JSM Bioinformatics, Genomics and Proteomics

Original Research

Use of Bioinformatics Techniques in the Characterization of Genes and Proteins Involved in the Transport of Polyamines from *Staphylococcus* Genus

Kaline Martins da Silva[1], Nicolas Gasparine Figueiredo[2] and Aline Sampaio Cremonesi[3]*

[1]*São Francisco University, Rua Waldemar César da Silveira, Campinas, Brazil*
[2]*Campinas State University, Cidade Universitária Zeferino Vaz, Campinas, Brazil.*
[3]*São Francisco University, Avenida São Francisco de Assis, Bragança Paulista, Brazil*

***Corresponding author**
Aline Sampaio Cremonesi, Avenida São Francisco de Assis, 218, Bragança Paulista, São Paulo, Brazil, Tel: (55 11) 2454-8000

Submitted: 24 July 2023
Accepted: 28 August 2023
Published: 31 August 2023
ISSN: 2576-1102

Keywords
- Protein structure
- Bacterial resistance
- *In silico* analysis
- Spermidine
- ABC transporter

Abstract

Polyamines are essential nitrogenous compounds for bacteria, being captured by ABC transporters that recognize and interact with the molecule to be transported. Given the specificity of this interaction, the present work sought to identify genes and proteins of polyamine transporters in *S. hyicus* using bioinformatics tools. Sequence searches and tests suggest that *S. hyicus* possesses a spermidine ABC transporter, whose genes form the potABCD operon. From structural homology modeling it was pos sible to identify the specific amino acids in the interaction between PotD and spermidine where the W^{260} residue, characteristic of spermidine transporters, is predicted to be displaced into the pocket. This shift allows interaction with a minor polyamine such as putrescine whose transporter was not found in the *S. hyicus* genome. An *in silico* cloning of the PotD protein was carried out in the pET28a, pET28aSUMO, and pET44aNusA vectors and by submitting them for analysis of the resulting ORFs, strongly suggesting that they have favorable physicochemical qualities and solubility. For the docking, presentations present in the literature that were admired in past studies of polyamine transport inhibitors were selected. We can conclude that the *S. hyicus* PotABCD complex is a possible spermidine transporter and the inscriptions: AcSpm, NpAcSpm, GLT-Spm, DFMO, and MGBG, are possible competitive inhibitors of spermidine uptake in PotD in *S. hyicus* bacteria.

INTRODUCTION

Polyamine

Polyamines are essential substances for the reproduction of bacteria, related to the neutralization of the negative charge in the nucleic acid and playing a significant role in DNA transcription processes. Thus, putrescine, spermidine, and spermine are important and universal substances for most prokaryotic cells and are involved in the biosynthesis of biomolecules, in addition to measuring cell growth and proliferation [1].

The intracellular concentration of polyamines is controlled by the active transport of ABC carriers (ATP binding cassette) and intracellular synthesis [1,2]. The biosynthesis of spermidine in prokaryotes occurs by several enzymatic reactions catalyzed by agmatine that originated from the decarboxylation of arginine. The importance of studying the polyamine spermidine is due to the development of inhibitors capable of interfering with the enzymes present in its biosynthesis, causing the loss of this cationic substance, enabling its therapeutic application since the loss of the polyamine proved to be essential to determine the regulation of processes with they may be involved [3].

Spermidine is related to the processes that stimulate the synthesis of nucleic acids and peptides, in addition to having the ability to act as a ligand in several molecular sites of DNA, RNA, proteins, phospholipids, and nucleotide triphosphate, with consistent importance for cell growth through of transport [4].

ABC transporters

Knowing bacterial metabolism is important to understand the mechanisms of bacterial resistance and develop new drugs to fight the infection caused by these microorganisms. Bacteria of the *Staphylococcus* genus have strains that have demonstrated resistance to antibiotics and also express ABC-type transporter (ATP binding cassette) for nutrient uptake. Using energy from the hydrolysis of ATP (adenosine triphosphate) signals, transporters belonging to the ABC family currently constitute of one of the largest superfamilies proteins known and projected due to their importance to organisms that have learned their functionalities [5]. They are usually formed by four domains; two hydrophobic membrane-spanning domains associated with the plasma membrane (MSD - hydrophobic membrane-spanning domains, also called TMD - transmembrane domains), which form the

translocation channel of substances through the membrane, also known as permeases, and two hydrophilic cytoplasmic (NBD nucleotide binding domains), which hydrolyze ATP molecules. A third group of a protein called substrate binding protein (SBP), or periplasmic, is exclusive in complex ABCs of a prokaryotic cell, being a structure that can evidence the transport made by the complex due to its specificity between substrate and site [5,6]. SBPs have a common structure, consisting of two globular domains interconnected by a hinge region, which can be formed in three different types. Based on the number of connections between the domains, they classify the SBP into types, in which type I is formed by three beta strands in its hinge, type II is formed by two beta strands in the hinge between the domains, and type III is formed by a single alpha helix. Knowledge of the structural architecture of substrate-binding proteins is important, as it is in this hinge region that the ligands interact with the protein, causing the domains to approach, resulting in the movement known as the "Venus trap"[7].

ABC-type transporters are associated with complex virulence factors, participate in the secretion of toxins, and the expulsion of antimicrobial agents and antibiotics out of the cell, in addition to being responsible for the bacterial cell [5,6]. Active transport intermediated by these transporters also plays a major role in bacterial cell nutrition and is directly or indirectly linked to more complex metabolic processes and cell growth [2,5]. The ABC transport systems, as well as the entire proteome, present in bacteria, reflect both references to the microorganism's habitat and may be associated with the revelation of relevant characteristics of pathogenic species, being capable of incorporating a wide variety of substances [2,5]. Among the most prominent physiologically-transported substances by the ABC complexes are polyamines [2].

Previous studies have shown that wild strains of Streptococcus pneumoniae maintained their growth with the addition of extracellular polyamines, both in the presence of the polyamine synthesis inhibitor and in its absence, maintaining the level of cell growth found [8]. The uptake of polyamines by the strains in the carrier out by Pot-type ABC transporters, guarantees the survival of the organism even in the presence of the polyamine synthesis inhibitor, showing how the transport system is more vital than synthesis for cell growth. This demonstrates the importance of ABC transporters in the homeostasis of physiological levels of polyamines within the bacterial cell, allowing its reproduction [2,8]. On the other hand, very high levels of polyamines can cause the tendency of reproduction and cell growth, as well as precarious levels [2]. Due to this function of ABC transporters for the regulation of intracellular polyamines, this structure can become a target for the action of drugs in bacteria that are resistant to antibiotics.

