged at 11 000 x g for 2 minutes. The flow-through was discarded and 500 µL of Wash Buffer I was added followed by centrifuge at 11 000 x g for 1 minute. The filtrate was discarded and the RTA Spin Filter was placed into a new RTA Receiver Tube. 600 µL of Wash Buffer II was added and the sample was centrifuged at 11 000 x g for 1 minute, the filtrate was discarded and the centrifuge was repeated once more at 11 000 x g for 4 minutes. The Spin Filter was placed into a new 1.5 mL Receiver Tube, 80 µL of Elution Buffer was added and the samples were left for 5 minutes. The Elution Buffer was previously preheated at 50°C. The samples were centrifuged for 1 minute at 11 000 x g.

### 2.3.3. Fragment preparation by PCR

Delivered primers were diluted ten fold to the final volume of 100 µL in nuclease-free water. Master mix for PCR was prepared by mixing 70 µL of 5X Phusion HF Buffer, 7 µL of 10 mM dNTPs, 7 µL of template, 3.5 µL of Phusion DNA Polymerase, and 226.5 µL of nuclease-free water and distributed into six 0.2 mL tubes. After the addition of 2.5 µL specific 10 µM forward and reverse primers into each tube the PCR reaction was started using the following protocol: initial denaturation at 98°C for 1 minute, denaturation at 98°C for 20 seconds, annealing at gradient temperature for each individual sample (APT=66.5°C, OPT=55.3°C, XPT=60.5°C,

HPT=57.4°C, UPT=60.3°C) for 20 seconds, extension at 72°C for 3 minutes, final extension at 72°C for 5 minutes and finally 4°C until removal.

In the next step 150 mL of 1.2% agarose gel was prepared with 1.8 g of agarose and dissolving it in 150 mL of 1X TAE buffer with the addition of 7.5 µL of Red Safe™ nucleic acid staining solution. 5 µL of ladder (1kbp) and then 4 µL of loading dye mixed with 2 µL of sample DNA was poured into the wells and electrophoresis was run at 90 V for 45 minutes.

Plasmid DNA was linearized by PCR reaction according to the following protocol: 10 µL of 5X HF Buffer, 1 µL of 10 mM dNTPs, 2.5 µL of 10 µM forward and reverse primers each, 1 µL of template, 0.5 µL of Phusion® DNA polymerase, 1.5 µL of DMSO (final concentration 3%) and nuclease free water to the final volume of 50 µL. Four samples were prepared to obtain a better yield. PCR conditions were the same as for the gene fragments but the annealing temperature was set to 62.3°C.

### 2.3.4. Fragment purification

Gene fragments were purified using NucleoSpin® Gel and PCR Clean-up kit. One volume of sample was mixed with 2 volumes of NTI buffer. 700 µL of this mixture was loaded into the column and centrifuged for 30 seconds at 11 000 x g. After discarding the flow-through and placing the column back into the collection tube, the membrane was washed with 700 µL of NT3 buffer and centrifuged for 30 seconds at 11 000 x g. The last step was repeated one more time. The columns were centrifuged for 3 minutes at 11 000 x g to remove residual ethanol. The columns were placed into new 1.5 mL collection tubes. 30 µL of preheated NE buffer (50°C) was added and the samples were incubated for 5 minutes before they were centrifuged for 1 minute at 11 000 x g.

For the purification of vector DNA gel NucleoSpin® Gel and PCR Clean-up protocol for gel excision was used. Volumes of each sample were measured with a pipette and total volume of 191.8 µL was obtained, and 38.36 µL of dye was added to the sample. 0.8% agarose gel was prepared by weighing 1.6 g of agarose, adding 10 µL of Red Safe™ and adding 200 mL of 1X TAE buffer. The **s**ample was loaded onto the gel and the electrophoresis was set to 80 V and 50 minutes. Meanwhile 1.5 mL collection tubes were weighed before gel excision. After cutting the gel into small cubes and weighing them following masses were obtained $m_1$=0.561 g, $m_2$= 0.543 g, $m_3$= 0.581 g, $m_4$= 0.459 g. The following volumes of buffer were added $V_1$=1122 µL, $V_2$=1086 µL, $V_3$=1162 µL, $V_4$=0.918 µL (200 µL of buffer per 100 mg of excised agarose gel). The mixtures were incubated at 50°C until the gel was completely dissolved. 750 µL of sample were added per tube and centrifuged for 30 seconds at 11 000 x g. This was repeated for all samples and until the whole volume was centrifuged. The samples were then washed with 700 µL of NT3 buffer, centrifuged for 30 seconds at 11 000 x g, and the step was repeated. The columns were centrifuged again at 11 000 x g to completely remove NT3 buffer. The columns were placed into new 1.5 mL tubes and 30 µL of previously preheated NE buffer (50°C) was added to the membrane. The samples were incubated for 5 minutes and then centrifuged for 1 minute at 11 000 x g. Gene fragments were checked by electrophoresis. 1.2% agarose gel with Red Safe™ was used, 5 µL 1 ladder (1 kbp) and a mixture containing 4 µL of loading dye and 2 µL of sample DNA were loaded onto the gel. Electrophoresis settings were adjusted to 90 V and 45 minutes. The resulting bands looked as expected.

### 2.3.5. Gibson assembly

Cloning of the inserts into a pET28b(+) vector was accomplished using a DNA cloning method called the Gibson assembly. There are three possible approaches to cloning using this method which differ in enzyme types and the number of thermocycle steps. The selected method employed T5 exonuclease, Phusion® DNA polymerase, and Taq ligase in a one-step isothermal reaction (Gibson 2011).

5 mixtures of insert and vector were prepared in molar excess of 5:1 in favor of the insert.

6 µL of vector and insert mixtures were added into the 13 µL Gibson-Mastermix solution and 1 µL of Taq-ligase (40 U/ µL) was added into each reaction tube. The reaction took place in the thermocycler under following conditions:

i) 50°C for 60 minutes

ii) 72°C for 15 minutes

Once the reaction was over 1 µL of DpnI was added and the samples were incubated for 1 hour at 37°C (Gibson 2011).

### 2.3.6. Transformation of bacterial cells

Competent cells were produced from a bacterial preculture by transferring bacteria from a single colony into fresh LB medium. Primary culture was produced by inoculation of 100 mL of fresh LB medium in 1:1000 ratio. After reaching the $OD_{600}$= 0.4 - 0.55, the culture was cooled on ice for 10 – 15 minutes, after which it was centrifuged at 2700 x g for 10 minutes at 4°C. LB medium was discarded, and leftover fluid was removed with a pipette. The pellet was resuspended by adding 33 mL of ice cold RF1 solution which consisted of 0.4 g of RbCl, 0.27 g of $MnCl_2 \times 2H_2O$, 0.098 g of $CH_3COOK$, 0.049g of $CaCl_2 \times 2H_2O$ and 5 g of glycerin. The suspension was incubated on ice for 15 minutes and centrifuged at 580 x g for 15 minutes on 4°C. The excess fluid was discarded and 4 mL of ice cold RF2 was added. RF2 solution was prepared by dissolving 0.105 g of MOPS, 0.06 g of RbCl, 0.55 g of $CaCl_2 \times 2\ H_2O$ and 7.5 g of glycerin in 50 mL of water. The pellet was resuspended with a pipette and incubated on ice for 15 minutes. Finally, 100 µL aliquots were prepared, frozen in liquid nitrogen and stored at -80°C. Transformation efficiency was calculated to be 1.68 X $10^6$/1µL of plasmid for BL21 (DE3) strain and 1.28 X $10^6$/1µL of plasmid for NEB -10 beta strain.

Two mixtures of medium were prepared, one containing 5.5 g of agarose, 7.62 g of LB medium and 305 mL of WDC water, and the other containing 10 g of LB medium (25g/L, Roth) dissolved in 400 mL of WDC. Both bottles were autoclaved for 20 minutes at 120°C. 305 µL of kanamycin was added into LB + agar medium which was then poured into Petri dishes.

3 µL of Gibson assembly reaction mixture was gently added into 5 tubes containing NEB -10 beta strain of *E.coli*, mixture was swirled slowly and incubated on ice for 15 minutes. Cells were then put on 42°C for exactly 30 seconds, and immediately after 300 µL of NEB- 10 beta/ Stable outgrow medium was added. Cells were placed on 37°C for 40 minutes. The cells were centrifuged for 2 minutes in 7000 rpm, planted on the Petri dishes and then incubated at 37°C over night.

### 2.3.7. Colony PCR

Six desirable colonies were selected and marked from each Petri dish. New Petri dishes were taken and divided into six equal regions. Each selected colony was transferred into the new plate and planted onto the designated areas using a toothpick. Before planting the colony the toothpick was shortly inserted into low osmolarity miliQ water in a tube. Tubes were heated for 10 minutes on 95°C and spun down for 2 minutes. 2 µL of the supernatant was used as a template for colony PCR. 30 colony PCR reaction mixtures each contained the following: 5 µL of 5X One Taq Standard Reaction Buffer, 0.5 µL of 10 mM dNTPs, 0.5 µL of 10 µM forward and 0.5 µL of 10 µM reverse primers, 0.13 µL of One Taq DNA Polymerase, 2 µL of template and nuclease free water to 25 µL. PCR reaction conditions were: 95°C for 3 minutes and 30 cycles of 95°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute and 72°C for 5 minutes. Once the PCR was over, 5 µL of 6X dye was added into each tube and 10 µL of this mixture from each sample was transferred onto the gel. The electrophoresis was run at 100 V and for 45 minutes.

### 2.3.8. Digestion analysis

5 flasks each containing 10 mL of LB medium and 10 µL of kanamycin were prepared and inoculated with bacteria from the lane number 6 of each individual gene fragment and incubated over night at 37°C. The next day 3.6 mL (2x 1.8 mL) of the culture were taken and centrifuged for 30 seconds at 11 000 x g. Supernatant was discarded and 250 µL of A1 buffer was added and the pellet was resuspended. A2 buffer was added and the sample was incubated for 3 minutes after which 300 µL of A3 buffer was added, followed by the gentle inverting of the tubes until discoloration of the content. The samples were centrifuged at 11 000 x g for 10 minutes. 750 µL of the supernatant was added in the spin column, centrifuged for 1 minute at 11 000 x g and the flow-through was discarded. The membrane was washed with 500 µL of AW, centrifuged for 1 minute at 11 000 x g and 600 µL of A4 buffer was added. The sample was centrifuged again to remove residual solvents for 1 minute at 11 000 x g. The spin column was transferred to the new collection tube. AE buffer was preheated at 50°C and 50 µL were used to wash the membranes. After 5 minutes of incubation the samples were centrifuged at 11 000 x g for 1 minute. Protocol for the digestion was the following: 1 µL of each restriction enzyme (NdeI and BamHI), DNA template 5 µL, 5 µL of 1X NEB Buffer and nuclease-free water to the total reaction volume of 50 µL. The reaction took place at 37°C for 3 hours.

### 2.3.9. Sequencing

7.5 µL from each of the samples was taken and 2.5 µL of the pET28b(+) reverse primer was added and transferred into the 1.5 ml tube. The tubes were labeled with bar codes provided by the manufacturer and sent for sequencing. The sequencing service was provided by GATC Biotech™. The results of both DNA sequencing and the amino acid sequence were compared with the sequences in our database using BLAST® algorithm to verify the frames and to exclude the presence of any mutations.

### 2.3.10. Expression optimization

Preculture for expression was prepared by filling five 100 mL flasks with 10 mL of LB medium (25 g/L, Roth) and adding 10 µL of kanamycin in each. Bacterial colonies were selected and transferred into the flasks using a pipette which were then incubated at 37°C over night. The next day $OD_{600}$ values were measured by pipetting 100 µL of the over-night culture and mixing

them with 900 µL of LB medium for dilution. Five flasks were prepared and filled with 50 mL of the LB medium and 50 µL of kanamycin. Volumes of the samples were calculated so that the new suspension has an approximate $OD_{600}$ = 0.15. The volumes were V(APT) = 1.2 mL, V(OPT) = 1.2 mL, V(XPT) = 1.5 mL, V(HPT) = 1.55 mL, V(UPT) = 1.5 mL. The flasks were incubated at 37°C on a shaker starting at 10:45h. Small aliquots were taken at following times and their $OD_{600}$ was measured. Expression was induced by adding 50 µL of IPTG. Samples containing $OD_{600}$ volume equivalent to 1 were prepared and filled with 50 mM TRIS buffer (pH = 8.0) to 1 mL and stored for the following stages: induction phase, 4 hours after induction and over-night stage.

