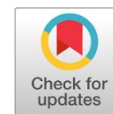


**Title: Structure-based virtual screening of inhibitors targeting thiamine phosphate synthase in *Streptococcus pneumoniae***



Authors: Guido Michael Sánchez-Rojas, Antero Enrique Yacarini-Martínez and Erick Giancarlo Suclupe-Farro

DOI: **10.17533/udea.redin.20250986**

To appear in: *Revista Facultad de Ingeniería Universidad de Antioquia*

Received: July 09, 2024

Accepted: September 19, 2025

Available Online: September 19, 2025

This is the PDF version of an unedited article that has been peer-reviewed and accepted for publication. It is an early version, to our customers; however, the content is the same as the published article, but it does not have the final copy-editing, formatting, typesetting and other editing done by the publisher before the final published version. During this editing process, some errors might be discovered which could affect the content, besides all legal disclaimers that apply to this journal.

Please cite this article as: G. M. Sánchez-Rojas, A. E. Yacarini-Martínez and E. G. Suclupe-Farro. Structure-based virtual screening of inhibitors targeting thiamine phosphate synthase in *Streptococcus pneumoniae*, *Revista Facultad de Ingeniería Universidad de Antioquia*, Sep. 2025 [Online]. Available: <https://www.doi.org/10.17533/udea.redin.20250986>

## Structure-based virtual screening of inhibitors targeting thiamine phosphate synthase in *streptococcus pneumoniae*

### Cribado virtual basado en estructuras de inhibidores dirigidos a la tiamina fosfato sintasa en *streptococcus pneumoniae*

Guido Michael Sánchez-Rojas<sup>1</sup> <https://orcid.org/0009-0008-7542-2458> Antero Enrique Yacarini-Martínez<sup>2</sup>  
<https://orcid.org/0000-0003-4716-4371> Erick Giancarlo Suclupe-Farro<sup>2\*</sup> <https://orcid.org/0000-0002-0334-2191>

<sup>1</sup> Facultad de Ciencias Biológicas, Universidad Nacional Pedro Ruíz Gallo, Calle Juan XXIII N° 391. C. P. 14013. Lambayeque, Perú.

<sup>2</sup> Programa de Medicina Humana, Universidad San Martín de Porres. Avenida Circunvalación del Club G, Urbanización Club Golf Los Incas 170. C. P. 150141. Santiago de Surco, Lima, Perú.

Corresponding author: Erick Giancarlo Suclupe-Farro

E-mail: [esuclupf@usmp.pe](mailto:esuclupf@usmp.pe)

#### Abstract

Antimicrobial resistance in *Streptococcus pneumoniae* poses a growing challenge to global public health. Although this bacterium has been widely studied, the enzyme thiamine phosphate synthase, essential for its metabolism, lacks an experimentally resolved three-dimensional structure. This study aimed to identify compounds with inhibitory potential against ThiL using structure prediction and molecular docking tools. The 3D structure was modeled using AlphaFold2 and its active site was adjusted based on homologous templates. Bioactive compounds from the Zinc20 and PubChem databases were virtually screened and their binding affinities were evaluated using AutoDock Vina. The results highlighted Bindarit as a potential inhibitor, exhibiting higher binding affinity (-9.78 kcal/mol) than the natural ligands. These findings suggest its potential as a targeted antimicrobial therapy. Further validation through in vitro and in vivo experiments is recommended.

**Keywords:** Antimicrobial resistance, bioinformatics, drug resistance, protein modeling, enzyme inhibitors

#### Resumen

La resistencia antimicrobiana de *Streptococcus pneumoniae* representa un desafío creciente para la salud pública global. Aunque esta bacteria ha sido ampliamente estudiada, la enzima tiamina fosfato sintasa, esencial en su metabolismo, aún no cuenta con una estructura



tridimensional resuelta experimentalmente. Este estudio tuvo como objetivo identificar compuestos con potencial inhibitorio frente a ThiL utilizando herramientas de predicción estructural y acoplamiento molecular. Se predijo la estructura tridimensional mediante AlphaFold2 y se ajustó su sitio activo en base a estructuras homólogas. Posteriormente, se realizó un cribado virtual de compuestos bioactivos provenientes de las bases de datos Zinc20 y PubChem, y se evaluó su afinidad de unión mediante AutoDock Vina. Los resultados destacaron a Bindarit como un posible inhibidor, mostrando una energía de unión mayor (-9.78 kcal/mol) que la de los ligandos naturales. Este hallazgo sugiere su potencial como terapia antimicrobiana dirigida. Se recomienda validar experimentalmente estos resultados mediante estudios *in vitro* e *in vivo*.

**Palabras clave:** Resistencia antimicrobiana; Bioinformática; Resistencia a los medicamentos; Modelamiento de proteínas; Inhibidores de enzimas

## 1. Introduction

The growing antibiotic resistance of *Streptococcus pneumoniae* placed this bacterium among the five most resistant pathogens globally in 2019 [1]. It is responsible for serious diseases such as otitis media, meningitis, pneumonia, and septicemia [2]. Its high global incidence causes constant concern, affecting both developed and developing countries [3].

Antibiotic resistance, resulting from the irresponsible use of medications and adaptive processes driven by resistance genes, spreads both horizontally and vertically, limiting the effectiveness of current treatments [4, 5, 6]. The persistence of these mechanisms highlights the bacteria's adaptability and rapid response to the new environments they find themselves [7].

Despite the development and application of pneumococcal vaccines, pneumonia persists as a global health problem [8, 9, 10]. In 2015, it was estimated to have caused 15% of deaths in children under 5 years old, reaching an alarming figure of 920,136 deaths that same year [11]. In Peru, bacterial pneumonia has remained the leading cause of mortality over the past 30 years, with specific serotypes such as *Sp.* 19A, 6C, 19F, and 15 standing out in pathogenicity [12].

Efforts have been made to validate new molecules with regulatory, inhibitory, and/or antibiotic properties, focusing on five antibacterial pathways [13, 4]. In the search for new regulation and/or inhibition pathways, metabolic targets with synthesis routes for biochemical compounds unique to bacteria have been proposed [14]. The peculiarity of biochemical pathways in bacteria, specifically for *Sp.*, is that they suggest pharmacological targets of a specific nature, such as the biosynthesis of thiamine monophosphate, which is crucial for its metabolism and synthesis of vitamin B1 (thiamine). This pathway becomes a potential target for drugs with inhibitory potential [15].

Thiamine phosphate synthase, key in carbohydrate metabolism, catalyzes the formation of thiamine monophosphate, which is essential for numerous metabolic pathways. This enzyme, homodimeric and belonging to the transferase family, features a TIM barrel structural

characteristic. Its active site, surrounded by loop regions, includes critical residues such as ASP93 and ASP112, which coordinate a Mg<sup>2+</sup> cation. The lack of thiamine synthesis in humans offers a promising pathway for the development of specific drugs with reduced side effects [13, 16, 17]. Despite the importance of the thiamine phosphate synthase enzyme in *Sp.*, its three-dimensional structure has not yet been experimentally resolved. This work is based on the search and prior validation of the amino acid sequence, suggesting the need for a three-dimensional model [18, 13].

Access to multiple biological databases has driven the development of computational tools such as AlphaFold2 [19] and LiSiCa [20, 21]. These applications use deep learning to predict protein structures and perform virtual ligand screenings, respectively. These tools, as part of a descriptive and observational approach, are integrated into experimental research, contributing to the discovery of substances with antibiotic potential.