Staphylococcus hyicus

Staphylococcus hyicus is a facultatively anaerobic, gram-positive bacterium belonging to the Micrcoccaceae family. Like several other members of this family, *S. hyicus* are commensal cocci belonging to the systemic microbiota of several opportunistic infections. The most referenced pathology in the literature caused by *S. hyicus* is exudative epidermatitis, frequent in domestic pigs. According to Werckenthin et al. [9], the Danish Program (DANMAP) isolated *S. hyicus* from pigs in 1999, for a bacterial monitoring program, and 75% of the bacteria showed resistance to penicillin, 35% showed resistance to streptomycin, 24% to tetracycline, 22% trimethoprim and 15% erythromycin. Therefore, given that this species is related to infections of medical and veterinary interest, it becomes pertinent to carry out structural and genetic studies in search of alternative pharmacological targets. Several previous studies highlighted the presence of exfoliation toxins, which cause the virulence of bacteria of this genus, these toxins being ExhA, ExhB, ExhC, ExhD, ShetA, and ShetB. It was noted that these *Staphylococcus hyicus* toxins are largely related to the toxins produced by *Staphylococcus* aureus [10].

MATERIALS AND METHODS

Staphylococcus hyicus polyamine transport proteins were identified from a search in the Kegg database (Kyoto Encyclopedia of Genes and Genomes - https://www.genome.jp/kegg/). The nucleotide and amino acid sequences of each protein obtained in Kegg were analyzed by BLAST sequence alignment [Basic Local Alignment Search Tool - https://blast.ncbi.nlm.nih.gov/Blast.cgi - [11], and Clustal [https://www.ebi.ac.uk/Tools/msa/clustalo/ - 12] in order to compare the amino acids involved in the binding of polyamines and construction of the phylogenetic tree. The protein sequences used for alignment were obtained by Kegg or by the PDB (Protein Data Bank - https://www.rcsb.org/) for those proteins that have a resolved structure.

The structure of the operon was analyzed from the genomic map of *Staphylococcus hyicus* available on Kegg and the promoter region was identified from the Softberry BPROM platform (http://www.softberry.com/ - [13] from the upstream nucleotide analysis of the first operon gene. The BPROM program identifies supposed promoter regions in the nucleotide sequences, recognizing the -35 and -10 regions [13], that are conserved in promoters of prokaryotic organisms [14].

Membrane proteins and proteins secreted by cells need a region that signals their direction to the site of action, this region is the signal peptide and is usually found in the N-terminal region of proteins, being cleaved when the proteins reach their destination [15]. Bioinformatics analyses are more efficient when we use the amino acid sequence of the functional protein, so it is necessary to identify the signal peptide in the proteins to perform the analysis without this region of the sequence. For this, the amino acid sequences of the proteins were analyzed using the SignalP 6.0 program [https://services.healthtech.dtu.dk/service.php?SignalP – [16] to identify signal peptide regions.

The polyamine ABC transporter permease proteins were identified from the TMHMM online program (http://www.cbs.dtu.dk/services/TMHMM/ - [17], to evaluate the characteristics of the amino acids present in the FASTA sequence of proteins,

searching hydrophobic regions, characteristics of transmembrane regions [17].

The generate the three-dimensional models of the ABC transport proteins, the homology modeling method was used [18], throughout the Swiss Model program [https://swissmodel.expasy.org – [19]. The amino acid sequences of the proteins were used to search the PDB for proteins with a resolved structure and with greater sequence coverage and identify the proteins of interest to serve as a template for modeling the tertiary structure. The quality of the generated models was determined from the Ramachandran graph generated by the Swiss Modle itself and the models were visualized and compared in the PyMol program [https://pymol.org/2/ - [20].

Before starting the in silico cloning, the construction of the primers was carried out using the NebCutter 2.0 program (http://nc2.neb.com/NEBcutter2/), which allows analyzing which endonucleases are present in a given nucleotide sequence, It is possible to compare the endonucleases that cleave the target proteins sequence, PotD, with the endonucleases present in the pET28A, pET28aSUMO, and pET44aNusA vectors.

The strategy for cloning the SBP protein from the *S. hyicus* bacterium was carried out using the SnapGene program (https://www.snapgene.com/), which is a complete program that allows the visualization of the maps of the imported vectors, in addition to your database, the designed primers, generate PCR fragments, perform the cloning and generate the ORFs sequences (open reading frame) of the protein after cloning.

Several programs were used to analyze the quality of the ORFs generated after cloning, verifying the physical-chemical characteristics, solubility, unfolding of these, and carrying out their three-dimensional modeling.

The FoldIndex program (https://fold.proteopedia.org/cgi-bin/findex) makes it possible, through the Uversky algorithm and collaborators, to verify, of a given protein sequence, its native unfolding under physiological conditions, based on the average hydrophobicity of the residue and net charge of the sequence generating a graph that shows the folded and unfolded region of the protein and how many are these regions [21].

The CamSol program was developed by the University of Cambridge, it is necessary to register on the site to use its resources. The most suitable for our research to predict the solubility of the sequence of ORFs was CamSol intrinsic (https://www-cohsoftware.ch.cam.ac.uk/index.php/camsolintrinsic), which through a calculation with scores of general solubility index, which is assigned to all sequences, a solubility profile is produced where regions with scores greater than 1 denote highly soluble regions, while scores less than 1 are poorly soluble [22].

Finally, the 3D modeling of the ORFs was carried out using the HHPred tool (https://toolkit.tuebingen.mpg.de/tools/hhpred), which is fast served to detect the homology of sensitive proteins, function, and structure prediction, allowing the protein triangle models selected by the user from the research carried out. molded in the homologous 3D structural modeling from them [23]. In addition to being one of the first to implement pairwise comparison of hidden Markov profile models. Produces, through the MODELLER software, 3D models of already recognized identical proteins [24].

AutoDock Vina software (https://vina.scripps.edu/), is a molecular docking and virtual screening program that features a fully empirical scoring function, including finite repulsion terms, linear hydrophobic interaction, hydrogen bonding, and an entropic term according to the number of rotating bonds [25]. It uses Monte Carlo theory, belonging to the class of stochastic methods, for this reason, it incorporates annealing, evolutionary, genetic, and Lamarckian genetic algorithm methods to model ligand flexibility while keeping the receptor rigid [26]. Its program scoring function is based on the AMBER force field, including Van der Waals, hydrogen bonds, electrostatic interactions, conformational entropy, and desolvation terms. In addition, it is able to model the flexibility of the receptor, allowing the side chains to move [26,27].