The procedure was repeated in the same way one more time, but this time expression was induced at 15°C. The calculated volumes for new suspensions with $OD_{600}$ were V(APT) = 1.35 mL, V(OPT) = 1.25 mL, V(XPT) = 1.41 mL, V(HPT) = 1.33 mL, V(UPT) = 1.23 mL from each of the over-night cultures. Samples were incubated at 37°C on a shaker starting at 11:45 h and small aliquots were taken at following times and their $OD_{600}$ was measured.

Expression was induced at 13:15 by using 50 µL of IPTG. Prior to induction, the temperature of the shaker was reduced to 15°C. Samples containing $OD_{600}$ volume equivalent to 1 were prepared and filled with 50 mM TRIS buffer (pH = 8.0) to 1 mL and stored for the following stages: induction phase, 4 hours after induction and overnight stage.

800 µL of each stored sample was taken and sonicated 3 times for 30 seconds with at least 1 minute of pause in between sonication repeats. Samples were then spun down at 18 000 x g for 10 minutes at 4°C. 700 µL of supernatant was kept in the freezer.

### 2.3.11. SDS-PAGE

The mixture for four separation gels (12%) contained 10 mL of miliQ water, 7.5 mL of 1.5 M TRIS-HCl (pH = 8.8), 12 mL of 30% acrylamide, 300 µL of 10% SDS, 300 µL of 10% APS and 6 µL of TEMED. The mixture for four collection gels (5%) contained 8.3 mL of miliQ water, 1.5 mL of 1M TRIS-HCl (pH= 6.8), 2 mL of 30% acrylamide, 60 µL of 10% SDS, 60 µL of 10% APS and 6 µL of TEMED. The separation gel had been prepared first and then isopropanol was added to the top to prevent reaction with oxygen. Once the separation gel solidified, isopropanol was poured out, APS and TEMED were added to the collection gel (5%) mixture and the solution was poured on top of the separation gel and finally comb was inserted.

30 µL of supernatant obtained after sonication and centrifugation of expression cultures at three different stages and at two different temperatures was mixed with 10 µL of 4X SDS loading dye. Probes were incubated at 95°C for 10 minutes on a heater. 25 samples were loaded onto 3 gels by pipetting 15 µL of each sample and adding 5 µL of the protein marker. The electrophoresis was run at 180 V for 1h and 15 minutes. At the end of electrophoresis**,** the gels were placed in the staining solution for 15 minutes, washed with water a few times and returned to the shaker. The results indicated that the best expression profile was present in the samples from over-night culture at 37°C.

### 2.3.12. Expression

Protein expression of APRT and OPRT was performed according to the following protocol. Four flasks were prepared, each containing 350 mL of LB medium and 350 µL of kanamycin. Two flasks for the expression of each protein were inoculated with the corresponding over-night culture and the volumes were calculated so that the final $OD_{600}$ = 0.15. Following volumes

were added: V(APRT) = 8.48 mL, V(OPRT) = 8.20 mL, V(XPRT) = 8.75 mL, V(HPRT) = 9.92 mL, V(UPRT) = 8.31 mL. The expression cultures were incubated at 37°C at 150 rpm starting at 10:50 a.m. and the $OD_{600}$ was measured. The expression was induced at 13:10 pm with 350 µL of IPTG and the samples were left in the incubator over night at 37°C at 150 rpm. The same was done with the other expression cultures that were placed in the incubator at 37°C at 150 rpm starting at 11:00 a.m. and the $OD_{600}$ was measured.

Protein expression of XPRT, HPRT and UPRT was performed the next day by induction at 12:40 pm with 350 µL of IPTG. Samples were left in the incubator over night at 37°C at 150 rpm. Flasks with expression cultures were placed on ice, transferred into bottles for centrifugation, balanced on the scale and centrifuged for 20 minutes at 8000 g at 4°C. The supernatant was poured back into the flasks and the precipitate was transferred into falcon tubes and frozen at -20°C.

### 2.3.13. Stock solution preparation

100 mL of 58% (V/V) glycerol solution was prepared and autoclaved. 10 flasks with 10 mL of LB medium for over-night cultures were prepared and inoculated with NEB10 and BL21 *E.coli* strains containing the corresponding plasmid vectors. 30 tubes were prepared with 500 µL of 58% glycerol solution and 500 µL of BL21 and NEB10 over-night cultures, mixed properly and frozen at -80°C.

### 2.3.14. Protein purification

Three buffers were prepared day prior to purification process. Washing buffer consisted of 1 L solution with 50 mM Tris-HCl, 300 mM NaCl and 20 mM imidazole. Elution buffer consisted of 0.5 L solution of 50 mM Tris-HCl, 300 mM NaCl and 500 mM imidazole. Buffer used for size exclusion chromatography consisted of 50 mM Tris-HCl and 300 mM NaCl. The samples were resuspended in 30 mL of wash buffer with 1 mM PMSF in which a spatula tip of DNase I was added. After complete resuspension 30 mg of lysozyme was added and the sample was left on a magnetic stirrer for 1 hour. The samples were sonicated on ice three times for 30 seconds (amplitude = 100, cycles 0.6). The samples were centrifuged at 15 000 x g for 1 hour at 4°C. Proteins were purified using the Äkta system on His Trap™ Crude 5 cm column, by injecting the sample and then washing it with washing buffer and then elution buffer followed by collecting of the fractions and then concentrating them if necessary to volume of approximately 2 mL in concentration tubes using centrifuge. The samples were then further purified using size exclusion chromatography by injecting the concentrate and using the previously prepared size exclusion buffer. Measured final protein concentrations are shown in the table 3.10.

## 2.4. Protein assays

### 2.4.1. Method establishment

For method establishment for use in HPLC, three solutions with the following composition were prepared: 100 µL of 1 mM adenine, 100 µL of 1 mM AMP and 100 µL of both 1 mM adenine and AMP. 100 µl of 100% methanol and 400 µL of water were added into each probe to the total volume of 600 µL. The same solutions were prepared for guanine/GMP, uracil UMP, cytosine/CMP, orotate/OMP. Method for efficient base/NMP separation is shown in the table 2.1.

**Table 2.1.** HPLC method for separation of bases and nucleotides

| Time (minutes) | Flow (mL/min) | 5% methanol (%) | 30% methanol (%) |
|---|---|---|---|
| 0 | 0,2 | 100 | 0 |
| 7,00 | 0,2 | 0 | 100 |
| 9,00 | 0,2 | 0 | 100 |
| 9,10 | 0,2 | 100 | 0 |
| 13,00 | 0,2 | 100 | 0 |

### 2.4.2 Enzyme activity assays

The first step in protein characterization involved determining the adequate protein working concentration. Several protein dilutions (10x, 20x, 30x, 100x, 1000x and 10 000x) were prepared for each enzyme in low binding tubes by mixing the enzyme with a buffer solution that contained 50mM HEPES, 150 mM NaCl and 2 mM $MgCl_2$. The pH of the solution was adjusted to 7.0. Each tube contained a total reaction volume of 50 µL with the following composition: 15.3 µL miliQ water, 10 µL HEPES pH 7.0 buffer, 5 µL 20 mM $MgCl_2$ , 5 µL PRPP, 10 µL of appropriate nitrogen base and 4.7 µL of the appropriate enzyme. The reactions were stopped with 50 µL 100% methanol after 10 minutes and incubated at 70˚C for 10 minutes more. 200 µL of miliQ water was added into the mixture and the samples were centrifuged for 10 minutes at 17 000 x g. 200 µL of the content was transferred into the HPLC vials after which the samples were analyzed.

An appropriate pH buffer was chosen between phosphate buffer, MES, HEPES and TRIS with pH value equal to 7. The reaction was tested in each of them to determine the best suitable one.

After the working concentration was determined, a time course was prepared with several time points in order to make a kinetics assay. The probes were prepared in triplicate and filled in advance with 50 µL of 100% methanol. Master mix was prepared with the appropriate buffer solution, and 50 µL were taken at each time point, transferred into the methanol containing tubes and incubated at 70 ˚C for 10 minutes. 200 µL of miliQ water were added into the samples that were centrifuged at 17 000 x g for 10 minutes. 200 µL of the content was transferred into the HPLC vials after which the samples were analyzed. Time curve was made so that the appropriate time point lies within the linear region.

pH assay to determine the optimal pH value was performed by preparing the tubes in triplicate and filling them with 100% methanol. Nine master mixes were prepared containing buffer solutions with the pH values of 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5 and 9.0. The reaction took place for the amount of time determined by the kinetics assay after which 50 µL were taken, transferred into the methanol containing tubes and incubated at 70 ˚C for 10 minutes. 200 µL of miliQ water were added into the samples that were centrifuged at 17 000 x g for 10 minutes. 200 µL of the content was transferred into the HPLC vials, after which the samples were analyzed.

Temperature assay to determine the optimal catalytic temperature was performed by preparing the tubes in and filling them with 100% methanol. Nine master mixes were prepared containing buffer solutions with the pH value of 7.0 at room temperature, 30˚C, 40˚C, 50˚C, 60˚C, 70˚C

and 80˚C. The reaction lasted the same amount of time as had been determined by the kinetics assay after which 50 µL were taken, transferred into the methanol containing tubes and incubated at 70 ˚C for 10 minutes. 200 µL of miliQ water were added into the samples that were centrifuged at 17 000 x g for 10 minutes. 200 µL of the content was transferred into the HPLC vials after which the samples were analyzed.

## 3. RESULTS

### 3.1. Cloning of phosphoribozyltransferase genes

Extracted *E.coli* DNA concentration and was measured from 1.5 µL of sample using NanoDrop™. DNA concentration was determined to be 141.7 ng/µL. PCR yielded satisfactory quantities with good purity with the exception of HPRT which shows increased level of impurities (Figures 3.1 and 3.2). Fragments were positioned at expected locations in the agarose gel compared with 1 kDa DNA ladder, without visible impurities (Figure 3.3).

Following the purification step DNA concentrations and purity was measured. The results are shown in the Figure 3.1 and 3.2.

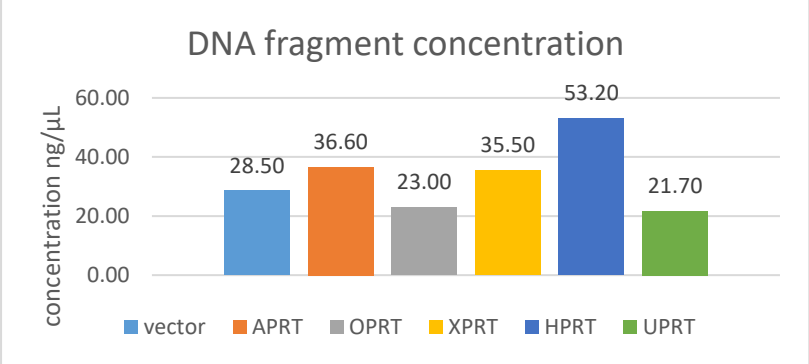



**Figure 3.1.** Fragment concentrations of empty vector and five *E.coli* phophoribozyl transferases after PCR amplification.

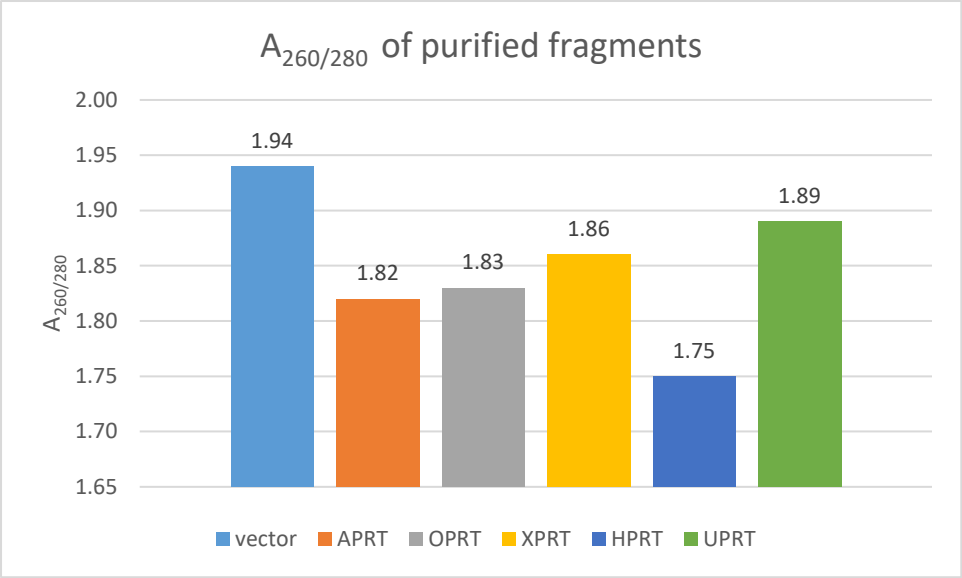



**Figure 3.2.** Purity expressed as $A_{260/280}$ of empty vector and five *E.coli* phophoribozyl transferase DNA fragments after PCR amplification.

