IOSUD – „DUNĂREA DE JOS" UNIVERSITY OF GALAȚI

Doctoral School of Mechanical and Industrial Engineering

ABSTRACT OF Ph.D. THESIS

AUTOMATION METHODS FOR THE CHARACTERIZATION AND DETECTION OF BIOAVAILABLE SUBSTANCES

PhD candidate,

Eng. Cristina NEDELCU (PARASCHIV)

Scientific Coordinator,

Prof. univ. dr. Mirela PRAISLER

Seria I 4 Industrial Engineering Nr. 100

GALAȚI

2024

IOSUD –„DUNĂREA DE JOS" UNIVERSITY OF GALAȚI

Doctoral School of Mechanical and Industrial Engineering

ABSTRACT OF Ph.D. THESIS

AUTOMATION METHODS FOR THE CHARACTERIZATION AND DETECTION OF BIOAVAILABLE SUBSTANCES

PhD candidate

Eng. Cristina NEDELCU (PARASCHIV)

Chair	Prof. univ. dr. eng. Cătălin FETECĂU, „Dunărea de Jos" University of Galați
Scientific Coordinator	Prof. univ. dr. Mirela PRAISLER, „Dunărea de Jos" University of Galați
Scientific referees	Prof. univ. dr. eng. Gheorghe NAGÎȚ, Technical University „Gh. Asachi" of Iași
	Prof. univ. dr. eng. Remus ZĂGAN, Maritime University of Constanța
	Prof. univ. dr. Dana TUTUNARU, „Dunărea de Jos" University of Galați

Seria I 4 Industrial Engineering Nr. 100

GALAȚI

2024

Seriile tezelor de doctorat susținute public în UDJG începând cu 1 octombrie 2013 sunt:

Domeniul fundamental ȘTIINȚE INGINEREȘTI

Seria I 1: **Biotehnologii**
Seria I 2: **Calculatoare și tehnologia informației**
Seria I 3: **Inginerie electrică**
Seria I 4: **Inginerie industrială**
Seria I 5: **Ingineria materialelor**
Seria I 6: **Inginerie mecanică**
Seria I 7: **Ingineria produselor alimentare**
Seria I 8: **Ingineria sistemelor**
Seria I 9: **Inginerie și management în agicultură și dezvoltare rurală**

Domeniul fundamental ȘTIINȚE SOCIALE

Seria E 1: **Economie**
Seria E 2: **Management**
Seria E 3: **Marketing**
Seria SSEF: **Știința sportului și educației fizice**
Seria SJ: **Drept**

Domeniul fundamental ȘTIINȚE UMANISTE

Seria U 1: **Filologie- Engleză**
Seria U 2: **Filologie- Română**
Seria U 3: **Istorie**
Seria U 4: **Filologie - Franceză**

Domeniul fundamental MATEMATICĂ ȘI ȘTIINȚE ALE NATURII

Seria C: **Chimie**

Domeniul fundamental ȘTIINȚE BIOMEDICALE

Seria M: **Medicină**
Seria F: **Farmacie**

Foreword (Acknowledgments)

At the completion of this significant stage in my life, I wish to express my gratitude to those who have guided me and supported me with their constant assistance and encouragement throughout this doctoral work.

First of all, I wish to express my deep gratitude to my scientific advisor, **prof. dr. Mirela PRAISLER**, for her guidance, patience, and unwavering support throughout the entire research and writing process. I am grateful for her accepting to be my mentor. I would like to emphasize that without her expertise, encouragement, and support, this work would not have been possible.

I owe special gratitude to **lect. dr. Camelia FRIGIOIU** for the trust, support, and encouragement given during the difficult moments of this „academic journey".

A special gratitude goes to **conf. dr. Steluța GOSAV** for the precious time she devoted, the valuable scientific advice, and for her competent and ongoing guidance throughout the research activity that led to the publication of various scientific articles, as well as of the doctoral thesis.

Next, I wish to express my gratitude to the other members of the advisory committee, namely **prof. dr. eng. Luminița Moraru** and **prof. dr. habil. eng. Antoaneta Ene**, for the advice and suggestions provided during the years of study and research.

I would especially like to thank my fellow doctoral colleagues, with whom I enjoyed a pleasant environment and mutual support.

I would also like to especially thank my parents, Ion and Ioana, my son, Andrei, and my extended family for always supporting me in my professional and personal development.

With special gratitude and love, I dedicate this thesis to the memory of my father, as well as to my family and my friends, who have supported me with affection and patience throughout my doctoral studies.

To all of you, once again, thank you!

Galați, 2024 Eng. Cristina Nedelcu (Paraschiv)

Table of contents

Introduction

Bioavailability is defined as the fraction of a substance that is efficiently absorbed from the gastrointestinal tract and enters the bloodstream [1,2]. Several prestigious global organizations, such as the World Health Organization (WHO), the Food and Drug Administration (FDA), and the American Pharmaceutical Association (APA), have provided definitions for this commonly encountered concept in pharmacology, nutrition, and toxicology, as it determines how effectively a nutrient or drug can exert its effects in the body.

As pollution levels increase, people become more susceptible to illness; therefore, it is necessary to evaluate and optimize the bioavailability of both medications and nutrients, as these are crucial for the effectiveness of drug therapy and for ensuring adequate nutrition.