For this reason, the protein of interest (receptor) is added and its preparation is carried out in the program, in which polar hydrogens are added, which are important for the protein-ligand interaction, it is also necessary to verify the Kollman charge, which is a calculation that verifies the amount of electrostatic charge present in the force field between the point-charged atoms in the protein [28], and saving in PDBQT format. When inserting the ligand in the program, a message is generated on the screen showing the added Gasteiger charges and the amount of active rotations detected, this information is relevant because of the theory of flexibility of the ligand over the receptor [26], It is vitally important to consider the flexibility of the ligand and the receptor since both end up changing their conformations to form a perfect fit of minimum energy [26]. Finally, the linker is saved in PDBQT format, together with the receiver. If the receiver and ligands are prepared and saved in PDBQT format correctly, AutoDock Vina automatically determines the grid box values, which are saved in Txt format. The results are processed after writing the commands in the command prompt present on the computer and when obtaining the results, they are directed to PyMol, to observe the location of the ligand positions in the protein and analyze their interaction with the amino acids of the binding pocket of the target protein, if they are in the pocket, it means that these ligands can be possible competitive inhibitors [29].

The enzyme ornithine decarboxylase (ODC) is involved in the biosynthesis of polyamines, catalyzing their production from their precursor, ornithine. There is an association between the increase in ODC activities and cell growth of tumor cells [30]. A study carried out by Burns and collaborators, followed possible referrals conjugated with spermine, capable of inhibiting the transport of polyamines inside tumor cells, cataloging some of these occurrences and carrying out their molecular docking under the target proteins, PotD, in then being analyzed the results generated in order to observe if they may be possible to inhibit the competitive ones with the polyamines captured by PotD.

RESULTS AND DISCUSSION

The genus *Staphylococcus* presents a characteristic mechanism of transports of spermidine, but not of putrescine

Four proteins described as involved in the uptake and transport of polyamines from the *Staphylococcus hyicus* species were identified in KEGG, namely: PotA (NBD), PotB, and PotC (TMD), and PotD (SBP), organized in the genome of *S. hyicus* in the form of an operon, potABCD, along with a transcriptional regulator (RT) (code: SHYC_08820) under the regulation of a promoter located in the intergenic region 616 bases to the amount of the PotA gene according to BPROM analysis (Figure 1A). The organization of ABC transporter genes in the form of an operon has already been described in bacteria such as Pseudomonas aeroginosa (1), *Escherichia coli* [31], and Klebsiella pneumoniae [32], among others. This organization allows for the reduction of the size of the genome by reducing the number of promoters, in this way, genes that are involved in metabolic processes activated by the same stimulus, are close in the genome and regulated by a single promoter [14].

The amino acid sequences of these proteins were used in the search for proteins with resolved structures and deposited in the PDB from a BLAST that indicated spermidine ABC transporter proteins from different bacterial species (data not shown). This first analysis corroborates KEGG's description that these *S. hyicus* proteins are involved in spermidine uptake. In E.coli, at least three polyamines were identified, among which two are ABC-type systems, one for spermidine (PotABCD), and one for putrescine (PotFGHI) [33,34], and a third system composed of Puu proteins for putrescine absorption, especially in the absence of glucose [35,34]. On the other hand, in *S. hyicus,* no proteins possibly involved in putrescine uptake were identified, which makes the PotABCD transporter a unique polyamine uptaker and a promising target for drug development, since its inhibition would significantly compromise the obtainment of compounds nitrogenous by the bacteria.

Analyzes of transmembrane regions performed by TMHMM corroborated the KEGG annotation. PotB and PotC present hydrophobic helices that fit the transmembrane profile, PotA does not present transmembrane domains since it is a cytoplasmic ATPase and PotD present a small transmembrane portion in the

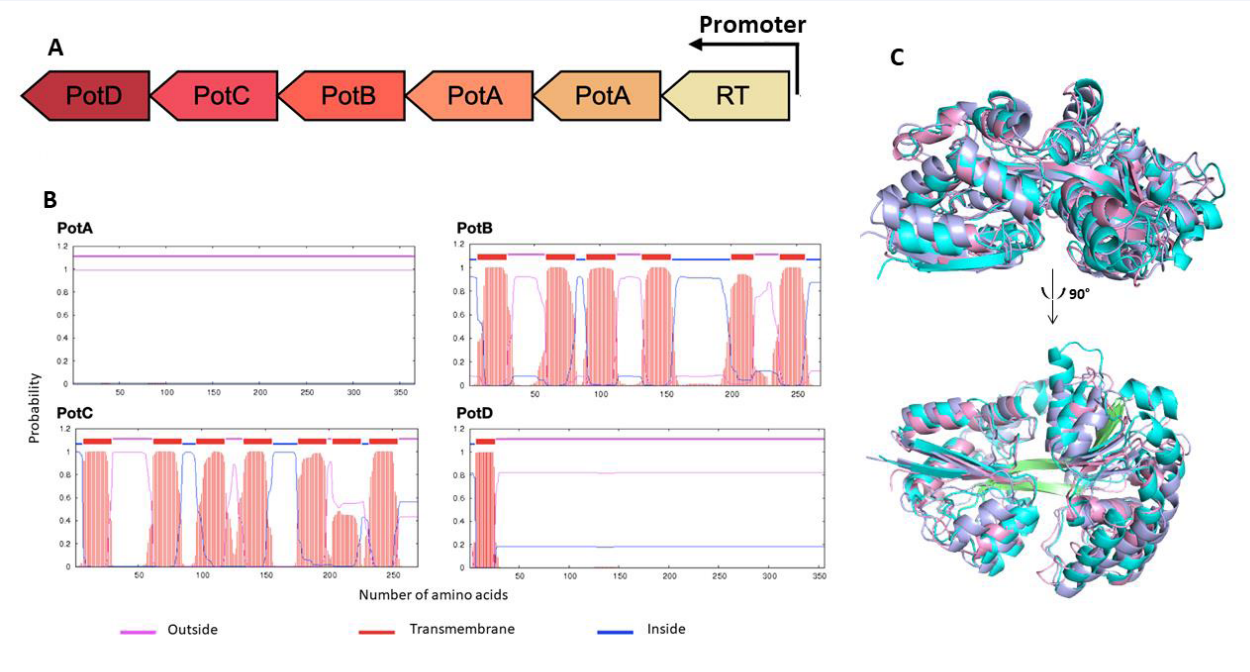

Figure 1 Genome characterization and proteins involved in spermidine uptake. **A:** The genes of the pot ABCD operon of *Staphylococcus hyicus* are represented by arrows whose direction indicates the direction of transcription of the gene. The black arrow indicates the promoter position of this operon as identified by the BPROM program and the direction of expression. **B:** the amino acid sequence of the proteins was analyzed by the TMHMM program that identifies possible transmembrane regions in proteins PotA is described as a cytoplasmic protein that lacks membrane-interacting regions and; PotB, and PotC are described of the transporter complex. The curves indicate transmembrane regions spaced by cytoplasmic or periplasmic regions, characterized by membrane proteins. PotD is described as the transporter substrate-binding protein. The N-terminal region of PotD has a transmembrane peak, however, although some periplasmic proteins can anchor in the membrane, this N-terminal region coincides with the signal peptide region predicted by the SignalP program. **C:** Model generated by homology from the Swiss Model program, using proteins of greater identity as models, namely: *Listeria monocytogenes* (PDB: 4GL0) in light pink, *Escherichia coli* (PDB: 1A99) in light blue and *Pseudomonas aeruginosa* (PDB: 3TTK) in light purple. The SBP protein present in *Staphylococcus hyicus* is classified as a type II SBP, as it presents two beta-sheets between the N-terminal and C-terminal ends, which in the figure is colored in light green. The ABC transporter substrate-binding protein is the protein that recognizes the substrate to be transported, giving specificity to the transporter (DAWSON and LOCHER 2006), thus, comparison of the PotD protein amino acid sequence with PDB polyamine transporter proteins allowed identify the amino acids of the ligand interaction pocket.

