Universidad Europea MADRID

Trabajo de Fin de Grado

Grado en Biotecnología

PRODUCTION, PURIFICATION AND BIOCHEMICAL CHARACTERIZATION OF ADENINE PHOSPHORIBOSYLTRANSFERASE FROM *ESCHERICHIA COLI*

Autor: **Laura Pilar Saiz Álvarez**

Villaviciosa de Odón, *Junio 2023*

ANEXO IX

Título del trabajo: **PRODUCTION, PURIFICATION AND BIOCHEMICAL CHARACTERIZATION OF ADENINE PHOSPHORIBOSYLTRANSFERASE FROM *ESCHERICHIA COLI.***

Este trabajo ha sido realizado en el campus de Villaviciosa de Odón de la Universidad Europea de Madrid (UEM); concretamente, en el Grupo de Biotecnología Aplicada (GBA), dirigido por el Dr. Jon del Arco Arrieta.

Tutor: Dr. Jon del Arco Arrieta.

Table of contents

Abstract

Currently, nucleoside analogues are molecules widely used as drugs. Due to their structural similarity to physiological nucleosides, they are used as anticancer, antibacterial, antiviral and even antiparasitic drugs. Traditionally, their synthesis has been carried out by chemical synthesis, which is accompanied by the use of substances that are harmful to the environment due to the solvents and chemicals used, as well as involving multi-step reactions that reduce the yield of the reaction due to the secondary products generated. More recently, their synthesis has been carried out by biocatalysis, which not only makes the process more environmentally friendly but also facilitates the obtaining of specific products as the enzymes used are regio- and stereoselective.

Different multienzyme cascades have been recently designed to increase the yield of the generation of the desired product by shifting the equilibrium of the reaction towards its formation. This work makes use of the enzyme adenine phosphoribosyltransferase from *Escherichia coli* (*Ec*APRT) to recover the adenine released by the enzyme nucleoside 2'-deoxyribosyltransferase from *Lactobacillus delbruekii* (*Ld*NDT) which is responsible for catalyzing the reaction that would give rise to the nucleoside analog. Furthermore, an additional cascade has been proposed using the enzymes *Mr*PPK (*Meiothermus ruber* polyphosphate kinase) and PRPP synthetase, with the purpose of regenerating PRPP, the necessary substrate for *Ec*APRT.

In order to implement *Ec*APRT in the multienzymatic system, in this work the protein has been produced, purified and biochemically characterized to know the optimal operational conditions. In addition, the experimental assays were studied bioinformatically by carrying out molecular dynamics simulations, in which spatial conformations and distances that the protein acquires were studied.

The results obtained revealed the structural and biochemical characteristics of the enzyme.

Resumen

En la actualidad, los análogos de nucleósidos son moléculas usadas ampliamente como fármacos. Debido a su similitud estructural con los nucleósidos fisiológicos, estos son utilizados como fármacos anticancerígenos, antibacterianos, antivirales e incluso antiparasitarios. Tradicionalmente su síntesis se ha llevado a cabo mediante síntesis química, lo cual va acompañado del uso de sustancias dañinas para el medio ambiente por los disolventes y químicos utilizados, además de requerir múltiples pasos de reacción, los cuales disminuyen el rendimiento del proceso por los productos secundarios generados. Más recientemente, su síntesis se ha llevado a cabo mediante biocatálisis, lo cual hace no solo el proceso más ecológico sino también facilita la obtención de productos específicos al ser las enzimas usadas regio- y estereoselectivas.

Recientemente se han diseñado diferentes cascadas multienzimáticas para mejorar el rendimiento del proceso, aumentando la cantidad de producto obtenido al desplazar el equilibrio de la reacción hacia su formación. Este trabajo hace uso de la enzima adenina fosforribosiltransferasa de *Escherichia coli* (*Ec*APRT) para recuperar la adenina liberada por la enzima nucleósido 2'-desoxirribosiltransferasa de *Lactobacillus delbruekii* (*Ld*NDT) la cual es la responsable de catalizar la reacción que daría lugar al análogo de nucleósido. Además, se ha propuesto una cascada adicional con el uso de las enzimas *Mr*PPK (*Meiothermus ruber* polifosfato quinasa) y PRPP sintetasa, con el propósito de regenerar el PRPP, sustrato necesario para *Ec*APRT.

Con el propósito de implementar la enzima *Ec*APRT en el sistema multienzimático, en este trabajo la proteina fue producida, purificada y caracterizada bioquímicamente para conocer las condiciones óptimas de reacción. Además, los ensayos experimentales fueron estudiados bioinformáticamente al llevar a cabo simulaciones de dinámicas molecular, en los que se estudiaron conformaciones espaciales y distancias que la proteína adquiere.

Los resultados obtenidos dieron a conocer las características estructurales y bioquímicas de la enzima.

Glossary of non- and standard abbreviations and acronyms

2-ClAde-NMP: 2-choloroadenosine nucleoside-5´-monophosphate

2-ClAde: 2-chloroadenine

2-FAde-NMP: 2-fluoroadenosine nucleoside-5'-monophosphate

2-FAde: 2-fluoroadenine

2,6-DAP-NMP: 2,6-diaminopurine nucleoside-5'-monophosphate

2,6-DAP: 2,6-diaminopurine

3-MetAde-NMP: 3-methyladenine nucleoside-5'-monophosphate

3-MetAde: 3-methyladenine

6-DMP-NMP: 6-(dimethylamino)purine 5´-monophosphate

6-DMP: 6-(dimethylamino)purine

N6-MetAde-NMP: N6-methyladenine nucleoside-5'-monophosphate

N6-MetAde: N6-methyladenine

Ade: Adenine

AMP: Adenosine 5'-monophosphate

APRT: Adenine phosphoribosyltransferase **DO_{600}:** Optical density at 600 nm

***Ec*APRT:** *Escherichia coli* adenine phosphoribosyltransferase

HPLC: High performance liquid chromatography

IPTG: Isopropyl-b-D-1-thiogalactopyranoside

IU: International units

LB: Luria-Bertani medium

Mg: Magnesium

$MgCl_2$: Magnesium chloride

PDB: Protein data bank

PPi: Pyrophosphate

PRPP: Phosphoribosyl pyrophosphate

PRTase: Phosphoribosyltransferase

RMSD: Root-Mean-Square-Deviation

RMSF: Root-Mean-Square-Fluctuations

RPM: Revolutions per minute

SDS-PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SDS: Sodium Dodecyl Sulfate

TEMED: Tetramethylethylenediamine

1. Introduction

1.1. Nucleosides and nucleotides

1.1.1. General concepts

Nucleotides are the basic components of nucleic acids. They have a specific structure (Figure 1A) composed of a pentose, a nitrogenous base covalently joined (N-β-glycosidic bond) the N9 for purine (adenine or guanine) (Figure 1B) or N1 for pyrimidine (cytosine, thymine, or uracil) (Figure 1B) to the C1 of the ribose and one or more phosphate groups, joined by a phosphoester bond to the C5 of the pentose, forming nucleotides mono-, di- or thiphosphate. This structure, without the phosphate group(s), is called nucleoside (Figure 1B) (1).

Specifically, nucleotides are the main components of nucleic acids, both deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), playing an essential role in the storage and transmission of genetic information. Furthermore, they are involved in several metabolic processes, such as energy transmission (e.g., ATP), signal transduction (e.g., cAMP, adenosine), enzyme catalysis. (e.g., coenzymes) and nucleotide coenzymes (e.g., NAD+) (2,3).

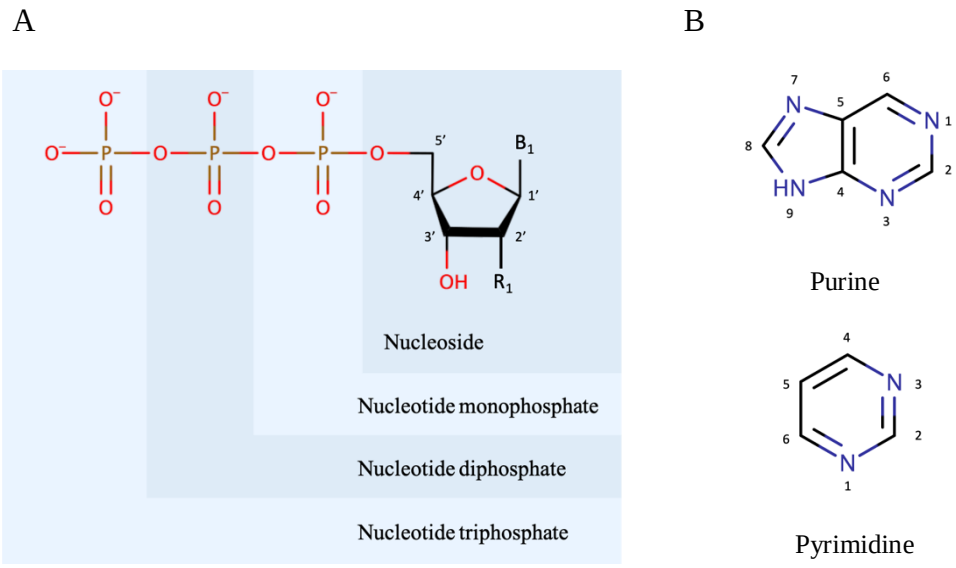


Figure 1. (A) *General structure of nucleotides and nucleosides with conventional pentose ring numbering (R_1 = H or OH, B1 = purine or pyrimidine nucleobase).* ***(B)*** *Basic ring structure of purine and pyrimidine nucleobases. Created on ChemSketch (4).*

1.1.2. Nucleotide synthesis

There are two types of pathways leading to the formation of nucleotides. *De novo* synthesis is one of them. Purine nucleobases (Figure 2A) are formed from ribose-5-P and adenosine-5'-triphosphate (ATP). As an intermediary molecule, phosphoribosyl pyrophosphate (PRPP) is created, process catalyzed by PRPP synthetase. The purine ring is then made from the following components: aspartic acid (Asp), tetrahydrofolate (THF), glutamine (Gln), glycine (Gly), and CO_2 thus forming inosine-5'-monophosphate (IMP). Guanosine-5'-monophosphate (GMP) is produced when IMP is combined with glutamine, and adenosine 5'-monophosphate (AMP) when combined with aspartic acid. Guanosine-5'-triphosphate (GTP) and ATP are produced when both are phosphorylated. Pyrimidine production is quite distinct (Figure 2B). The base is the first component to be made from NH_3 and CO_2, obtaining carbamoyl phosphate. Orotate is then produced through a sequence of enzymatic processes in which aspartic acid is used as a precursor. This, along with PRPP, produces uridine-5'-monophosphate (UMP), the precursor of uridine-5'-triphosphate (UTP), cytidine-5'-triphosphate (CTP) and thymidylate (dTMP) (1,5).

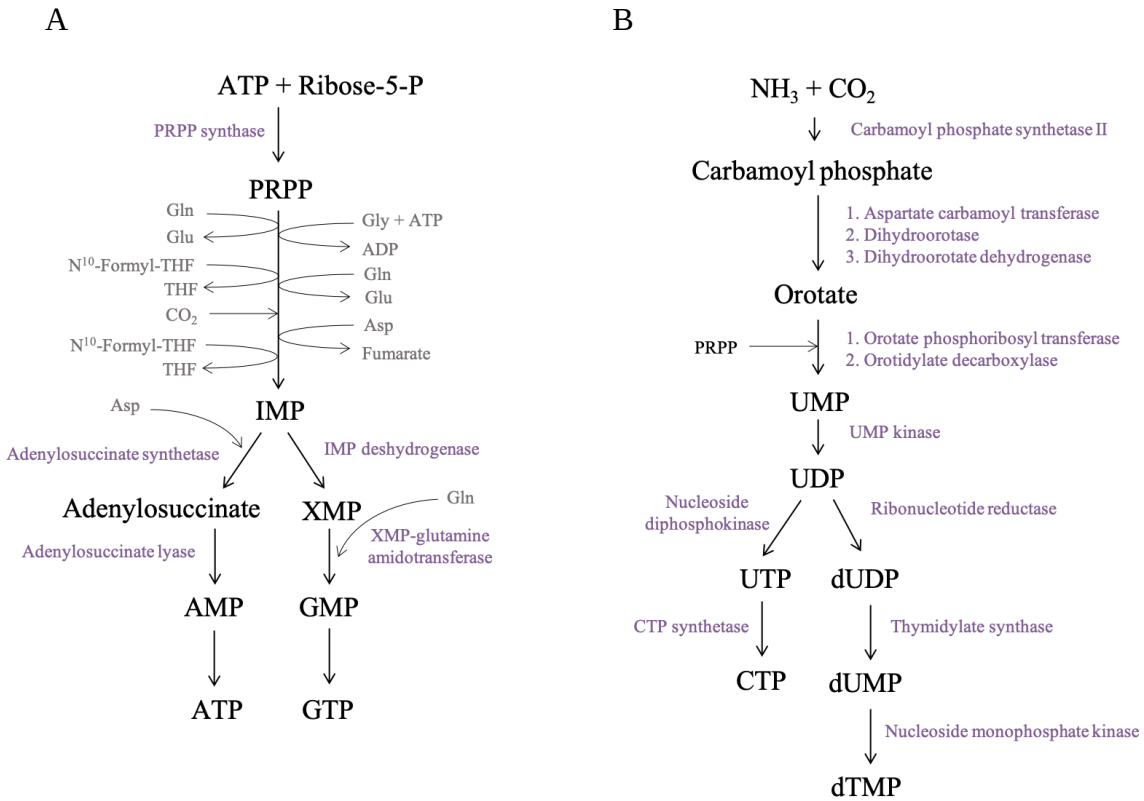


Figure 2. (A) *De novo synthesis for purine nucleobases. Gln: glutamine, Glu: glutamate, Gly: glycine, N^{10}-formyl-THF: 10-formyltetrahydrofolate, THF: tetrahydrofolate, Asp: aspartic acid.* ***(B)*** *De novo synthesis for pyrimidine synthesis. The enzymes involved in the reaction are shown in dark purple. Figure adapted from (6).*

In the case of the salvage pathway, it is carried out in periods of rapid cell growth such as embryogenesis or tumoral processes, being therefore studied as a target for chemotherapy (7,8). The free heterocyclic bases of the degraded nucleotides and nucleosides from both endo- and exogenous sources are reused in this process for the generation of new nucleotides (9). Purines are synthesized directly from phosphoribosyltransferases (PRTases) forming nucleotide monophosphates from phosphoribosyl pyrophosphate (PRPP) and the corresponding nucleobase (10) (Figure 3A). On the one hand, adenine phosphoribosyltransferase (APRT) catalyzes the formation of AMP, while the 6-oxo PRTs (guanine phosphoribosyltransferase (GPRT), xanthine phosphoribosyltransferase (XPRT), hypoxanthine phosphoribosyltransferase (HPRT), etc.) catalyze the formation of GMP (from guanine), or XMP (from xanthine) or IMP (from hypoxanthine). On the other hand, in the synthesis of pyrimidines (Figure 3B), firstly, the uracil is first obtained from cytosine or deoxyuridine. Then, uridine is generated being this the precursor of UMP, from which phosphorylated thymine and cytosine can be formed.

A B

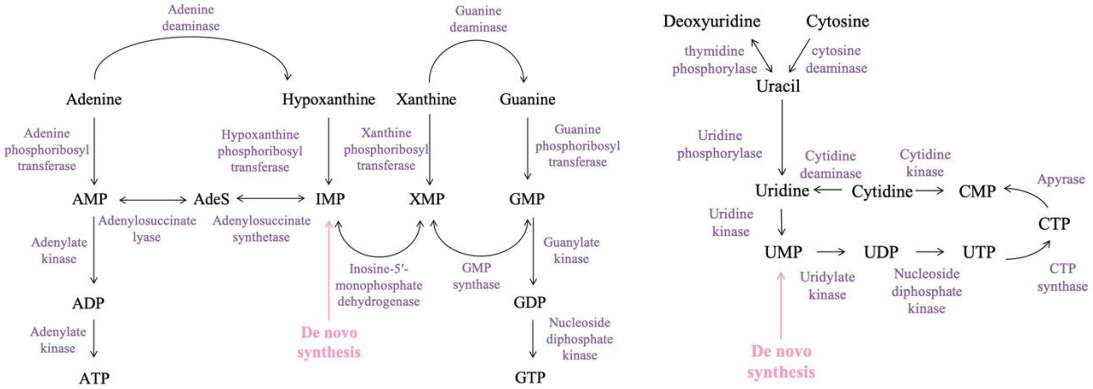


Figure 3. (A) *Purine salvage synthesis.* ***(B)*** *Pyrimidine salvage synthesis. The enzymes involved in the reaction are shown in dark purple. Figure adapted from (11,12) respectively.*

The absence or failure of these enzymes causes different disorders such as Lesch-Nyhan syndrome caused by a severe deficit in the hypoxanthine guanine phosphoribosyltransferase (HGPRT) enzyme (13). Other enzymes involved in this synthesis route are nucleoside kinase (NKs, nucleotide monophosphates are formed from ATP and a nucleoside), nucleoside 2'-deoxyribosyltransferase and nucleoside phosphorylases (NDTs and NPs respectively, catalyzes the purine and/or pyrimidine nucleobase interchange of the 2'-deoxynucleoside by a transglycosylation reaction in one or two steps, respectively), among others (14).