One of the key techniques most commonly used to theoretically understand binding interactions is molecular docking. This method allows the computational prediction and calculation of the most favorable interaction position between a ligand and a target (usually a protein). Additionally, this technique enables the understanding of molecular stability between macromolecules and the identification of organic compounds with biological activity. It uses negative binding free energy calculations, following the principles of the second law of thermodynamics, to predict if the reaction is spontaneous [22, 23].

Bioinformatics software designed for molecular docking studies use free energy calculations. AutoDock Vina is one of the interaction calculation software, notable for its capacity to generate negative free energy scores, which are key to interpreting results [22, 24, 25]. This open-source software has already been validated in numerous previous studies that evaluated the interaction of drugs aimed at treating obesity, such as Lorcaserin, Phentermine, and Liraglutide, with proteins for pharmacological repositioning [26].

This study focused on identifying potential drugs using *in silico* techniques, such as screening and molecular docking, employing free-access pharmacological databases that offer valuable resources for research in chemistry and biology [27, 28, 29]. The aim was to analyze the binding free energy via molecular docking of drugs with inhibitory capacity against the thiamine phosphate synthase enzyme from *Streptococcus pneumoniae*.

## 2. Materials and Methods

### 2.1 Identification of Similar Structures

The amino acid sequence of the enzyme with identification code [AFC94415.1](#) from *Streptococcus pneumoniae* ST556 (GenBank CP003357.2) [30] was used as a query in a BLASTp search in the Protein Data Bank (PDB) database to identify crystal structures that showed the greatest similarity in terms of functionality, identity, and coverage with our target sequence.

## 2.2 Structure Prediction

The amino acid sequence of the thiamine phosphate synthase from *Streptococcus pneumoniae* (accession AFC94415.1) was used to predict its three-dimensional structure using AlphaFold2 [19]. This tool was selected due to its high performance in the CASP14 competition and its capacity to provide structural models with accuracy comparable to experimental techniques. To model the homodimer nature of the enzyme, the sequence was input twice, following the recommended practices for simulating symmetric assemblies in AlphaFold2 predictions. Default settings were used to generate the structure through the AlphaFold2 Colab notebook (<https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb>). The quality of the predicted model was evaluated using the Predicted Local Distance Difference Test (pLDDT) and the Predicted Aligned Error (PAE) matrix, which assessed the confidence within domains and between chains, respectively.

## 2.3 Selection of Organic Ligands

The substrates of thiamine phosphate synthase are thiazole (TZ) and pyrimidine (PIR), which, after the reaction, yield thiamine monophosphate (TMP) as a product. To obtain ligands with possible inhibitory potential, a comparison based on structure was initiated using the substrates and the product. The PyMOL software [31] and its LiSiCa-GUI extension [20] were used in the elements downloaded from online databases, ZINC20 (containing drugs approved in major health jurisdictions, including the FDA [27], with 5,903 drugs with three-dimensional structures as of 05-01-2023) and PubChem [28] (with 9,801 drugs with three-dimensional structures as of 05-01-2023). In both cases, the molecules were obtained with the MOL2 extension.

## 2.4 Adjustment of Modeled Protein

The predicted enzyme structure lacked any bound ligand and the visual inspection revealed that the active site appeared closed—a known limitation when using AlphaFold2 in ligand-free conformations. Since AutoDock Vina treats the protein as a rigid body during docking simulations, this closed conformation could prevent accurate docking of ligands into the active site.

To overcome this limitation, the AlphaFold2-predicted model (AFC94415) was structurally aligned with the crystallized structure 1G6C from *Bacillus subtilis*, which contains natural ligands in its catalytic site. This alignment, done using PyMOL [31], enabled the manual transfer of ligands from 1G6C into the predicted model to induce an open conformation.

Following the alignment, the predicted structure underwent energy minimization using the YASARA Minimization Server [32] to relieve steric clashes and stabilize the modified active site. This protocol has been previously validated in modeling workflows to refine protein-ligand conformations and improve structural realism.

Although this approach is qualitative, it provided a practical solution to approximate the open conformation of the enzyme, which was then used for docking studies. Future work may include molecular dynamics simulations to validate the stability of the open conformation over time.

## 2.5 Identification of the Active Site

The online program DoGSite3 [33, 34, 35] was used as a fundamental part of the process for identifying the active site of the predicted enzyme to detect potential binding sites in the 3D structure of macromolecules. The program determined the global properties of the active site, including size, shape, and chemical characteristics of the predicted pockets. Additionally, this methodological step was crucial for identifying and verifying the results of molecular docking, thereby establishing an integral connection between the identification of the active site and the interpretation of molecular interaction in the predicted enzyme.

## 2.6 Preparation of Elements for Molecular Docking

The ligands selected by LiSiCa [20] were prepared using Open Babel [36] and AutoDock Tools [37] programs, adjusting their pH to 7.4 and applying the MMFF94 force field, and were then converted to a compatible format (PDBQT). The receiver, together with the functions provided by the AutoDock suite and considering the parameters set by Forli [38], was also converted to the compatible PDBQT format.

## 2.7 Molecular Docking

Molecular docking was performed using AutoDock Vina v.1.2.5 [22, 39], where the code was modified to perform multiple molecular docking. To guarantee the reliability of the results, this procedure was performed three times and an exhaustivity of 32 (which reduces the parallelism of runs, improving the local optimization of the results) was used to validate the results and prevent possible errors in the search for the best affinity free energy.

## 2.8 Interpretation of Results

Molecular interactions conducive to enzyme inhibition were analyzed with the online program PLIP [40]. The most salient theoretical results were detailed, addressing the binding sites and the amino acid residues involved in complex formation with enzyme ligands. This analysis provided a comprehensive view of the structural features and specific interactions that determine the inhibitory potential of the ligands under study.

## 3. Results

### Comparison with Similar Structures

When carrying out the BLASTp analysis, detailed in Table 1, a wide coverage, equal to 98% in the amino acid sequences, was observed. However, the highest percentage of similarity was only 41.75% for both structure cod. PDB 1G4T and structure cod. PDB 3O15, both from the organism *Bacillus subtilis*, with 227 and 235 amino acids, respectively. These results reveal variable levels of similarity, confirming the absence of the structure of the protein of interest in the PDB database. This justifies the choice and use of AlphaFold2 as the predictive tool to