N-terminal region that refers tothe signal peptide (Figure 1B), confirmed by reviews on SignalP.

For a better understanding of the structure and functioning of the amino acid sequences of the proteins was used for structural modeling by homology using the Swiss Model program. As a product of the modeling, it was possible to observe the organization of the spermidine transporter protein SBP from *S. hyicus* (Figure 1C), and the proteins encoded by the operon form the entire protein complex characteristic of an importer-type ABC transporter as described in the literature [6]. The PyMol analyses of the generated models made it possible to understand that the PotD protein from *S. hyicus* is a possible type II SBP (Figure 1C) because it has two beta-sheets pleated between the two N- and C-terminal domains [7].

PotD protein from *S. hyicus* presents conserved amino acids for interaction with spermidine

Based on the literature and from the generated models, it was possible to identify in PotD the amino acids that are part of the ligand binding pocket. The identification and recognition of the amino acids present in this binding region were possible by means of two *in silico* processes: 1) alignment of the PotD structural model generated by homology by Swiss-Model from the structure of the SpuD polyamine binding proteins linked to putrescine (PDB: 3TTM), and SpuE bound to spermidine (PDB: 3TTN) from *Pseudomonas aeruginosa,* PotD (PDB: 1POT), and PotF (PDB: 1A99) from *Escherichia coli* linkers od spermidine and putrescine, respectively, and finally the binding protein of spermidine/putrescine (PDB: 4GL0) from *Listeria monocytogenes* which was used as a template for constructing the PotD model due to greater sequence identity and coverage; 2) CLUSTAL amino acid sequence alignment between *S. hyicus* polyamine binding proteins and PotD.

Wu et al., [1], identified the amino acids involved in the interaction between periplasmic proteins and spermidine or putrescine in *Pseudomonas aeruginosa*. According to the authors, five regions are essential for the interaction with the ligand, in these regions are found, respectively, the amino acids W, Y, E, W/Y, and W/F, the latter being determinant for the interaction with the ligand. The W amino acid in this last position is characteristic of a spermidine-binding protein, while the F amino acid in this position refers to a putrescine-binding protein. The amino acids described by Wu et al. [1], were identified and labeled in the sequence (Figure 2A). Among the transporters compared, it was possible to note the conservation of the five amino acids facing the pocket and the possible variations described by the authors. These analyses also allowed the construction of a phylogenetic tree (Figure 2B) where spermidine-binding proteins from *S. hyicus*, *E. coli*, and *L. monocytogenes* are found in the same branch. All of them have the amino acid tryptophan located in the fifth region of the pocket. The putrescine-binding proteins from *E. coli*, and *P. aeruginosa* are also together in one branch, leaving only SpuE, spermidine binding from *P. aeruginosa* in a separate branch. Putrescine-binding proteins have the F residue in the fifth region of the pocket. These data corroborate the analysis by Wu and coauthors [1], by pointing out the homology between proteins that capture the same polyamine and the conservation of amino acid residues involved in the interaction with the ligand.

To identify the structural similarity of these transporters, the PotD protein from *S. hyicus* was compared with the structures of the SBPs mentioned above deposited in the PDB. The analysis of the pocket of the model, when compared with the pocket of the spermidine binding proteins, indicates a significant conservation not only of the residues involved but also in the position and direction of these residues. Comparing *S. hyicus* PotD with spermidine binding SBP (Figure 2C) it was possible to observe that all, including *S. hyicus* PotD, have a tryptophan (W) in the fifth position of the pocket, strongly indicating that *S. hyicus* PotD is a spermidine linker. These findings point to a complex system of spermidine uptake and import in *S. hyicus* to obtain nitrogenous compounds for the bacteria. Because no other polyamine transporters have been described in *S. hyicus*, the PotABCD transporter is an important mechanism for capturing nitrogenous compounds into the cell and can be exploited as a target for drug development. Its inhibition may compromise bacterial metabolism and proliferation.

When comparing the pocket amino acids with the putrescine transporters, it is possible to notice that all five amino acids are described as responsible for the interaction with the ligand and have their position conserved in the structure and although some are different, their fundamental characteristics are conserved as in the last two positions: amino acids can be tryptophan, phenylalanine or tyrosine, however, both are aromatic amino acids.

It is important to emphasize that the amino acid W^{236} in PotD of *S. hyicus* was found displaced towards the interior of the pocket, making it smaller. This conformation does not seem to impair the interaction with spermidine, however, it may allow interaction with a smaller molecule, such as putrescine. As previously described, other bacteria have different uptake systems for putrescine and spermidine, however, as putrescine transporters have not been identified in *S. hyicus* and, in addition to the closure of its PotD pocket, it is possible that *S. hyicus* has only a single carrier for both polyamines.

When aligning the PotD structure of *S. hyicus* with the putrescine-uptake SBPs (Figure 2D), it is noted that W^{236} is positioned very close to putrescine, indicating a possible interaction with this polyamine. Although experimental analyzes are needed to confirm this interaction, these findings allow us to direct the studies.