1.1.3. Nucleoside and nucleotide analogs

Nucleic acids derivatives, NADs (nucleosides, nucleotides, and nucleobases) have a wide range of applications, such as antiparasitic (i.e., treatment against protozoans or helminths) (15), antiviral (i.e., treatment of human immunodeficiency virus, herpesvirus, or SARS-CoV-2) (16,17), antibacterial (gram-positive and gram-negative) (18) and antitumoral (activity in solid tumors and malignant disorders of the blood) (19) activity. These molecules can both inhibit enzymes such as the reverse transcriptase or HIV protease, or interfere with nucleic acid synthesis by incorporating these molecules during synthesis, as in the case of cancer (20,21).

Their synthesis has historically been accomplished by chemical synthesis. As this process attempts to phosphorylate the precursor nucleoside and subsequently glycosylates it, the functional groups must be protected and deprotected and the reactions are carried out using environmentally risky and expensive substances, such as organic solvents or chemicals reagents as phosphorus pentoxide (P_2O_5) or phosphoryl chloride ($POCl_3$). It is important to mention the pollution generated by the disposal of this waste. Furthermore, due to the complexity of the process, several by-products are generated, lowering the yield and efficiency (22,23). They can, however, be produced by enzymatic processes (biocatalysis) by white biotechnology (involved in industrial processes) through simple and efficient methodologies. This have a significant benefit over chemical synthesis since most of these are performed in a single step, having regio- and stereoselectivity thus correctly forming the N-β-glycosidic bond. Most importantly, this process does not use the hazardous substances mentioned above, so it is environmentally benign (24). However, their inability to catalyze reactions in severe reaction conditions (extreme pH, high temperature, or organic solvents) often used at industrial scale, for avoiding contamination, to improve the solubility of substrates or to avoid diffusional problems, is one of their limitations, which can be solved by protein engineering, using enzymes from extremophilic organisms or enzyme immobilization (14,25).

Considering the purpose of this work, different AMP analogues (Figure 4) have been produced when working with modified adenine nucleobases, in order to characterize the specificity of the *Ec*APRT enzyme for its future use in the multienzymatic cascade, described below.

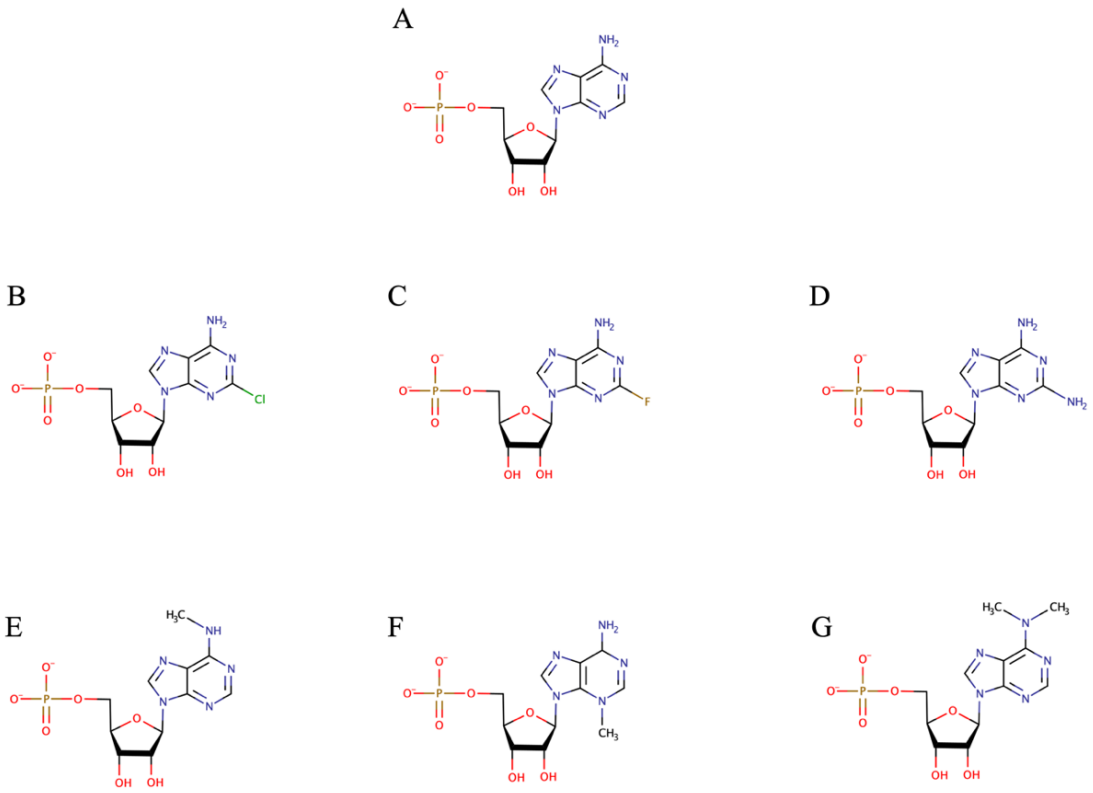


Figure 4. *AMP and AMP analogs.* ***(A)*** *Adenosine-5'-monophosphate (AMP),* ***(B)*** *2-choloroadenosine nucleoside-5´-monophosphate (2-ClAde-NMP),* ***(C)*** *2-fluoroadenosine nucleoside-5'-monophosphate (2-FAde-NMP),* ***(D)*** *2,6-diaminopurine nucleoside-5'-monophosphate (2,6-DAP-NMP),* ***(E)*** *N6-methyladenine nucleoside-5'-monophosphate (N6-MetAde-NMP),* ***(F)*** *3-methyladenine nucleoside-5'-monophosphate (3-MetAde-NMP),* ***(G)*** *6-(Dimethylamino)purine 5´-monophosphate (6-DMP-NMP). Abbreviated according to the recommendations of the IUPAC-IUB Commission on Biochemical Nomenclature. Created on ChemSketch (4).*

1.2. Purine phosphoribosyltransferase

1.2.1. General concepts

Phosphoribosyltransferase enzymes (PRTases), as mentioned above, are involved in the purine synthesis via the salvage pathway, in addition to vitamin and amino acid synthesis (26).

It catalyzes the reaction that transfers an α-1'-pyrophosphate group (from PRPP, donor) to a nucleobase (acceptor), thus forming the corresponding nucleotide and releasing a pyrophosphate (PPi) (26,27). Moreover, the reaction is metal dependent (28). The divalent cation, usually magnesium (Mg^{2+}) is essential for the reaction to happen as it has been reported to be associated with the PRPP (29). Adenine phosphoribosyltransferases (APRTases) from several species have been studied and found to form an octahedral

coordination complex with Mg^{2+}, coordinated by two water molecules present in the active site and four oxygen atoms from the PRPP (Figure S 1), with binding distances ranging from 2.0 to 2.3 Å (7,30).

Regarding PRTase classification, they are divided into two types based on their substrate specificity. Class 1 PRTs (APRT) are specific for 6-aminopurine bases and its analogues such as 2-chloroadenine, 2-fluoroadenine, and 2,6-diaminopurine, among others. Class 2 PRTs (XPRT, HPRT, GPRT, HGPRT, etc.) target 6-oxopurines such as hypoxanthine, xanthine, and guanine, as well as their analogues (31).

1.2.2. Adenine phosphoribosyltransferase

This work is focused on the class 1 PRTases (EC 2.4.2.7) from *Escherichia coli* (*Ec*APRT), which, as shown in figure 5, catalyzes a reversible reaction that transfers the ribose 5-phosphate group (from the PRPP) to the adenine nucleobase thereby forming AMP and releasing a PPi (28,30,32).

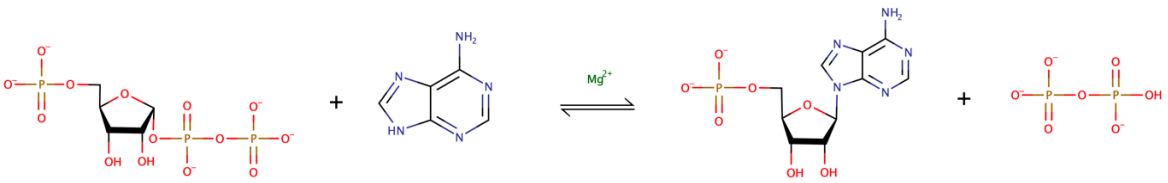


Figure 5. *Reaction catalyzed by APRTase. PRPP: phosphoribosyl pyrophosphate, Ade: adenine, AMP: adenine monophosphate, PPi: pyrophosphate. Created on ChemSketch (4).*

APRTases display conserved elements in their structure in all organisms (Figure 6). In this sense, *Ec*APRT contain a core with six α-strands encircling nine β-strands. This core contains the hood subdomain, where the nucleobase binds, so it is the responsible for the substrate specificity (does not recognize 6-oxopurines); a flexible loop that takes on a closed conformation for the reaction to take place; a PRPP binding motif that binds the ribose ring and a PPi loop. Therefore, the active center is located between these four aforementioned structures. As residues of other subunits are involved in substrate recognition, no monomeric PRTases have been found, so the presence of at least two monomers is essential for catalysis (7,32–34).

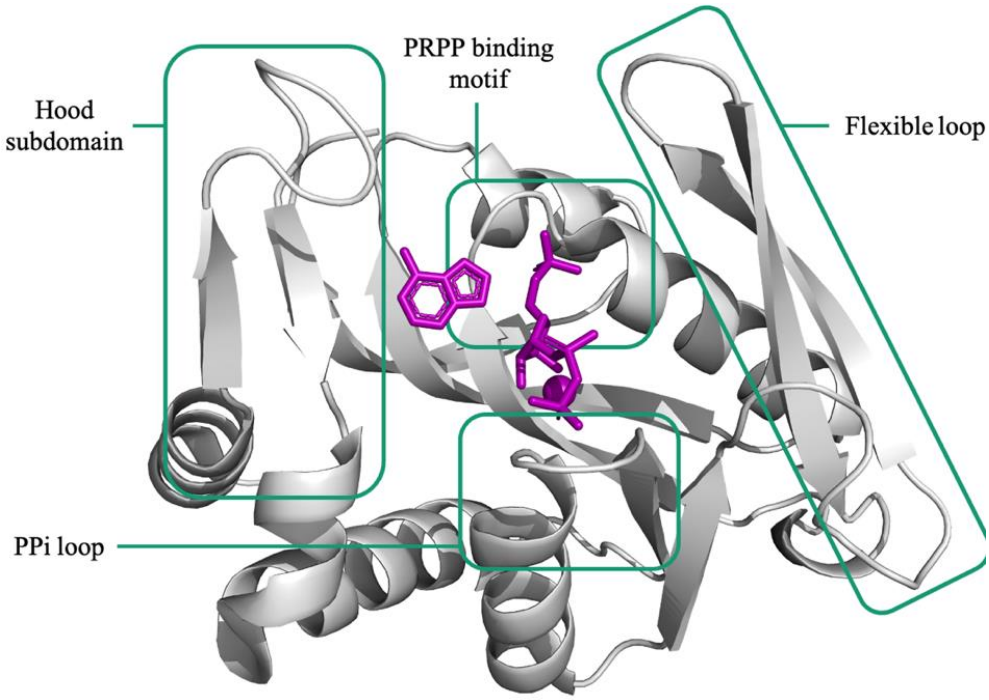


Figure 6. *Cartoon representation of EcAPRT monomer structure (PDB id 2DY0) in grey. Ligands are shown (adenine and PRPP) in purple sticks and Mg^{2+} ion is shown as a purple sphere. Created on PyMOL (35).*

Regarding its reaction mechanism, a substitution nucleophilic unimolecular (Sn1) reaction has been proposed (Figure 7). Firstly (Figure 7A), the PPi would be spontaneously released from PRPP, being stabilized by Mg^{2+}. As an intermediate molecule, a carbocation is formed on the C1 of the ribose from PRPP, stabilized by the carbonyl group of Arg66 (residue numbering according to *Ec*APRT, UNIPROT id P69503). Next (Figure 7B), the OE (charge -1) of the Glu103 (residue numbering according to *Ec*APRT, UniProt id P69503) side chain would perform a nucleophilic attack on the hydrogen of the N7, so the N9 binds to the C1 carbocation. Finally (Figure 7C), the hydrogen from OE of the Glu103 residue is released and trapped by the PPi. The reaction products are AMP and PPi (Figure 7D) (30,36–38).

However, a substitution nucleophilic bimolecular (Sn2) mechanism has also been described in which the enzyme would perform the nucleophilic attack simultaneously with the formation of the product (38).

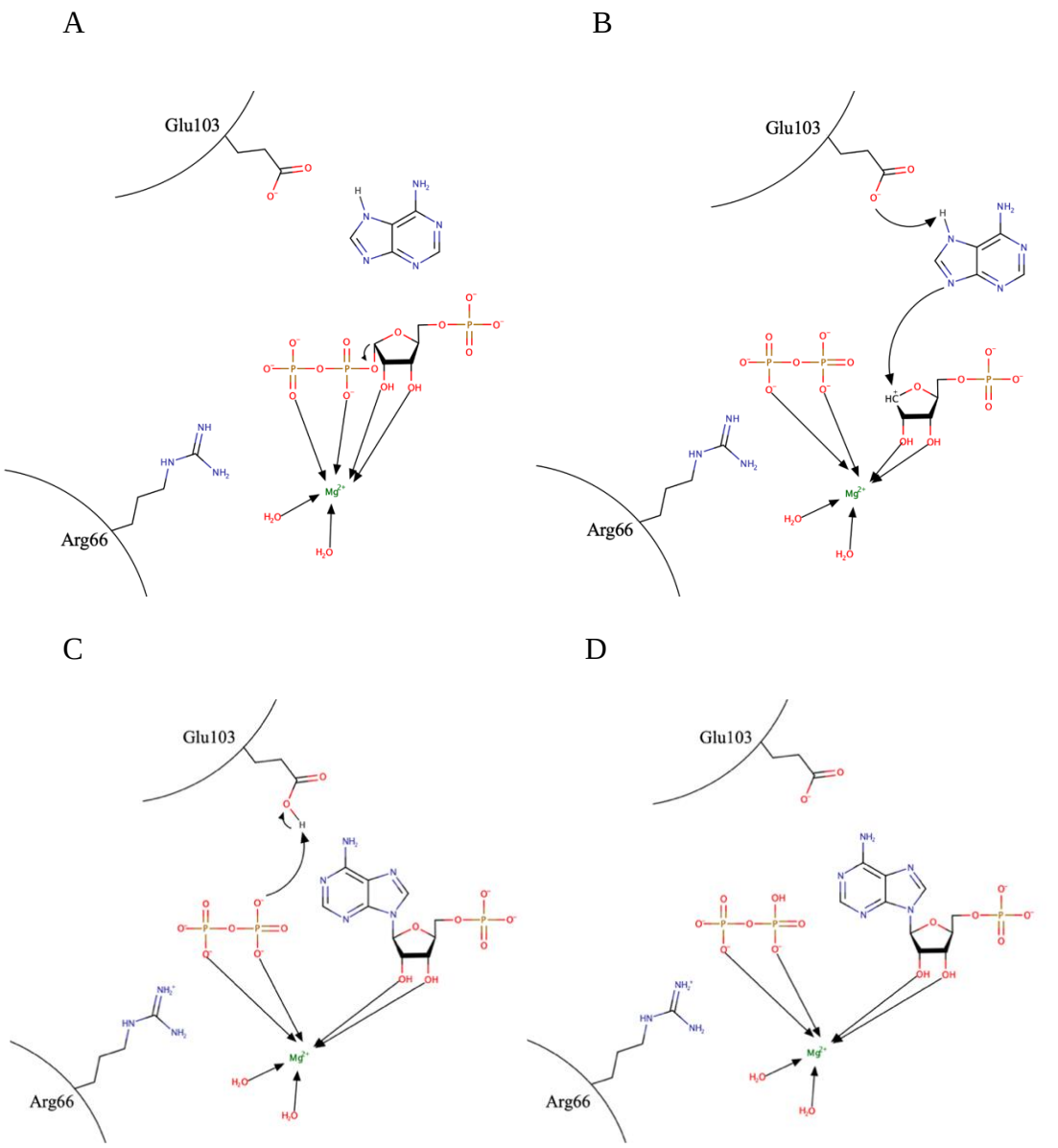


Figure 7. *Sn1 reaction mechanism of APRT.* ***(A)*** *Carbocation formation,* ***(B)*** *nucleophilic attack of the nucleobase on the carbocation,* ***(C)*** *recovery of the negative charge of the glutamic OE from the PPi,* ***(D)*** *reaction products. Created on ChemSketch (4).*

1.3. Molecular dynamics

Atoms are composed of a nucleus (with mass and positive charge) and electrons (with negligible mass and negative charge). The Born-Oppenheimer approximation describes the separation of the nuclear and electron motions, so depending on where we focus on the atom, we will have a classical model (Newtonian mechanics or molecular mechanics, MM) that considers just the nuclear motion or a quantum model (quantum mechanics, QM) that considers the electron motion (39–41).