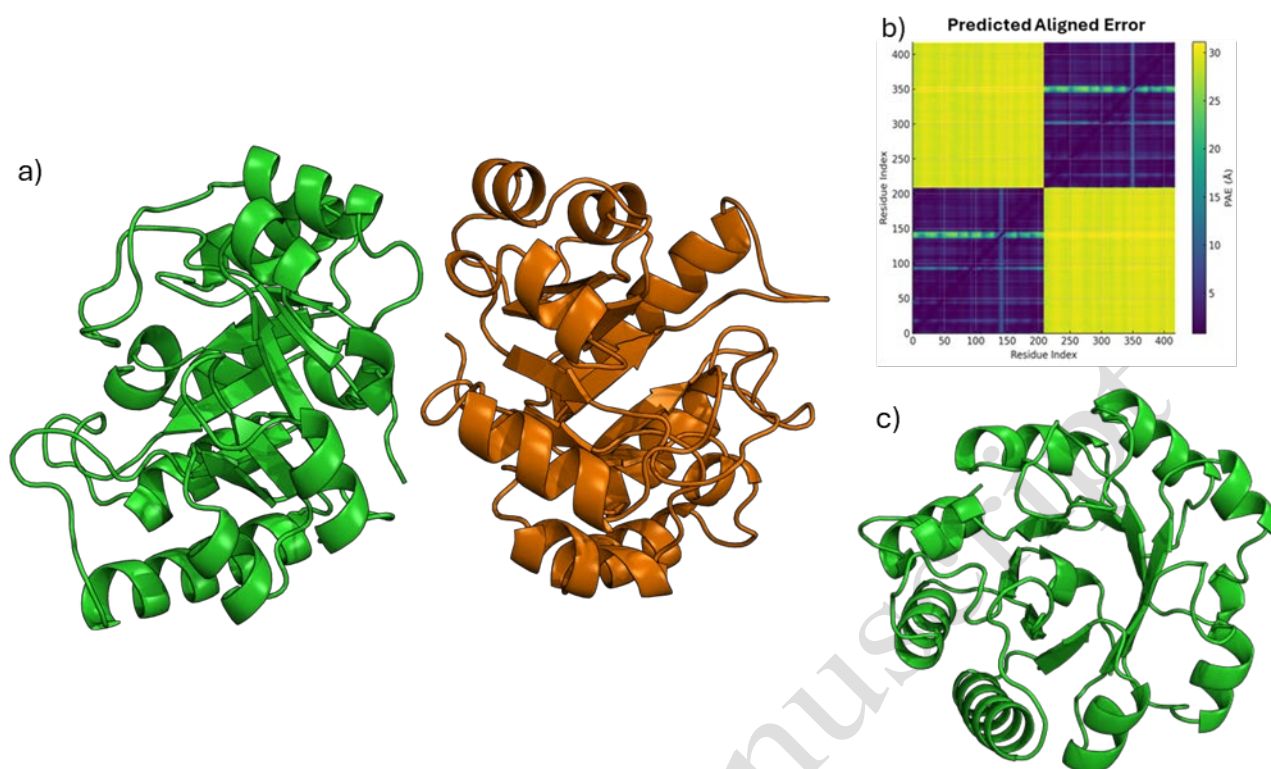
model its three-dimensional structure.

**Table 1:** PDB code with sequences similar to the target sequence (data extracted from the official website BLASTp).

Description	Scientific Name	Query Cover	Per. Ident	Acc. Len	Accession
Chain A, Thiamin Phosphate Synthase	<i>Bacillus subtilis</i>	98%	41.75%	227	1G4T_A
Chain A, Thiamine-Phosphate Pyrophosphorylase	<i>Bacillus subtilis</i>	98%	41.75%	235	3O15_A
Chain A, Thiamin Phosphate Synthase	<i>Bacillus subtilis</i>	98%	41.26%	225	1G67_A
Chain A, Thiamin Phosphate Synthase	<i>Bacillus subtilis</i>	98%	41.26%	228	1G69_A
Chain A, Thiamin Phosphate Synthase	<i>Bacillus subtilis</i>	98%	41.26%	227	1G4E_A
Chain A, Thiamin Phosphate Synthase	<i>Bacillus subtilis</i>	98%	41.26%	226	1G4P_A
Chain A, Thiamine-Phosphate Pyrophosphorylase	<i>Bacillus subtilis</i>	98%	41.26%	235	3O16_A
Chain A, Thiamine-Phosphate Pyrophosphorylase	<i>Pyrococcus furiosus</i>	95%	37.13%	215	1XI3_A
Chain A, Thiamine Biosynthetic Bifunctional Enzyme	<i>Nakaseomyces glabratus</i>	88%	39.22%	540	3NL2_A
Chain A, Thiamine Biosynthetic Bifunctional Enzyme	<i>Nakaseomyces glabratus</i>	88%	39.22%	540	3NM1_A
Chain A, Probable Thiamine-Phosphate Pyrophosphorylase	<i>Mycobacterium tuberculosis</i>	93%	26.57%	243	3O63_A
Chain A, Regulatory Protein tenI	<i>Bacillus subtilis</i>	81%	26.59%	221	1YAD_A

### 3.2 Structure of the Thiamine Phosphate Synthase Enzyme Modeled by AlphaFold

The predicted structure of thiamine phosphate synthase (AFC94415) was modeled using AlphaFold2 and exhibited a globular fold typical of the TIM barrel family [41], arranged in a homodimeric conformation (Figure 1a). The model's confidence was evaluated using the Predicted Aligned Error (PAE) matrix (Figure 1b), indicating low inter-chain alignment error, supporting the feasibility of a biologically relevant dimer.



**Figure 1. Predicted dimeric structure of thiamine phosphate synthase model AFC94415 with ligand and cofactor positioning.** (a) Homodimeric structure of thiamine phosphate synthase from *Streptococcus pneumoniae* predicted by AlphaFold2, with each monomer colored separately (green and orange). (b) Predicted Aligned Error (PAE) plot illustrating inter- and intra-chain confidence levels. Dark blue areas correspond to low predicted error (<5 Å). (c) Single monomer structure highlighting the TIM barrel fold.

### 3.3 Identification of Ligands Based on Structure

The drugs from the selected databases PubChem and Zinc20 consisted of a total of 9,801 and 5,903 compounds, respectively. From this initial total, the results were reduced to 54 drugs, initially 30 from each of the databases (10 results per reference ligand). Additionally, 4 duplicate drugs were eliminated: ID: 1051; ID: ZINC000008215517; ID: ZINC000001532514; ID: ZINC000001531009 (shaded in yellow), and 2 phosphoric acid drugs: ID: 1023; ID: ZINC000006827695 (shaded in green), as shown in Table 2. Furthermore, their original coding is retained for subsequent identification and evaluation.

**Table 2:** Results of the LiSiCa program in the virtual screening of the PubChem and Zinc20 databases against the substrates Pyrimidine and Thiazole, as well as thiamine monophosphate

No.	DB*	Thiamine monophosphate		Pyrimidine		Thiazole	
		ID	Similarity	ID	Similarity	ID	Similarity
1	PubChem	1130	0.65	1051	0.55	10783	0.57
2		2320	0.47	1023	0.50	6274	0.50
3		3032771	0.47	6076	0.48	64969	0.47