Implementation of the PotD protein cloning strategy

One of the most valuable tools in the field of molecular biology is restriction endonucleases, they are enzymes that recognize short DNA sequences, ranging from 4 to 8 nucleotides, cutting close to their recognition sites [36, 37]. By submitting the nucleotide sequence of the PotD protein, the program involved into a schematic diagram of the gene sequence with its respective

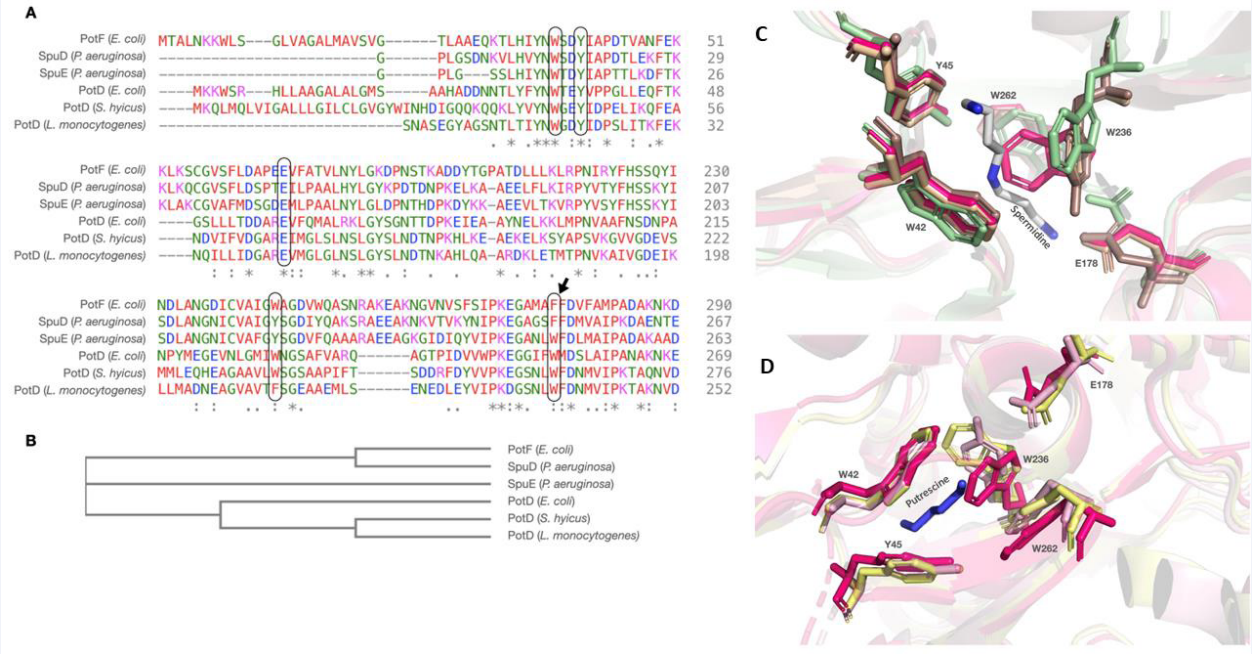

Figure 2 Sequence alignment and phylogenetic tree of ABC transporters of polyamines from different bacterial species and structural analysis of the pocket of interaction with the ligand. The amino acid sequence of polyamines ABC transporters from *Pseudomonas aeruginosa*, *Escherichia coli*, *Listeria monocytogenes*, and *Staphylococcus hyicus* were compared. They are described as spermidine transporters: SpuE and PotD, and putrescine transporters: SpuD and PotF. **A:** The amino acid sequences were requested by the Clustal program that allowed identifying regions of high identity, mainly the amino acids experienced as essential for interaction with the ligand by Wu et al. (2012). These amino acids are circled in black in the image. The authors also point out the fifth region of interaction with the ligand (black arrow) as being crucial for the identification of the polyamine to be transported. They state that the proteins linked to putrescine present a phenylalanine (F) in this region, while those linked to spermidine present a tryptophan (W). Given this analysis, it is possible to confirm that PotD from *S. hyicus* is probably a spermidine-bound protein. **B:** The phylogenetic tree allows comparing the evolution of the transporter across species also performed by Clustal, the tree indicates greater homology between the sequences of putrescine-binding proteins (PotF and SpuD) and between spermidine-binding proteins (PotD) from *E. coli*, *L. monocytogenes*, and *S. hyicus*, however, SpuE is found on a separate branch. The *S. hyicus* PotD model generated by the Swiss Model was aligned in the PyMol program with the structures of the other polyamine transporters. The proteins are represented in cartoons and the amino acids felt to be involved in the interaction with the ligand are represented in sticks. **C:** PotD from *S. hyicus* positioned with spermidine-bound proteins, showing conservation in amino acid positions and a displacement of W236 that makes the binding pocket smaller. W262 remains preserved in all constructed proteins. **D:** PotD from *S. hyicus* has bound putrescine proteins, indicating that there is a possibility for putrescine to fit into the pocket. The numbered amino acids refer to the PotD amino acid sequence of *S. hyicus*, the codes are W (Tryptophan), Y (Tryrosine), and E (Glutamate). Protein cores: SpuD (pink) and SpuE (bronze) from *P. aeruginosa*, PotD from *S. hyicus* (pink), PotD from *L. monocytogenes* (beige), and PotD from *E. coli* (green).

restriction endonucleases, being compared with the restriction enzymes present in vectors pET28a, pET28aSUMO (small ubiquitin-like modifier) and pET44aNusA (N-using substance A), in which the main endonucleases close to the multiple cloning site were selected, then allowing the creation of primers. For the creation of the primers, some rules were followed, namely the removal of the signal peptide from the protein sequence, before creating the forward primer, on the other hand, in the reverse primer, the stop codon sequence must be added between the endonuclease and DNA sequence of the proteins (Table 1).

The SnapGene software was used, in which the sequence of the PotD protein was imported, being presented in a linear way and with its restriction enzymes, then the created primers were added to the database of the program, denoting the amount of guanine and cytosine (GC) and the melting temperature it contains in the primers. The program performs a PCR (polymerase chain reaction) *in silico*, generating the reaction product fragments used for ligation in the plasmid of interest. after cloning, the cloned vector is displayed, highlighting the places where the forward and reverse primer are found together with the target gene, it is also possible to observe the presence of the polyhistidine tail used for transformation, the fusion proteins present in the vectors and the location of the lac I operon, which is a regulatory operon that encodes enzymes that participate in the absorption and metabolism of lactose, when there is no glucose in the medium for the growth of the bacterial strain [38]. With each of the vectors chosen to obtain the ORFs (open reading frames) whose sequences were followed.