Molecular dynamics (MD) is a powerful computer simulation technique to study the motion of a molecular system and to explore the conformational space of large molecules such as proteins (42).

1.3.1. Molecular dynamics simulation

To carry out a MD simulation, we need to specify the initial positions, the distribution of velocities and the accelerations of each of the atoms in the system under study. Velocities are usually assigned randomly and taken from a Maxwell-Boltzman distribution (43). Particles interact with each other through potentials functions, described below, that cause their position to vary. Then, the motion of the particles is propagated according to the Newton's second law, defined in equation 1A, in the simplest case by keeping the acceleration constant, given by equation 1B (39,44).

A $$F_i = m_i \cdot a_i = \frac{1}{2} \cdot a \cdot t^2 + v_0 \cdot t + x_0$$

B $$a_i = -\frac{1}{m}\frac{dE}{dr}$$

Equation 1. (A) *Newton´s second law. F: Force on each particle of the system (i), m_i: mass of each particle, a_i: acceleration of each particle, t: time set for the dynamics, x_0: initial position, v_0: initial velocity.* ***(B)*** *Acceleration obtained from the derivative of potential energy (E) with respect to atomic positions (r).*

The potential energy is obtained from equation 2, in which the energy of the molecule is calculated as the contribution of directly bonded atoms (bonding terms), bonds (involving two atoms), angles (involving three atoms) or dihedrals (involving four atoms). Those that are not bonded (non-bonding terms) are also considered, such as electrostatic and van der Waals interactions. The complete set of equations and parameters are also known as a force field (45,46).

$$E = E_{bonded} + E_{non-bonding} = \sum_i E_{bonds} + \sum_i E_{angles} + \sum_i E_{dihedral} + \sum_i (E_{elec} + E_{vdW})$$

Equation 2. *Equation to obtain the potential energy of all particles in the system.*

Therefore, during the MD simulation, at each step, the forces and accelerations of each particle are calculated, obtaining new coordinates within a time interval, giving rise, finally, to what is known as a trajectory.

1.3.2. Molecular dynamics simulation protocol and analysis

MD simulation consists of three successive steps. Firstly heating, starting from an initial temperature (usually -173 ºC) the system is brought to the working temperature (usually 27 ºC). After setting the working temperature, the system is equilibrated at a constant pressure and finally the production step, carried out with the final conditions of the equilibration (45,46).

Regarding MD analysis, different calculations (geometrical, solvation and energetic, among others) can be made and compared to experimental data. For example, the analysis of the structural stability is given by the Root-Mean-Square-Fluctuations (RMSF, measuring individual residue fluctuation relative to the mean system fluctuation) and the Root-Mean-Square-Deviation (RMSD, measuring the average of the entire structure flexibility over time, relative to a reference structure). In addition, the total energy of the system can be analyzed and decomposed into different components to observe the energy distribution among the active center residues, for example, in order to study the interaction energies with a ligand (45,46).

Actually, significant and rapid progresses have been made, allowing the study of larger systems during longer simulation times (42). However, to reach simulation times in the range of reality, the main limit is still the computational cost involved. As an alternative, coarse-grained force fields can be used instead, consisting of decreasing the complexity of the simulation system by grouping certain atoms into virtual particles, then reducing the size of the system. As a result, simulations are quicker (47–49).

1.4. Multienzymatic system to produce nucleoside analogs

Le Chatelier's principle states that an equilibrium will shift in the opposing way, in order to compensate for a change in pressure, temperature, or the concentration of products or reactants (50). This has led to the design of enzymatic cascades to minimize by-products, and simplify the downstream process (purification), to finally obtain a higher final yield (50,51). As a subsequent step to the presented study, we propose the following multienzymatic system (Figure 8) with the aim of improving the production yield of nucleoside analogues.

Similar to the previous multienzymatic cascades described (50–53), this proposed multienzymatic cascade would involve the use of two enzymes. First, an NDT (reaction 1) will catalyze a transglycosylation reaction to produce the desired nucleoside analogs. Subsequently, in order to improve the conversion rates of this reaction, a second enzyme is used. In this case, APRT (reaction 2) forms AMP from the released adenine of the first reaction. In this sense, the structural and biochemical characterization of APRTase is important in order to optimally apply it to the cascade. One of the main disadvantages when using PRTases is the high cost of PRPP, essential substrate for the reaction. So, the proposed cascade can be further improved to regenerate the PRPP by using other two enzymes, PPK (polyphosphate kinase, reaction 3 and 4) and PRPP synthase (reaction 5) (Figure 8).

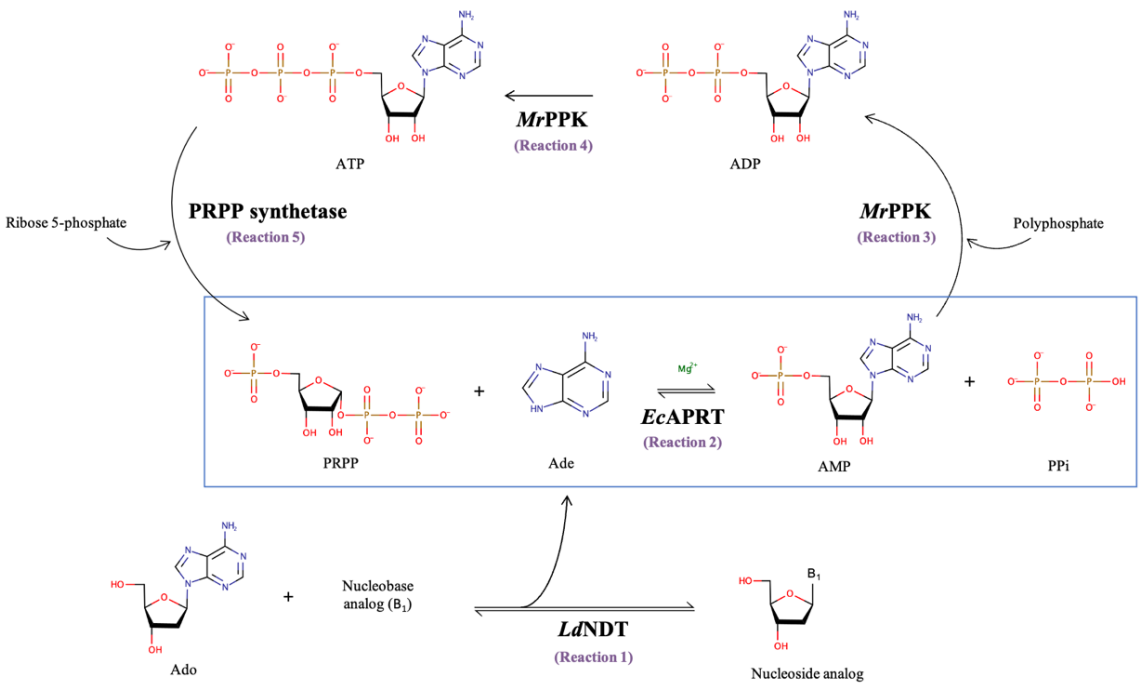


Figure 8. *Schematic representation of the proposed multienzymatic system to produce a nucleoside analog.* ***Reaction 1:*** *LdNDT (Lactobacillus delbrueckii nucleoside 2'-deoxyribosyltransferase).* ***Reaction 2:*** *EcAPRT.* ***Reactions 3 and 4:*** *MrPPK (Meiothermus ruber polyphosphate kinase).* ***Reaction 5:*** *PRPP synthetase. PRPP: Phosphoribosyl pyrophosphate, Ade: adenine, AMP: adenine monophosphate, ADP: Adenine diphosphate, ATP: adenine triphosphate. Created on ChemSketch (4).*

2. Objectives

2.1. General Objectives

As stated above, due to the usefulness of *Ec*APRT, the aim of this work is to characterize the enzyme thus determine its optimal conditions in order to establish the conditions and substrate scope for carrying out the enzymatic cascade.

2.2. Specific objectives

- Production and purification of recombinant adenine phosphoribosyltransferase from *E. coli* (*Ec*APRT).
- Full biochemical characterization of *Ec*APRT.
- Application of *Ec*APRT in the synthesis of non-natural nucleoside-5'-monophosphates.

3. Materials and methods

3.1. Reagents

Luria Bertani (LB) cell culture media reagents are from Difco (St. Louis, MO, USA). All additional organic solvents and reagents were given by Symta (Madrid, Spain). Carbosynth Ltd. (Compton, UK) supplied both natural and synthetic nucleobases.

3.2. Analysis of the primary and secondary structure

APRT enzyme sequences from other organisms were obtained by performing a blastp (54) of *Ec*APRT (UniProt id P69503) against the UnipProtKB/Swiss-Prot (swissprot) database (55). The search was performed with the BLOSUM62 matrix with a word size of 5. Sequences with the least sequence homology were selected to perform a multiple sequence alignment with Clustal Ω (56) . This was further processed with ESPript (57).

3.3. Molecular dynamics

3.3.1. Model construction

The tertiary structure of *Ec*APRT was previously elucidated by X-ray diffraction corresponding to the PDB id 2DY0. To ensure the quaternary structure, protein-protein docking studies were carried out using ClusPro (58), obtaining 500 poses. The PDB file was adapted to make use of the Amber20 (59) software package. To evaluate the protonation state of the receptor, a constant pH molecular dynamics simulation (CpHMD) was performed. In the initial protonation state Asp and Glu are deprotonated; Lys, Tyr and Cys are protonated while His is doubled protonated. The system was prepared for the simulation using the ff10 and frcmod.constph force fields from Amber20. The protonated receptor was introduced in an implicit water solvent (60). Once the system was generated, its energy was minimized for 1000 cycles with a 100 mM of salt concentration, maintaining constant the protonation state. For MD simulation, it was firstly heated form -173 ºC to 27 ºC using the Langevin dynamics thermostat with a γ = 5.0 collision frequency. After equilibration for 2 ns, the production step was carried out (1-60 ns), allowing the modification of the protonation state in the 4-11 pH range. The trajectories

were analyzed with cpptraj (61) obtaining the pK_a of each residue for obtaining the protonation percentage at each pH.

3.3.2. Standard molecular dynamics assay

MD simulations were performed using the *pmemd.cuda* module of Amber20 in the Single-Precision-Fixed-Precision (SPFP) mode. The system was embedded in a box of TIP3P water molecules that extended 12.0 Å away from any solute atom, and neutralized with 8 Na^+ ions. The system was relaxed by energy minimization in three consecutive steps (hydrogen minimization, dissolvent minimization and whole system minimization). The heat step consisted of increasing the temperature from -173 ºC to 27 ºC using the Langevin dynamics thermostat with a γ = 5.0 collision frequency at constant volume periodic boundaries. Next, the system was equilibrated in six steps. On the one hand, in the first three steps (1 ns) the volume was maintained constant (NVT) and the harmonic potential was decreasing in order to gradually eliminate the motion restrain. On the other hand, in the last three steps (1 ns) the pressure was maintained constant (NPV) using the Berendsen barostat isotropic position scale. Finally, the production step was carried out maintaining constant the pressure for 300 ns using the ff19SB force field (62). The trajectories were analyzed with cpptraj (61) and displayed in PyMOL (35).

3.3.3. Coarse grained molecular dynamic

For exploring the protein conformational space, a coarse grained (CG) molecular dynamic was carried out. SIRAH force field (63) implemented in Amber20 was applied for protein parametrization, letting a 1 μs simulation time. The system was embedded by a preestablished water box (WT4 molecules) and its electroneutrality was established by replacing WT4 molecules for electrolytes. The MD was carried out following the standard MD protocol explained above with 5 ns of equilibration. For transforming the CG coordinates to fine grained, the *sirah_vmdtk.tcl* plugin (from VMD) was used. The MD was also carried out following the standard MD protocol explained above to evaluate the consistency of the system. Finally, the open (*Ec*$APRT_{op}$) and closed (*Ec*$APRT_{cl}$) conformations were obtained and validated on a Ramachandran plot.

3.3.4. Molecular dynamics of the enzyme-substrate complex

The complex was done using *Ec*APRT$_{cl}$. The ligands (adenine, PRPP and Mg^{2+}) were positioned on the active center by manual docking. To this end, *Ec*APRT$_{cl}$ was overlapped with the three-dimensional structure APRT from *Homo sapiens* (PDB id 6FCI) as it has a closed conformation with the ligands correctly positioned. The ligand structures were refined by semiempirical quantum mechanics (sqm) implemented in antechamber (64) of Amber20. The protein was refined using ff19SB force field. Additional force fields with the missing parameters were calculated using parmchk2 (64) of Amber20. The polyphosphate parameters were obtained from the study of Carlson and coworkers (65). Molecular dynamics were run following the MD standard assay but neutralizing the system with 14 Na^+ ions.

3.4. Transformation and protein Expression

The encoding mutant *prt* gen for adenine phosphoribosyltransferase from *Escherichia coli* (NCBI Reference Sequence: NP_415002.1) was subcloned as a *Nde*I-*Eco*RI fragment into the expression vector pET28b (+). *E. coli* BL21 cells were transformed with the recombinant plasmid by thermal shock, and LB-agar plaques containing *E. coli* BL21 (DE3) pET28b (+)$_{EcAPRT}$ and supplemented with kanamycin 50 µg/mL were grown at 37 °C overnight. An isolated colony was selected, further grown in LB medium containing kanamycin 50 µg/mL and stored in 20% glycerol at - 80 ºC.

Cells were obtained from this cell stock in order to generate a pre-inoculum in 6 mL of LB medium supplemented with 50 µg/mL kanamycin. This was incubated overnight at 37 ºC and 220 rpm. It was then transferred to fresh medium (600 mL) and incubated at 37 ºC and 130 rpm, measuring the OD_{600} every 30 minutes. Once reached an OD_{600} 0.9 - 1.1, protein overexpression was induced by adding 0.5 mM isopropyl-b-D-1-thiogalactopyranoside (IPTG) to the medium and cells were further grown overnight at 37 ºC and 130 rpm. Cells were harvested by centrifugation at 3,600 × *g* at 4 ºC for 30 minutes. The obtained pellet was resuspended with Tris-HCl 20 mM buffer, pH 8.0, and centrifuged with the same method. This cell pellet was stored at -20 °C until the cell lysis.

3.5. Protein purification

The previous cell pellet was resuspended in 20 mM Tris-HCl buffer, pH 8.0. Crude extracts were prepared by cellular disruption using a digital sonifier. This lysis method was carried out using ten 30-second pulses with amplitudes of 2x40%, 3x30%, and 5x20%. Then, the lysates were centrifuged at 16,500 × *g* for 30 min at 4 °C to remove cellular debris.

Cleared lysates were loaded (1 mL/min) onto a 5-mL HisTrap FF column (GE Healthcare), using a BioLogic LP (Low Pressure, Bio-Rad) unit. The column was previously equilibrated with binding buffer (20 mM Tris-HCl buffer, pH 8.0, with 100 mM NaCl and 20 mM imidazole). Bound proteins were eluted (1 mL/min) using a linear gradient of imidazole (from 20 to 500 mM). Fractions containing *Ec*APRT were collected using a collector and identified by SDS-PAGE. All the process was detected by measuring the absorbance at 280 nm through LP Data View software. Fractions were dialyzed and concentrated using Amicon® filters at 3,500 x *g* at 4 °C with 20 mM Tris-HCl buffer, pH 8.0. The final purity of the sample was also evaluated by SDS-PAGE.

Electrophoresis separation was carried out on a discontinuous (stacking and a running gel) (Table 1), retrieved from Laemmli (66). It was carried out at room temperature (RT) with a constant amperage of 25 mA per gel. After the electrophoresis, the gels were stained for 30 minutes with a staining solution (Coomassie Brilliant Blue G250 0.012% (w/v), 50% methanol, and 10% glacial acetic acid). The staining solution was removed and the gels were rinsed with water and treated overnight with a distaining solution (40% methanol and 10% glacial acetic acid). Finally, the gel visualization was accomplished using the Bio-Rad ChemiDoc unit and Image Lab TM software. Moreover, the precision Plus Protein Dual Color Standards 10-250 kDa (Bio-Rad; California, USA) was used as the molecular weight marker.