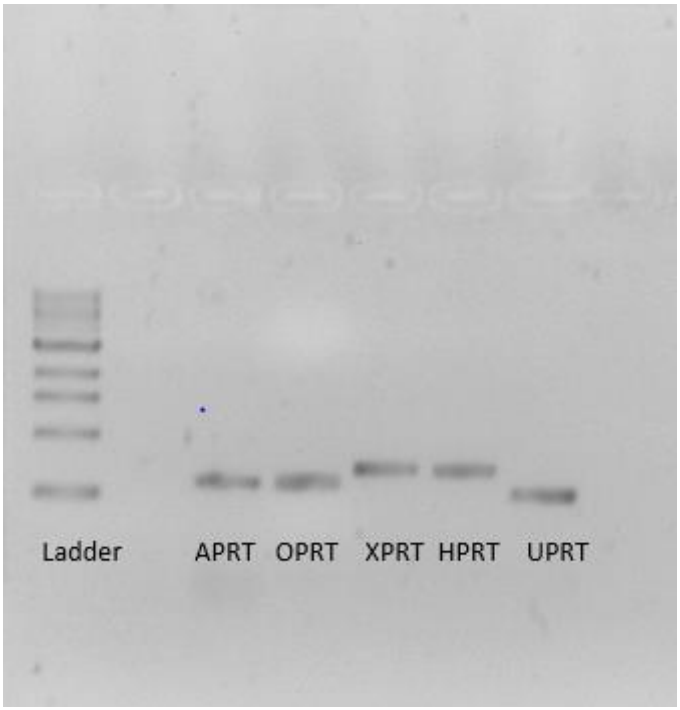



**Figure 3.3.** Visualization of APRT, OPRT, XPRT, HPRT and UPRT fragments in agarose gel after PCR amplification. The picture is shown in negative for better visualization.

After PCR amplification of phosphoribozyltransferases genes, fragments were inserted into a vector by Gibson assembly procedure. 5 mixtures of insert and vector were prepared in molar excess of 5:1 in favor of the insert.

Mixtures contained:

APT: $V_{(vector)}$= 3.7 µL , $V_{(insert)}$= 2.3 µL

OPT: $V_{(vector)}$= 4 µL , $V_{(insert)}$= 2 µL

XPT: $V_{(vector)}$= 3.9 µL , $V_{(insert)}$= 2.1 µL

HPT: $V_{(vector)}$= 3.1 µL , $V_{(insert)}$= 2.9 µL

UPT: $V_{(vector)}$= 4.2 µL , $V_{(insert)}$= 1.8 µL

Investigation of successful transformation was conducted on bacterial cells by colony PCR. Selected colonies showed successful overall insertion for APRT and XPRT. Only three of six selected colonies transformed with recombinant vectors carrying UPRT and HPRT fragments were successfully transformed, possibly due to lower purity and quantity of generated fragments (Fig. 3.4).

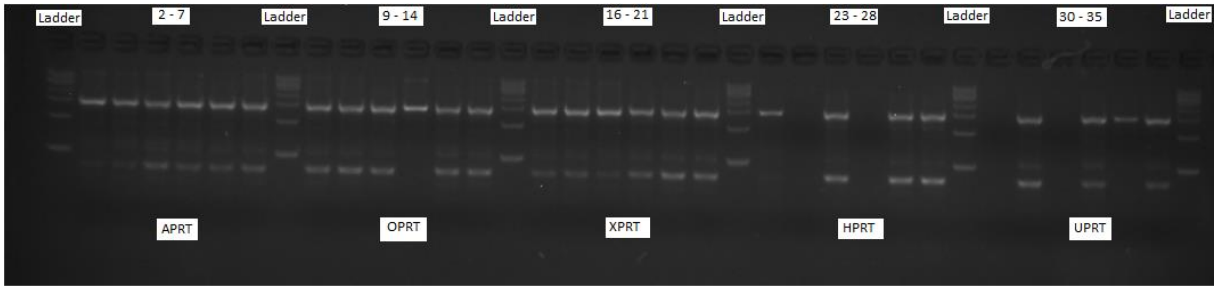



**Figure 3.4.** Gel electrophoresis of six selected colonies from five bacterial cultures, each containing different gene insert. Wells 2-7 contain APRT, 9-14 OPRT, 16-21 XPRT, 23-28 HPRT, 30-35 UPRT amplicons.

Further confirmation was obtained by digestion analysis. Concentration and purity of extracted DNA was measured using NanoDrop™. Purity and quantity of DNA fragments before digestion was satisfactory, with particularly high values for APRT, OPRT and UPRT. The results are shown in figure 3.5 and 3.6 respectively.

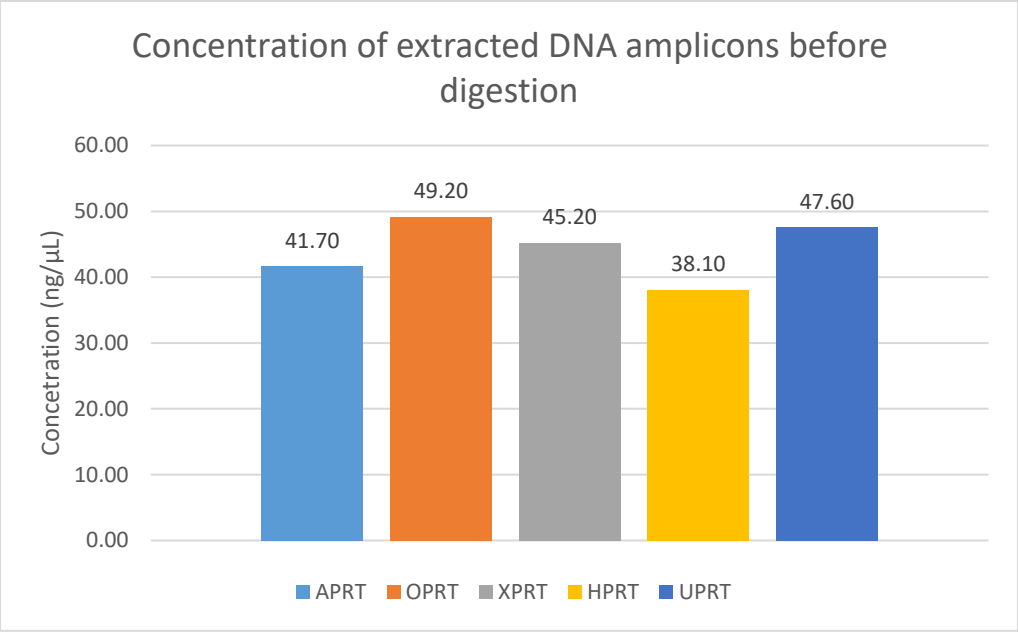



**Figure 3.5.** Fragment concentrations of five *E.coli* phophoribozyl transferases prior to digestion.

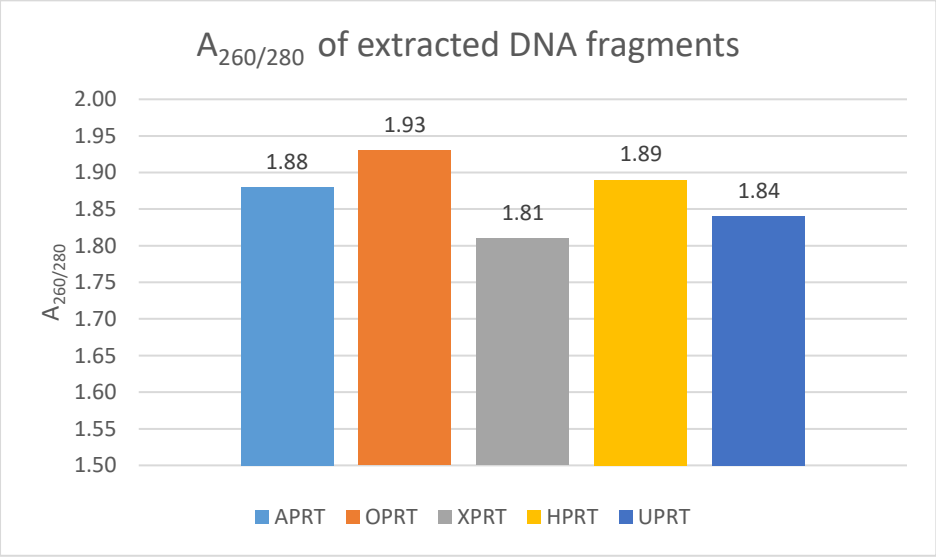



**Figure 3.6.** Purity expressed as $A_{260/280}$ of five *E.coli* phophoribozyltransferase DNA fragments prior to digestion.

Results of digestion analysis indicate proper insertion into vectors. Results are shown in figures 3.7 and 3.8

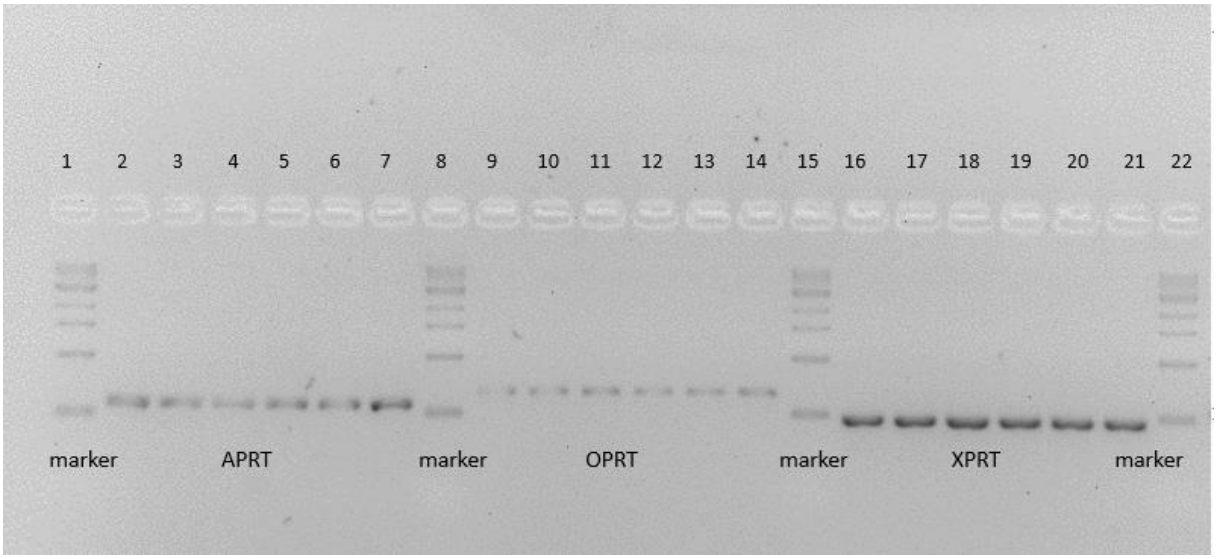



**Figure 3.7.** Gel electrophoresis of six selected colonies from five bacterial cultures, each containing different gene insert. Wells 2-7 contain plasmids with APRT, 9-14 wtih OPRT, 16-21 with XPRT amplicons. Picture is shown in negative for better visualization.

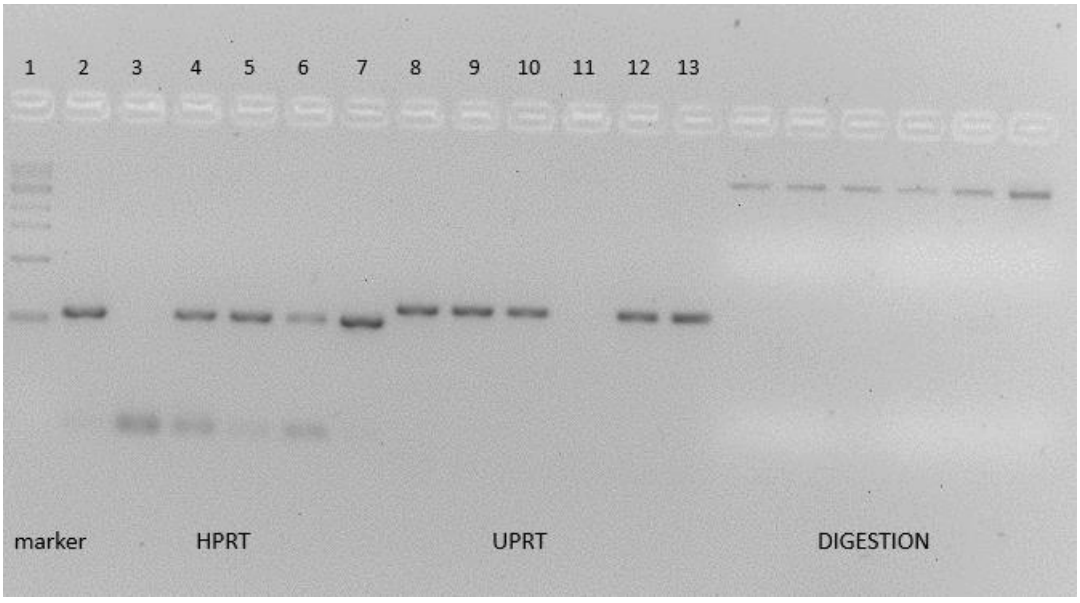



**Figure 3.8.** Gel electrophoresis of six selected colonies from five bacterial cultures, each containing different gene insert. Wells 2-7 contain plasmids with HPRT, 8-13 plasmids with UPRT amplicons. Picture is shown in negative for better visualization.