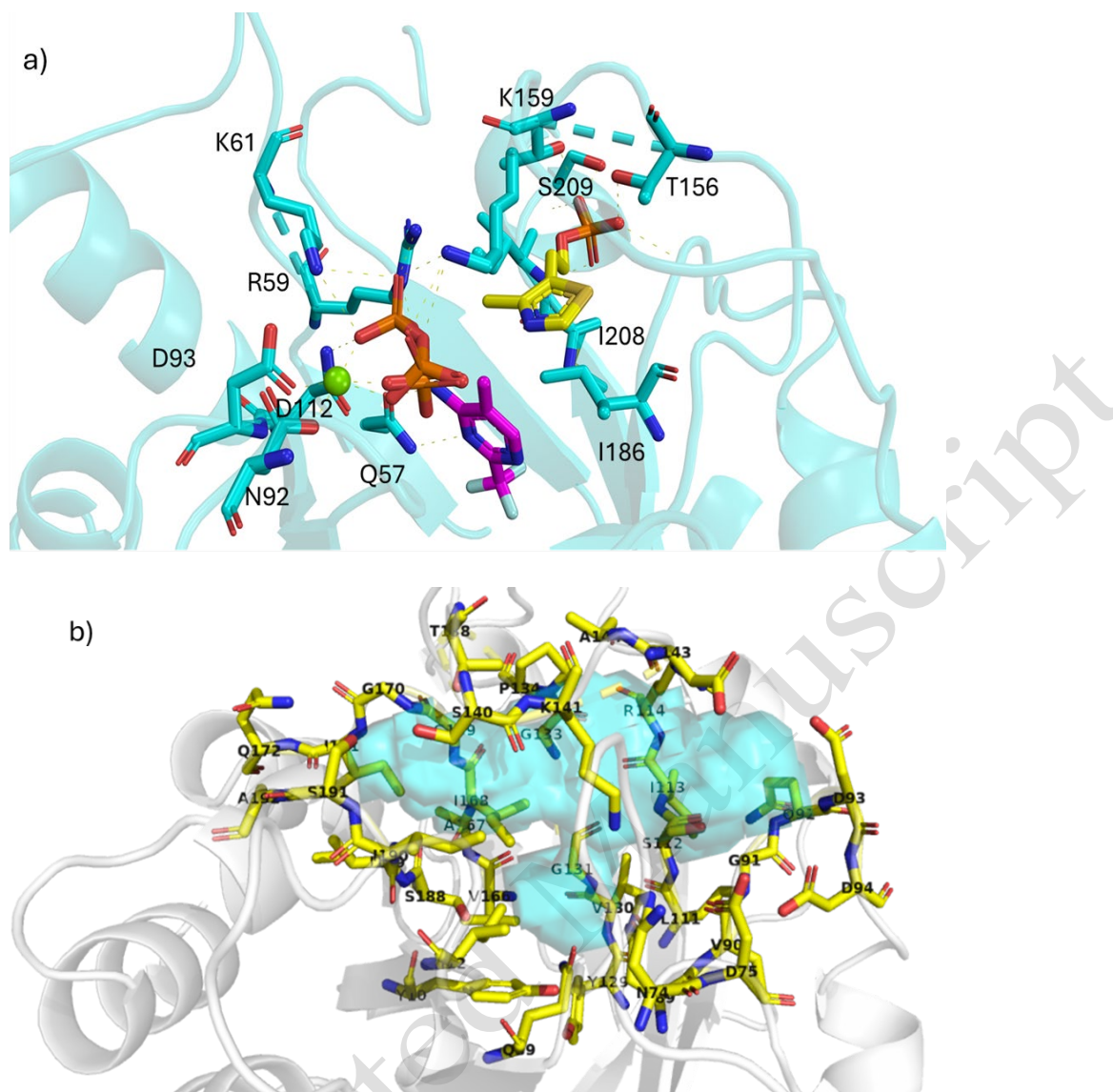


4	5282168	0.47	27304	0.47	92105	0.47
5	6083	0.45	100299	0.46	3614	0.47
6	43860	0.44	5245	0.46	1021	0.45
7	56032	0.44	1195	0.45	1051	0.45
8	3368	0.44	3305	0.45	68740	0.45
9	30751	0.44	39214	0.44	1221	0.44
10	71354	0.44	60172	0.44	75810	0.44
11	ZINC00000821551 7	0.70	ZINC00000153251 4	0.55	ZINC00000016079 0	0.69
12	ZINC00000004915 3	0.65	ZINC00000682769 5	0.50	ZINC00000666122 7	0.50
13	ZINC00000201555 8	0.47	ZINC00007787420 8	0.48	ZINC00000153251 4	0.45
14	ZINC00000201555 9	0.47	ZINC00002474049 2	0.47	ZINC00000380365 2	0.45
15	ZINC00003073129 0	0.47	ZINC00000153100 9	0.46	ZINC00000153100 9	0.43
16	ZINC00009561859 9	0.46	ZINC00000383081 3	0.45	ZINC00000378166 4	0.42
17	ZINC00009561859 8	0.46	ZINC00000397915 6	0.44	ZINC00000386958 0	0.42
18	ZINC00000359079 0	0.46	ZINC00000153060 0	0.44	ZINC00000383135 2	0.42
19	ZINC00000258363 1	0.46	ZINC00009561886 4	0.43	ZINC00009590887 5	0.42
20	ZINC00000386015 6	0.45	ZINC00000821551 7	0.42	ZINC00000015456 4	0.41

\* DB: Databases.

### 3.4 Active Site of the Enzyme AFC94415

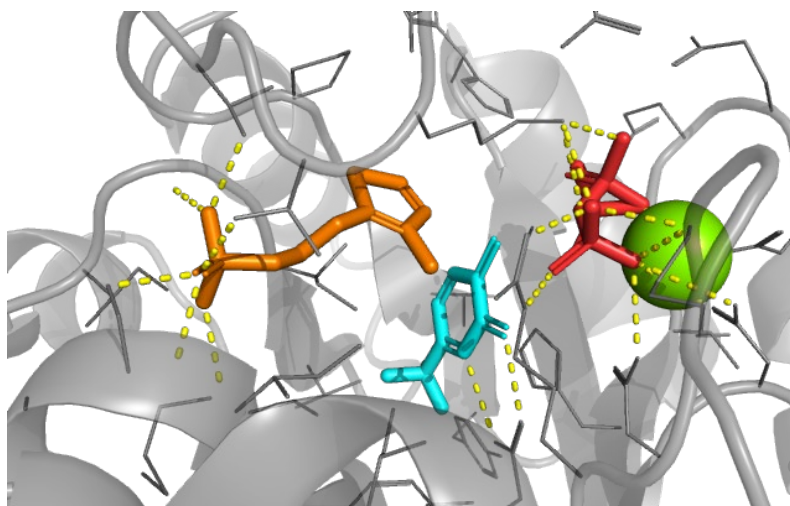
The active site identified by DoGSite3 consisted of 37 residues with specific roles in ligand binding. These residues can be grouped into two main categories: (1) those associated with substrate (thiazole and pyrimidine) interaction and (2) those associated with cofactor/magnesium binding (e.g., ASP93, ASP94, LYS141). The residues composing the active site were functionally categorized and color-coded in Figure 2 to distinguish those involved in substrate recognition from those contributing to cofactor stabilization. This visual clarity supports the subsequent analysis of ligand interactions through docking.



**Figure 2. Comparison of the active site between the crystal structure 1G6C and the predicted model AFC94415.** (a) View of the active site of the 1G6C crystal structure showing key catalytic residues (cyan sticks) interacting with substrate analogs (orange and magenta sticks), along with the  $Mg^{2+}$  ion (green sphere). Relevant hydrogen bonds and coordination interactions are depicted as dashed lines. (b) Predicted active site cavity in the AFC94415 model. The binding cavity is shown as a cyan surface, and surrounding residues are displayed as yellow sticks. A notable conformational difference can be observed when compared to 1G6C, suggesting a more closed or restricted binding site in the predicted model.