Statistics on cancer cases vary depending on the region, type of cancer, and various demographic factors. Thus, globally, cancer is still the leading cause of mortality, with approximately 10 million deaths recorded in 2020 [3]. On the other hand, according to data presented in the 2023 European Cancer Inequalities Registry, developed by the Organization for Economic Co-operation and Development (OECD), cancer is the second leading cause of death in Romania, with a rate higher than the EU average. Lung, colorectal, and breast cancer are the leading types of cancer-causing death in Romania. Although lung cancer mortality has decreased in recent years, the incidence of breast cancer is on the rise [4]. The process of discovering and developing new drugs, also known as "*drug design*," involves identifying chemical compounds that have the potential to treat specific diseases or conditions. This process engages a series of advanced stages and techniques, including *target identification*, *compound design*, *synthesis* and *optimization*, *in vitro* and *in vivo testing*, *clinical studies*, *approval*, and *market* launch. The drug discovery process is a multidisciplinary field that involves biology, chemistry, pharmacology, and bioinformatics, with the goal of developing new, effective, and safe therapies for various diseases, but it is also a significant consumer of time and financial resources.

As such, the discovery of new bioavailable substances is crucial in the fight against cancer, as it can lead to more effective and safer treatments, reducing toxicity and improving patients' quality of life. Additionally, this process supports continuous innovation in the field of oncology, contributing to the development of new and personalized solutions for combating cancer. New bioavailable substances can be tailored to meet the specific needs of each patient, based on the genetic and biological characteristics of their tumor, thereby contributing to the advancement of personalized medicine.

The identification and characterization of bioavailable substances necessitate the optimization and integration of advanced technologies in industrial engineering, especially in the development process of drugs and other synthetic substances. Molecular docking employs advanced computational tools, thereby contributing to the reduction of physical experiments required to identify and optimize potential compounds, ultimately conserving material and human resources. By utilizing these techniques, more efficient processes can be designed, risks associated with the development of new drugs are minimized, and drugs can be produced in large quantities with controlled costs and high-quality standards.

Summary and structure of scientific work

The PhD thesis, entitled "*Automation Methods for Characterization and Detection of Bioavailable Substances*", is structured into three chapters, which are accompanied by an introduction related to the studied topic, and followed by general conclusions and future research directions.

The introduction provides a brief overview of the evolution of the main causes of mortality globally and nationally, along with cancer prevention strategies. Chapter 1 presents the general characterization of the bioavailable substances studied, highlighting the two categories examined: cannabinoids and flavonoids. Chapter 2 describes characterization and identification methods that have proven to be highly effective for the analyzed bioavailable substances, respectively identification methods using the principles of IR spectrometry, characterization methods such as DFT, and molecular docking. Chapter 3 outlines the author's own scientific achievements about the inhibitory efficacy of resokaempferol and tectochrysin on PI3Kα and H1047R. PI3Kα contributes to cell proliferation while H1047R is an oncogenic mutation associated with increased enzymatic activity and, consequently, with uncontrolled growth of tumor cells. In this chapter, the electronic parameters of the two compounds were evaluated, their physico-chemical profiles were presented, and molecular interactions were analyzed through molecular docking. In this chapter, the electronic parameters for four flavonoids were evaluated, their physico-chemical profiles were presented, and the molecular interactions were analyzed through molecular docking for three compounds. Regarding the compounds belonging to the class of cannabinoids, they were characterized in terms of bioavailability. Several methods that can be used for their identification using ATR-FTIR spectra were also presented.

The conclusions and future directions section highlights the innovative character of the obtained results and outlines the main research directions. The PhD thesis concludes with a list of scientific papers presented at national or international conferences, as well as those published in scientific journals.

Motivation of choosing the research theme

The contemporary world faces a range of significant health issues, reflecting the complexity and diversity of global challenges, including emerging infectious diseases (pandemics), chronic non-communicable diseases (cardiovascular diseases, diabetes, cancer), mental health issues, health problems related to climate change, nutrition and lifestyle (obesity and unhealthy diets), as well as public health policies whose correct implementation can contribute to disease prevention and health promotion.

Studying the properties of bioavailable substances, both synthetic and natural compounds, is essential to assure the efficacy, safety and accessibility of pharmaceutical treatments, contributing to the overall improvement of public health. Identification of natural compounds allows the inclusion of various fruits and vegetables in daily consumption, serving as adjuncts in the treatment of various conditions.

Among the positive effects of discovering new bioavailable substances are: maximizing the therapeutic effect by determining the proportion and rate at which an anticancer drug reaches tumor cells; personalizing the treatment by adapting dosages and administration forms to the individual

needs of patients; developing targeted therapies by identifying substances that specifically attack cancer cells; and accelerating the approval and commercialization process by using chemometric techniques to analyze inhibitory potential, which aids in their evaluation in clinical studies.

From my perspective as an educator, it is essential to promote a healthy lifestyle among young people, as well as to raise their awareness about the need for developing healthy eating habits. Students who learn about nutrition in school can bring this knowledge home, positively influencing their families' eating habits and contributing to a culture of health in their communities. Identifying, especially natural bioavailable compounds, would add value to our daily diet by including specific fruits and vegetables tailored to our health conditions.

The use of advanced computational techniques contributes to the rapid identification of compounds with high therapeutic potential, thereby reducing the time and costs required for experimental testing and streamlining industrial processes for the production and development of new drugs.