### 3.2. Protein expression and optimization

$OD_{600}$ profile of bacterial precultures grown at 37°C and 15°C showed a similar pattern of increase with time, however last two measurements seem to indicate that bacteria grown at 37°C reached the stationary phase faster and either stagnated or decreased in number during the night. That was not the case with bacteria grown at 15°C which grew more steadily due to slower rate of metabolism at lower temperature (Fig. 3.10 and 3.11)

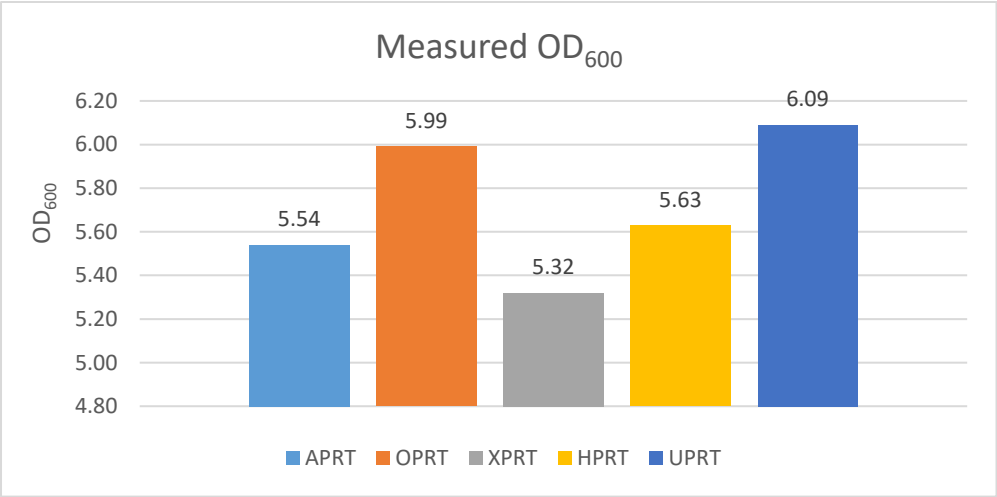



**Figure 3.9.** Measured values of 100 µL of over-night cultures mixed with 900 µL of LB medium.

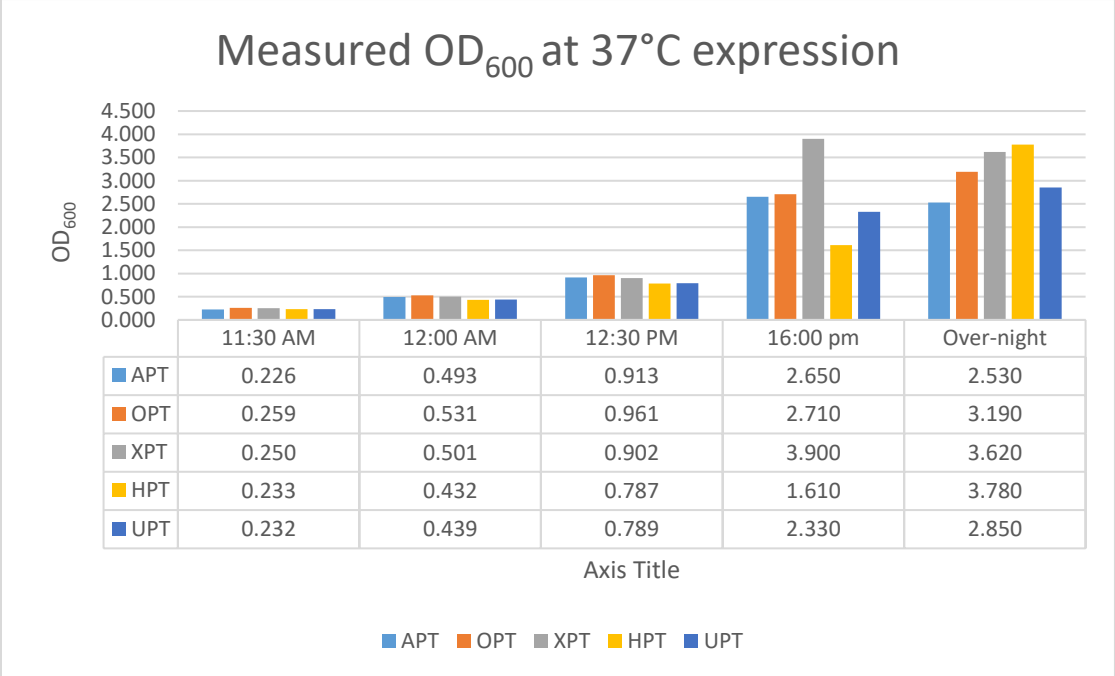



**Figure 3.10.** $OD_{600}$ measurements of *E.coli* samples grown at 37° C containing recombinant plasmids with appropriate gene inserts.

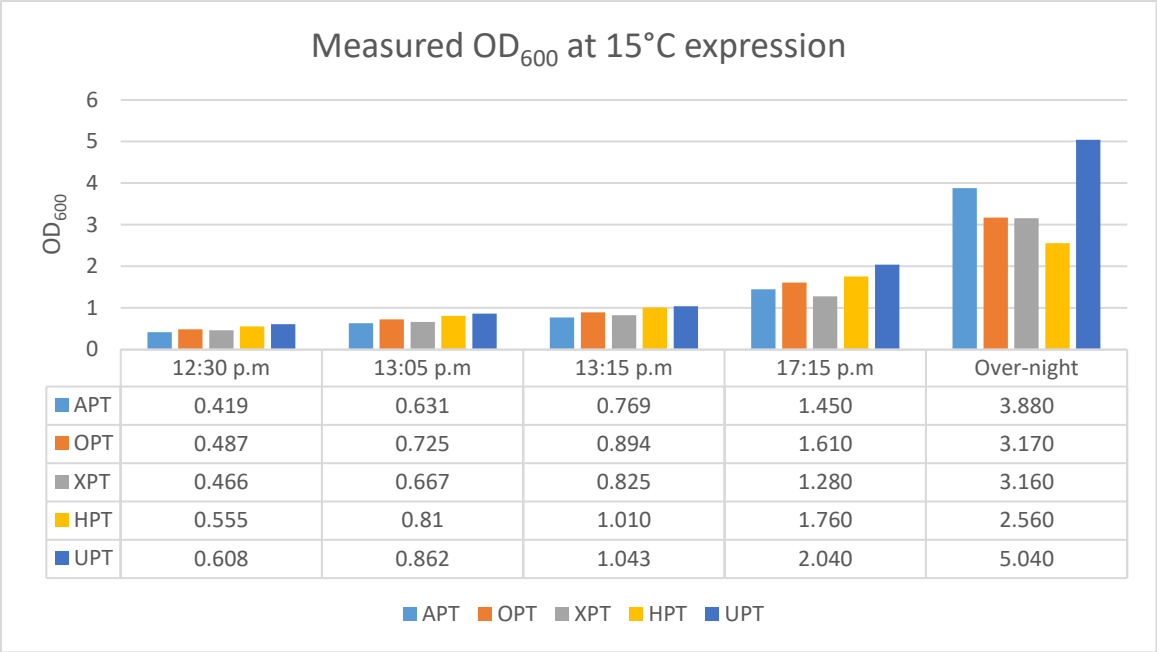



**Figure 3.11.** $OD_{600}$ measurements of *E.coli* samples grown at 15° C containing recombinant plasmids with appropriate gene inserts.

Protein yields after SDS-PAGE are shown below (Figures 3.12, 3.13 and 3.14).

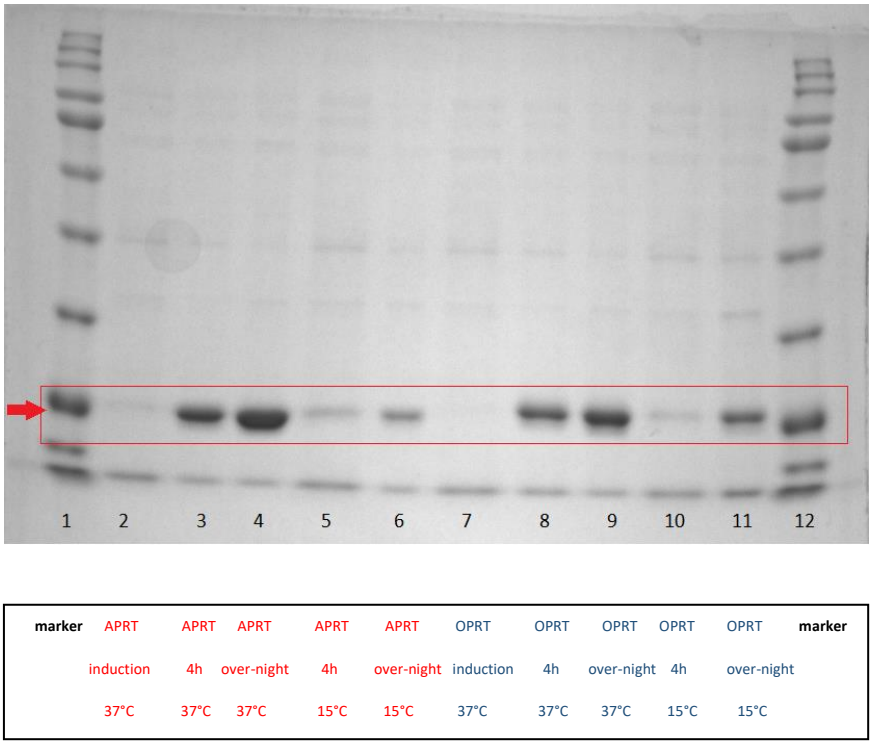
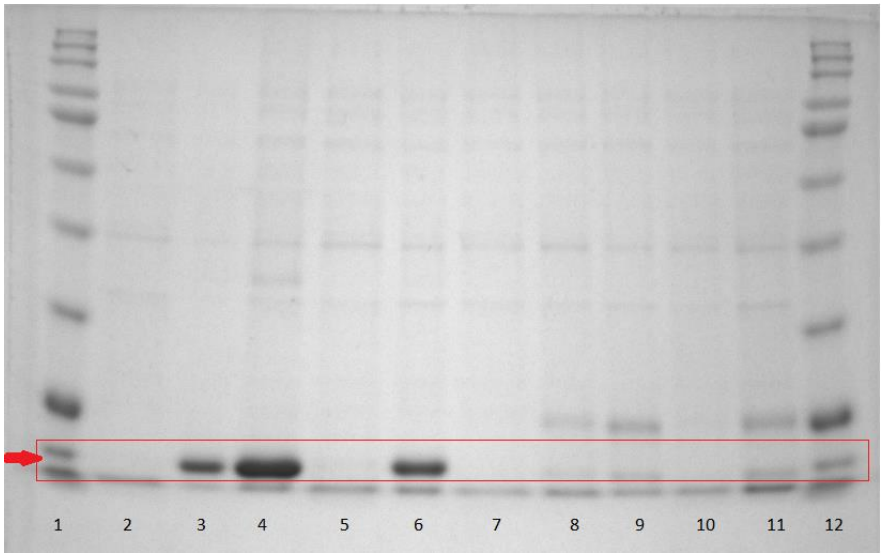
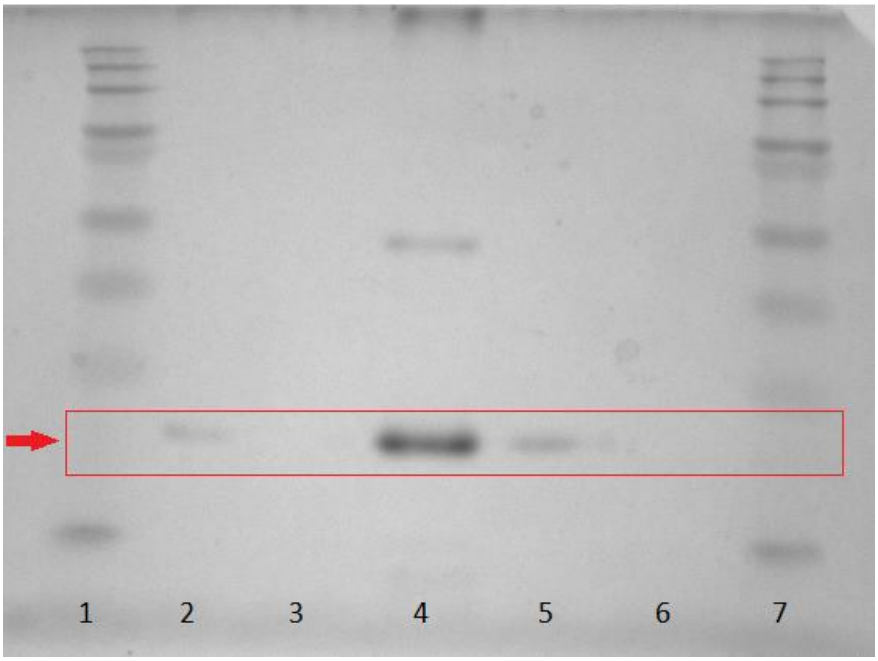