### 3.5 Superimposition and Preparation of AFC94415 for Molecular Docking

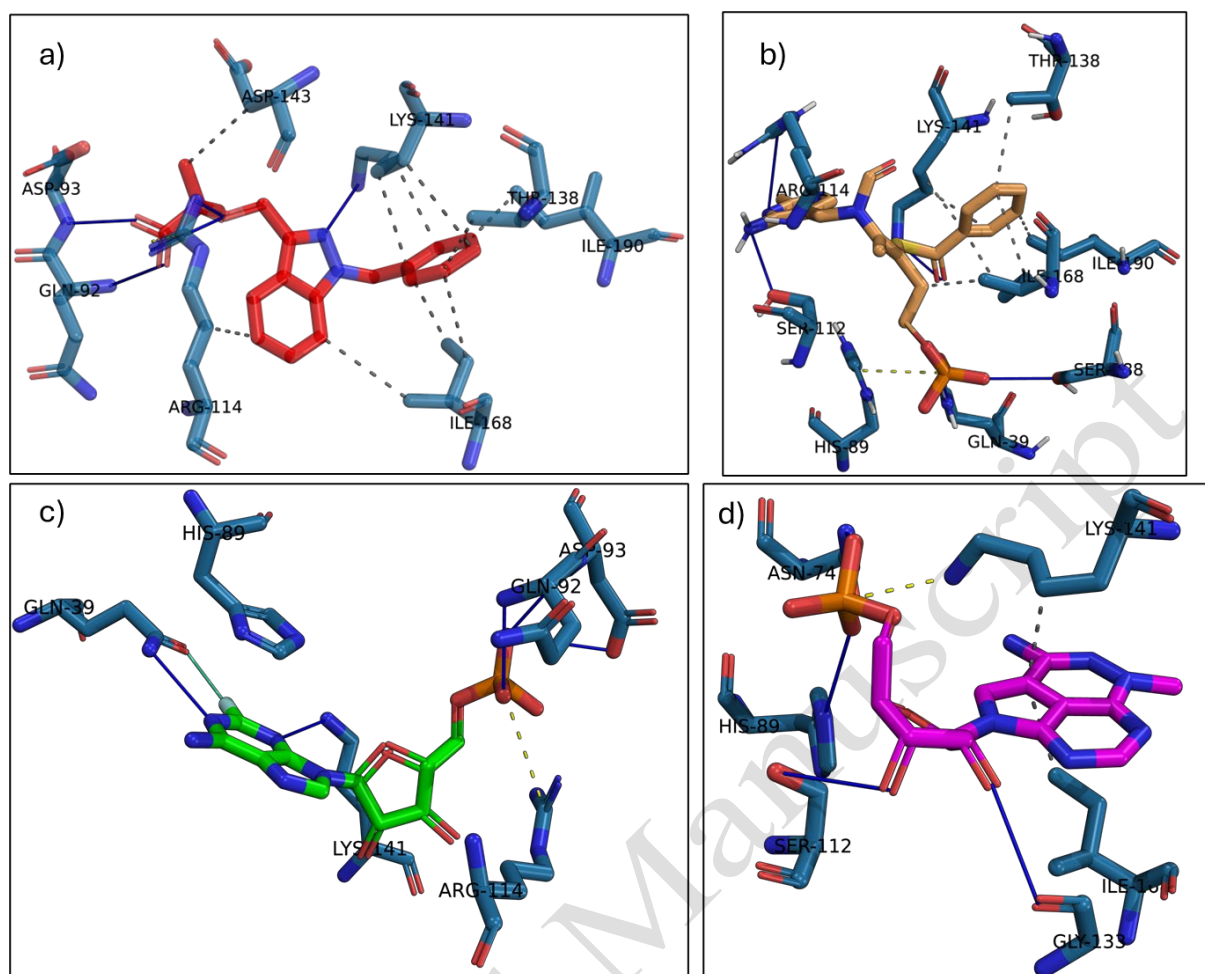
It is important to highlight that both the proteins coded as 1G4T and 1G6C belong to the same research and are derived from the same bacterial species, *Bacillus subtilis*, making it possible to substitute the protein 1G4T with 1G6C. Despite the relevance of 1G4T, the focus of the research is particularly on the structure represented by the code 1G6C due to the specific ligands it contains, as shown in Figure 3.



**Figure 3.** Monomer of enzyme 1G6C, showing the ligands and their intermolecular interactions. Active site, showing the substrates: thiazole (orange), pyrimidine (cyan), phosphate (red), and the magnesium ion (green).

### 3.6 Molecular Docking with AutoDock Vina

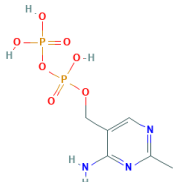
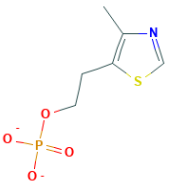
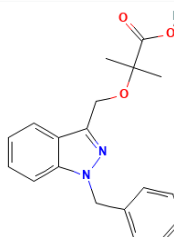
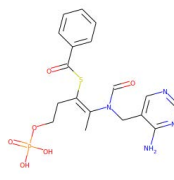
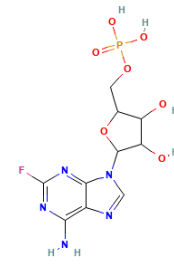
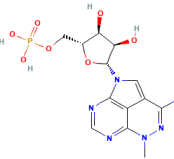
The results obtained with the AutoDock Vina program are shown in Table 3 and the interactions with the predicted model AFC94415 are illustrated in Figure 4. The values of the binding free energy (Kcal/mol) from lowest to highest for the AFC94415 structure and the 1G6C structure are presented. Additionally, the native ligands (thiamine monophosphate, pyrimidine, and thiazole) are detailed.



**Figure 4.** Molecular interactions of selected ligands docked into the predicted active site pocket of thiamine phosphate synthase (model AFC94415). (a) Bindarit, (b) Benfotiamine, (c) Fludarabine (Fludura), and (d) Triciribine phosphate. The protein is shown as blue sticks, and each ligand is represented in a distinct color. The blue lines indicate hydrogen bonds, the yellow dashed lines represent salt bridges, and the pale-yellow dashed lines correspond to hydrophobic interactions. The ligands contact with key catalytic residues, suggesting potential binding affinity and compatibility with the predicted binding site geometry.

**Table 3.** Molecular docking results between proteins and ligands.

Ligand: Database ID	2D Structure	Ligand Commercial Name	AFC94415			1G6C		
			Binding Energy (Kcal/mol)	Amino Acids and Types of Bonds		Binding Energy (Kcal/mol)	Amino Acids and Types of Bonds	
				Hydrogen Bond Interactions	Hydrophobic Interaction		Hydrogen Bond Interactions	Hydrophobic Interaction
PubChem: <a href="#">1131</a>		Thiamine monophosphate (product)	-9.75	Gln39, Asn74, Ser140, Lys141, Gly169, Gly170, Ile171,	Tyr10, Lys141, Ile190, His89 <sup>a</sup>	-8.61	Gln49, Asn84, Thr148, Lys151, Gly179, Val176, Gly180, His99 <sup>a</sup> Ile181,	

			Ile190, Ser191		Ile200, Ser201
PubChem: <a href="#">217</a>	 <i>Pyrimidine (substrate)</i>	-7.60	Gln39, Gln92, Lys141, Arg114+, Lys141 <sup>c</sup>	Tyr10, His89, His89 <sup>a</sup>	-7.59 Gln49, Arg51, Asn84, Val176, Arg51+, Ile178, Lys53+, His99 <sup>a</sup> , Lys151+
PubChem: <a href="#">25245616</a>	 <i>Thiazole (substrate)</i>	-6.90	Ser140, Lys141, Gly169, Gly170, Ile171, Ile190, Ser191	Ile12, Lys141, Ile168	-6.46 Thr148, Gly179, Gly180, Ile181, Ile200, Ser201, Thr148, Lys151, Ile178
PubChem: <a href="#">71354</a>	 Bindarit	-9.78	Gln92, Asp93, Arg114, Lys141, Arg114 <sup>c</sup>	Arg114, Thr138, Lys141, Asp143, Ile168, Ile190,	-8.47 Ala122, Pro144, Thr148, Lys151, Thr154, Ile178, Met199, Ile200, Arg51, Asn84, Arg51 <sup>c</sup> , His99 <sup>c</sup> , Lys151 <sup>c</sup>
Zinc20: <a href="#">ZINC000002015559</a>	 Benfotiamine	-9.43	Gln39, Gln92, Ser112, Arg114, Lys141, Ser188, His89 <sup>c</sup>	Thr138, Lys141, Ile168, Ile190,	-7.57 Gln49, Arg51, Glu103, His124, Lys151, Ser198, His99 <sup>c</sup> , Thr148, Lys151, Thr154, Ile178, Ile200, Lys151 <sup>d</sup>
PubChem: <a href="#">3368</a>	 Fludura	-8.92	Gln39, Gln92, Asp93, Lys141, Arg114 <sup>c</sup> , Gln39 <sup>b</sup>	His89 <sup>a</sup>	-7.79 Gln49, Arg51, Lys53, Asn84, Lys151, Ser198, His99+
PubChem: <a href="#">43860</a>	 Triciribine Phosphate	-8.79	Asn74, His89, Ser112, Gly133, Lys141, Lys141 <sup>c</sup>	Lys141, Ile168	-7.26 Arg51, Asn84, Lys151, Lys151 <sup>d</sup> , Ile178,