Quality analysis of ORFs resulting from cloning in vectors

The result of the analysis in the FoldIndex of the sequences of the ORFs indicated that PotD_pET44aNusA has 849 amino acids of which 336 residues are disordered, on the other hand, PotD_pET28a presented 365 amino acids in its sequence and 203 disordered residues and PotD_pET28aSUMO presented

Table 1: Design of primers made for *in silico* cloning of SBP from *Staphylococcus hyicus*
Table representing the design of the primers used for cloning and their melting temperatures (Tm) and guanine-cytosine (GC) concentration taken from the SnapGene program, the GC concentration are alow than adequate, since the PotD sequence has a greater amount of adenine and thymine. The sequence highlighted in bold is the Stop Codon sequence, with the aim of slightly increasing the concentration of GC. We used different endonucleases for the pET44aNusA vector, as a method to circumvent the problem of the fusion protein, NusA, being found in the opposite direction in the vector. Source: the authors.

pET	Endonuclease	Sequence	Tm and GC
28a	Forward: *Eco*RI- GAA TTC	GAT ATT GGT CAA CAA AAA CAA CAA	57°C / 30%
28a	Reverse: *Pci*I- ACA TGT	**TGA** TCA TAA TCC CAT TTT AAA GTT TAG A	57°C / 26%
28aSUMO	Forward: *Eco*RI- GAA TTC	GAT ATT GGT CAA CAA AAA CAA CAA	57°C / 30%
28aSUMO	Reverse: *Pci*I- ACA TGT	**TGA** TCA TAA TCC CAT TTT AAA GTT TAG A	57°C / 26%
44aNusA	Forward: *Bam*HI- GGA TCC	GAT ATT GGT CAA CAA AAA CAA CAA	59°C / 37%
44aNusA	Reverse: *Eco*RI- GAA TTC	**TGA** TCA TAA TCC CAT TTT AAA GTT TAG A	56°C e/26%

249 disordered residues of the total of 453 amino acids in its sequences.

The solubility of a protein is the amount that it dissolves in a solution, this is given by the amount of take present in its structure, being in a soluble state under specific conditions [39]. There are two challenges when it comes to producing heterologous proteins in Escherichia coli, namely: the expression of a protein being poor or low, and the exact processing of the expressed protein into insoluble aggregates called inclusion bodies. To circumvent these challenges, he began to attach the highly translated native gene as a fusion in the N-terminal region, in order to improve the yield [40].

In the present work, we used vectors that have fusion proteins in their gene, namely, pET28aSUMO and pET44aNusA, for the reason that the NusA proteins is a phage that mediates the termination of the transition, in addition to increasing the solubility of insoluble proteins [41], as for the SUMO proteins, it was observed that it increases expression and facilitates purification with NiNTA chromatography [42].

The solubility of the ORFs was evaluated by the CamSol program, which provided the calculated solubility profile where regions with a value greater than 1 denote highly soluble regions. When submitting the ORFs sequence, the results were cataloged, namely: PotD_pET28a presented a score of 1.25, PotD_pET28aSUMO presented a score of 1.58 and PotD_pET44aNusA presented a score of 1.55. By analyzing the results obtained, it is possible to understand that all of them appear to be soluble, but PotD_pET28aSUMO is the one with the best solubility, thus proving to be stable and allowing experiments and results to be carried out on the bench. The models of each ORF (Figure 3), generated by HHPred, demonstrated structural conservation even in the presence of different fusion proteins in some ORFs, suggesting that the fusion does not necessarily influence the structure and the target protein.

The PotD proteins are subjected to molecular docking with polyamine and inhibitors selected from the literature

A study written by Burns and coauthors [30], was used as a basis for the search for ligands, in which molecules of simple conjugation with spermine are biologically evaluated, facilitating their entry into the pocket of protein binding, since spermine has the biosynthesis process, through the enzyme ornithine decarboxylase, also putrescine and spermidine. It was selected through simple and basic criteria, the ligands which were similar structurally to the spermidine and that obtained a size similar to it, being then cataloged these possible ligands.

The PyMol program was used to visualize the position of the ligands and whether they interact with these amino acids, first analyzing the polyamines, putrescine, and spermidine, to get an idea of which of the two could possibly be transported by the SBP of *Staphylococcus hyicus.* From this visualization, we were able to assume that supposedly spermidine is the polyamine actively transported by PotD (Figure 4A). From then on, the file of ligands resulting from molecular docking began to be imported, to see if they would be possible inhibitors of this transport. Two of these selected ligands are drugs used for the treatment of cancers, studies describe them as presenting interaction with the enzyme ornithine decarboxylase, therefore, causing the inhibition of the synthesis of polyamines, namely, Difluoromethylornithine (DFMO) and Methylglyoxalbisguanylhydrazone (MGBG) [43]. The other three ligands, Acetyl Spermine (AcSpm), Naphthyl-Acetyl Spermine (NpAcSpm), and Glutathionyl-spermine (GLT-Spm), acted as potent inhibitors of polyamine transport in the breast cancer line [30].

AutoDock has two important components, namely a good positioning algorithm and a robust classification or scoring system, i.e., a wide sampling of the conformational space of the ligand is required to correctly position it at the receptor binding site. , obtaining various numbers of potential ligand orientations and conformations under the binding pocket. It presents all possible connection modes, while the scoring system ranks all solutions and identifies the most likely and energetically most favorable connection mode [44]. When analyzing the interaction of the ligand with the PotD protein binding pocket of *S. hyicus* using PyMol, it was possible to observe the 9 positions of the ligands intrinsically in the PotD protein binding pocket that the docking provided, selecting the best positions, taking into account the program's scoring function that estimates free energy, so there is the participation of ligands with amino acids in the binding pocket. Being: Acetyl-spermine presented free energy of -5.5 kcal/mol (Figure 4B), -5.4 kcal/mol, and -5.3 kcal/mol, Naphthyl-acetyl-spermine presented free energy of -6.1 kcal/mol, -6.2 kcal/mol and -6.3 kcal/mol (Figure 4C), Glutathionyl-spermine had free energy of -4.9 (Figure 4D), -4.8 and -4.7, MGBG had free energy of -5.1 kcal/mol (Figure 4E), -5.0 kcal/mol and -4.8 kcal/

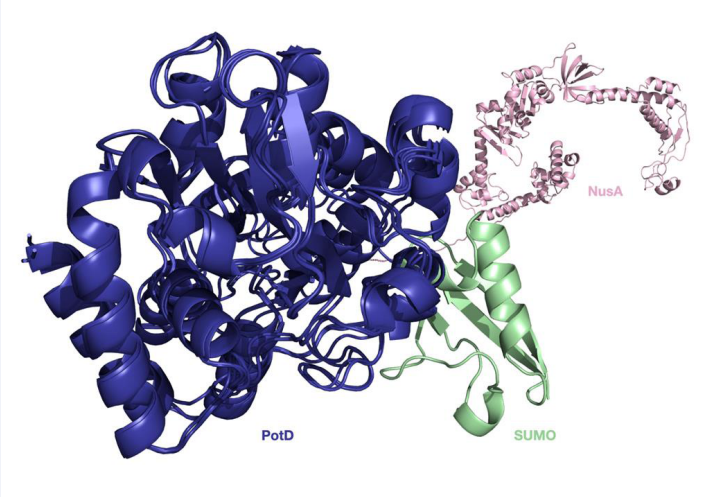

Figure 3 Alignment of the three-dimensional modeling of the ORFs resulting from the *in silico* cloning of the PotD protein.
The protein colored in dark blue is the cloned region of PotD in pET28a, pET28aSUMO, pETa44aNusA, presents the RMSD values: PotD_pET28a and PotD_pET28aSUMO, 0.208; PotD_pET28a and PotD_pET44aNusA, 1.330. The protein in green is the portion of the SUMO fusion protein, transcribed in the pET28aSUMO cloning and the portion of the protein in light pink is the portion of the NusA fusion protein, transcribed in the pETa44aNusA cloning.