**Figure 3.12.** Visualization of APRT and OPRT proteins by SDS-PAGE after expression at 37°C and 15°C and at three time points (at induction, 4 hours after induction and overnight).




| marker | XPRT induction 37°C | XPRT 4h 37°C | XPRT over-night 37°C | XPRT 4h 15°C | XPRT over-night 15°C | HPRT induction 37°C | HPRT 4h 37°C | HPRT over-night 37°C | HPRT 4h 15°C | HPRT over-night 15°C | marker |
|---|---|---|---|---|---|---|---|---|---|---|---|

**Figure 3.13.** Visualization of XPRT and HPRT proteins by SDS-PAGE after expression at 37°C and 15°C and at three time points (at induction, 4 hours after induction and overnight).




| marker | UPRT induction 37°C | UPRT 4h 37°C | UPRT over-night 37°C | UPRT 4h 15°C | UPRT over-night 15°C | marker |
|---|---|---|---|---|---|---|

**Figure 3.14.** Visualization of UPRT protein by SDS-PAGE after expression at 37°C and 15°C and at three time points (at induction, 4 hours after induction and overnight).

All proteins showed the best expression profile at 37°C and overnight duration of production. These conditions were selected for production of proteins to be studied in the following steps.

Protein expression results are shown in the following section:

**Table 3.1.** $OD_{600}$ measurements of *E.coli* samples grown at 37° C containing recombinant plasmids with appropriate gene inserts, at different time intervals prior to induction with IPTG

| time | $A_1$ | $A_2$ | $O_1$ | $O_2$ |
|---|---|---|---|---|
| START | 0.137 | 0.135 | 0.142 | 0.149 |
| 1h | 0.615 | 0.601 | 0.637 | 0.659 |
| 1h 20 min | 1.160 | 1.140 | 1.130 | 1.250 |

**Table 3.2.** $OD_{600}$ measurements of *E.coli* samples grown at 37° C containing recombinant plasmids with appropriate gene inserts, , at different time intervals prior to induction with IPTG

| time | $X_1$ | $X_2$ | $H_1$ | $H_2$ | $U_1$ | $U_2$ |
|---|---|---|---|---|---|---|
| START | 0.379 | 0.366 | 0.419 | 0.434 | 0.344 | 0.512 |
| 40 min | 0.871 | 0.969 | 1.006 | 0.950 | 1.075 | 1.237 |

Purification of expressed proteins using His-Tag affinity chromatography and size-exclusion chromatography yielded results shown in Figure 3.15. Measurement of purified protein concentrations was performed once. XPRT concentration after purification was the highest and that of HPRT the lowest.

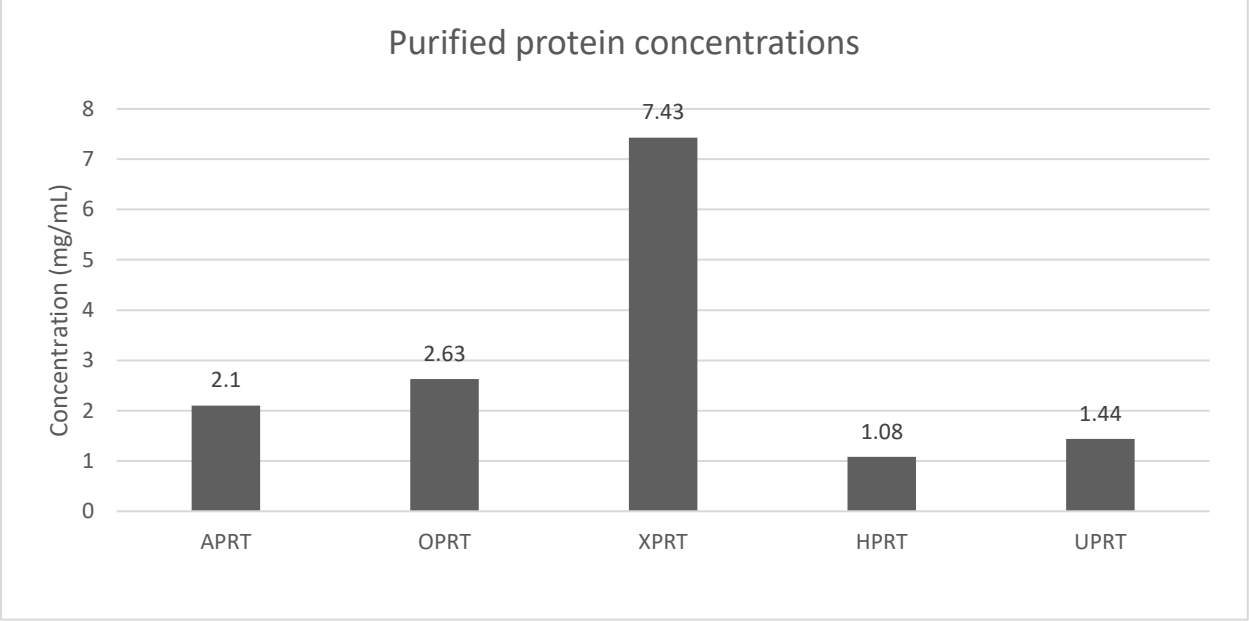



**Figure 3.15.** Measured protein concentrations after two-step protein purification with His-tag and size-exclusion columns.

### 3.3. Protein assays

#### 3.3.1. Calibration curves

Calibration curves of UPRT and HPRT were made using 100 fold dilutions with both enzymes to test the linearity and establish a relationship between AUC and proper UMP concentration (Figure 3.16 and 3.17). AUC (area under curve) is a measurement unit that corresponds to HPLC response to analyte presence. Calibration curves for other enzymes are not shown due to problems encountered with HPLC columns.

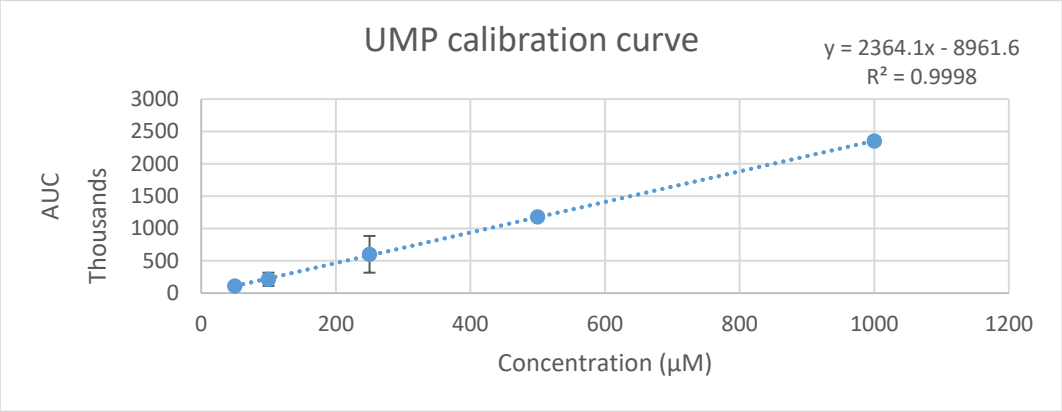



**Figure 3.16.** Uracil monophosphate calibration curve.

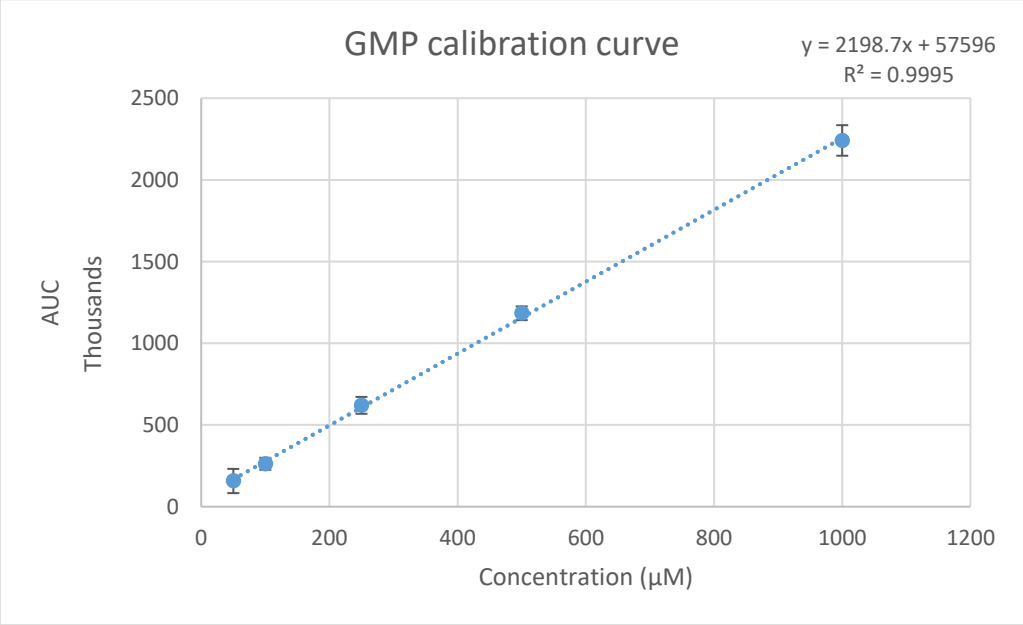



**Figure 3.17.** Guanosine monophosphate calibration curve.

### 3.3.2. APRT

The purified APRT was characterized for the parameters in buffer system and enzyme activity over 10 minutes (Figure 3.19). According to premliminary experiments a 1000 fold dilution of the enzyme was applied in the assays. Activity was determined to be the highest in HEPES buffer, which was used in subsequent characterizations (Figure 3.18).

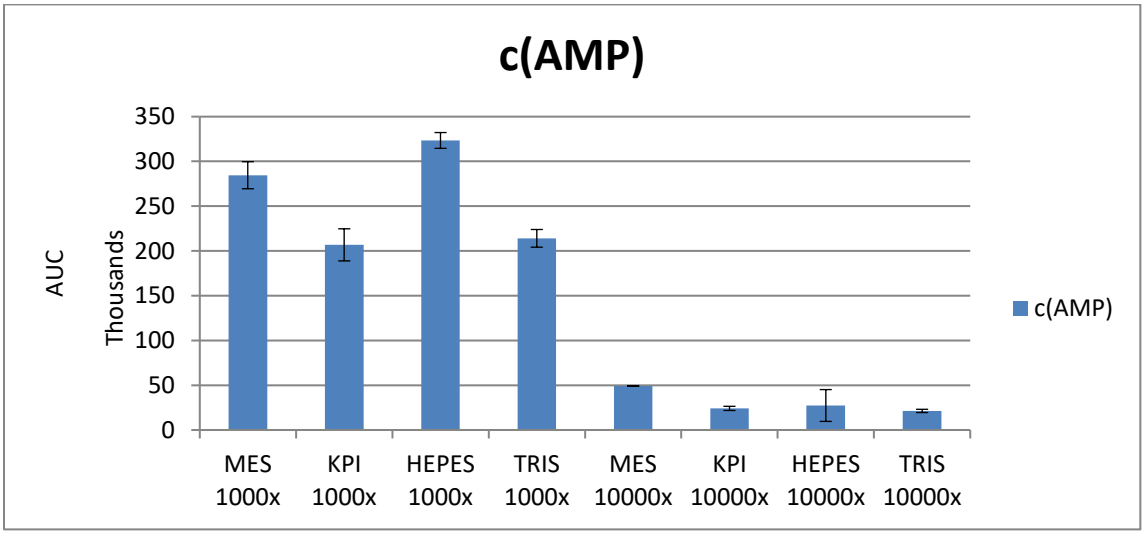



**Figure 3.18.** APRT production of adenosine monophosphate in different buffers with 1000 fold and 10000 fold protein dilutions in 10 minutes.