Table 1. *Components for preparation of SDS-PAGE.*

Components	**Stacking gel (5%)**	**Running gel (12.5%)**
Water	3.85 mL	2.4 mL
Acrylamide 30% (w/w) / Bisacrylamide 0.8% (w/w)	0.81 mL	4.65 mL
TRIS-HCl buffer	0.25 mL (2.5 M, pH 6.8)	4.2 mL (1 M, pH 8.8)
SDS 20%	25 µL	57 µL
APS 10%	25 µL	60 µL
TEMED	10 µL	12 µL

Protein concentration was determined, on the one hand by measuring its absorbance (spectrophotometrically, with a NanoDrop® spectrophotometer) at 280 nm (67) and using its molar extinction coefficient ε_{280} = 10,430 M^{-1} cm^{-1} (obtained from the *Ec*APRT aminoacidic sequence), following the Lambert-Beer law (Equation 3). Additionally, the protein concentration was also determined by densitometric analysis, using as a standard curve the enzyme *Ec*HPRT at different known concentrations.

$$A = c \cdot \varepsilon \cdot l$$

Equation 3. *Lambert-Beer law in which the absorbance (A) is proportional to the protein concentration (c), molar extinction coefficient (ε) and optical path length (l).*

3.6. Standard enzyme activity assay

The standard activity assay was carried out by incubating 0.03 µg of pure enzyme in a reaction volume of 40 µL containing 10 mM adenine and 10 mM PRPP in 12 mM Tris-HCl buffer pH 8.0, in presence of 12 mM $MgCl_2$. The reaction mixture was incubated at 50 °C and 300 rpm in an orbital shaker for 3 minutes. The enzyme was inactivated by adding 40 µL of cold methanol in an ice bath and then heating for 5 min at 95 °C. After centrifugation at 9000 × *g* for 5 minutes, samples were half-diluted with water and frozen at -20 °C. Nucleoside production was analyzed through HPLC to quantitatively measure the reaction products, as described below. All determinations were carried out in triplicate and the maximum error was less than 5%. Under such conditions, one international activity unit (IU) was defined as the amount of substrate (µmol) converted per unit time (minutes) and per milligram of total enzyme. All determinations were carried out in triplicate and the maximum error was less than 5%.

3.7. Influence of temperature and pH

*Ec*APRT activity was determined using the standard enzyme activity assay across a 20-90 ºC temperature range. Similarly, the effect of the pH was studied carrying out the standard assay with sodium citrate (pH 4-6), MES (5.5-7), sodium phosphate (pH 6-8), Tris-HCl (pH 7-9) and sodium borate (pH 8-11) as reaction buffers (50 mM). All determinations were carried out in triplicate and the maximum error was less than 5%.

3.8. Thermal stability assay

Thermal stability of *Ec*APRT and was studied by incubating 30 µg of pure enzyme in 20 mM Tris-HCl buffer, pH 8.0 in a 40-70 ºC temperature range for 2 h. Samples were taken periodically, placed in an ice bath and thereupon enzymatic activity was evaluated following the standard assay described above. All determinations were carried out in triplicate and the maximum error was less than 5%. Thermal stability was defined as the relative activity between the first and successive reactions.

3.9. Effect of divalent cations

The effect of the ions on the enzyme activity was determined following the standard enzyme activity assay described above using the optimal pH and temperature. Several divalent cations (Mg^{2+}, Mn^{2+}, Co^{2+}, Zn^{2+}, and Ca^{2+}) were tested at different concentrations (all of them tested at 2 mM, 5 mM, 12 mM, and 20 mM). All determinations were carried out in triplicate and the maximum error was less than 5%.

3.10. Effect of organic solvents

For studying the influence of organic solvents on *Ec*APRT activity, the standard enzyme activity assay was followed, considering the optimal conditions described in the previous assays (pH, temperature and concentration of divalent cations). Aprotic polar solvents (acetonitrile, acetone, chloroform, DMF, DMSO and ethyl acetate), alcohols (methanol, ethanol, isopropanol and tertbutanol) and polyalcohols (glycerol and propylene glycol) were tested at 20% (v/v) of the total reaction volume (40 µL). In addition, as a control, an additional reaction was carried out in absence of organic solvents. All determinations were carried out in triplicate and the maximum error was less than 5%.

3.11. *Ec*APRT activity on nucleobase analogs

Enzymatic production of non-natural nucleotides was carried using 2-ClAde, 2-FAde, 2,6-DAP, 6-DMP, 3-MetAde and N6-MetAde. Reactions were carried out incubating 0.3 µg of *Ec*APRT. With respect to the reaction conditions, they were carried out with 1 mM PRPP and 1 mM nucleobase analogs, in 12 mM Tris-HCl buffer pH 8.0 in presence of 1.2 mM $MgCl_2$, at 50 ºC in a final volume of 40 µL, with 300 rpm orbital shaking for 3 minutes. Thereafter, reactions were stopped following the standard assay described above. All determinations were carried out in triplicate and the maximum error was less than 5%. Enzymatic activity was quantified by HPLC as described in analytical methods.

3.12. Analytical methods

HPLC was used to identify and quantify reaction products, with the Agilent 1100 series HPLC equipment. The used chromatographic column was the ACE EXCEL 5 µm CN-ES column 250 mm x 4.6 mm (Advanced Chromatography Technologies). The column was equilibrated with 100% triethylammonium acetate at a continuous flow rate of 0.8 mL/min and reaction products were eluted into the diode array detector for quantification at 254 nm. Retention times for substrates and reaction products are shown in table 2. Two distinct elution techniques were used. For adenine (natural base), an isocratic elution was carried out (100% triethylammonium acetate) for 30 minutes. For non-natural bases, a discontinuous gradient was done (100–90% triethylammonium acetate and 0–10% acetonitrile for 10 minutes, and 90–100% triethylammonium acetate and 10–0% acetonitrile in the following 10 minutes) for 20 minutes, followed by an isocratic elution (100% triethylammonium acetate) for 10 minutes.

Table 2. *HPLC retention times of the substrates and reaction products.*

	Nucleobase compound	**Retention times (min)**	**Nucleotide compound**	**Retention times (min)**
Natural	Adenine	8.0	AMP	2.8
Non-natural	2- ClAde	11.3	2- ClAde NMP	4.3
	2-FAde	9.0	2-Fade NMP	5.0
	2,6- DAP	7.8	2,6- DAP NMP	4.4
	N6-MetAde	9.0	N6-MetAde NMP	3.8
	3-MetAde	7.1	3-MetAde NMP	3.9
	6-DMP	13.3	6-DMP NMP	3.8

4. Results and discussion

4.1. Structural analysis of *Ec*APRT

*Ec*APRT has been cloned into the pET28b (+) vector, containing the following amino acid sequence:

MGSS**HHHHHH**SSGLVPRGSHMTATAQQLEYLKNSIKSIQDYPKPGILFRDVTSLLEDPKAYALS
IDLLVERYKNAGITKVVGTEARGFLFGAPVALGLGVGFVPVRKPGKLPRETISETYDLEYGTDQ
LEIHVDAIKPGDK**VLVVDDLLATGGT**IEATVKLIRRLGGEVADAAFIINLFDLGGEQRLEKQGI
TSYSLVPFPGH

This sequence includes a His-Tag (highlighted in dark purple) at the aminoacidic terminal (N-t) enabling a purification by affinity chromatography (explained above). It contains 203 amino acids with a theoretical pI of 6.30 (68), containing approximately the same amount of basic and acidic residues (14.77% and 11.82%, respectively). The sequence was analyzed with different tools (69) in order to identify conserved patterns (Figure 9).

[LIVMFYWCTA]-[LIVM]-[LIVMA]-[LIVMFC]-[DE]-D-[LIVMS]-[LIVM]-
[STAVD]-[STAR]-[GAC]-x-[STAR]

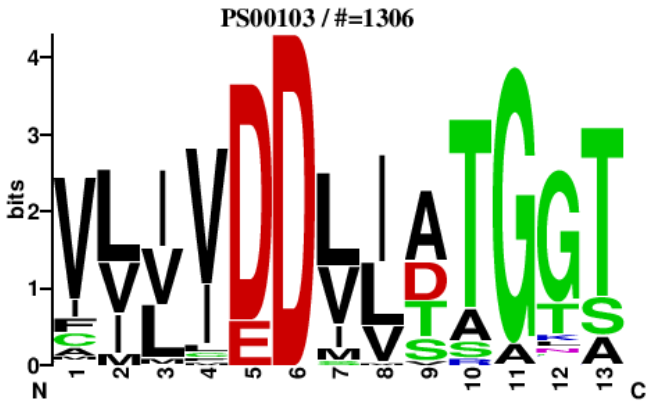


Figure 9. *Sequence logo of the pattern.*

The pattern shown in the figure 10 (highlighted in the sequence in dark green) is conserved in the family of purine/pyrimidine phosphoribosyltransferases, which coincides with the PRPP binding motif. As can be seen, the pattern is constituted by four hydrophobic residues followed by two negatively charged amino acids, mostly aspartic acid, then two hydrophobic residues are followed by one amino acid which can be either charged or uncharged and finally four small residues.

The structure of *Homo sapiens* adenine phosphoribosyltransferase (*Hs*APRT) has been previously studied (33), and the amino acids forming the main domains were identified. These structures are "hood subdomain" or "purine pocket", (residues 7-35) and "flexible loop" (residues 93-112). When comparing the amino acid range with the sequence alignment shown (Figure 10), Glu103 and Tyr104 residues contained in the flexible catalytic loop are highly conserved in all species, highlighting their importance. Glu103 is the residue responsible for performing the nucleophilic attack and Tyr104 residue has been reported to be essential for maintaining the active conformation. In this sense, Tyr104 forms a hydrogen bond with Arg66 (located in the PPi loop), preventing the solute to access to the N-9 and C1, where the reaction takes place (7). Regarding the hood subdomain, Phe28 and Arg29 interacts with the adenine via van der Waals and hydrogen bonds, respectively, reason why they are conserved.

Other regions that appear to be highly conserved are those between amino acids 122-134. This matches with the PRPP binding motif, which corresponds to the previous consensus sequence pattern for PRTases (Figure 9). Asp126, Asp127, Ala130, Thyr131, Gly132 and Thr134 are conserved in all species (Figure 10) due to the fact that they interact with the PRPP via hydrogen bonds. Additionally, Leu128 residue interacts with adenine through van der Waals forces. Moreover, the region between amino acids 60-71, is named PPi loop, as determined in other study (29). Arg66 is very well-preserved in all species due to the aforementioned interaction with Tyr104 and the additional interaction with the PPi. Ala65 residue is also well conserved because it is essential for PPi recognition (31,70,71).

As already seen in other studies (30,37), PRTases are composed of alternative β-sheets and α-helices, constituting the core (Figure 11). *Ec*APRT is formed of 9 β-sheets and 6 α-helices, thus constituting the core domain, being characteristic of purine PRTs. As previously reported (33), the hood subdomain is formed by two α-helices (α1 and α2) and a β turn located between a two stranded β-sheet (β1 and β2). The flexible loop also includes a β turn between two β-strands (β5 and β6). It should be noted that the two amino acids that coincide with the β turn are Glu103 and Tyr104. Moreover, the PPi loop is flanked by one β-sheet (β3) and an α-helix (α4) similar to the PRPP binding motif (β7 and α5).

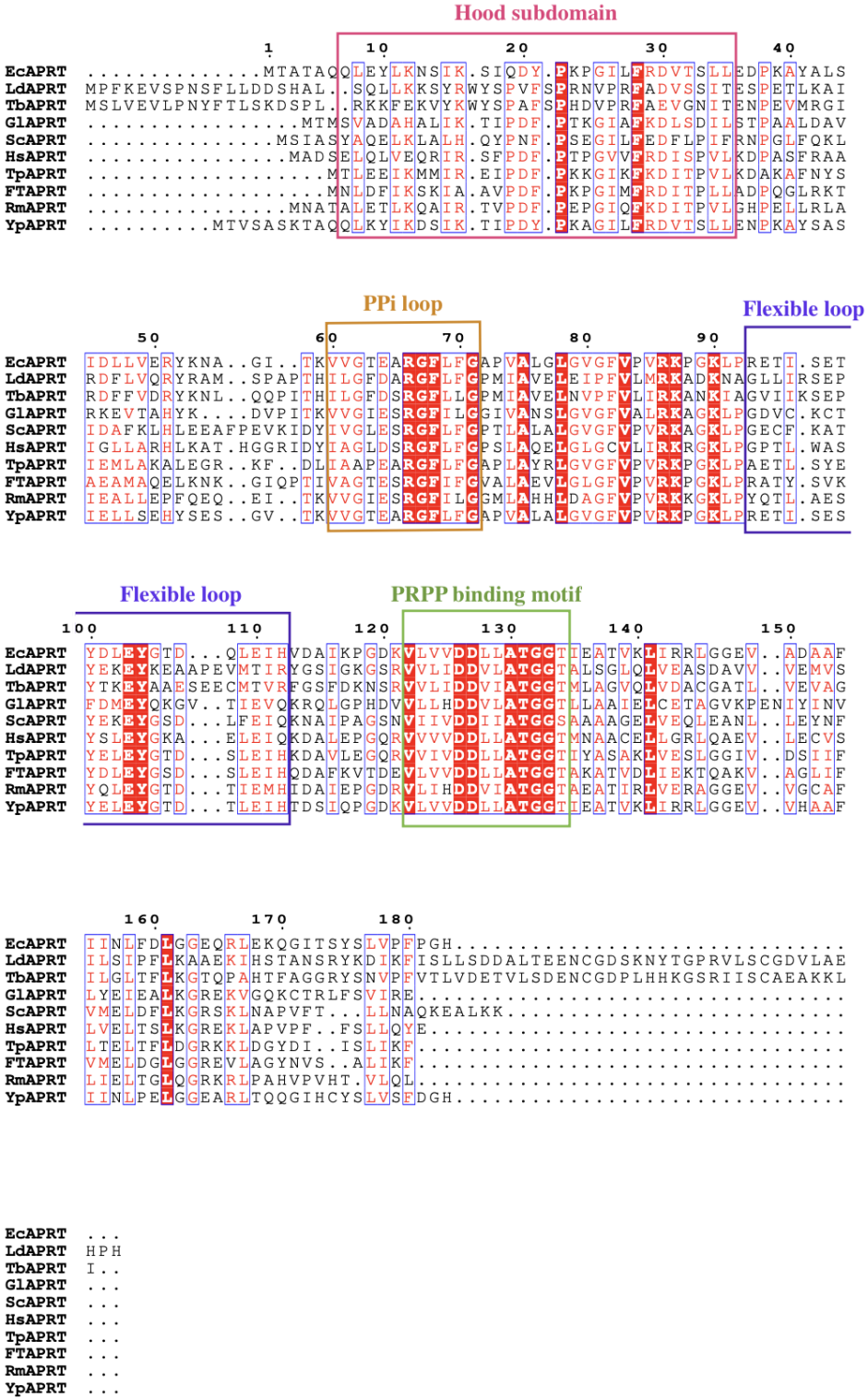


Figure 10. *Multiple aminoacidic sequence alignment. EcAPRT (Escherichia coli, UniProt id: P69503), LdAPRT (Leishmania donovani, UniProt id: Q27679), TbAPRT (Trypanosoma brucei, UniProt id: Q57V32), GlAPRT (Giardia lamblia, UniProt id: Q967M2), ScAPRT (Saccharomyces cerevisiae, UniProt id: P49435), HsAPRT (Homo sapiens, UniProt id: P07741), TpAPRT (Thermoanaerobacter pseudethanolicus, UniProt id: B0K969), FtAPRT (Francisella tularensis, UniProt id: Q5NII9), RmAPRT (Rhodothermus marinus, UniProt id: D0MGF4), YpAPRT (Yersinia pseudotuberculosis, UniProt id: Q66DQ2). Amino acids underlined and white lettering indicate fully conserved residues and amino acids with red lettering indicate conservation of weak groups. Sequences were aligned by CLUSTAL Ω (56) and edited by ESPript (57).*

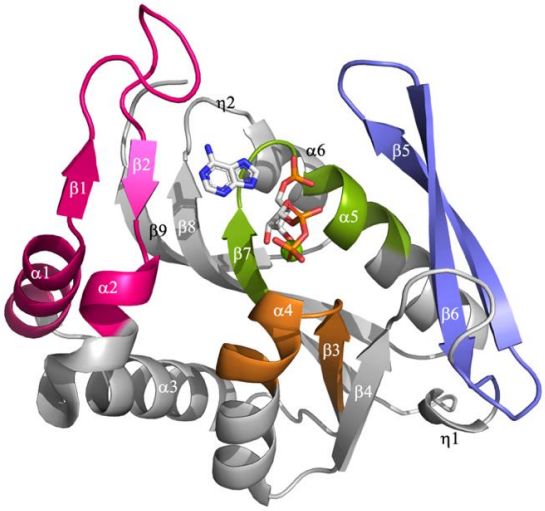


Figure 11. *Cartoon representation of EcAPRT monomer structure (PDB id 2DY0). Hood structure (pink), PPi loop (orange), flexible loop (blue) and PRPP binding* motif *(green). Ligands are shown (adenine and PRPP) in sticks and* Mg^{2+} *ion is shown as a green sphere. Created on PyMOL (35).*

4.2. Modelization of *Ec*APRT$_{cl}$

The three-dimensional structure of *Ec*APRT was previously published in the PDB database (id: 2DY0), hereinafter referred to as *Ec*APRT$_{op}$. This structure was studied and due to its open conformation, it was further refined to finally obtain its closed conformation, referred to as *Ec*APRT$_{cl}$.