- + Data on molecular interactions obtained from AutoDock Vina calculations for the modeled structure and the crystallized structure 1G6C.
  - + The table shows the comparison between the modeled structure AFC94415 and the crystallized structure 1G6C. Additionally, the reference ligands have been included.
  - + The selection by the online program PLIP includes various types of intermolecular interactions.
- <sup>a</sup>:  $\pi$ -Stacking      <sup>b</sup>: Halogen Bond      <sup>c</sup>: Salt Bridges      <sup>d</sup>: Cation- $\pi$  Interaction

### 3. Discussion

The comparative analysis between the amino acid sequence of thiamine phosphate synthase from *Streptococcus pneumoniae* (accession AFC94415.1) and crystallized structures from other bacterial species revealed a high coverage of 98%, but only moderate sequence identity (41.75%) with structures such as PDB 1G4T and PDB 3O15 from *Bacillus subtilis* [30]. Other entries, including 1G67, 1G69, 1G4E, 1G4P, and 3O16, showed a slightly lower identity of 41.26%. More distant structures such as 1XI3, 3NL2, and 3NM1 had identities below 40%, highlighting the evolutionary divergence of this enzyme across species. These findings confirmed the absence of a resolved crystal structure for thiamine phosphate synthase from *S. pneumoniae* in the Protein Data Bank (PDB), justifying the use of AlphaFold2 for predictive structural modeling [19].

The AlphaFold2-predicted structure exhibited the typical TIM barrel fold [41] and a homodimeric arrangement consistent with the enzyme's catalytic function. However, visual inspection revealed a closed conformation of the active site, a known limitation of ligand-free predictions. To simulate a physiologically open active site suitable for docking, the model was aligned with PDB 1G6C from *Bacillus subtilis*, which contains natural substrates and cofactors in its crystal structure [17]. This alignment allowed ligand transfer and structural adjustment, followed by energy minimization using the YASARA Minimization Server [32], thus enabling accurate docking simulations with a properly exposed active site.

For the ligand selection, the Zinc20 [27] and PubChem [28] databases were used, extracting a total of 5,903 and 9,801 compounds, respectively. A ligand-based virtual screening was performed using the LiSiCa extension [20], selecting molecules with structural similarity to the natural substrates (thiazole and pyrimidine) and the product (thiamine monophosphate) of the enzyme. After removing duplicates and phosphoric acid derivatives, a final list of 54 compounds was established for docking analysis.

Molecular docking using AutoDock Vina [22] revealed that several compounds had strong binding affinity to both the AlphaFold2-predicted structure (AFC94415) and the crystallized reference structure (1G6C). Among these, Bindarit demonstrated the highest binding affinity (−9.78 kcal/mol with AFC94415 and −8.47 kcal/mol with 1G6C), surpassing the binding energy of the natural product thiamine monophosphate (−9.75 kcal/mol) [22], Table 3. Binding interactions involved hydrogen bonds with key residues (GLN92, ASP93, ARG114, and LYS141), as well as hydrophobic contacts with residues such as THR138, ASP143, ILE168, and ILE190—all localized within the active pocket identified by DoGSite3 [33].



The biological relevance of targeting thiamine phosphate synthase lies in its essential role in the biosynthesis of thiamine monophosphate, a precursor for vitamin B1 [15, 16]. Since humans lack the thiamine biosynthetic pathway and rely on dietary intake, inhibiting this bacterial enzyme offers a high degree of selectivity, minimizing the risk of cross-reactivity and adverse effects. This pathway has been proposed previously as a promising target for antibiotics development [14, 15], yet specific studies on thiamine phosphate synthase from *S. pneumoniae* remain limited.

This study advances the state-of-the-art by proposing a reproducible *in silico* pipeline for early identification of inhibitors against a critical bacterial enzyme. The integration of AlphaFold2 structural modeling, ligand-based filtering, and docking simulations provides a robust foundation for rational drug discovery. Although these findings are theoretical, they offer a convincing argument for experimental validation through molecular dynamics simulations and enzymatic inhibition trials. Particularly, given its pharmacological profile and binding properties observed here [42, 43], Bindarit emerges as a strong candidate for future *in vitro* and *in vivo* testing.

#### 4. Conclusions

This study used a structure-based virtual screening approach to identify potential inhibitors of thiamine phosphate synthase from *Streptococcus pneumoniae*, using predictive modeling and molecular docking. The absence of an experimentally determined structure for the enzyme was addressed by generating a high-confidence model with AlphaFold2, adjusted to expose the active site through structural alignment with homologous templates.

A total of 54 bioactive compounds were screened and six candidates demonstrated promising binding affinity. Notably, Bindarit exhibited the strongest binding energy, suggesting its potential as a selective inhibitor of thiamine phosphate synthase.

These findings highlight the relevance of thiamine phosphate synthase as a strategic antimicrobial target, given its essential role in bacterial metabolism and absence in humans. The computational pipeline presented here lays a foundation for the rational discovery of novel antibiotics against multidrug-resistant *S. pneumoniae* strains. Future studies should focus on validating these computational predictions through *in vitro* enzymatic trials and *in vivo* infection models, aiming to confirm the inhibitory effect of Bindarit and develop it as a potential antimicrobial therapeutic agent.

#### 5. Declaration of competing interest

We declare that we have no significant competing interests including financial or non-financial, professional, or personal interests interfering with the full and objective presentation of the work described in this manuscript.

## 6. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

## 7. Author contributions

G. M. Sánchez-Rojas: Conceptualization, methodology, formal analysis, scientific literature review, statistical analysis and interpretation of data, prepared the text. A. E. Yacarini-Martínez: Data curation, interpretation of data, prepared the text. E. G. Suclupe-Farro: Conceptualization, methodology, formal analysis, research, data curation, writing- original draft preparation, prepared the text and project administration.

## 8. Data availability statement

The authors confirm that the data supporting the findings of this study are available within the article or its supplementary materials.