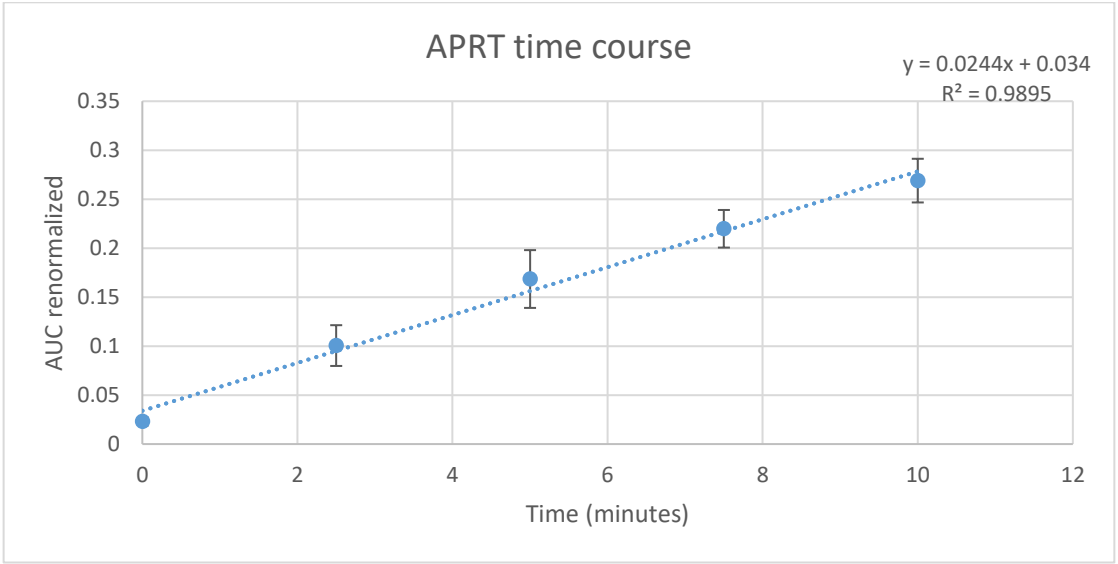



**Figure 3.19.** APRT production of adenosine monophosphate in HEPES buffer with 1000 fold protein dilution over 10 minutes.

### 3.3.3. XPRT

The purified XPRT was characterized for enzyme activity over 10 minute interval (Figure 3.20). According to premliminary experiments (data not shown) a 100 fold dilution of the enzyme was applied in the assays. XPRT activity was tested for the buffer systems MES, KPI HEPES and Tris (data not shown) and it was established that the best turnover rate is present in HEPES buffer.

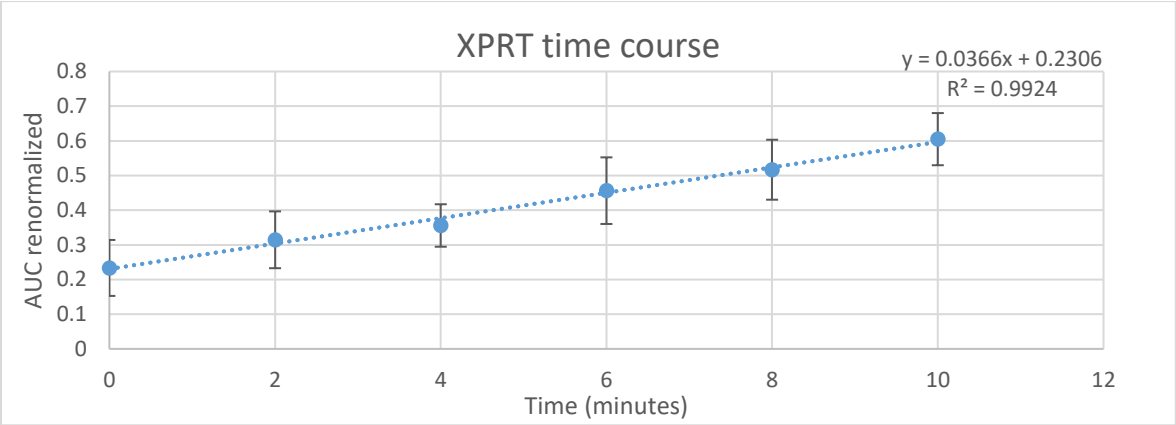



**Figure 3.20.** XPRT production of xanthine monophosphate in HEPES buffer with 100 fold protein dilution over 10 minutes.

### 3.3.4. HPRT

The purified HPRT was characterized for the parameters in buffer system, reaction temperature and pH value. According to premliminary experiments (data not shown) a 100 fold dilution of the enzyme was applied in the assays. HPRT activity was tested for the buffer systems MES, KPI HEPES and Tris (Figure 3.21). Activity was comparable for all three amine-based buffer, only the activity in KPI buffer was six fold lower. Due to the best performance of the enzyme in HEPES buffer all subsequent steps were performed in solutions containing HEPES. In the next step linearity of the enzyme assay was tested (Figure 3.22) followed by temperature and pH assays (Figure 3.23 and 3.24). HPRT catalyzed production of GMP was tested in HEPES and MES buffers with 100 fold protein dilution over 10 minutes between pH=5 and pH=9. HPRT activity on different temperatures in HEPES buffer was tested with 100 fold protein dilution over 10 minutes.

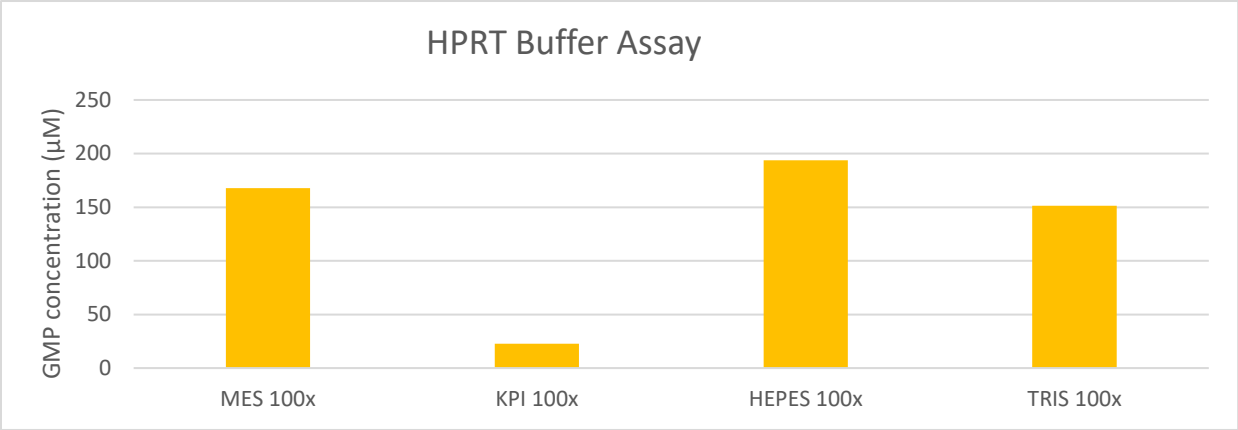



**Figure 3.21.** Guanosine monophosphate production by HPRT in 10 minutes in different buffers. All measurements were performed once.

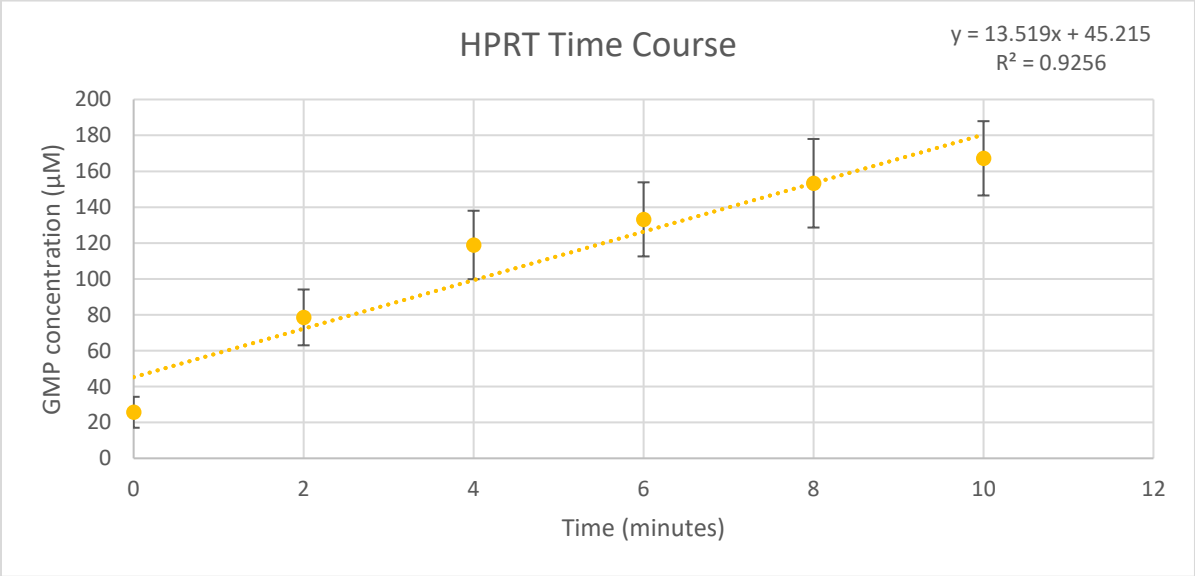



**Figure 3.22**. Guanosine monophosphate production by HPRT in HEPES buffer with 100 fold protein dilution over 10 minutes. Error bars represent standard error (SE) after three measurements.

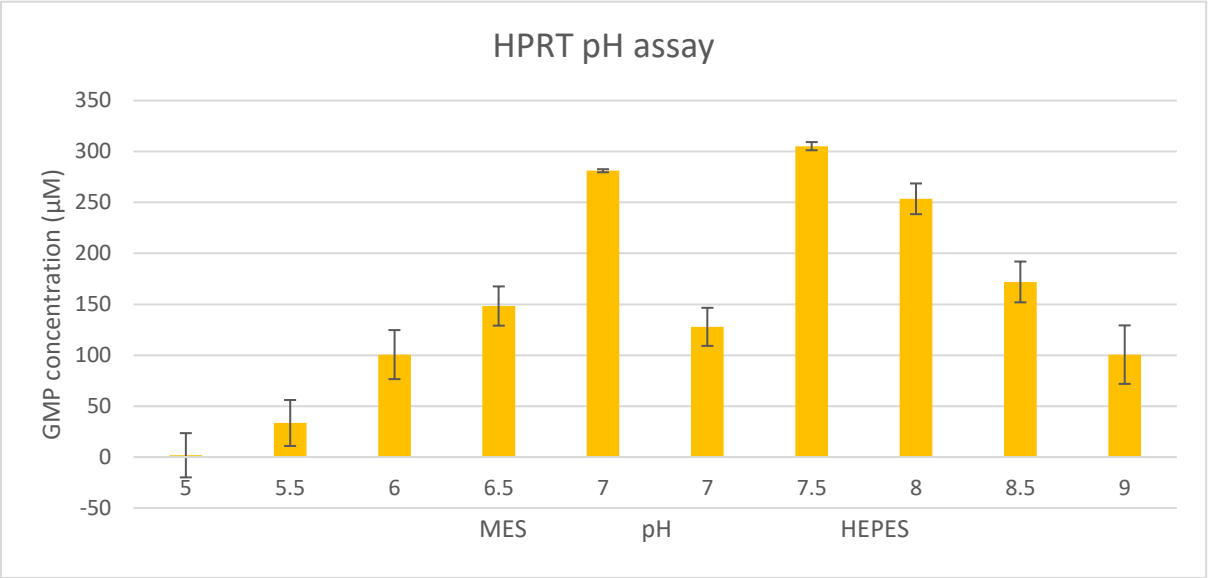



**Figure 3.23.** HPRT production of GMP in HEPES and MES buffers with 100 fold protein dilution over 10 minutes between pH=5 and pH=9. Error bars represent standard error (SE) after three measurements.