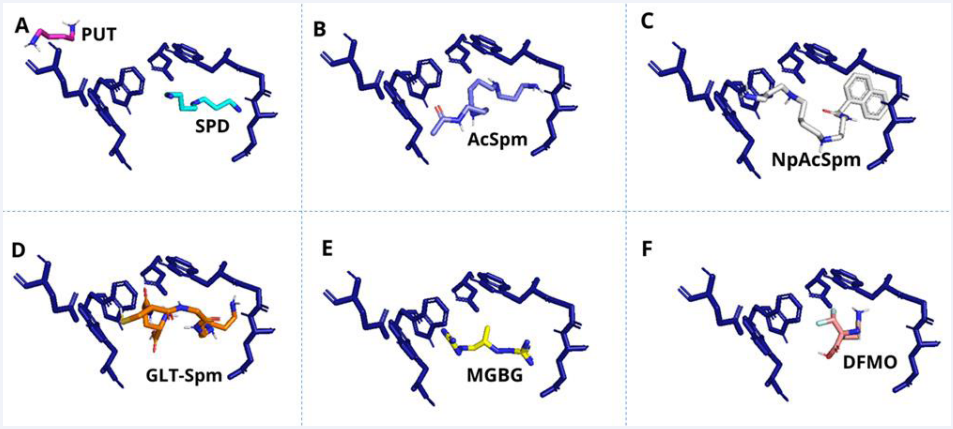

Figure 4 Positioning of ligands in the PotD binding pocket resulting from molecular docking. **A:** SPD: acronym for the spermidine polyamine interacting with the potD binding pocket; a PUT: acronym for the polyamine putrescine shown to be outside the binding pocket and not interacting with the pocket amino acids. With this knowledge, it is possible to understand that this periplasm is theoretically a spermidine carrier; **B:** the Acetyl-spermine molecule, present in the interaction with the amino acids of the binding pocket; **C:** the Naphthyl molecule conjugated with acetyl-spermine, receiving the name of Naphthyl Acetyl-spermine, present in the interaction with the amino acids of the binding pocket; **D:** the Glutathionyl molecule conjugated with spermine, called Glutathionyl-spermine, present in the interaction with the amino acids of the binding pocket; **E:** drug MGBG interacting with potD binding pocket amino acids; **F:** DFMO drug interacting with potD binding pocket amino acids.

mol, and finally, DFMO had free energy of -5.7 kcal/mol (Figure 4F), -5.4 kcal/mol and -5.1 kcal/mol. However, we can presume that the ligands proposed in the study can act as competitive inhibitors since it is known that competitive inhibitors must remain within the binding pocket of the target protein [29].

CONCLUSION

In the present work, the in silico analysis carried out allowed the identification and prediction of the operon and the protein structures that constitute the spermidine ABC transporter of the bacterium *Staphylococcus hyicus*, quickly and with a high degree of reliability. The genes that encode this transporter are organized in the form of an operon called potABCD. whose promoter is located upstream of the potA gene. These genes encode the PotA protein (ATPapase), the PotB and PotC permeases, and the SBP PotD whose functions were confirmed by analyzes of transmembrane regions, signal peptide, and sequence alignment by specialized algorithms.

The PotD protein shows conservation in the amino acid sequence and in the characteristic structure of a periplasmic polyamine scavenger. The amino acids involved in the interaction with spermidine are conserved, but tryptophan 236 is displaced into the binding pocket, which suggests the possible interaction of lower molecular weight polyamines such as putrescine since the transporter for this polyamine was not identified in the study. *Staphylococcus hyicus* genome.

These findings allow us to understand the structure and functioning of the spermidine ABC transporter of *Staphylococcus hyicus* in order to identify its importance for microbial growth and metabolism, which will serve as a basis for future studies on these mechanisms in order to find new targets for development. of new drugs to control this and other bacteria resistant to current microorganisms.

The cloning strategy allows a prediction of events that could happen *in vitro*. *In silico* cloning resulted in a recombinant nucleotide sequence, containing the sequence of the PotD protein, the fusion protein, and the His-Tag sequence present in the vector. It was able to catalog the amino acid sequence of the ORFs, being submitted in specific programs for solubility analysis and identification of deregulated residues, through bioinformatics analysis, allowing the visualization, if possibly the bacterium would be able to express the target protein adequately, or would start to generate inclusion corpuscles, allowing, again, the creation of strategies to overcome this fact. The modeling carried out of the ORFs resulted in the observation of the structure of the protein and its fusion with the fusion proteins.

The inhibition of polyamine uptake can compromise the growth of bacteria and consequently their virulence factors such as capsule formation, and biofilms, among others, for this reason, we selected some molecules described in the literature as possible inhibitors of polyamine transport in other bacteria and the from molecular docking showed promise as they were able to interact with the pocket amino acids, strongly suggesting that these molecules, mainly the Naphthyl Acetyl-spermine ligand, which showed free energy of -6.3 kcal/mol, indicating that it was the ligand that shows with greater affinity and can guarantee greater efficiency in inhibiting the action of PotD in *Staphylococcus hyicus*, causing growth to be inhibited. new drugs capable of being antibiotics, therefore, more studies must be carried out to identify new inhibitors of the action of ABC transporters present in bacteria resistant to antibiotics.

ACKNOWLEDGEMENTS

We would like to thank Universidade São Francisco for the scientific initiation program and support in carrying out this project.