Firstly, analyzing the RMSD of *Ec*APRT$_{op}$, the system showed to be stable for 300 ns (data not shown). In addition, since proteins are dynamic entities (72), a depiction of its flexibility was obtained. For this purpose, the RMSF by residue of *Ec*APRT$_{op}$ enzyme was analyzed (Figure 12).

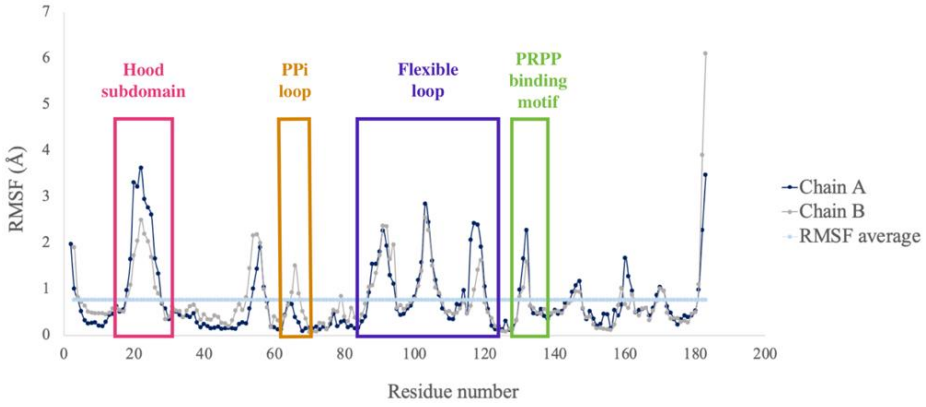


Figure 12. *RMSF of EcAPRT$_{op}$ residues (PDB id 2DY0). The fluctuations of chain A are shown in dark blue and chain B in gray. The average of the fluctuations is shown as a straight light blue line. Each structure is highlighted, hood subdomain (pink), PPi loop (orange), flexible loop (blue) and PRPP binding motif (green).*

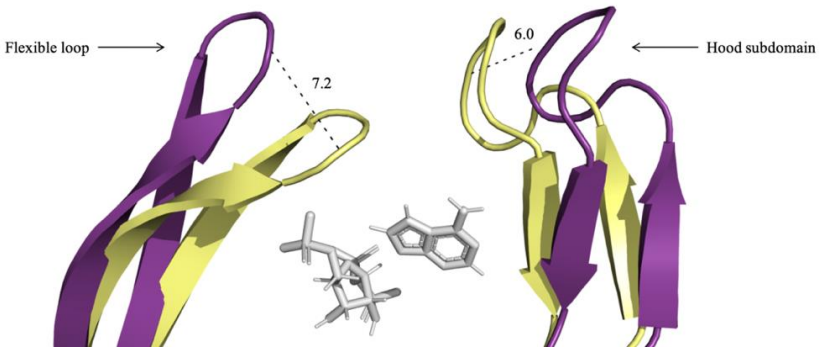
As seen in figure 12, a straight line (light blue) represents the average of the fluctuations, so the residues above it are considered the most flexible. It is important to point out that chain A and B follow a very similar trend.

The most flexible regions coincide with the aforementioned structures involved in substrate recognition and catalysis. Firstly, a significant fluctuation is observed between amino acids 18-27, which coincides with the previously defined hood subdomain. Between residues 62-67, the structure, which was previously referred to as the PPi loop, does not have as much flexibility as the hood subdomain. In addition, three large fluctuations are observed between residues 86-120, which comprises the flexible loop. Considering the secondary structure mentioned above (Figure 11), the first and second peaks are separated by β5 while the second and third by β6, making this region more rigid. The Glu103 and Tyr105 residues are located in the middle peak, and therefore, the loop in which they are located takes part in the catalytic reaction, making it the most flexible of the three. Finally, the residues that are part of the last flexible region are those that form the PRPP binding motif, 129-133 residues. Additionally, there is at least one more flexible region between residues 53-59, thus coinciding with a loop which is not related with catalysis or substrate recognition. Finally, two more peaks are seen at the beginning and end of the sequence due to the inherit flexibility of N-terminal and C-terminal.

To further study the possible conformations of *Ec*APRT$_{op}$ structure, a coarse grained force field MD was made in order to reduce the complexity of the system. As a result, two different types of conformations (open and closed) were observed for the protein, over time (Figure 13). According to the literature (29,34), the open conformation allows the active site to be accessible for substrate binding and product release, while the catalysis is performed on the closed conformation as the residues are correctly positioned. When comparing *Ec*APRT$_{op}$ and *Ec*APRT$_{cl}$ structures, a difference of 6.0 and 7.2 Å (Figure 13A) is observed for the hood subdomain and flexible loop, respectively. In addition, *Ec*APRT$_{cl}$ places the Glu103 residue at an appropriate distance to carry out the nucleophilic attack on the N-7 hydrogen of the adenine (Figure 14B), supporting the previously described catalytic role of the closed conformation. Moreover, this conformation leads the residues of the "hood subdomain" (Leu27 and Arg29, through

their backbone) to interact with the adenine (Figure 13B), keeping it stable in the active center, since they have shorter bond distances than in the open conformation.

A



B

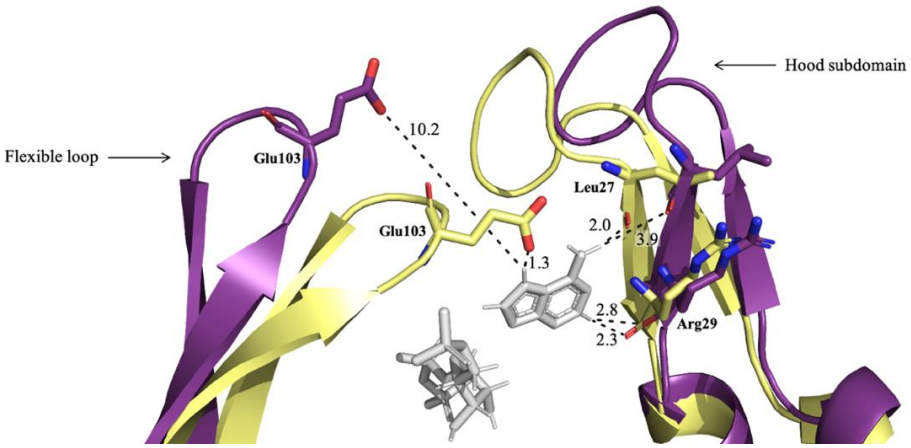


Figure 13. *Comparison of EcAPRT$_{op}$ (PDB id: 2DY0, purple) and EcAPRT$_{cl}$ (frame from the MD simulations, yellow), shown in cartoon and ligands (PRPP and adenine) in white sticks.* ***(A)*** *Measured distances between EcAPRT$_{op}$ (purple) and EcAPRT$_{cl}$ (yellow) conformations.* ***(B)*** *Measured distances comparison of catalytic residues from EcAPRT$_{op}$ (purple) and cAPRT$_{cl}$ (yellow) conformations. Measured distances are in Å. Created on PyMOL (35).*

To further validate the constructed open and closed *Ec*APRT models a Ramachandran plot (Figure 16) was made. On the one hand, *Ec*APRT$_{op}$ (Figure S 2A), has 95.4% of favored regions and 4% of residues in additional allowed regions. On the other hand, *Ec*APRT$_{cl}$ (Figure S 2B), has 94.1% of favored regions and 4.6% of residues in additional allowed regions. Therefore, by checking that the structure complies with the Ramachandran plot, the *Ec*APRT$_{cl}$ structure can be validated.

Regarding its quaternary structure, a protein-protein docking has been done in order to verify its oligomerization state. As already said, no monomeric APRTases has been reported so allostery is essential for its catalysis. Among all the positions explored, the monomer consistently occupies a particular location (Figure 14) with significantly higher

frequency. This finding strongly suggests the existence of a unique dimerization site, with a probability exceeding 90%.

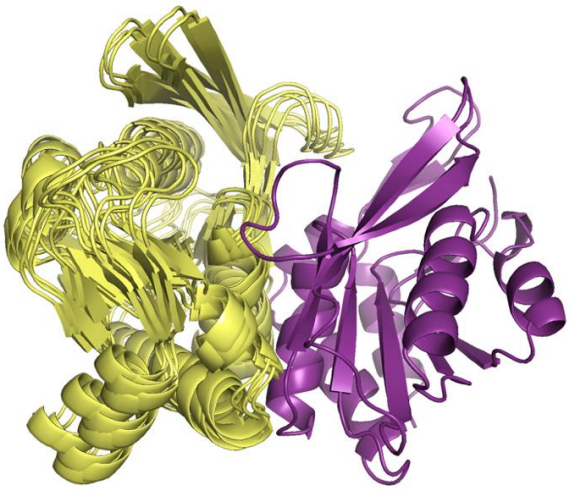


Figure 14. *Oligomeric state of EcAPRT$_{op}$ obtained by protein-protein docking. Best poses are shown in yellow cartoon and the receptor is shown as purple cartoon. Created on PyMOL (35).*

4.3. Cloning, Expression and Protein Purification

As previously mentioned, after protein overexpression, *E. coli* BL21 recombinant cells were collected and further submitted to cell disruption. After removing cell debris by centrifugation, the cleared lysate was introduced into the HisTrap column for affinity chromatography. The His-tag present in the N-terminal of the recombinant protein has high affinity for the Ni^{2+} groups of the column. The whole purification process is followed by UV spectroscopy at 280 nm. The first signal belongs to non-affinity proteins (flowthrough) and then, only when the signal returns to the baseline, the elution phase begins, and the recombinant protein is harvested (Figure S 3). These final fractions, corresponding to the elution of the protein of interest, were identified by SDS-PAGE (data not shown) and brought to a concentration of 1.55 mg/mL.

To further study the purification process, all the steps were analyzed. Based on the SDS-PAGE analysis presented (Figure 15), affinity chromatography brought to a purer *Ec*APRT sample, observing a single protein band corresponding to *Ec*APRT with an estimated molecular mass of 22 kDa, which is in agreement with the molecular mass calculated from the amino acid sequence 21.89 kDa.

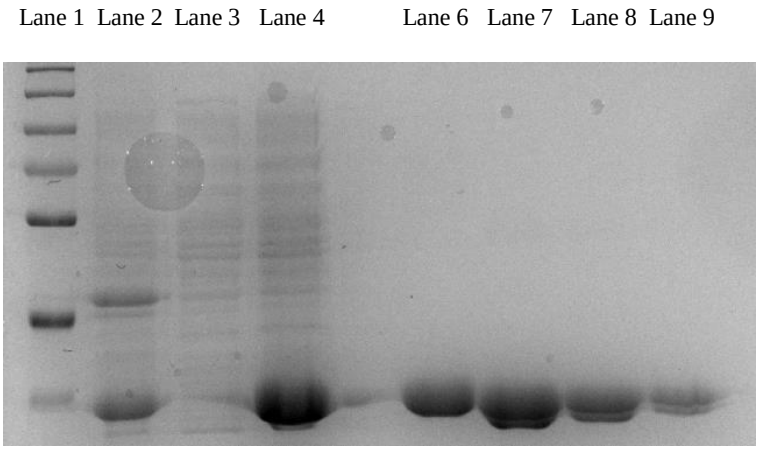


***Figure 15**. SDS-PAGE analysis of the affinity purification. Lane 1: pre-stained protein molecular weight ladder. Lane 2: EcAPRT cell culture pellet. Lane 3: non-affinity proteins (affinity chromatography flowthrough). Lane 4: cleared lysate. Lane 6: purified EcAPRT. Lanes 7-9 represent the standard curve, lane 7: EcHPRT 2.32 mg/mL, lane 8: EcHPRT 1.16 mg/mL and lane 9: EcHPRT 0.58 mg/mL.*

4.4. Enzyme activity evaluation

Different amount of enzyme and reaction times have been assessed (Figure 16) to determine the reaction conditions under which *Ec*APRT has the highest activity but does not reach the enzymatic saturation.

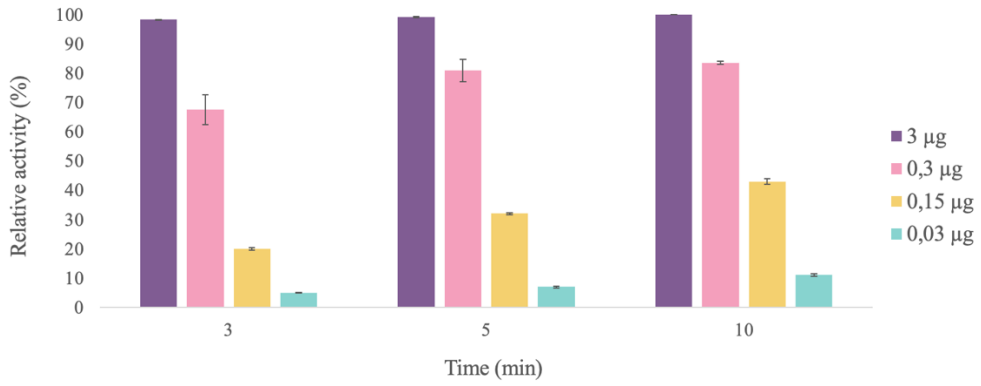


***Figure 16.** Comparative of relative conversion values obtained with different enzyme concentrations and reaction times. Conversion rates are relativized to the highest value obtained with 3 µg of enzyme and 10 minutes reaction time.*

When compared the relative conversion of 3 µg of enzyme at different reaction times, very little difference is observed between the values obtained, therefore, it is estimated that the enzyme saturation is reached quickly. For this reason, the assay was repeated with a smaller amount of enzyme. Moreover, with 0.3 µg and 0.15 µg of enzyme the substrate conversion is still high at a 3-minute reaction time (more than 60% of relative conversion

for 0.3 µg). After testing different conditions, with 0.03 µg of enzyme at a 3-minute reaction time the conversion was not enough to reach the enzyme saturation. In this condition, *Ec*APRT displayed a specific activity value of 611.69 IU, very similar to the reported by Loderer and coworkers (53), 586 IU.

4.5. Biochemical characterization

The effects of temperature and pH on *Ec*APRT activity were studied and compared to those results previously reported.

In terms of temperature (Figure 17), *Ec*APRT displayed more than 50% relative activity in a 50-80 ºC temperature range, with the maximum activity at 70 ºC. A similar trend was already observed by Loderer and coworkers (53), where the activity increases linearly in the 20-50 ºC temperature range. As expected, a remarkable drop is observed in the 70-90 ºC range, showing residual enzymatic activity at 90 ºC (less than 10% relative activity). Similar results have been reported for APRTases from different mesophilic organisms such as APRT from *Arabidopsis thaliana* (73) or *Brassica juncea* (74), with optimum reaction temperatures at 65 ºC and a rapid activity drop from 70 ºC onwards.

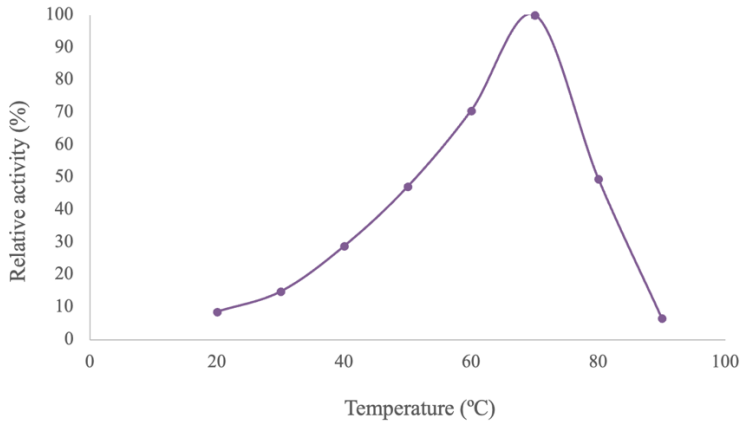


Figure 17. *Effect of temperature (20 – 90 ºC) on EcAPRT activity. Activity values are relativized to the highest activity reported at 70 ºC.*

With respect to the pH profile (Figure 18), *Ec*APRT displayed more than 60% relative activity in a 6-9 pH range, with the maximum activity at pH 8.0, as previously reported for *E. coli* (53,75,76) and other species such as *Triticum aestivum* (77) and *Arabidopsis thaliana* (73). It should be noted that at low pH (pH 4-5) barely any activity is detected, while at basic pH (pH 9-11) ~50% relative activity is observed. In addition, it is important to note that the O2B oxygen of the PPi (from the PRPP) remains deprotonated according

to the reaction mechanism proposed in the literature. Therefore, taking into account its pK_a (7.24), pH lower than this will disfavor catalysis as there will be less catalytically active PRPP.