Pursued research objectives

The main scientific research objectives pursued during the elaboration of this PhD thesis were:

- Conducting a literature review on:
  - types of bioavailable substances, especially cannabinoids and flavonoids;
  - proteins responsible for the multiplication of cancer cells;
  - IR spectra;
  - spectral methods for characterization and identification of bioavailable compounds;
  - methods for characterization of bioavailable compounds;
  - algorithms for molecular docking;
  - methods for automating the detection of cannabinoids.

- Exploring the inhibitory efficacy of resokaempferol and tectochrysin on PI3Kα and H1047R
  - calculation of electronic parameters;
  - the description of physico-chemical and pharmaco-kinetic profiles;
  - analysis of the molecular interactions between ligands (bioavailable compounds) and the PI3Kα protein, as well as the oncogenic mutant form of H1047R;
  - molecular docking validation and MolProbity analysis, which provides a rigorous and detailed evaluation of molecular structures and ensures that the models are of high quality and exact.

- Exploring the inhibitory efficacy of kaempferol on PI3Kα
  - calculation of electronic parameters;
  - the description of physico-chemical and pharmaco-kinetic profiles;
  - analysis of molecular interactions ligands (bioavailable compounds) and the PI3Kα protein;
  - molecular docking validation in comparison with the standard compound, alpelisib, which is a drug already used in cancer therapy.

- Methods for automating the detection of cannabinoids based on ATR-FTIR spectra.

Chapter 1. Theoretical aspects of bioavailable compounds

1.1 Bioavailable compounds: general notions

Bioavailable substances are those substances that, after being administered in various forms, are absorbed by the circulatory system to be used by the body. Among the types of bioavailable substances, the following can be mentioned: *essential nutrients* (vitamins, minerals), *amino acids*, *fatty acids*, *phytonutrients* (polyphenols, flavonoids, carotenoids), *probiotics* and *prebiotics*, *enzymes, medications*.

The classification of bioavailable substances can be made based on several factors, such as water solubility and permeability through biological membranes. One of the most widely used classifications is the Biopharmaceutics Classification System (BCS), which was developed in 1995 and has become a standard in the regulation of bioequivalence for oral pharmaceutical forms [7,8].

1.2 Presentation of the studied bioavailable substances

1.2.1 Cannabinoids

Cannabinoids are chemical compounds that interact with cannabinoid receptors in the body. There are several types of cannabinoids, including endogenous cannabinoids (produced by the human body), phyto-cannabinoids (found in plants, especially in cannabis), and synthetic cannabinoids (created in a laboratory) [11].

Endogenous cannabinoids, known as endocannabinoids, are naturally produced by the human body and play an essential role in regulating physiological functions by interacting with cannabinoid receptors (CB1 and CB2). Two main groups have been identified: amides of arachidonic acid, which include anandamide, and esters of arachidonic acid with glycerol, which include 2-arachidonoylglycerol, whose chemical structures are presented in figure 1.1.[11–15].

a) anandamide b) 2- arachidonoylglycerol

Figure 0.1. *Main types of endocannabinoids [12,15].*

Natural cannabinoids are terpenophenolic compounds found in the resin secreted by the flowers and leaves of the cannabis plant. These compounds have been found in small amounts in the stem, but are absent from the seeds. To date, over 100 phyto-cannabinoids have been discovered, the most psychoactive being Δ^9 THC [16].

Synthetic cannabinoids have psychoactive effects similar to those of natural cannabinoids. The effects of cannabinoids on the human body are due to the stimulation of a membrane protein called a receptor. So far, two receptors have been fully discovered and characterized: CB1 (a receptor with implications for memory, perception, and mobility) and CB2 (located in the spleen, tonsils, and immune system cells) [12].

While cannabinoids have palliative effects on cancer patients by preventing nausea, vomiting, or pain, and by stimulating the appetite, they also inhibit the growth of tumor cells in cultures and animal models by modulating essential cellular signaling pathways. [11]. Research conducted on cannabinoids has shown that they have anticancer effects. For example, it has been proved that THC is useful in the treatment of breast cancer, liver carcinoma, prostate cancer, lung cancer, and glioblastoma. The synthetic cannabinoid JWH-015 is useful in treating prostate cancer, while WIN 55,212-2 has been successfully used in cases of glioblastoma and prostate cancer [11,12].

1.2.2 Flavonoids

Flavonoids are a distinct class of natural phenolic compounds that are widespread in the plant kingdom. They are known for their contribution to the color, aroma, and taste of fruits, vegetables, and other plants, and are also found in cereals, bark, roots, stems, flowers, tea, and wine. These bioactive compounds, with varied chemical structures, are essential in the human diet due to their antioxidant, anti-inflammatory, anticancer, and cardioprotective effects. They are present in fruits (apples, citrus fruits, berries), vegetables (onions, peppers), teas (green, black), and red wine [2,20–24].

Flavonoids have a basic structure (see figure 1.2) of the flavan type, which includes a skeleton of 15 carbon atoms organized in a model of 2 benzene rings ((A), (B)) linked by a pyranic heterocycle (C). This basic structure is divided into several subgroups, each having variations in the attached functional groups, such as hydroxyl, methoxy, or carbonyl groups [25].