## 9. References

- [1] Antimicrobial Resistance Collaborators, "Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis," *Lancet* (London, England), vol. 399, no. 10325, pp. 629–655, 2022. [Online]. Available: [https://doi.org/10.1016/S0140-6736\(21\)02724-0](https://doi.org/10.1016/S0140-6736(21)02724-0)
- [2] S. Guan, K. Zhu, Y. Dong, H. Li, S. Yang, S. Wang, and Y. Shan, "Exploration of Binding Mechanism of a Potential *Streptococcus pneumoniae* Neuraminidase Inhibitor from Herbaceous Plants by Molecular Simulation," *International Journal of Molecular Sciences*, vol. 21, no. 3, p. 1003, 2020.
- [3] Organización Mundial de la Salud, "Datos recientes revelan los altos niveles de resistencia a los antibióticos en todo el mundo," 29 de enero de 2018. [Online]. Available: <https://www.who.int/es/news/item/29-01-2018-high-levels-of-antibiotic-resistance-found-worldwide-new-data-shows>
- [4] C. T. Walsh and T. A. Wencewicz, "Prospects for new antibiotics: a molecule-centered perspective," *The Journal of Antibiotics*, vol. 67, no. 1, pp. 7-22, 2014. [Online]. Available: <https://doi.org/10.1038/ja.2013.49>
- [5] J. M. Munita and C. A. Arias, "Mechanisms of Antibiotic Resistance," *Microbiol Spectrum*, vol. 4, no. 2, pp. 485-511, 2016. [Online]. Available: <https://doi.org/10.1128/microbiolspec.VMBF-0016-2015>
- [6] D. L. Nelson and M. M. Cox, *Principios de bioquímica Lehninger*, 7a ed., C. M. Cuchillo Foix, trans., Ediciones Omega, S.A., 2018.
- [7] R. Utili, "Le infezioni da germi gram-positivi resistenti ai trattamenti antibiotici" [Gram-positive bacterial infections resistant to antibiotic treatment], *Annali italiani di medicina interna : organo ufficiale della Società italiana di medicina interna*, vol. 16, no. 4, pp.

- 205–219, 2001. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/11799629>
- [8] J. Aceil y F. Y. Avci, "Pneumococcal Surface Proteins as Virulence Factors, Immunogens, and Conserved Vaccine Targets," *Frontiers in cellular and infection microbiology*, vol. 12, p. 832254, 2022. [Online]. Available: <https://doi.org/10.3389/fcimb.2022.832254>
- [9] A. M. Correa, M. A. Onieva-García, I. López, y N. Montiel, "Enfermedad neumocócica invasiva en el Hospital Costa del Sol: emergencia de serotipos no vacunables" [Invasive pneumococcal disease in Costa del Sol Hospital: replacement by non-vaccinable serotypes], *Revista española de salud pública*, vol. 92, 2018. [Online]. Available: [https://www.sanidad.gob.es/biblioPublic/publicaciones/recursos\\_propios/resp/revista\\_c\\_drom/VOL92/ORIGINALES/RS92C\\_201806034.pdf](https://www.sanidad.gob.es/biblioPublic/publicaciones/recursos_propios/resp/revista_c_drom/VOL92/ORIGINALES/RS92C_201806034.pdf)
- [10] N. Sosa Delgado, D. Martínez Rojas, y J. Lugo, "Serotipos vacunales y no vacunales de *Streptococcus pneumoniae* en niños de Latinoamérica: revisión del último reporte SIREVA II," *CES Medicina*, vol. 34, no. 3, pp. 179–187, 2020. [Online]. Available: <https://doi.org/10.21615/cesmedicina.34.3.1>
- [11] Organización Mundial de la Salud, "Neumonía," 11 de noviembre de 2021. [Online]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/pneumonia>
- [12] G. D. Vásquez Ludeña, "Análisis y situación de salud: Situación epidemiológica de las neumonías bacterianas en el Perú, 2018 – 2022," *Boletín Epidemiológico*, vol. 31, no. SE 35, pp. 1551–1555, 2022. [Online]. Available: [https://www.dge.gob.pe/epipublic/uploads/boletin/boletin\\_202235\\_02\\_115216.pdf](https://www.dge.gob.pe/epipublic/uploads/boletin/boletin_202235_02_115216.pdf)
- [13] A. L. Barra, L. G. Morão, R. F. Gutierrez, I. Polikarpov, C. Wrenger, A. S. Nascimento, and L. Dantas, "Essential Metabolic Routes as a Way to ESKAPE From Antibiotic Resistance," *Frontiers in public health*, vol. 8, no. 26, 2020. [Online]. Available: <https://doi.org/10.3389/fpubh.2020.00026>
- [14] Q. Du, H. Wang, and J. Xie, "Thiamin (vitamin B1) biosynthesis and regulation: a rich source of antimicrobial drug targets?" *International Journal of Biological Sciences*, vol. 7, no. 1, pp. 41–52, 2011. [Online]. Available: <https://doi.org/10.7150/ijbs.7.41>
- [15] S. Singh, B. K. Malik, and D. K. Sharma, "Metabolic pathway analysis of *S. pneumoniae*: an in silico approach towards drug-design," *Journal of bioinformatics and computational biology*, vol. 5, no. 1, pp. 135–153, 2007. [Online]. Available: <https://doi.org/10.1142/s0219720007002564>
- [16] T. P. Begley, D. M. Downs, S. E. Ealick, F. W. McLafferty, A. P. Van Loon, S. Taylor, N. Campobasso, H. J. Chiu, C. Kinsland, J. J. Reddick, and J. Xi, "Thiamin biosynthesis in prokaryotes," *Archives of Microbiology*, vol. 171, no. 5, pp. 293–300, 1999. [Online]. Available: <https://doi.org/10.1007/s002030050713>
- [17] H. J. Chiu, J. J. Reddick, T. P. Begley, and S. E. Ealick, "Crystal structure of thiamin phosphate synthase from *Bacillus subtilis* at 1.25 Å resolution," *Biochemistry*, vol. 38, no. 20, pp. 6460–6470, 1999. [Online]. Available: <https://doi.org/10.1021/bi982903z>
- [18] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990. [Online]. Available: [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- [19] J. Jumper *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, pp. 583–589, 2021. [Online]. Available: <https://doi.org/10.1038/s41586-021->

03819-2

- [20] A. Dilip, S. Lešnik, T. Štular, D. Janežič, and J. Konc, "Ligand-based virtual screening interface between PyMOL and LiSiCA," *Journal of Cheminformatics*, vol. 8, no. 46, 2016. [Online]. Available: <https://doi.org/10.1186/s13321-016-0157-z>
- [21] S. Lešnik, T. Štular, B. Brus, D. Knez, S. Gobec, D. Janežič, and J. Konc, "LiSiCA: A Software for Ligand-Based Virtual Screening and Its Application for the Discovery of Butyrylcholinesterase Inhibitors," *Journal of Chemical Information and Modeling*, vol. 55, no. 8, pp. 1521–1528, 2015. [Online]. Available: <https://doi.org/10.1021/acs.jcim.5b00136>
- [22] J. Eberhardt, D. Santos-Martins, A. F. Tillack, and S. Forli, "AutoDock Vina 1.2.0: New docking methods, expanded force field, and Python bindings," *Journal of Chemical Information and Modeling*, vol. 61, no. 8, pp. 3891–3898, 2021. [Online]. Available: <https://doi.org/10.1021/acs.jcim.1c0020>
- [23] F. Stanzione, I. Giangreco, and J. C. Cole, "Use of molecular docking computational tools in drug discovery," *Progress in Medicinal Chemistry*, vol. 60, pp. 273–343, 2021. [Online]. Available: <https://doi.org/10.1016/bs.pmch.2021.01.004>
- [24] D. B. Kitchen, H. Decornez, J. R. Furr, and J. Bajorat, "Docking and scoring in virtual screening for drug discovery: methods and applications," *Nature Reviews Drug Discovery*, vol. 3, no. 11, pp. 935–949, 2004. [Online]. Available: <https://doi.org/10.1038/nrd1549>
- [25] F. D. Prieto-Martínez, M. Arciniega, and J. L. Medina-Franco, "Acoplamiento Molecular: Avances Recientes y Retos," *TIP revista especializada en ciencias químico-biológicas*, vol. 21, Supl. 1, pp. 65-87, 2018. [Online]. Available: <https://doi.org/10.22201/fesz.23958723e.2018.0.143>
- [26] I. Lanchero Barrios and F. López Vallejo, "Estudio in silico e in vitro de compuestos inhibidores de la enzima lipasa pancreática: una contribución al reposicionamiento de fármacos antiobesidad," *Tesis de Maestría en Ciencias*, Universidad Nacional de Colombia Sede Bogotá Facultad de Ciencias Departamento de Química, 2016. [Online]. Available: <https://repositorio.unal.edu.co/handle/unal/59555?show=full>
- [27] J. J. Irwin, K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J. Mayfield, and R. A. Sayle, "ZINC20—A free ultralarge-scale chemical database for ligand discovery," *Journal of Chemical Information and Modeling*, vol. 60, no. 12, pp. 6065–6073, 2020. [Online]. Available: <https://doi.org/10.1021/acs.jcim.0c00675>
- [28] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, and E. E. Bolton, "PubChem 2023 update," *Nucleic Acids Research*, vol. 51, no. D1, pp. D1373-D1380, 2023. [Online]. Available: <https://doi.org/10.1093/nar/gkac956>
- [29] V. Eskandari, "Repurposing the natural compounds as potential therapeutic agents for COVID-19 based on the molecular docking study of the main protease and the receptor-binding domain of spike protein," *Journal of Molecular Modeling*, vol. 18, no. 153, 2022. [Online]. Available: <https://doi.org/10.1007/s00894-022-05138-3>
- [30] G. Li, F. Z. Hu, X. Yang, Y. Cui, J. Yang, F. Qu, G. F. Gao, and J.-R. Zhang, "Complete genome sequence of *Streptococcus pneumoniae* strain ST556, a multidrug-resistant