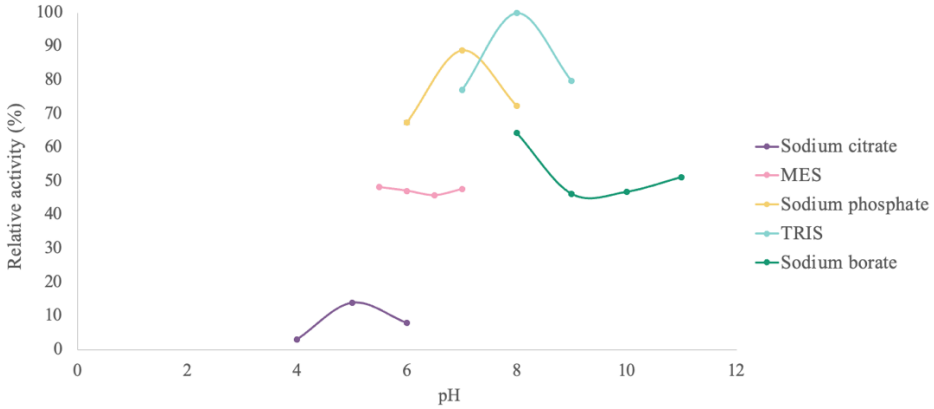


Figure 18. *Effect of pH on EcAPRT activity. Sodium citrate 50 mM (pH 4-6, dark purple), MES 50 mM (pH 5.5-7, pink), sodium phosphate 50 mM (pH 6-8, yellow), Tris-HCl 50mM (pH 7-9, light blue) and sodium borate (pH 8-11, dark green). Activity values are relativized to the highest acivity reported with Tris-HCl pH 8.0.*

Molecular dynamics titration of the amino acids of the protein was performed by molecular dynamics (Figure 19) in order to shed a light on *Ec*APRT activity at different pH conditions. This study was carried out at different simulation times (1-60 ns). It is noteworthy that from 1 to 8 ns the pK_a is observed to have large variations while from 8 to 60 ns the pK_a of the amino acids remains relatively constant.

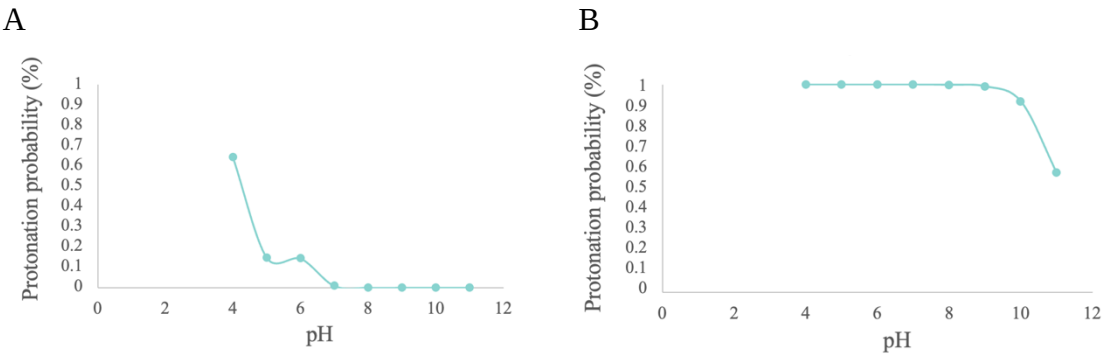


Figure 19*. Amino acid simulated protonation profile.* ***(A)*** *Simulated protonation profile of Glu103 residue.* ***(B)*** *Simulated protonation profile of Tyr104 residue.*

The Glu103 residue, as mentioned above, has a fundamental role in the reaction mechanism. Observing its protonation probability in an open conformation (Figure 19A), at low pH values, it is very likely to be protonated, which would explain the experimental results (low activity at pH less than 7.0), whereas at pH greater than 8.0, the residue maintains its deprotonation, which could explain higher relative activity at basic pH

conditions. Moreover, as already said, the Tyr104 residue is essential due to the hydrogen bond formed with the Arg66 (7). Therefore (Figure 19B), at pH above 10.0, the protonation probability for OH of Tyr104 is lower, thus hindering the interaction with Arg66.

4.6. Thermostability

The operational stability of *Ec*APRT was tested at 40-70 ºC temperature range for 120 minutes (Figure 20).

At 40 and 50 ºC the enzymatic activity drops a 30% of the initial activity on the first 10 minutes, then the enzyme mantains a 70% of relative activity during the whole incubation period. On the contrary, a total loss of activity is observed after the first 5 minutes of incubation at 60 and 70 ºC. A similar tendency has been previously reported by Srivastava and Beutler (78), who observed a severe loss of activity at 60 ºC, while it slightly decreases at 50 and 40 ºC (APRT isolated from human erythrocytes). However, crude extracts of *E. coli* has been found to be reasonably stable at 60 ºC (5 minutes of incubation) when the substrate is added for stabilizing the enzyme during heat treatment (76).

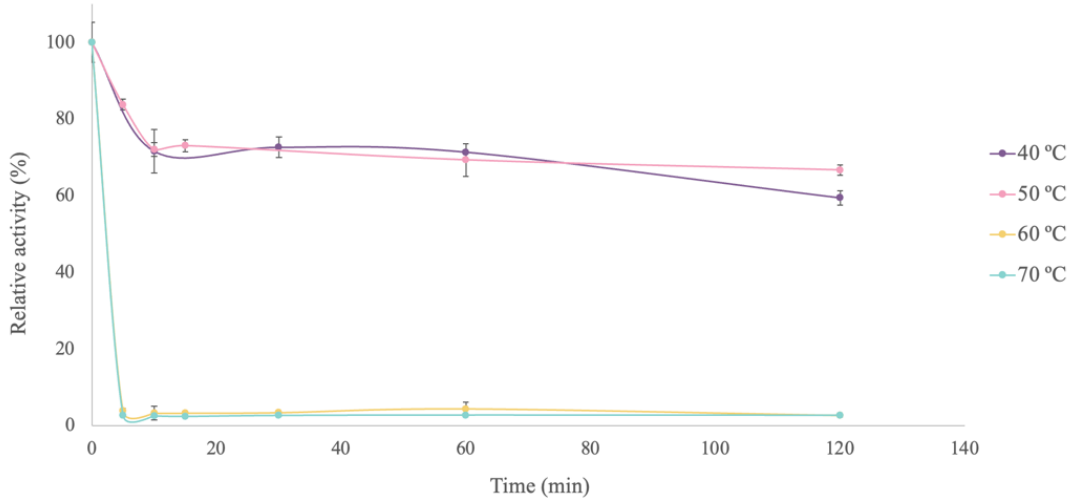


Figure 20. *Thermal stability of EcAPRT from 5 to 120 minutes in the 40-70 ºC temperature range. Activities are relativized to the highest acivity reported from the standard activity assay.*

Despite the observed optimum temperature for *Ec*APRT is 70 ºC, the enzyme displays very low stability at this condition. Therefore, 50 ºC was selected as the reaction temperature for further assays.

4.7. Influence of divalent cations on *Ec*APRT activity

Once the optimal reaction conditions were established (pH 8.0 and 50 ºC), the effect of divalent cations on enzyme activity was studied. As mentioned above, it is reported that the presence of metals is essential for the reaction to occur as it is associated with PRPP.

The maximum activity for APRTase of bacteria (*E. coli, Francisella tularensis)* (75,79), mamals (*Homo sapiens, Mus musculus)* (75,80) and plants (*Brassica juncea, Triticum aestivum)* (74,77) have been documented with the use of Mg^{2+} as an essential cofactor. As expected, the same happened in this study (Figure 21), registering its maximum activity at a concentration of 12 mM Mg^{2+}. For *Schizosaccharomyces pombe* (81), the same tendency is observed. High enzymatic activity was also observed with 5 mM Mn^{2+} (> 80% relative activity), as also described for APRT from *Triticum aestivum* (77). However, the use of other divalent cations causes severe drop on the enzymatic activity. On the one hand, Zn^{2+} generates a moderate inhibition, since at a concentration of 2 mM, 60% of relative activity is observed. On the other hand, cations such as Co^{2+} or Ca^{2+} have greater inhibitory effect, for example, Ca^{2+} at a concentration of 2mM reports approximately 5% relative activity. Inhibition by these cations has also been reported previously in APRT from other species such as *Plasmodium falciparum* (82), *Artemia* (83) or *Homo sapiens* (75).

It is important to note that for the second period metals (Mg^{2+} and Ca^{2+}) the activity increases along with their concentration. The other way around occurs with the transition metals (Mn^{2+}, Zn^{2+} and Co^{2+}), where the activity decreases when increasing the cofactor concentration.

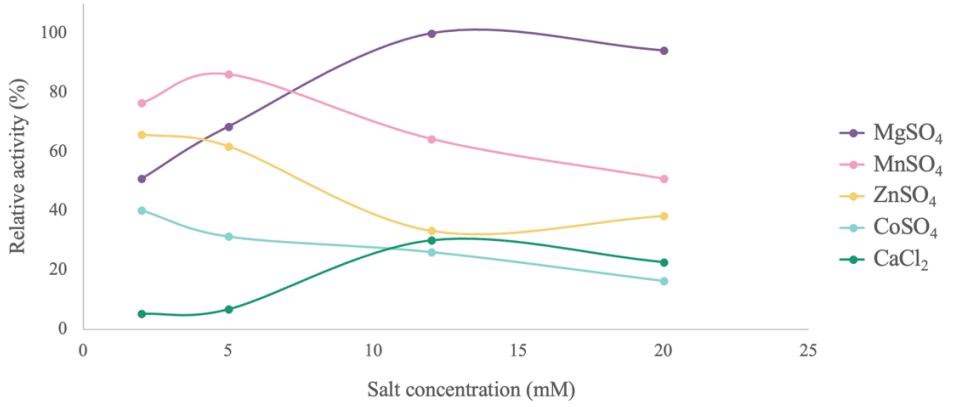


***Figure 21**. Influence of divalent cations on EcAPRT activity. Activity values are relativized to the highest reported with 12 mM Mg^{2+}.*

4.8. Effect of organic solvents on *Ec*APRT activity

The alteration of enzyme activity in the presence of 20% (v/v) organic solvents was studied employing the previously defined optimal conditions (pH 8.0, 50 ºC and 12 mM Mg^{2+}) (Figure 22). Polar protic solvents (methanol, ethanol, isopropanol, propylene glycol, glycerol and tertbutanol) and alcohols and polyols (acetonitrile, chloroform, DMF, DMSO, acetone and ethyl acetate) were tested.

On the one hand, regarding alcohols and polyalcohols (Figure 22A), a 30-40% activity drop is observed in most of them when comparing the activity values obtained in the aqueous buffer. However, the lowest activity is recorded with polyols, being the lowest activity with glycerol, which contains the most hydroxyl groups. Furthermore, considering the isoelectric constant, the glycerol reports the highest (ε = 47) while the lowest is recorded for ethanol (ε = 25) and isopropanol (ε = 20).

On the other hand, in the presence of polar protic solvents (Figure 22B), the maximum activities reported are with acetonitrile and ethyl acetate (20% and 6% higher than that recorded with water, respectively). Both with chloroform and DMF, the same relative activity is recorded (72% relative activity). A lower activity is reported in the presence of acetone while with DMSO the relative activity decreases by 50% regarding to that obtained in the aqueous buffer. Considering the log P of the solvents, the lowest activities are recorded with those with log P < 0 (hydrophilic), while the highest activities are recorded with those with log P > 0 (hydrophobic). However, it should be noted that acetonitrile has a log P = - 0.3 and is the one with which the highest activity has been obtained.

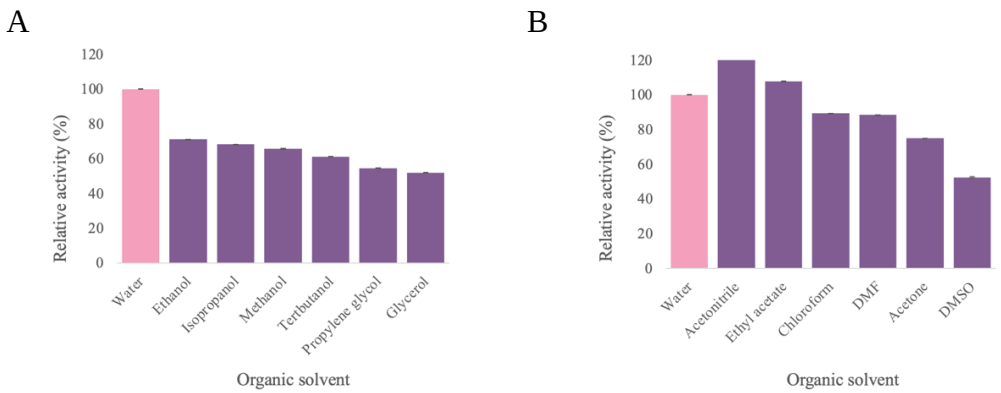


Figure 22. *Influence of organic solvents on EcAPRT activity.* ***(A)*** *Alcohols and polyalcohols.* ***(B)*** *Aprotic polar solvent. DMF: N,N-dimethylformamide, DMSO: dimethyl sulfoxide. Activity values are relativized to that observed with the aqueous buffer.*

4.9. Enzymatic activity with non-natural bases

With an eye towards including *Ec*APRT in the multienzymatic cascade for the generation nucleoside analogs, enzymatic activity with non-natural bases has been studied and compared with previous studies. In this sense, different non-natural bases have been assessed to determine the specificity of *Ec*APRT in order to design the multienzyme cascade.

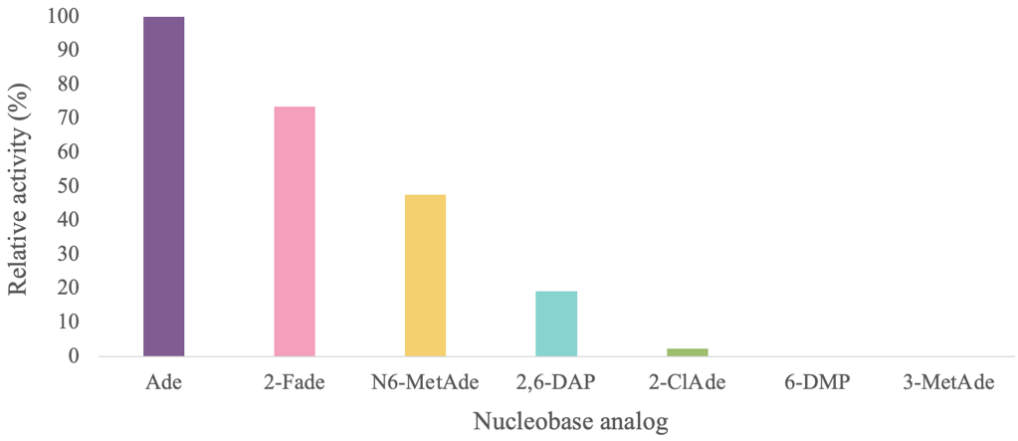


Figure 23. *EcAPRT activity with non-natural bases. Specific activity values are relativized to the highest values obtained with Adenine.*

*Ec*APRT has strong preference for 2-fluoroadenine, but it also shows slightly activity for 2,6-diaminopurine and no activity towards 6-dimethylpurine and 3-methyladenine (Figure 23). According to (84), high activity was also obtained with 2-fluoroadenine, and authors consider it to be comparable to the activity obtained with adenine. Some activity was also detected with 2-chloroadenine. In previous studies (53), the amount of enzyme was adjusted 100X and much longer reaction times (compared to the initial conditions in which adenine was used as substrate) were used to obtain a good substrate conversion, as just residual activity was detected with this substrate. Additionally, in this study approximately 50% of relative activity is reached with N6-methyladenine.

Bioinformatically, *Ec*APRT$_{cl}$ in complex with the three natural substrates (adenine, PRPP and Mg^{2+}) was constructed (referred to as *Ec*APRT$_{cl}$-Ade). After molecular dynamics simulations, the stability of *Ec*APRT-Ade was analyzed by studying the RMSD (Figure 24). It's observed that the complex is stable from 50 ns up to 300 ns, enough time to evaluate the molecular movements.

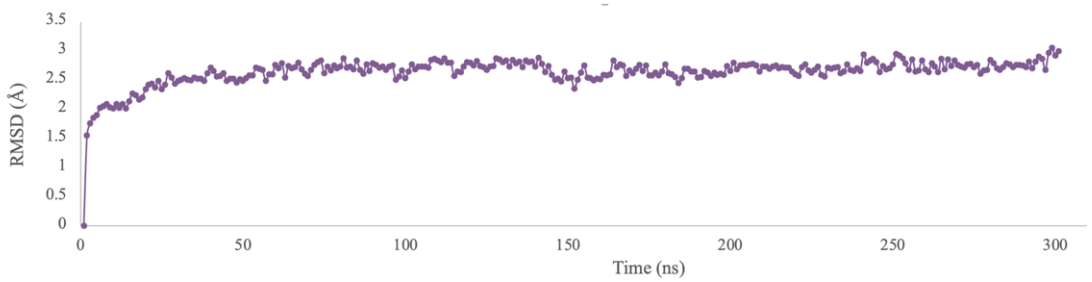


Figure 24. *RMSD of EcAPRT$_{cl}$-Ade.*

The *Ec*APRT$_{cl}$-Ade stability was also analyzed by measuring the distances between OE1 and OE2 of the catalytic Glu103 to the N7 hydrogen of adenine. Additionally, the distances between the N9 of adenine to the C1 of the PRPP (Figure 25) were also monitorized. Firstly, the physical rotation of Glu103 can be observed. In this regard, both OE1 and OE2 of Glu103 are positioned alternately at 4 Å and 1.7 Å (approximately) (Figure 26), being this last distance the optimal to perform the nucleophilic attack on the N7 hydrogen of adenine. The distance between N9 of adenine to C1 of PRPP remains constant at 3.4 Å.