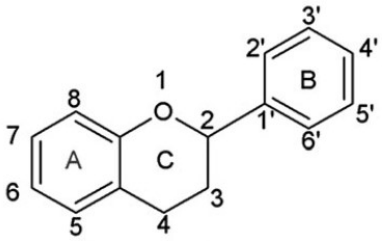

Figure 0.2. *The fundamental chemical structure of flavonoids [21].*

Flavonoids are classified into six groups, based on the involved heterocycle: flavonols, flavones, isoflavones, flavanones, anthocyanidins, and flavanols (catechins and proanthocyanidins) [21,23]. The correspondence between the types of flavonoids and the colors of fruits and vegetables is displayed in figure 1.3.

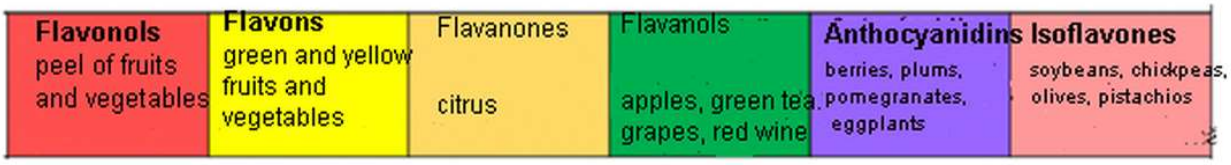

Figure 0.3. *Different colors of polyphenols and the corresponding fruits and vegetables [26].*

Chapter 2. Methods for the characterization and identification of bioavailable compounds

2.1 Spectral methods for characterization and identification of bioavailable compounds

Spectroscopic methods provide precise and versatile tools for identifying bioavailable substances. Each technique has specific advantages and can be used complementarily to achieve a complete characterization of the substances of interest. Depending on the nature of the samples and the required information, the choice of the appropriate spectroscopic method is essential for accurate and efficient analysis. Among the most common spectroscopic methods are Infrared spectroscopy (IR), Nuclear Magnetic Resonance spectroscopy (NMR), the combination of Gas Chromatography (GC) with Fourier-Transform Infrared spectroscopy (GC/FTIR) and Mass spectrometry (MS)[46].

2.1.1 Infrared (IR) spectroscopic analysis

Infrared spectroscopy (IR) is a powerful analytical technique used to study and identify molecules by analyzing the absorption of infrared radiation. This technique is based on the interaction of infrared radiation with molecules, leading to characteristic vibrations of chemical bonds. Understanding the fundamentals of IR is essential for harnessing its capabilities in chemical analysis as a tool for obtaining detailed information about the molecular composition of various substances, molecular vibrations, and chemical bonds [51,52]. IR spectroscopy is used for both qualitative and quantitative analyses, although the latter is not as precise as other techniques such as GC or liquid chromatography (LC). In quantitative analysis, the Lambert–Beer law is used to determine the concentration of a substance.

2.1.2 Presentation of the spectra of bioavailable substances

An example of the absorption spectrum for the synthetic cannabinoid, JWH-019, is presented in figure 2.6. It was recorded by using a diamond ATR (attenuated total reflection) accessory (ATR 3 – diamond 3 Bounce) at 32 scans and a resolution of 4 cm^{-1} [57].

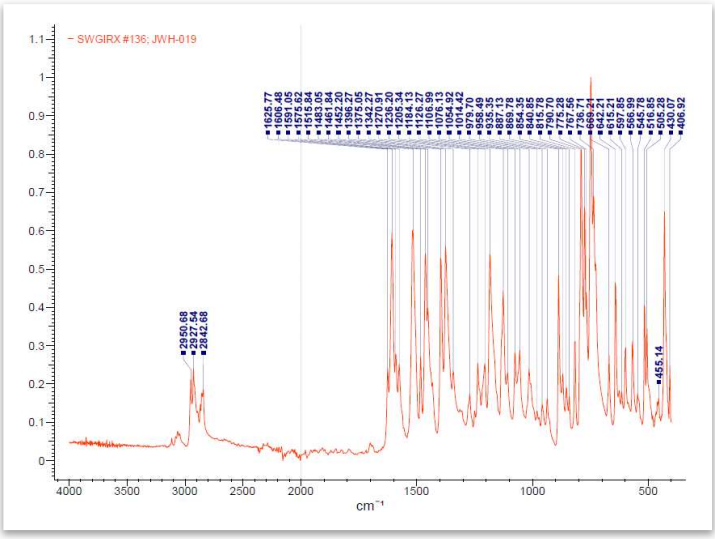

Figure 0.1. *ATR-FTIR spectra for JWH-019 [57].*

Comparison of the theoretical and experimental IR spectra of resokaempferol

We are presenting in figure 2.7 the theoretical IR spectrum of flavonol in comparison with the experimental IR spectrum obtained after structure optimization. The theoretical spectrum was obtained from the online SpectraBase platform [58], then scaled and compared with the experimental one. It was recorded in solid phase from 4000 to 400 cm^{-1} using a Bio-Rad FTS instrument equipped with an ATR-Neat (DuraSamplIR II) technique.