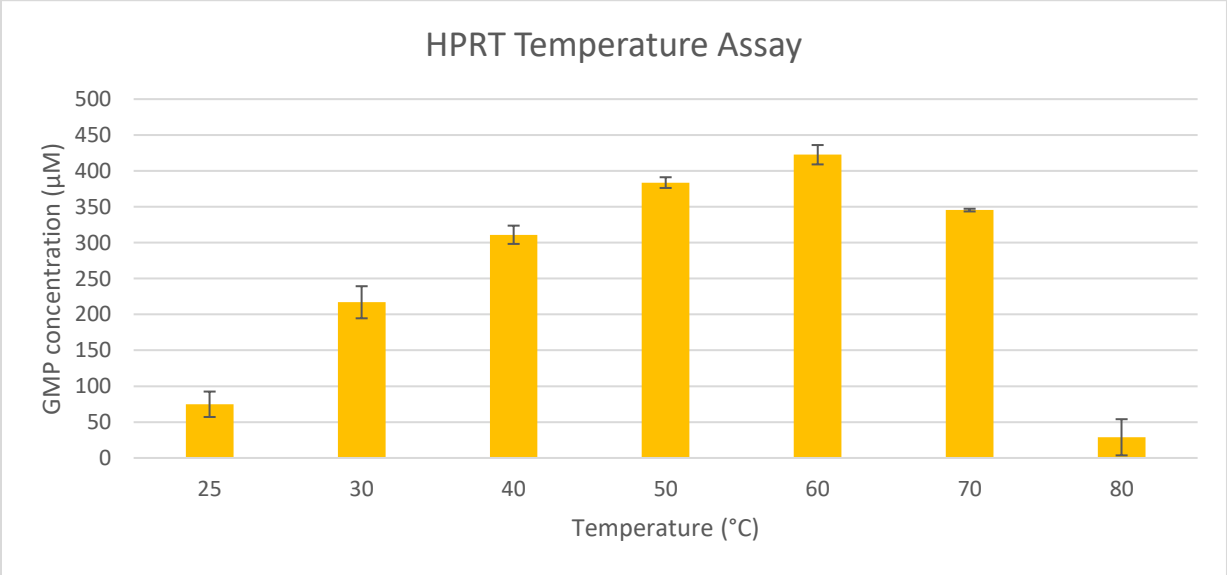



**Figure 3.24.** HPRT production of GMP in HEPES buffer with 100 fold protein dilution over 10 minutes on different temperatures. Error bars represent standard error (SE) after three measurements.

### 3.3.5. UPRT

The purified UPRT was characterized for the parameters of reaction temperature (Figure 3.27), enzyme activity over 10 minutes (Figure 3.25) and pH value (Figure 3.26). According to premliminary experiments (data not shown) a 100 fold dilution of the enzyme was applied in the assays. Due to the best performance of the enzyme in HEPES buffer all subsequent steps were performed in solutions containing HEPES buffer.

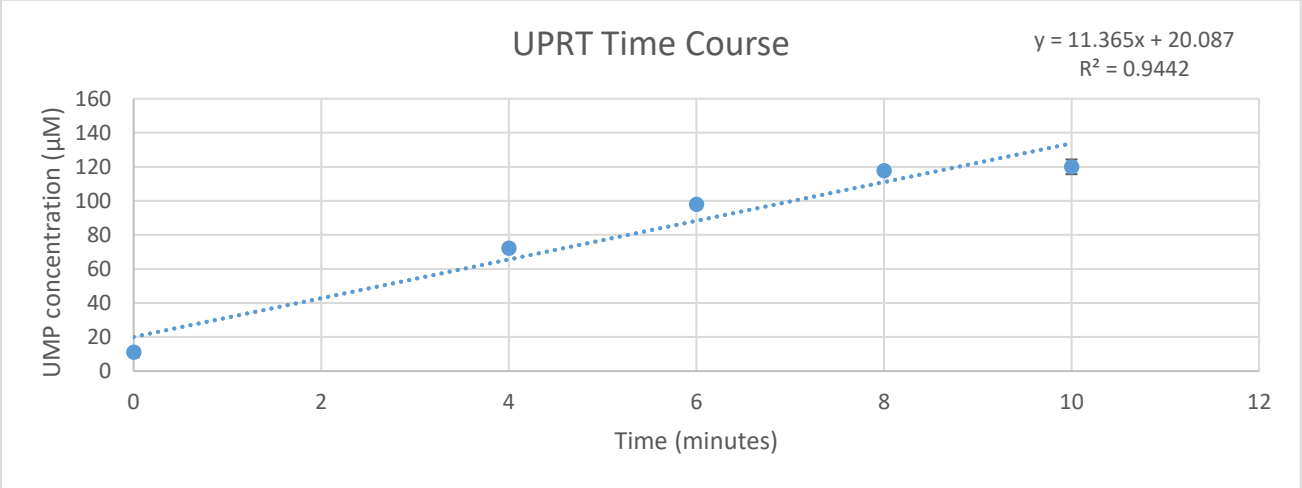



**Figure 3.25.** UPRT production of uracil monophosphate in HEPES buffer with 100 fold protein dilution over 10 minutes. Error bars represent standard error (SE) after three measurements.

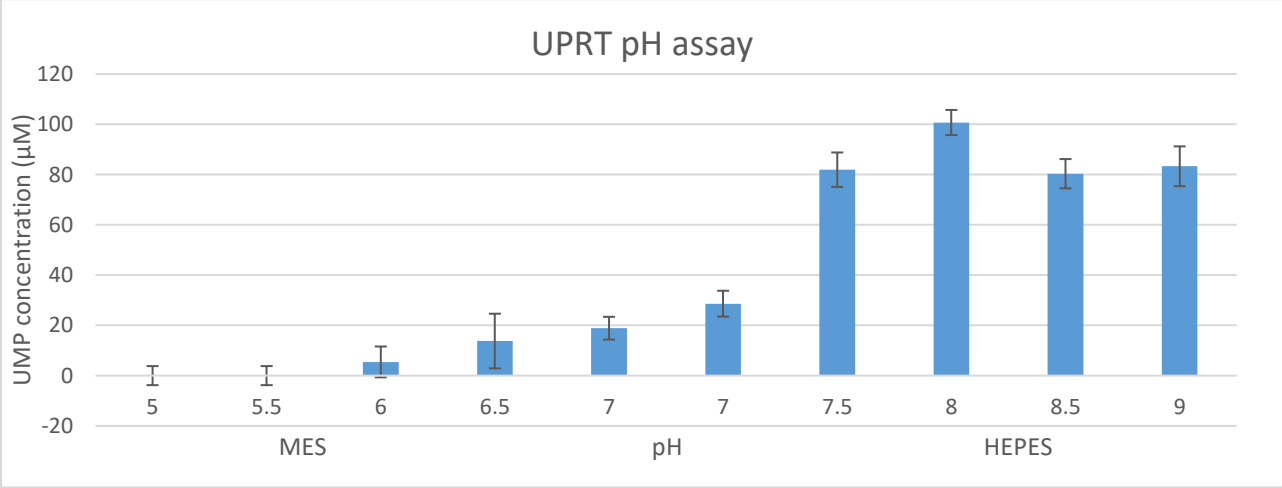



**Figure 3.26**. UPRT production of UMP in HEPES and MES buffers with 100 fold protein dilution over 10 minutes between pH=5 and pH=9. Error bars represent standard error (SE) after three measurements.

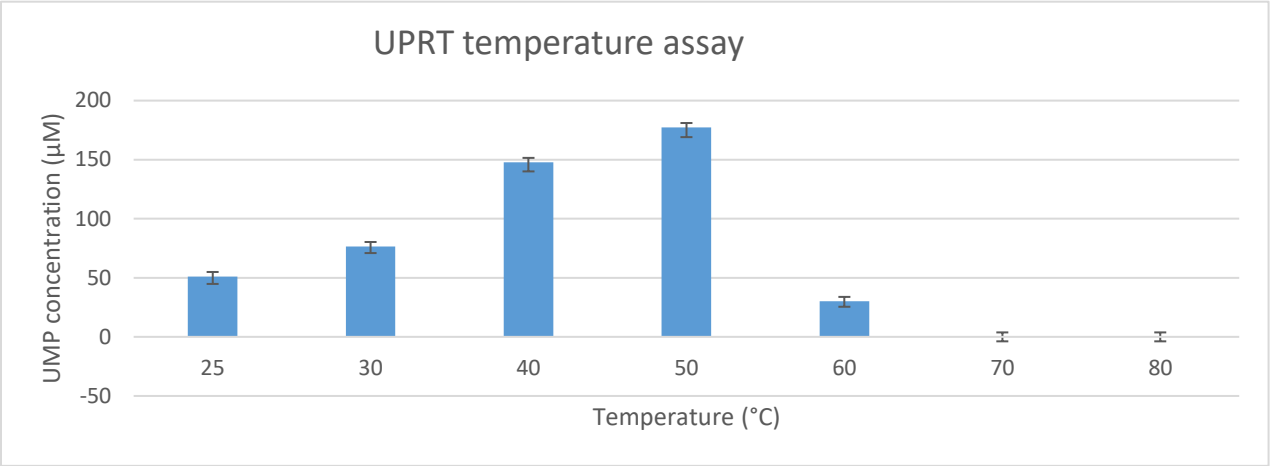



**Figure 3.27.** UPRT production of UMP in HEPES buffer with 100 fold protein dilution over 10 minutes on different temperatures. Error bars represent standard error (SE) after three measurements.

### 3.3.6. Non-natural nucleotides

All purified enzymes were tested for their ability to process non-natural nucleotides. XPRT, HPRT and APRT were tested for the purine analogue 6-mercaptopurine, and OPRT and UPRT were tested for 5-fluorouracil. All reactions involved appropriate enzyme dilutions in HEPES buffer at room temperature in pH=7. Product could be observed in reactions catalyzed by XPRT, HPRT and UPRT (Figure 3.28 and 3.29).

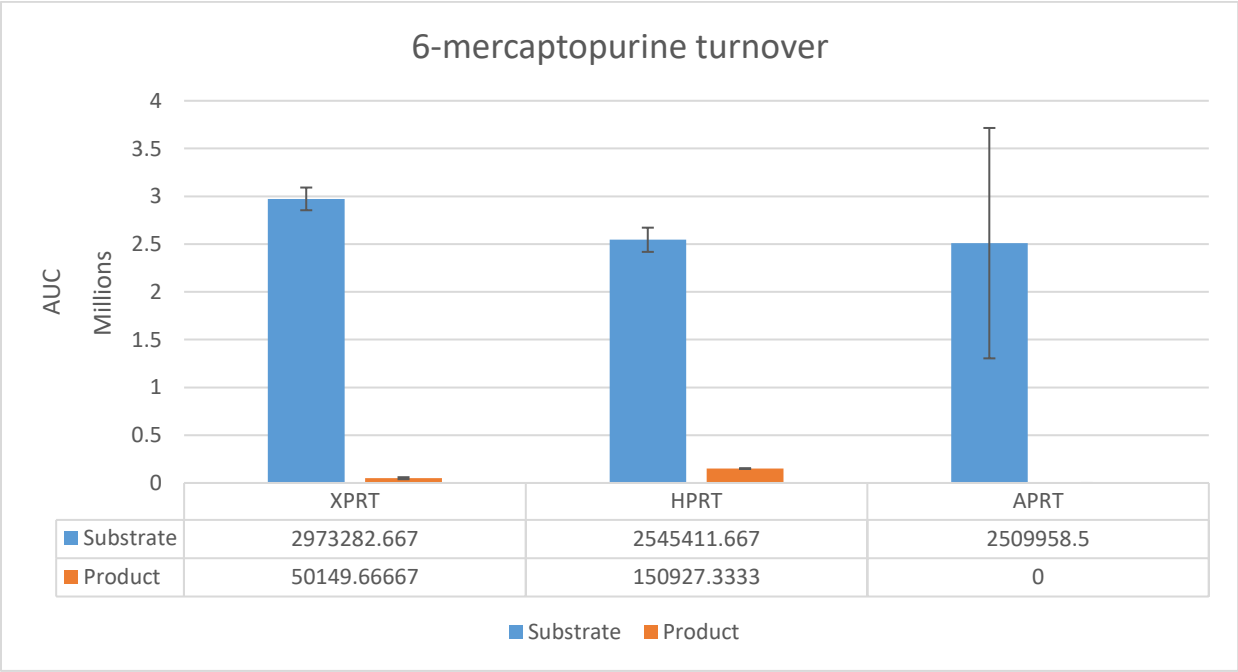



**Figure 3.28.** Production of non-natural nucleotides from 6-mercaptopurine by XPRT, HPRT and APRT in HEPES buffer with 100 fold protein dilutions over 10 minutes. Error bars represent standard error (SE) after three measurements