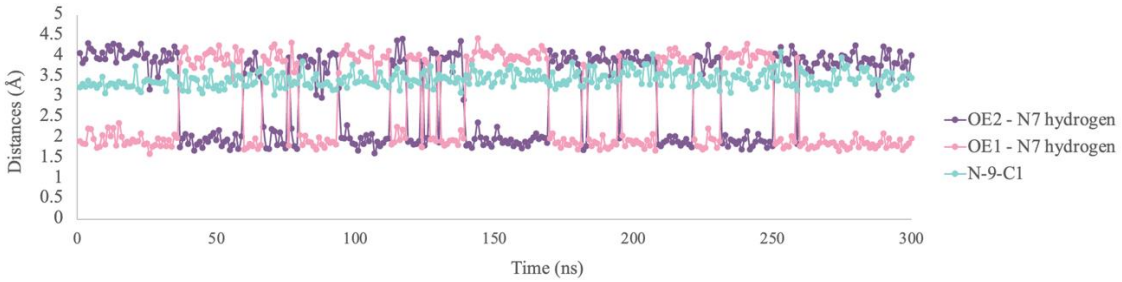
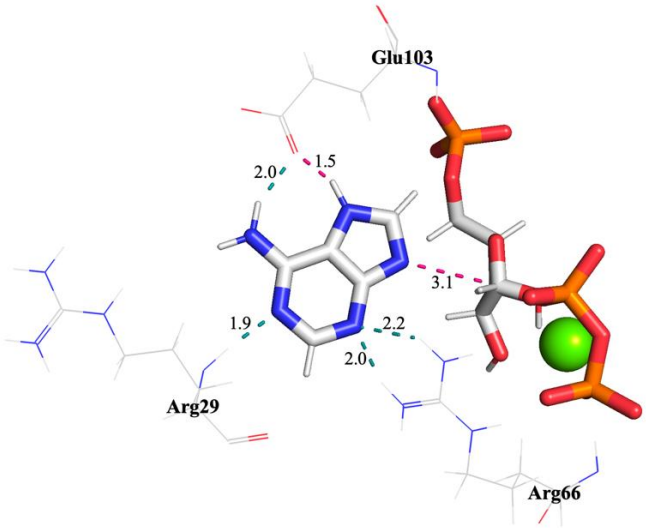


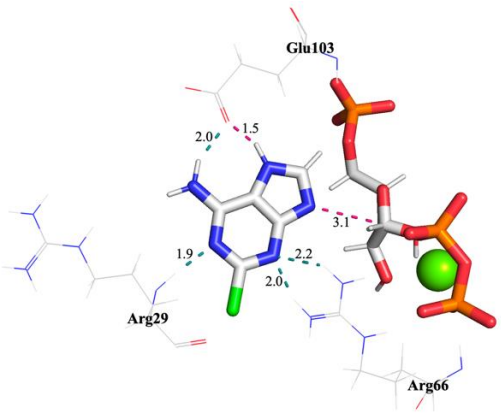
Figure 25*. Measured distances during a MD of EcAPRT$_{cl}$-Ade over time. Glu103 (OE1, pink and OE2, dark purple) to the hydrogen N7 (from adenine) and from N9 (from adenine, light blue) to C1 (from PRPP).*

Achieving a stable complex with the natural substrates (Figure 26A) is important to form complexes with experimentally tested nucleobase analogs. By manual docking, nucleobase analogs were positioned in the active center and they were analyzed.

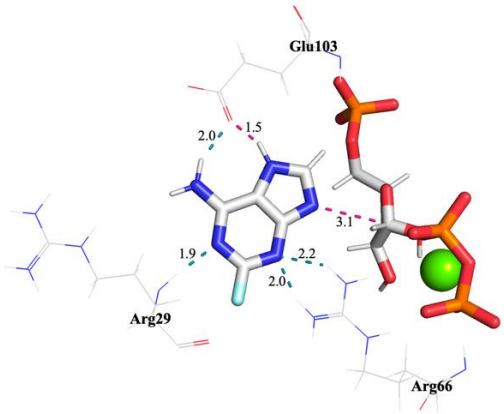
A


Glu103
2.0
1.5
3.1
1.9
2.2
2.0
Arg29
Arg66

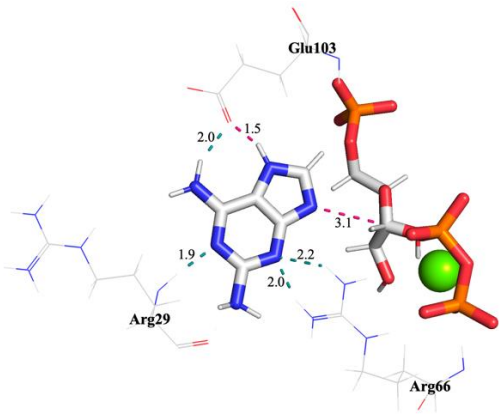
B


Glu103
2.0
1.5
3.1
1.9
2.2
2.0
Arg29
Arg66

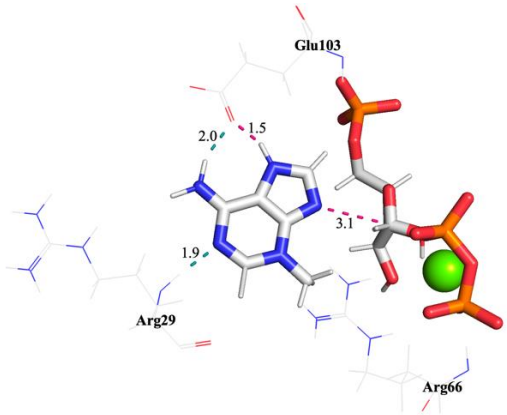
C


Glu103
2.0
1.5
3.1
1.9
2.2
2.0
Arg29
Arg66

D


Glu103
2.0
1.5
3.1
1.9
2.2
2.0
Arg29
Arg66

E


Glu103
2.0
1.5
3.1
1.9
Arg29
Arg66

F G

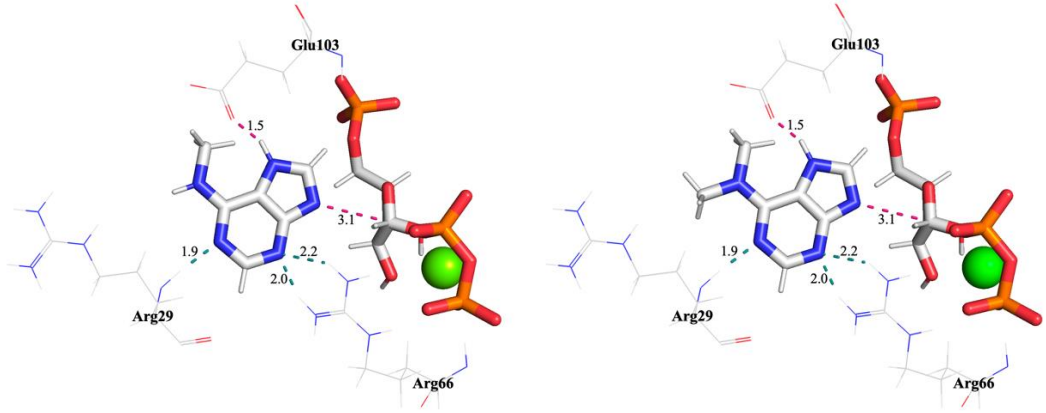


Figure 26. *Adenine and nucleobase analogs positioned in the active center.* ***(A)*** *Adenine is shown in white sticks.* ***(B)*** *2-Chloroadenine is shown in white sticks.* ***(C)*** *2-Fluoroadenine is shown in white sticks.* ***(D)*** *2,6-Diaminopurine is shown in white sticks.* ***(E)*** *3-Methyladenine is shown in white sticks.* ***(F)*** *N6-methyladenine is shown in white sticks.* ***(G)*** *6-(dimethylamino)purine is shown in white sticks. $EcAPRT_{cl}$ residues (Glu103, Arg29 and Arg66) are shown in white lines, PRPP in white sticks and Mg^{2+} in a green sphere. The distance of the hydrogen bonds are shown in green, the distances of the reaction mechanism are in pink. Measured distances are in Å. Created on PyMOL (35).*

Although it may initially appear that the presence of substituents at the N2 position has no impact on ligand recognition (Figure 26B, 26C, and 26D), it should be noted that these substituents may introduce some steric hindrance. This hindrance arises from the increasing atomic radius of the substituents, which follows the order: H < F < N < Cl, corresponding to the increasing periods of the periodic table. The experimental data also confirms this trend of diminishing activity (Figure 23). With respect to substituents at the N3 position (Figure 26E), the loss of two hydrogen bonds compromises the stability of the substrate within the active center. This observation aligns with the experimental findings of no activity, confirming the correlation between the absence of activity and the destabilization caused by the missing hydrogen bonds. Lastly, when it comes to substituents at the N6 position (Figure 26F and 26G), the presence of such substituents hinders substrate recognition due to the loss of the hydrogen bond with Glu103. Furthermore, in the case of 6-(dimethylamino)purine, no activity is detected as the presence of an additional methyl group causes steric hindrance with Leu154.

5. Conclusion

Based on the findings of this study, the following conclusions may be drawn:

1. *Ec*APRT shows its maximal activity at 70 ºC, Tris-HCl pH 8.0 and Mg^{2+} 12 mM.
2. Due its low thermal stability at 70 ºC, 50 ºC was chosen as optimal temperature for enzymatic reactions
3. In presence of acetonitrile and ethyl acetate, no loss of activity is detected.
4. Comparing to its bacterial counterparts *Ec*APRT shows higher phosphoribosyltransferase activity on adenine.
5. When tested on modified nucleobases, *Ec*APRT shows remarkable specificity for 2-fluoroadeadenine. Lower activity was detected with N6-methyladenine, 2,6-diaminopurine and 2-fluoroadenine.
6. A Michaelis-Menten complex of *Ec*APRT in complex with adenine, PRPP and Mg^{2+} has been proposed by means of molecular modeling techniques. It has allowed us to thoroughly analyze the substrate specificity of *Ec*APRT.
7. *Ec*APRT is a good candidate for the multienzymatic cascade for the production of nucleoside analogs.

6. Bibliography

1. Lehninger AL, Nelson DL, Cox MM. Principios de bioquímica. 5{487} ed. Barcelona: Omega; 2009.
2. Ferr S, FuxE K. Adenosine-dopamine interactions in the brain.
3. Eltzschig HK, Weissmüller T, Mager A, Eckle T. Nucleotide Metabolism and Cell-Cell Interactions. In: Cell-Cell Interactions in Health and Disease [Internet]. New Jersey: Humana Press; 2006 [cited 2023 May 22]. p. 73–88. Available from: http://link.springer.com/10.1385/1-59745-113-4:73
4. ChemSketch [Internet]. ACD/Labs. Available from: www.acdlabs.com
5. Carver JD, Allan Walker W. The role of nucleotides in human nutrition. J Nutr Biochem. 1995 Feb;6(2):58–72.
6. Kaminski PA. Functional Cloning, Heterologous Expression, and Purification of Two Different N-Deoxyribosyltransferases fromLactobacillus helveticus. J Biol Chem. 2002 Apr;277(17):14400–7.
7. Huyet J, Ozeir M, Burgevin MC, Pinson B, Chesney F, Remy JM, et al. Structural Insights into the Forward and Reverse Enzymatic Reactions in Human Adenine Phosphoribosyltransferase. Cell Chem Biol. 2018 Jun;25(6):666-676.e4.
8. Weber G, Jayaram HN, Pillwein K, Natsumeda Y, Reardon MA, Zhen YS. Salvage pathways as targets of chemotherapy. Adv Enzyme Regul. 1987 Jan;26:335–52.
9. Nyhan WL. Nucleotide Synthesis via Salvage Pathway.
10. Pérez-Aguilar MC, Rondón-Mercado R. Adenosin deaminasa en la tripanosomiasis experimental: futuras implicaciones. Investig Clínica. 2015 Sep;56(3):309–19.
11. Walther T, Novo M, Rössger K, Létisse F, Loret M, Portais J, et al. Control of ATP homeostasis during the respiro-fermentative transition in yeast. Mol Syst Biol. 2010 Jan;6(1):344.
12. Nygaard P. Purine and Pyrimidine Salvage Pathways. In: Sonenshein AL, Hoch JA, Losick R, editors. Bacillus subtilis and Other Gram-Positive Bacteria [Internet]. Washington, DC, USA: ASM Press; 2014 [cited 2023 May 22]. p. 359–78. Available from: http://doi.wiley.com/10.1128/9781555818388.ch26
13. Jinnah HA, Ceballos-Picot I, Torres RJ, Visser JE, Schretlen DJ, Verdu A, et al. Attenuated variants of Lesch-Nyhan disease. Brain. 2010 Mar 1;133(3):671–89.
14. Del Arco J, Acosta J, Fernández-Lucas J. New trends in the biocatalytic production of nucleosidic active pharmaceutical ingredients using 2′-deoxyribosyltransferases. Biotechnol Adv. 2021 Nov;51:107701.
15. Lawton P. Purine analogues as antiparasitic agents. Expert Opin Ther Pat. 2005 Aug;15(8):987–94.

16. De Clercq E. Recent highlights in the development of new antiviral drugs. Curr Opin Microbiol. 2005 Oct;8(5):552–60.
17. Mahmoud S, Hasabelnaby S, Hammad S, Sakr T. Antiviral Nucleoside and Nucleotide Analogs: A Review. J Adv Pharm Res. 2018 Apr 1;2(2):73–88.
18. Thomson JM, Lamont IL. Nucleoside Analogues as Antibacterial Agents. Front Microbiol. 2019 May 22;10:952.
19. Galmarini CM, Mackey JR, Dumontet C. Nucleoside analogues and nucleobases in cancer treatment. Lancet Oncol. 2002 Jul;3(7):415–24.
20. Geraghty R, Aliota M, Bonnac L. Broad-Spectrum Antiviral Strategies and Nucleoside Analogues. Viruses. 2021 Apr 13;13(4):667.
21. Ewald B, Sampath D, Plunkett W. Nucleoside analogs: molecular mechanisms signaling cell death. Oncogene. 2008 Oct 27;27(50):6522–37.
22. Yoshikawa M, Kato T, Takenishi T. Studies of Phosphorilation. III. Selective Phosphorylation of Unprotected Nucleosides. 42(12).
23. Ichikawa E, Kato K. Sugar-Modified Nucleosides in Past 10 Years, A Review. Curr Med Chem. 2001 Mar 1;8(4):385–423.
24. Lapponi MJ, Rivero CW, Zinni MA, Britos CN, Trelles JA. New developments in nucleoside analogues biosynthesis: A review. J Mol Catal B Enzym. 2016 Nov;133:218–33.
25. Arco JD, Pérez E, Naitow H, Matsuura Y, Kunishima N, Fernández-Lucas J. Structural and functional characterization of thermostable biocatalysts for the synthesis of 6-aminopurine nucleoside-5′-monophospate analogues. Bioresour Technol. 2019 Mar;276:244–52.
26. Harris LD, Harijan RK, Ducati RG, Evans GB, Hirsch BM, Schramm VL. Synthesis of *bis* -Phosphate Iminoaltritol Enantiomers and Structural Characterization with Adenine Phosphoribosyltransferase. ACS Chem Biol. 2018 Jan 19;13(1):152–60.
27. Musick WmDL, Nyhan WL. Structural Features of the Phosphoribosyl-Transferases and Their Relationship to the Human Deficiency Disorders of Purine and Pyrimidine Metabolis. Crit Rev Biochem. 1981 Jan;11(1):1–34.
28. Jensen KF, Hansen MR, Jensen KS, Christoffersen S, Poulsen JCN, Mølgaard A, et al. Adenine Phosphoribosyltransferase from *Sulfolobus solfataricus* Is an Enzyme with Unusual Kinetic Properties and a Crystal Structure that Suggests It Evolved from a 6-Oxopurine Phosphoribosyltransferase. Biochemistry. 2015 Apr 14;54(14):2323–34.
29. Sinha S. The PRT protein family. Curr Opin Struct Biol. 2001 Dec 1;11(6):733–9.
30. Shi W, Sarver AE, Wang CC, Tanaka KSE, Almo SC, Schramm VL. Closed Site Complexes of Adenine Phosphoribosyltransferase fromGiardia lamblia Reveal a Mechanism of Ribosyl Migration. J Biol Chem. 2002 Oct;277(42):39981–8.
31. Del Arco J, Martinez M, Donday M, Clemente-Suarez VJ, Fernández-Lucas J. Cloning,

expression and biochemical characterization of xanthine and adenine phosphoribosyltransferases from *Thermus thermophilus* HB8. Biocatal Biotransformation. 2018 May 4;36(3):216–23.