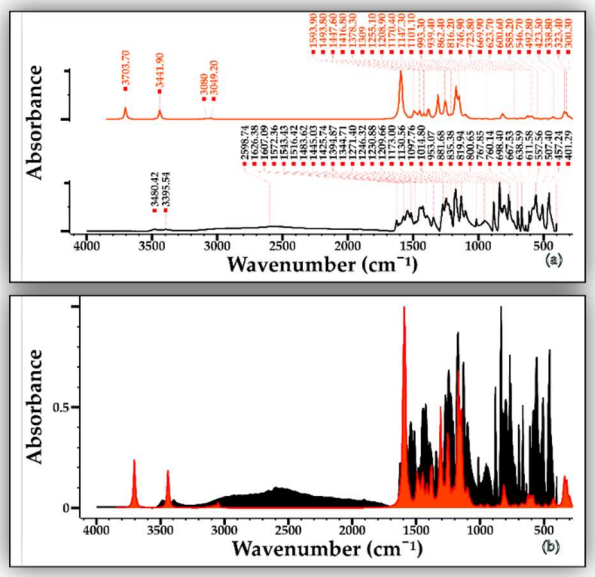

Figure 0.2. *Comparative infrared spectroscopic analysis of resokaempferol: (a) stacked and (b) overlay views of the scaled theoretical IR-derived spectrum (red) alongside the experimental IR spectrum (black).*

Comparison of the theoretical and experimental IR spectra of tectochrysin

We are presenting in figure 2.8 the theoretical IR spectrum of natural flavonoid in comparison with the experimental IR spectrum obtained after structure optimization. The theoretical spectrum was sourced from the online SpectraBase platform [58], then scaled and compared with the experimental one. It has been recorded in solid phase from 4000 to 400 cm^{-1} using a Bio-Rad FTS instrument equipped with an ATR-Neat (DuraSamplIR II) technique.

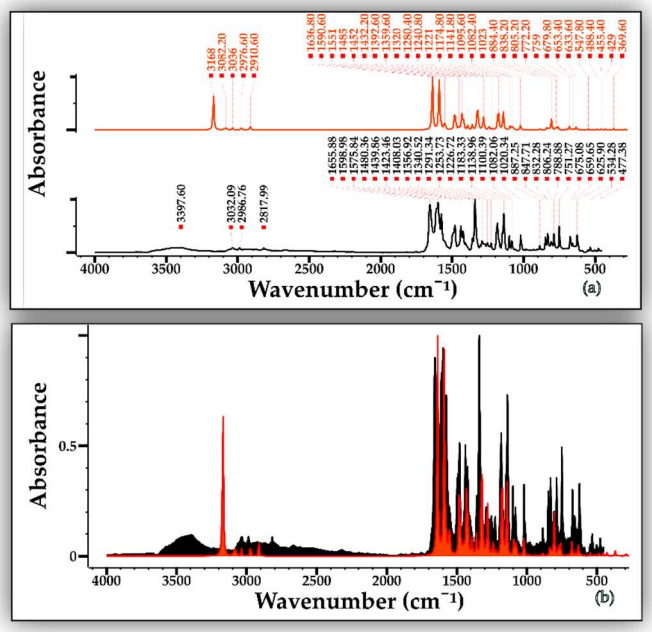

Figure 0.3. *Comparative infrared spectroscopic of tectochrysin: (a) stacked and (b) overlay presentations of the scaled theoretical IR-derived spectrum (red) with the experimental IR spectrum (black).*

Comparison of the theoretical IR and experimental spectra of the flavanone 3-(N-morpholinylcarbodithioate)-2-(4-methoxyphenyl)chroman-4-one

We are presenting in figure 2.9 the theoretical and experimental spectra of the flavanone 3-(N-morpholinylcarbodithioate)-2-(4-methoxyphenyl)chroman-4-one. It has been recorded on a Brucker Tensor 27 FT-IR spectrometer with an ATR device (Ge-ATR crystal), in the range of 4000–400 cm^{-1}, at room temperature, with a spectral resolution of 2.0 cm^{-1} and 32 added scans [59].

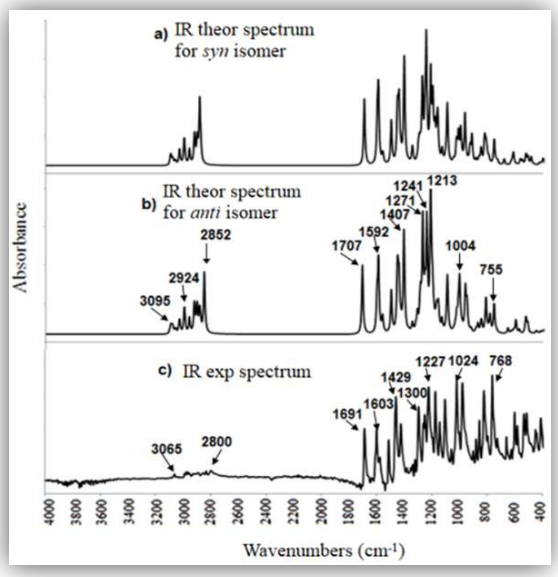

Figure 0.4. *Comparative infrared spectroscopic analysis of 3-(N-morpholinylcarbodithioate)-2-(4-methoxyphenyl)chroman-4-one [59].*

2.2 Analysis of the physicochemical properties of bioavailable substances based on molecular descriptors

QSAR (Quantitative Structure-Activity Relationship) modeling is a method used to predict the biological activity of chemical compounds based on their structural and physico-chemical characteristics. Adherence to standards and protocols, which impose restrictions on molecular properties such as molecular weight, partition coefficient (logP), hydrogen bond acceptors and donors, topological polar surface area, the number of rotatable and rigid bonds, as well as the ratio of heteroatoms to carbon, has increased the likelihood of identifying compounds with successful therapeutic applications [60,61].One of the tools used to evaluate the ADMET (*Absorption, Distribution, Metabolism, Excretion*, and *Toxicity*) properties of the studied compounds is the *radar chart*.