- isolate from an otitis media patient," *Journal of Bacteriology*, vol. 194, no. 12, pp. 3294-3295, 2012. [Online]. Available: <https://doi.org/10.1128/jb.00363-12>
- [31] Schrödinger, LLC, "The PyMOL Molecular Graphics System (version 2.5.5)," [Software]. [Online]. Available: <https://pymol.org/2/>
- [32] E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker, and K. Karplus, "Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8," *Proteins*, vol. 77 Suppl 9, Suppl. 9, pp. 114–122, 2009. [Online]. Available: <https://doi.org/10.1002/prot.2257>
- [33] J. Graef, C. Ehrt, and M. Rarey, "Binding site detection remastered: Enabling fast, robust, and reliable binding site detection and descriptor calculation with DoGSite3," *Journal of Chemical Information and Modeling*, vol. 63, no. 10, pp. 3128-3137, 2023. [Online]. Available: <https://doi.org/10.1021/acs.jcim.3c00336>
- [34] A. Volkamer, A. Griewel, T. Grombacher, and M. Rarey, "Analyzing the topology of active sites: On the prediction of pockets and subpockets," *Journal of Chemical Information and Modeling*, vol. 50, no. 11, pp. 2041-2052, 2010. [Online]. Available: <https://doi.org/10.1021/ci100241y>
- [35] A. Volkamer, D. Kuhn, T. Grombacher, F. Rippmann, and M. Rarey, "Combining global and local measures for structure-based druggability predictions," *J. Chem. Inf. Model.*, vol. 52, no. 2, pp. 360-372, 2012. [Online]. Available: <https://doi.org/10.1021/ci200454v>
- [36] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, "Open Babel: An open chemical toolbox," *J. Cheminf.*, vol. 3, no. 33, 2011. [Online]. Available: <https://doi.org/10.1186/1758-2946-3-33>
- [37] M. F. Sanner, "Python: a programming language for software integration and development," *J. Mol. Graphics Modell.*, vol. 17, no. 1, pp. 57-61, 1999.
- [38] S. Forli, R. Huey, M. E. Pique, M. F. Sanner, D. S. Goodsell, and A. J. Olson, "Computational protein–ligand docking and virtual drug screening with the AutoDock suite," *Nat. Protoc.*, vol. 11, no. 5, pp. 905–919, 2016. [Online]. Available: <https://doi.org/10.1038/nprot.2016.051>
- [39] O. Trott and A. J. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *J. Comput. Chem.*, vol. 31, no. 2, pp. 455-461, 2010. [Online]. Available: <https://doi.org/10.1002/jcc.2133>
- [40] M. F. Adasme, K. L. Linnemann, S. N. Bolz, F. Kaiser, S. Salentin, V. J. Haupt, and M. Schroeder, "PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA," *Nucleic Acids Res.*, vol. 49, no. W1, pp. W530–W534, 2021. [Online]. Available: <https://doi.org/10.1093/nar/gkab294>
- [41] R. K. Wierenga, "The TIM-barrel fold: a versatile framework for efficient enzymes," *FEBS Lett.*, vol. 492, no. 3, pp. 193–198, 2001. [Online]. Available: [https://doi.org/10.1016/s0014-5793\(01\)02236-0](https://doi.org/10.1016/s0014-5793(01)02236-0)
- [42] E. Iwasawa *et al.*, "The anti-inflammatory agent bindarit attenuates the impairment of neural development through suppression of microglial activation in a neonatal hydrocephalus mouse model," *J. Neurosci.*, vol. 42, no. 9, pp. 1820-1844, 2022. [Online]. Available: <https://doi.org/10.1523/jneurosci.1160-21.2021>



- [43] S. Oddi *et al.*, "The anti-inflammatory agent bindarit acts as a modulator of fatty acid-binding protein 4 in human monocytic cells," *Sci. Rep.*, vol. 9, no. 1, 2019. [Online]. Available: <https://doi.org/10.1038/s41598-019-51691-y>
- [44] C. H. Serezani *et al.*, "Cyclic AMP: Master regulator of innate immune cell function," *Am. J. Respir. Cell Mol. Biol.*, vol. 39, no. 2, pp. 127-132, 2008. [Online]. Available: <https://doi.org/10.1165/rcmb.2008-0091tr>
- [45] A. Stirban *et al.*, "Benfotiamine counteracts smoking-induced vascular dysfunction in healthy smokers," *Int. J. Vasc. Med.*, vol. 2012, pp. 1-7, 2012. [Online]. Available: <https://doi.org/10.1155/2012/968761>
- [46] K. R. Rai *et al.*, "Fludarabine compared with chlorambucil as primary therapy for chronic Lymphocytic leukemia," *New England J. Med.*, vol. 343, no. 24, pp. 1750-1757, 2000. [Online]. Available: <https://doi.org/10.1056/nejm200012143432402>
- [47] B. Schmitt *et al.*, "Fludarabine combination therapy for the treatment of chronic Lymphocytic leukemia," *Clin. Lymphoma*, vol. 3, no. 1, pp. 26-35, 2002. [Online]. Available: <https://doi.org/10.3816/clm.2002.n.008>
- [48] B. Faubert *et al.*, "The AMP-activated protein kinase (AMPK) and cancer: Many faces of a metabolic regulator," *Cancer Lett.*, vol. 356, no. 2, pp. 165-170, 2015. [Online]. Available: <https://doi.org/10.1016/j.canlet.2014.01.018>
- [49] F. Marín-Aguilar *et al.*, "Adenosine monophosphate (AMP)-activated protein kinase: A new target for nutraceutical compounds," *Int. J. Mol. Sci.*, vol. 18, no. 2, p. 288, 2017. [Online]. Available: <https://doi.org/10.3390/ijms18020288>
- [50] C. R. Garrett *et al.*, "Phase I pharmacokinetic and pharmacodynamic study of triciribine phosphate monohydrate, a small-molecule inhibitor of AKT phosphorylation, in adult subjects with solid tumors containing activated AKT," *Investing. New Drugs*, vol. 29, no. 6, pp. 1381-1389, 2011. [Online]. Available: <https://doi.org/10.1007/s10637-010-9479-2>