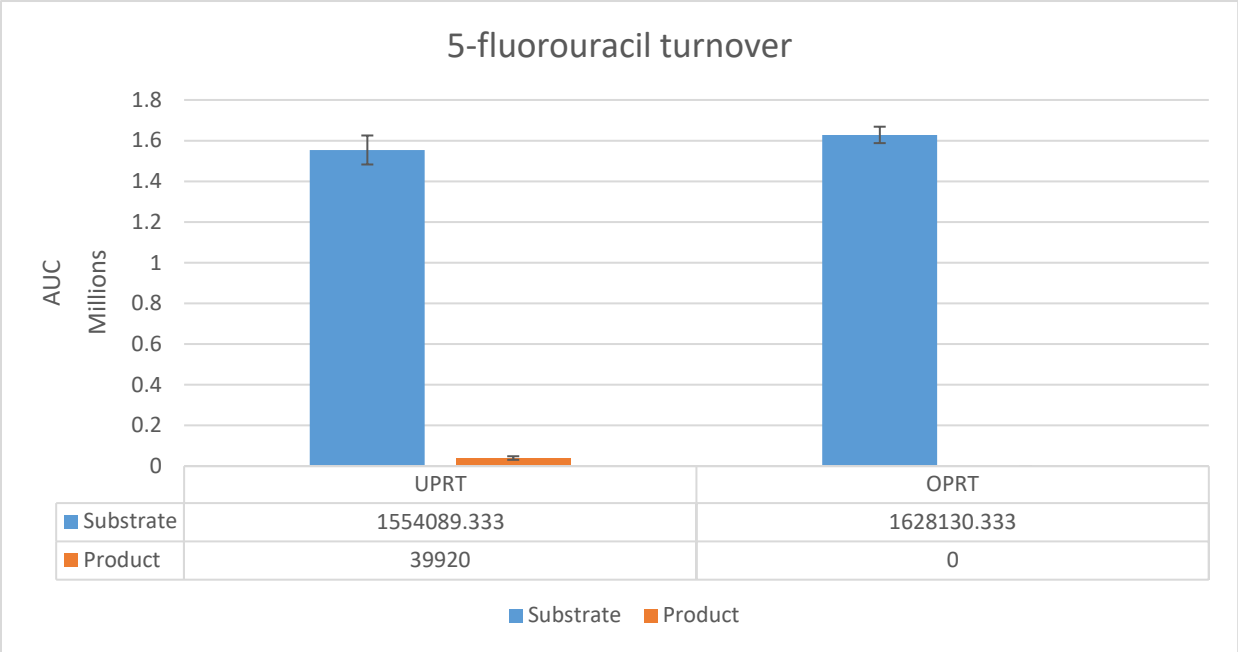



**Figure 3.29.** Production of non-natural nucleotides from 5-fluorouracil by UPRT and OPRT in HEPES buffer with 100 fold protein dilutions over 10 minutes. Error bars represent standard error (SE) after three measurements

# 4. DISCUSSION AND CONCLUSIONS

## 4.1. Cloning of gene fragments

PCR amplification yielded sufficient amount of amplified DNA to be used in subsequent steps. The highest yield was obtained with HPRT whereas lowest yield was obtained with UPRT. However, regarding the purity of respective fragments, the opposite proved to be the case. All five phosphoribozyltransferase genes were visualized by gel electrophoresis and all fragment sizes agreed with expectations. After the transformation of bacterial cells, colony PCR was used to verify the efficacy of transformation by randomly selecting six colonies formed after the transformation with recombinant plasmids, each carrying their respective phosphoribosyltransferase gene.

## 4.2. Protein expression

Proteins, in general, showed satisfying expression profiles, but the best yield for all proteins was obtained by cultivating bacteria at 37°C overnight. Protein yields were higher, as expected, at 37°C than at 15°C. All expressed proteins showed progressively increasing yields from induction to overnight measuring points. APRT, OPRT and XPRT bands were particularly pronounced. HPRT yield was very low when compared to other proteins. Bands were visible only after overnight production and barely visible in other cases (Figures 3.12, 3.13 and 3.14). These results were expected because PRTs of interest originate from the homologous organism. $OD_{600}$ values did not show any sign of decline, which would indicate issues with protein production or their potential toxicity. After His-Tag and size exclusion chromatography, the measured amount of proteins was satisfying in all cases. However, the best production yield was achieved with XPRT, whereas HPRT showed the lowest yield.

## 4.3. Natural substrate conversion

The best buffer for conducting protein assays for all enzymes was HEPES. UMP and GMP calibration curves were used to calculate concentrations of produced nucleotides. Calibration curves for other nucleotides could not be made due to problems with HPLC column, so the results are expressed as AUC. Both calibration curves show excellent linearity between various points and share similarity in terms of correlation between concentration and AUC. APRT characterization was rather unsuccessful because of inconsistencies in experimental values obtained when performing repeated measurements under different pH and temperature conditions. This could be attributed to problems experienced with HPLC column, that was frequently congesting in the course of performing protein assays. However, a curve of time dependency was successfully made, which showed linear relation between time of the reaction and renormalized AUC. Renormalization of AUC parameter was implemented to compensate for potential variations in initial substrate concentrations. The same was true with XPRT, which also showed linear relation between time of the reaction and renormalized AUC. OPRT results are not shown due to inability to measure product concentrations.

Optimal pH for HPRT catalysis was determined to be 7.5 in HEPES, while at pH 5 the protein showed almost no activity, probably due to denaturation in acidic environment. At pH over 7.5 the enzyme showed gradual loss of activity. The optimal temperature was determined to be 60°C. At room temperature enzyme activity is very low as well as at 80°C probably due to high temperature denaturation over 70°C. The reason for low activity at room temperature is unknown, and it could be an inherent property of the enzyme.

UPRT showed optimal activity at pH 8.0 with a significant decrease below pH 7.5. The enzyme showed high resistance to alkaline conditions with activity being pronounced even at pH 9.0. Optimal temperature was determined to be 50°C with a significant decline at higher temperatures and no activity at 70°C and 80°C.

### 4.4. Non-natural substrate conversion

6-mercaptopurine was taken up by XPRT and HPRT albeit with very small nucleotide production rate. APRT showed no affinity towards 6-mercaptopurine which could potentially be attributed to the absence of the $-NH_2$ group at position 6 in the aromatic ring which forms hydrogen bond with the neighboring amino acid residues. The aromatic structure is also disturbed by the substitution of $-NH_2$ group with the sulfur atom, which likely inhibits π-stacking between the aromatic ring of adenine and nearby phenylalanine moiety. HPRT active site seems to be able to process 6-mercaptopurine to a small extent probably because nitrogen atoms in the purine ring allow for limited binding to surrounding amino acid residues such as Ile-135 and Asp-92. Similarly, HPRT can accommodate 6-mercaptopurine in its active site because of the similar substrate size and interaction of N-7 atom and Asp-107. Inosine and 6-mercaptopurine structures are similar enough that sulfur, which substitutes oxygen in inosine, can form hydrogen bond with the nearby Lys-135.

5-fluorouracil was taken up by UPRT but not by OPRT. Lack of activity towards 5-FU in case of OPRT possibly has to do with the presence of electronegative fluorine atom that has difficulties entering into the active pocket or exerts unfavorable electrostatic influence on the surrounding amino acid residues. 5-FU also lacks the carboxyl group that could be important in forming hydrogen bonds with the adjacent functional groups of amino acids. All investigated enzymes require further genetic manipulation in order to produce unnatural nucleotides with higher efficiency.

## 5. LITERATURE


Becerra, A. & Lazcano, A., 1998. The role of gene duplication in the evolution of purine nucleotide salvage pathways. *Origins of life and evolution of the biosphere*, 28(1969), pp.539–553.

Craig, S.P. & Eakin, A.E., 2000. Purine Phosphoribosyltransferases. *JBC Papers*, 275, pp.20231–20235.

Deo, S.S. et al., 1985. Purification and characterization of Escherichia coli xanthine-guanine phosphoribosyltransferase produced by plasmid pSV2gpt. *BBA - General Subjects*, 839(3), pp.233–239.

Eakin, A. et al., 1997. Hypoxanthine Phosphoribosyltransferase from Trypanosoma cruzi as a Target for Structure-Based Inhibitor Design : Crystallization and Inhibition Studies with Purine Analogs. *ANTIMICROBIAL AGENTS AND CHEMOTHERAPY*, 41(8), pp.1686–1692.

Galmarini, C.M., Mackey, J.R. & Dumontet, C., 2002. Nucleoside analogues and nucleobases in cancer treatment. *Lancet Oncology*, 3(7), pp.415–424.

Gibson, D.G., 2011. Enzymatic assembly of overlapping DNA fragments. *Methods in Enzymology*, 498, pp.349–361.

Henriksen, A. et al., 1996. A flexible loop at the dimer interface is a part of the active site of the adjacent monomer of Escherichia coli orotate phosphoribosyltransferase. *Biochemistry*, 35(12), pp.3803–3809.

Hochstadt-Ozer, J., 1972. The Regulation of Purine Utilization in Bacteria. *The Journal of biological chemistry*, 247(8), pp.2419–2426.

Jinnah, H.A., 2009. Lesch-Nyhan disease: from mechanism to model and back again. *Disease Models and Mechanisms*, 2(3–4), pp.116–121.

Keough, D.T. et al., 2002. Crystal structures of free, IMP-, and GMP-bound Escherichia coli hypoxanthine phosphoribosyltransferase. *Protein Science*, 11(7), pp.1626–1638.

Keough, D.T. et al., 2006. Lead Compounds for Antimalarial Chemotherapy: Purine Base Analogs Discriminate between Human and P. Falciparum 6-Oxopurine Phosphoribosyltransferases. *J. Med. Chem*, 49, pp.7479–7486.

Lavergne, T. et al., 2017. A semisynthetic organism engineered for the stable expansion of the genetic alphabet. *Proceedings of the National Academy of Sciences*, 114(6), pp.1317–1322.

Mahmoud, S. et al., 2018. Antiviral Nucleoside and Nucleotide Analogs : A Review. , 547(2), pp.73–88.

Mulligan, R.C. & Berg, P., 2006. Selection for animal cells that express the Escherichia coli gene coding for xanthine-guanine phosphoribosyltransferase. *Proceedings of the National Academy of Sciences*, 78(4), pp.2072–2076.

Ochiai, T. et al., 2006. Prognostic impact of orotate phosphoribosyl transferase among 5-fluorouracil metabolic enzymes in resectable colorectal cancers treated by oral 5-fluorouracil-based adjuvant chemotherapy. *Int. J. Cancer*, 3088, pp.3084–3088.

Phillips, C.L. et al., 1999. Crystal structures of adenine phosphoribosyltransferase from Leishmania donovani. *The EMBO Journal*, 18(13), pp.3533–3545.

Rasmussen, U.B., Mygind, B. & Per, N., 1986. Purification and some properties of uracil phosphoribosyltransferase from Escherichia coli K12. *BBA - General Subjects*, 881(2), pp.268–275.

Ross, C.W., 1981. Biosynthesis of nucleotides, *Academic Press Inc.*, 6, pp.169-205.

Schmidt, M., 2010. Xenobiology: A new form of life as the ultimate biosafety tool. *BioEssays*, 32(4), pp.322–331.

Schumacher, M.A. et al., 1998. Crystal structures of Toxoplasma gondii uracil phosphoribosyltransferase reveal the atomic basis of pyrimidine discrimination and prodrug binding. *EMBO Journal*, 17(12), pp.3219–3232.

Scism, R.A., Stec, D.F. & Bachmann, B.O., 2007. Synthesis of nucleotide analogues by a promiscuous phosphoribosyltransferase. *Organic Letters*, 9(21), pp.4179–4182.

Somoza, J.R. et al., 1998. Rational Design of Novel Antimicrobials : Blocking Purine Salvage in a Parasitic Protozoan. *Biochemistry*, 37(97), pp.5344–5348.

Subbayya, I.N.S. et al., 2000. Unusual Substrate Specificity of a Chimeric Hypoxanthine – Guanine Phosphoribosyltransferase Containing Segments from the Plasmodium falciparum and Human Enzymes. *Biochemical and Biophysical Research Communications*, 272, pp.596–602.

Ullman, B. & Carter, D., 1997. Molecular and Biochemical Studies on the Hypoxanthine-guanine Phosphoribozyltransferases of the Pathogenic Haemoflagellates. *International Journal for Parasitology*, 27(2), pp.203–213.

Vos, S. et al., 1998. Structures of free and complexed forms of Escherichia coli xanthine-guanine phosphoribosyltransferase. *Journal of Molecular Biology*, 282(4), pp.875–889.

Wang, C.C., 1984. Parasite enzymes as potential targets for antiparasitic chemotherapy. *J Med Chem.*, 27(1), pp.1–9.

Zhang, Y. *et al.* (2017) 'A semisynthetic organism engineered for the stable expansion of the genetic alphabet', *Proceedings of the National Academy of Sciences*, 114(6), pp. 1317–1322.