2.2.1 Cannabinoids

To analyze certain physico-chemical properties, five cannabinoids were compared with THC (tetrahydrocannabinol), as it is one of the most studied and well-known compounds and is often used as a reference standard.

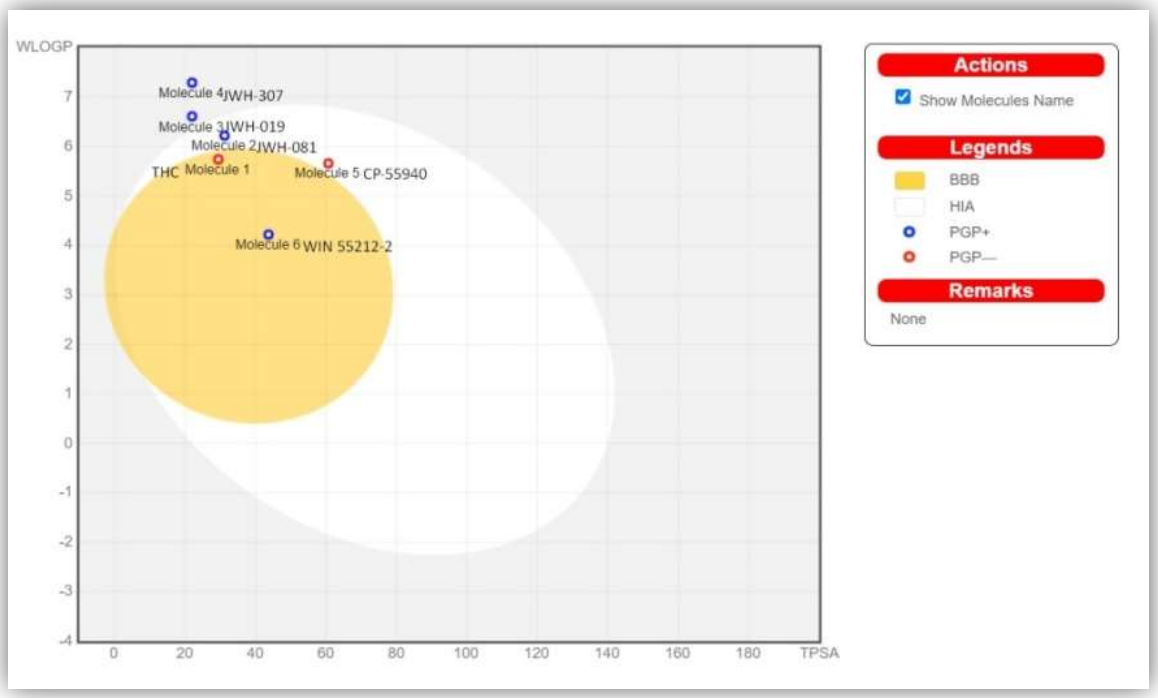

Figure 0.2. *Evaluation of the gastro-intestinal absorption and blood-brain barrier penetration for the studied synthetic cannabinoids [14].*

Figure 2.12 presents the radar charts that define the optimal physico-chemical space for the oral bioavailability of the five studied synthetic cannabinoids (CP-55940, WIN 55212-2, JWH-019, JWH-081, JWH-307).