32. Craig SP, Eakin AE. Purine Phosphoribosyltransferases. J Biol Chem. 2000 Jul;275(27):20231–4.
33. Silva M, Silva CHTDP, Iulek J, Thiemann OH. Three-Dimensional Structure of Human Adenine Phosphoribosyltransferase and Its Relation to DHA-Urolithiasis ’. Biochemistry. 2004 Jun 1;43(24):7663–71.
34. Phillips CL, Ullman B, Brennan RG, Hill CP. Crystal structures of adenine phosphoribosyltransferase from Leishmania donovani.
35. Schrödinger L, DeLano W. PyMOL | pymol.org [Internet]. 2020. Available from: http://www.pymol.org/pymol
36. Arco JD, Fernandez-Lucas J. Purine and Pyrimidine Phosphoribosyltransferases: A Versatile Tool for Enzymatic Synthesis of Nucleoside-5’-monophosphates. Curr Pharm Des. 2018 Feb 16;23(45):6898–912.
37. Shi W, Tanaka KSE, Crother TR, Taylor MW, Almo SC, Schramm VL. Structural Analysis of Adenine Phosphoribosyltransferase from *Saccharomyces cerevisiae* ’. Biochemistry. 2001 Sep 1;40(36):10800–9.
38. Silva CHTP, Silva M, Iulek J, Thiemann OH. Structural Complexes of Human Adenine Phosphoribosyltransferase Reveal Novel Features of the APRT Catalytic Mechanism. J Biomol Struct Dyn. 2008 Jun;25(6):589–97.
39. Vanommeslaeghe K, Guvench O, D. MacKerell Jr. A. Molecular Mechanics. Curr Pharm Des. 2014 May 1;20(20):3281–92.
40. Schwabl F. Quantum Mechanics. Springer Science & Business Media; 2007. 425 p.
41. Gao J, Truhlar DG. Quantum Mechanical Methods for Enzyme Kinetics. Annu Rev Phys Chem. 2002;53(1):467–505.
42. Hansson T, Oostenbrink C, van Gunsteren WF. Molecular dynamics simulations.
43. Rowlinson JS. The Maxwell–Boltzmann distribution. Mol Phys. 2005 Nov 10;103(21–23):2821–8.
44. Cárdenas Espinosa LP. Dinámica molecular como técnica de simulación. Rev Habitus Semilleros Investig. 2009 Jan 6;(1):29–32.
45. Leach A. Molecular modelling. Principles and applications. Inglaterra: Pearson Educational Limited; 2001.
46. Schlick T. Molecular Modeling and Simulation: An Interdisciplinary Guide: An Interdisciplinary Guide [Internet]. New York, NY: Springer New York; 2010 [cited 2023 May 23]. (Interdisciplinary Applied Mathematics; vol. 21). Available from: https://link.springer.com/10.1007/978-1-4419-6351-2

47. Barnoud J, Monticelli L. Coarse-Grained Force Fields for Molecular Simulations. In: Kukol A, editor. Molecular Modeling of Proteins [Internet]. New York, NY: Springer New York; 2015 [cited 2023 May 22]. p. 125–49. (Methods in Molecular Biology; vol. 1215). Available from: http://link.springer.com/10.1007/978-1-4939-1465-4_7
48. Kmiecik S, Gront D, Kolinski M, Wieteska L, Dawid AE, Kolinski A. Coarse-Grained Protein Models and Their Applications. Chem Rev. 2016 Jul 27;116(14):7898–936.
49. Singh N, Li W. Recent Advances in Coarse-Grained Models for Biomolecules and Their Applications. Int J Mol Sci. 2019 Aug 1;20(15):3774.
50. Cruz G, Saiz LP, Bilal M, Eltoukhy L, Loderer C, Fernández-Lucas J. Magnetic Multi-Enzymatic System for Cladribine Manufacturing. Int J Mol Sci. 2022 Nov 7;23(21):13634.
51. Acosta J, Del Arco J, Martinez-Pascual S, Clemente-Suárez V, Fernández-Lucas J. One-Pot Multi-Enzymatic Production of Purine Derivatives with Application in Pharmaceutical and Food Industry. Catalysts. 2018 Jan 1;8(1):9.
52. Eltoukhy L, Loderer C. A Multi-enzyme Cascade for the Biosynthesis of AICA Ribonucleoside Di- and Triphosphate. ChemBioChem [Internet]. 2022 Feb 4 [cited 2023 May 22];23(3). Available from: https://onlinelibrary.wiley.com/doi/10.1002/cbic.202100596
53. Frisch J, Maršić T, Loderer C. A Novel One-Pot Enzyme Cascade for the Biosynthesis of Cladribine Triphosphate. Biomolecules. 2021 Feb 25;11(3):346.
54. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 Oct;215(3):403–10.
55. The UniProt Consortium, Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023 Jan 6;51(D1):D523–31.
56. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011 Jan;7(1):539.
57. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014 Jul 1;42(W1):W320–4.
58. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein–protein docking. Nat Protoc. 2017 Feb;12(2):255–78.
59. Case DA, Belfon K, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham TE, III, Cruzeiro VWD, Darden TA, Duke RE, Giambasu G, Gilson MK, Gohlke H, Goetz AW, Harris R, Izadi S, Izmailov SA, Kasavajhala K, Kovalenko A, Krasny R, Kurtzman T, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Man V, Merz KM, Miao Y, Mikhailovskii O, Monard G, Nguyen H, Onufriev A, Pan F, Pantano S, Qi R, Roe DR, Roitberg A, Sagui C, Schott-Verdugo S, Shen J, Simmerling CL, Skrynnikov NR, Smith J, Swails J, Walker

RC, Wang J, Wilson L, Wolf RM, Wu X, Xiong Y, Xue Y, York DM and Kollman PA. AMBER 2020, University of California, San Francisco. 2020.

60. Mongan J, Case DA, McCammon JA. Constant pH molecular dynamics in generalized Born implicit solvent. J Comput Chem. 2004 Dec;25(16):2038–48.
61. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J Chem Theory Comput. 2013 Jul 9;9(7):3084–95.
62. Tian C, Kasavajhala K, Belfon KAA, Raguette L, Huang H, Migues AN, et al. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. J Chem Theory Comput. 2020 Jan 14;16(1):528–52.
63. Machado MR, Pantano S. SIRAH tools: mapping, backmapping and visualization of coarse-grained models. Bioinformatics. 2016 May 15;32(10):1568–70.
64. Wang J, Wang W, Kollman PA, Case DA. Antechamber, An Accessory Software Package For Molecular Mechanical Calculations.
65. Meagher KL, Redman LT, Carlson HA. Development of polyphosphate parameters for use with the AMBER force field. J Comput Chem. 2003 Jul 15;24(9):1016–25.
66. Laemmli UK. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature. 1970 Aug;227(5259):680–5.
67. Gill SC, Von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem. 1989 Nov;182(2):319–26.
68. Walker JM, editor. The proteomics protocols handbook. Totowa, NJ: Humana Press; 2005. 988 p.
69. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014 May 1;30(9):1236–40.
70. Acosta J, Del Arco J, Del Pozo ML, Herrera-Tapias B, Clemente-Suárez VJ, Berenguer J, et al. Hypoxanthine-Guanine Phosphoribosyltransferase/adenylate Kinase From Zobellia galactanivorans: A Bifunctional Catalyst for the Synthesis of Nucleoside-5′-Mono-, Di- and Triphosphates. Front Bioeng Biotechnol. 2020 Jun 24;8:677.
71. Hansen MR, Jensen KS, Rasmussen MS, Christoffersen S, Kadziola A, Jensen KF. Specificities and pH profiles of adenine and hypoxanthine–guanine–xanthine phosphoribosyltransferases (nucleotide synthases) of the thermoacidophile archaeon Sulfolobus solfataricus. Extremophiles. 2014 Jan;18(1):179–87.
72. Teilum K, Olsen JG, Kragelund BB. Functional aspects of protein flexibility. Cell Mol Life Sci CMLS. 2009 Jul;66(14):2231–47.
73. Lee D, Moffatt BA. Purification and characterization of adenine phosphoribosyltransferase from Arabidopsis thaliana. Physiol Plant. 1993 Apr;87(4):483–92.
74. Moffatt BA, Somerville CR. Purification of adenine phosphoribosyltransferase from Brassica juncea. Arch Biochem Biophys. 1990 Dec;283(2):484–90.

75. Hoffee PA, Jones ME, editors. Purine and pyrimidine nucleotide metabolism. New York: Academic Press; 1978. 654 p. (Methods in enzymology).
76. Hochstadt-Ozer J, Stadtman ER. The Regulation of Purine Utilization in Bacteria. J Biol Chem. 1971 Sep;246(17):5294–303.
77. Chen CM, Melitz DK, Clough FW. Metabolism of cytokinin: Phosphoribosylation of cytokinin bases by adenine phosphoribosyltransferase from wheat germ. Arch Biochem Biophys. 1982 Apr;214(2):634–41.
78. Srivastava SK, Beutler E. Purification and kinetic studies of adenine phosphoribosyltransferase from human erythrocytes. Arch Biochem Biophys. 1971 Feb;142(2):426–34.
79. Pavithra GC, Ramagopal UA. Crystal structures of APRT from *Francisella tularensis* – an N–H···N hydrogen bond imparts adenine specificity in adenine phosporibosyltransferases. FEBS J. 2018 Jun;285(12):2306–18.
80. Okada G, Kaneko I, Koyama H. Purification and characterization of adenine phosphoribosyltransferase from mouse mammary carcinoma FM3A cells in culture. Biochim Biophys Acta BBA - Gen Subj. 1986 Nov;884(2):304–10.
81. Nagy M, Ribet AM. Purification and Comparative Study of Adenine and Guanine Phosphoribosyltransferases from Schizosaccharomyces pombe. Eur J Biochem. 1977 Jul;77(1):77–85.
82. Queen SA, Vander Jagt DL, Reyes P. Characterization of adenine phosphoribosyltransferase from the human malaria parasite, Plasmodium falciparum. Biochim Biophys Acta BBA - Protein Struct Mol Enzymol. 1989 Jul;996(3):160–5.
83. Montero C, Llorente P. *Artemia* purine phosphoribosyltransferases. Purification and characterization. Biochem J. 1991 Apr 15;275(2):327–34.
84. Gadd REA, Henderson JF. Studies of the binding of adenine to adenine phosphoribosyltransferase. Can J Biochem. 1970 Mar 1;48(3):295–301.
85. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr. 1993 Apr 1;26(2):283–91.

Acknowledgments

I greatly appreciate Jesús Fernández Lucas and my lab co-workers Jon del Arco Arrieta and Javier Acosta Bueno for their trust and guidance from day one as well as the Universidad Europea de Madrid. I also thank my family for their unconditional support.

Supplementary material

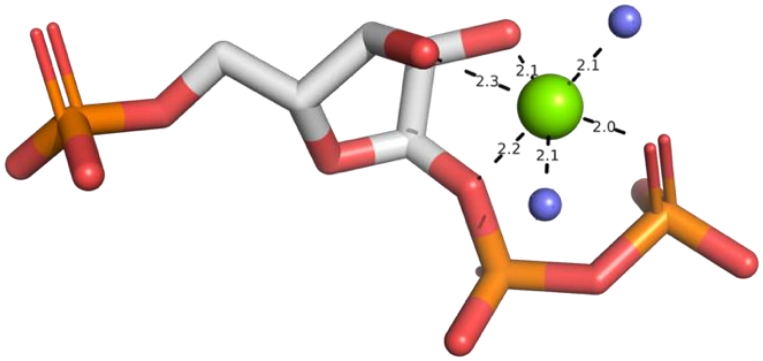
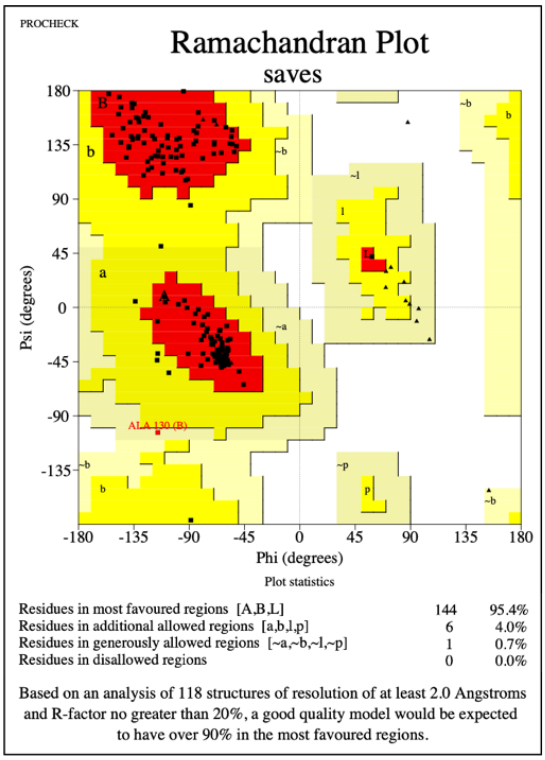


Figure S 1. *Coordination of the magnesium ion, forming an octahedral structure. Measured distances are in Å. PRPP is shown in sticks, H_2O is represented as blue spheres and Mg^{2+} as a green sphere. Created on PyMOL (35).*

A



B

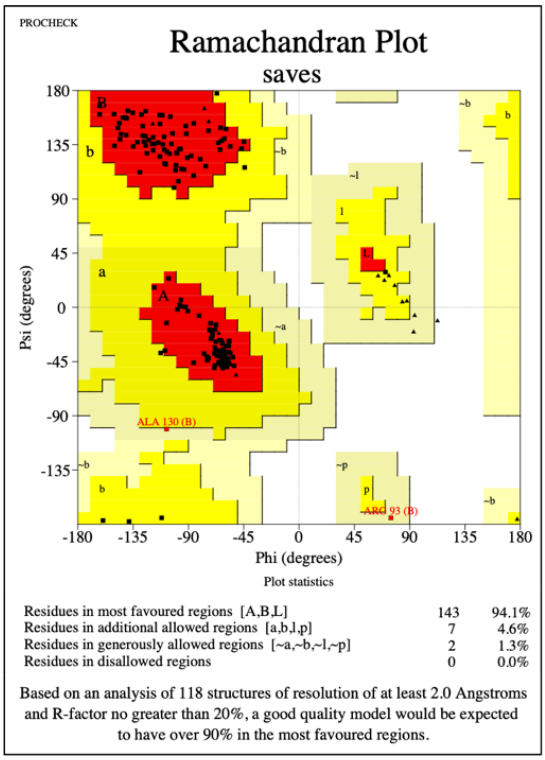


Figure S 2. *Ramachandran plot.* ***(A)*** *$EcAPRT_{op}$ Ramachandran plot.* ***(B)*** *$EcAPRT_{cl}$ Ramachandran plot. Glycines shown as triangles. Created with SAVES v6.0 (85).*

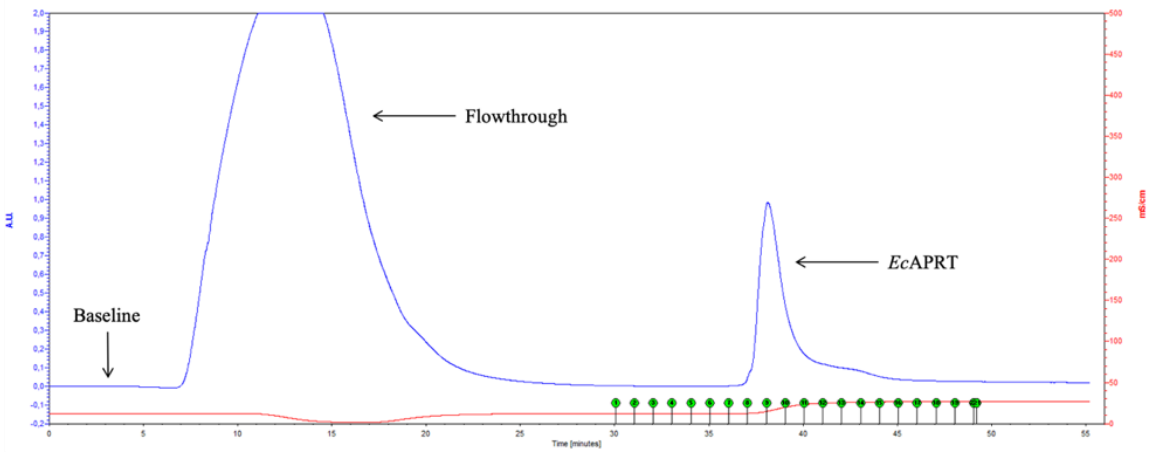


Figure S 3. *Absorbance profile (280 nm) of the affinity purification process registered on LP Data view. The first peak corresponds to the proteins without affinity (flowthrough) and the second peak corresponds to the protein of interest (EcAPRT). Conductivity shown as a red line (mS/cm), absorbance as a blue line (A.U.) and recollected fractions in green circles.*