REFERENCES

1. Donghui Wu, Siew Choo Lim, Yihu Dong, Jien Wu, Fei Tao, Lian Zhou, et al. Structural Basis of Substrate Binding Specificity Revealed by the Crystal Structures of Polyamine Receptors SpuD and SpuE from Pseudomonas aeruginosa. J Mol Biol. 2012; 416: 697-712.
2. LIMA, RN. Análise do papel dos transportadores ABC de oligopeptídeos, poliaminas, fosfato inorgânico, glutamato e glutamina na fisiologia e patogênese de bactérias do trato gastro-intestinal. Tese (Doutorado em Ciências). - Universidade de São Paulo, São Paulo, 2013
3. ARAÚJO MJSMPD. Síntese de poliaminas naturais selectivamente protegidas derivatização e síntese total. Tese (Doutorado em Química Orgânica) - Universidade do Porto, Porto. 1995.
4. Heather M Wallace, Alison V Fraser, Alun Hughes. A pesrpective of polyamine metabolism. Biochem J. 2003; 376: 1-14.
5. Amy L Davidson, Jue Chen. ATP-binding cassette transporters in bacteria. Annu Rev Biochem. 2004; 73: 241-268.
6. Roger J P Dawson, Kaspar P Locher. Structure of a bacterial multidrug ABC transporter. Nature. 2006; 443: 180-185.
7. Mohammad M Rahman, Mayra A Machuca , Anna Roujeinikova. Bioinformatics analysis and biochemical characterisation of ABC transporter-associated periplasmic substrate-binding proteins ModA and MetQ from Helicobacter pylori strain SS1. Biophys Chem. 2021
8. Ware D, Jiang Y, Lin W, Swiatlo E. Involvement of potD in Streptococcus pneumoniae polyamine transport and pathogenesis. Infect Immun. 2006; 74: 352-361.
9. Werckenthin C, Cardoso M, Martel JL, Schwarz S. Antimicrobial resistance in staphylococci from animals with particular reference to bovine Staphylococcus aureus, porcine Staphylococcus hyicus, and canine Staphylococcus intermedius. Vet Res. 2001; 32: 341-362.
10. Peter Ahrens, Lars Ole Andresen. Cloning and Sequence Analysis of Genes Encoding S hyicus Exfoliative Toxin Types A, B, C, and D. J Bacteriol. 2004; 186: 1833-1837.
11. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25: 3389-3402.
12. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23: 2947-2948.
13. Solovyev V, Salamov A. Automatic annotation of microbial genomes and metagenomic sequences. In LI, Robert W. (org). Metagenomics and its applications in agriculture, biomedicine and environmental studies. Hauppauge. 2011; 61-78.
14. LODISH H, BERK A, KAISER CA, KRIEGER M, BRETSCHER A, PLOEGH H, et al. Molecular Cell Biology. 7ª ed. cap. 7. Porto Alegre: Artmed, 2014.
15. G von Heijne. The signal peptide. J Membr Biol. 1990; 115: 195-201.
16. Armenteros JJA, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol. 2019; 37: 420-423.
17. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001; 305: 567-580.
18. FILHO OAS, ALENCASTRO RB. Modelagem de Proteínas por Homologia. Química Nova.2003; 26: 253-259.
19. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucle Acid Res. 2018; 46: 296-303.
20. DELANO WL. The PyMol Molecular Graphics System. In BALLARD, Charles; HOWARD-EALES, Maeri. CCP4 Newsletter on Protein Crystallography. Warrington. 2002; 44-53.
21. Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O,

Beckmann JS, et al. FoldIndex©: a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics. 2005; 21: 3435-3438.

22. Sormanni P, Vendruscolo M. Protein solubility predictions using the CamSol method in the study of protein homeostasis. Cold Spring Harb Perspect Biol. 2019; 11: a033845.

23. Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, et al. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol. 2018; 430: 2237-2243.

24. Soding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 2005; 33: W244-W248.

25. Gaillard T. Evaluation of AutoDock and AutoDock Vina on the CASF-2013 benchmark. J Chem Inf Model. 2018; 58: 1697-1706.

26. Xuan-Yu Meng, Hong-Xing Zhang, Mihaly Mezei, Meng Cui. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. Curr Comput Aided Drug Des. 2011; 7: 146-157.

27. Oleg Trott, Arthur J Olson. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010; 31: 455-461.

28. Yong Duan, Chun Wu, Shibasish Chowdhury, Mathew C Lee, Guoming Xiong, Wei Zhang, et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem. 2003; 24: 1999-2012.

29. MOTTA VT. Bioquímica Básica. Laboratório Autolab LTDA, capítulo 3, Enzimas. 2005

30. Burns MR, Carlson CL, Vanderwerf SM, Ziemer JR, Weeks RS, Cai F, et al. Amino acid/spermine conjugates: polyamine amides as potent spermidine uptake inhibitors. J Med Chem. 2001; 44: 3632-3644.

31. Sugiyama S, Vassylyev DG, Matsushima M, Kashiwagi K, Igarashi K, Morikawa K. Crystal Structure of PotD, the Primary Receptor of the Polyamine Transport System in Escherichia coli. J Biol Chem. 1996; 271: 9519-9525.

32. Mann E, Mallette E, Clarke BR, Kimber MS, Whitfield C. The Klebsiella pneumoniae O12 ATP-binding Cassette (ABC) Transporter Recognizes the Terminal Residue of Its O-antigen Polysaccharide Substrate. The J Biol Chem. 2016; 291: 9748-9761.

33. Kashiwagi K, Shibuya S, Tomitori H, Kuraishi A, Igarashi K. Excretion and Uptake of Putrescine by the PotE Protein in Escherichia coli. J Biol Chem. 1997; 272: 6318-6323.

34. Terui Y, Saroj SD, Sakamoto A, Yoshida T, Higashi K, Kurihara S, et al. Properties of putrescine uptake by PotFGHI and PuuP and their physiological significance in Escherichia coli. Amino Acids. 2014; 46: 661-670.

35. Kurihara S, Tsuboi Y, Oda S, Kim HG, Kumagai H, Suzuki H. The Putrescine Importer PuuP of Escherichia coli K-12. J Bacteriol. 2009; 191: 2776-2782.

36. Pingoud AE, Jeltsch A. Structure and function of type II restriction endonucleases. Nucleic Acids Res. 2001; 29: 3705-3727.

37. Romanuka J. NMR studies of the allosteric effectors of the lac operon. LAP Lambert Academic Publishing. 2009; 92.

38. Roberts RJ, Halford SE. Type II restrictions enzymes. In: Nucleases, 2nd Ed. Cold Spring Harbor Laboratory Press. 1993; 35-88.

39. Zayas JF. Functionality of Proteins in Food. Springer Berlin Heidelberg. 1997.

40. Malhotra A. Tagging for protein expression. Methods Enzymol. 2009; 463: 239-2.

41. Correa EM. Expressão heteróloga, purificação e caracterização das proteínas humanas DCRA (Down Syndrome Critical Region Gene A) e DSCR8 (Down Syndrome Critical Region Gene 8). 2004.

42. Marblestone JG, Edavettal SC, Lim Y, Lim P, Zuo X, Butt TR. Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. Protein science. 2006; 15: 182-189.

43. Bae DH, Lane DJR, Jansson PJ, Richardson DR. The old and new biochemistry of polyamines. Biochimica et Biophysica Acta (BBA)-General Subjects. 2018; 1862: 2053-2068.

44. Santos LH. Docagem Molecular: em busca do encaixe perfeito e acessível. BIOINFO-Revista brasileira de Bioinformatica. Edição 01. Julho, 2021.