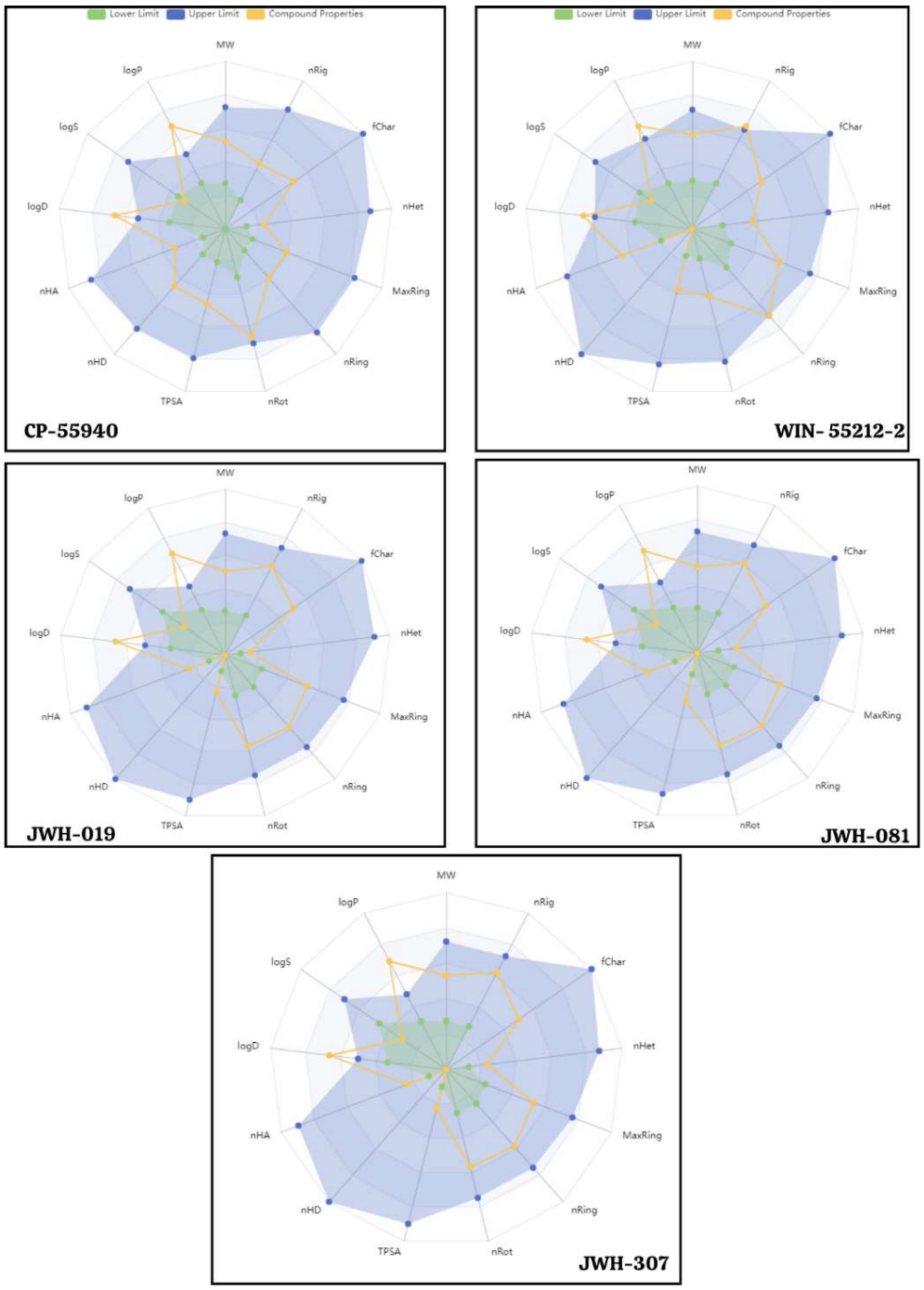

Figure 0.5. *Evaluation of the bioavailability of the five studied synthetic cannabinoids obtained with ADMETlab3.0 [81].*

2.2.2 Flavonoids

The radar charts for resokaempferol and tectochrysin are presented in Figure 2.13. They indicate that resokaempferol has more favorable properties linked to bioavailability than tectochrysin.

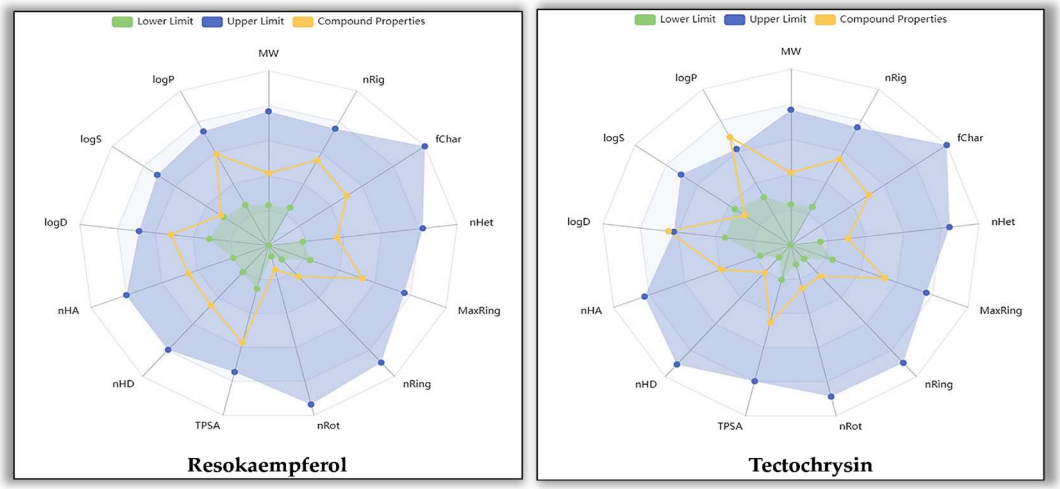

Figure 0.6. *Evaluation of the bioavailability of resokaempferol and tectochrysin obtained with ADMETlab3.0 [81,84].*

The assessment of kaempferol presented in Figure 2.14 indicates that this flavonoid also has properties specific to bioavailable substances.

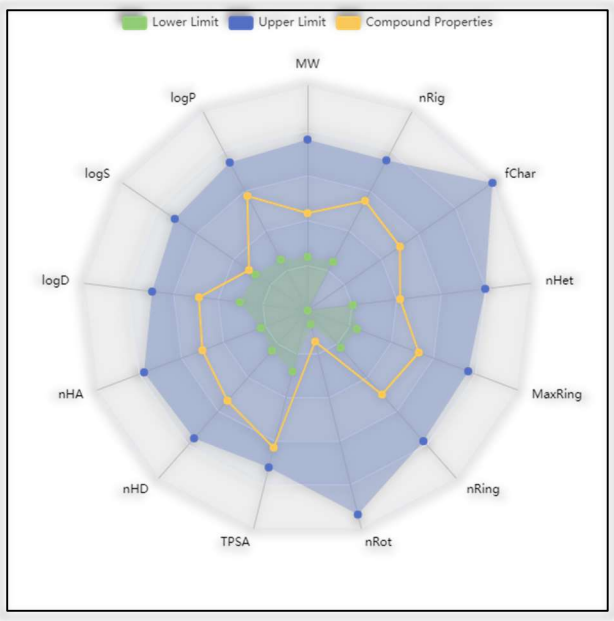

Figure 0.7*. Evaluation of the bioavailability of kaempferol obtained with ADMETlab3.0 [81].*

2.3 Application of the DFT methods and chemometric techniques in the characterization of bioavailable substances

One of the main computational methods used in this research is ***Density Functional Theory - DFT***.

2.3.1 Density Functional Theory – DFT

The functional theory of density, known as DFT (Density Functional Theory), is one of the most widely used methods in computational chemistry and condensed matter physics for studying the electronic properties of atoms, molecules, and solids. DFT is essential in various fields, including materials science, quantum chemistry, and biophysics.

2.3.2 Hohenberg and Kohn theorems

The beginnings of DFT (1926) are connected to the development of quantum mechanics and electronic structure theory. Early studies were influenced by the works of Schrödinger and Dirac. In 1964, Pierre Hohenberg and Walter Kohn laid the theoretical foundations of DFT through two fundamental theorems. The first is an existence theorem that presents the ground state energy of a system as a functional of the density. The second theorem provides a variational principle and, thus, at least in principle, a way to find the optimal density: searching for the one that yields the lowest energy, as has been known for decades in wave function-based quantum mechanics [86].

2.3.3 Kohn-Sham formalism

The Kohn-Sham approach involves replacing the real system of interacting electrons (classical and non-classical interactions) with an auxiliary system of non-interacting electrons that has the same density as the physical system and can be described by a single wave function with N orbitals [85].

2.3.4 B3LYP- hybrid functional

Over time, there have been several attempts to approximate the total energy of a system. The most widely used and known hybrid functional is ***B3LYP (Becke, 3-parameter, Lee-Yang-Parr)***, which provides a relatively inexpensive yet quite accurate way to predict molecular structures, energies and frequencies [22,142].

2.4 Molecular docking

Molecular docking is a computational technique used since the 1980s to study interactions between molecules, particularly to predict how two molecules, usually a ligand and a target protein, bind to each other. This method is essential in drug discovery and bioinformatics, with significant applications in drug design and optimization. The main goal of molecular docking is to understand and predict molecular recognition from both a structural perspective (i.e., finding possible binding modes) and an energetic perspective (i.e., predicting binding affinity) [89].

The 3D structures of the target macromolecule and the small molecule must first be selected, and then each structure must be prepared according to the requirements of the docking method used.

2.4.1. Preparing receptor and ligand

One of the objectives of this study is to investigate the inhibitory capacity of the *PI3Kα* kinase and the *H1047R* mutation by the native ligands *VYP* and *UE9*, as well as by the new compounds *resokaempferol* and *tectochrysin*.

In this research, the molecular docking analyses was performed by using three software applications: AutoDock, Vina, and Glide. Preparing receptor and ligand is a step that involves preparing the receptor and the ligand molecules as inputs for docking calculations, which predict the orientations of a ligand within the active site of a receptor. We utilized the tools available in Glide, and we employed its default framework, the Protein Preparation Workflow, which includes preprocessing, optimizing H-bond assignments, and cleanup steps. For the ligands, the preparation follows a generally simpler yet similar approach by using Glide LigPrep in conjunction with Epik from the Maestro Suite, which included corrections, conversions, and optimizations of the structures at. pH 7.0±2.0 [84,100–104].

2.4.2 Docking algorithms

AutoDock (AD4) uses a computational algorithm known as the *Lamarckian Genetic Algorithm* (LGA) to simulate and predict how a ligand might bind to a receptor [89,90,105–111]. The *Vina* software employs a global optimization algorithm known as a *Gradient-Based Local Search Genetic Algorithm* to predict the binding modes of small molecules to their protein targets [90,108,112–116].

The docking methodology (developed by Schrödinger) used by the *Glide XP* software incorporates a sophisticated hierarchy of filters to explore potential docking sites within the receptor's binding region. Glide employs the Emodel scoring function to differentiate between various protein-ligand complexes for a given ligand, and the GlideScore function to rank compounds. This complex scoring system provides a comprehensive assessment of potential inhibitors, highlighting those with optimal binding characteristics [96,105–107,117–120].

Chapter 3. Personal contributions regarding characterization and identification of flavonoids and cannabinoids

3.1 Exploring the inhibitory efficacy of resokaempferol and tectochrysin on PI3Kα and H1047R [84]

3.1.1 Analysis of the optimized molecular configurations

The molecular geometries of potential ligands investigated in this study were sourced from the *Metabolome Database* (HMDB) and the *PhytoChemical Interactions Database* (PCIDB) [121,122] and are the following:

- Flavonol: *3,7-dihydroxy-2-(4-hydroxyphenyl)chromen-4-one* (resokaempferol, ID: HMDB0034004)
- Flavone*: 5-hydroxy-7-methoxy-2-phenylchromen-4-one* (tectochrysin*, ID: C00003795)*

Figure 3.1 presents optimized structure of resokaempferol and tectochrysin. The numbers in the 3D structure denote the specific atoms within each molecule. These numerical labels are used to identify and differentiate between the various atoms for clarity in chemical structure analysis. Specifically, each number corresponds to a unique position within the molecule, helping in the precise understanding of their spatial arrangement and bonding interactions.

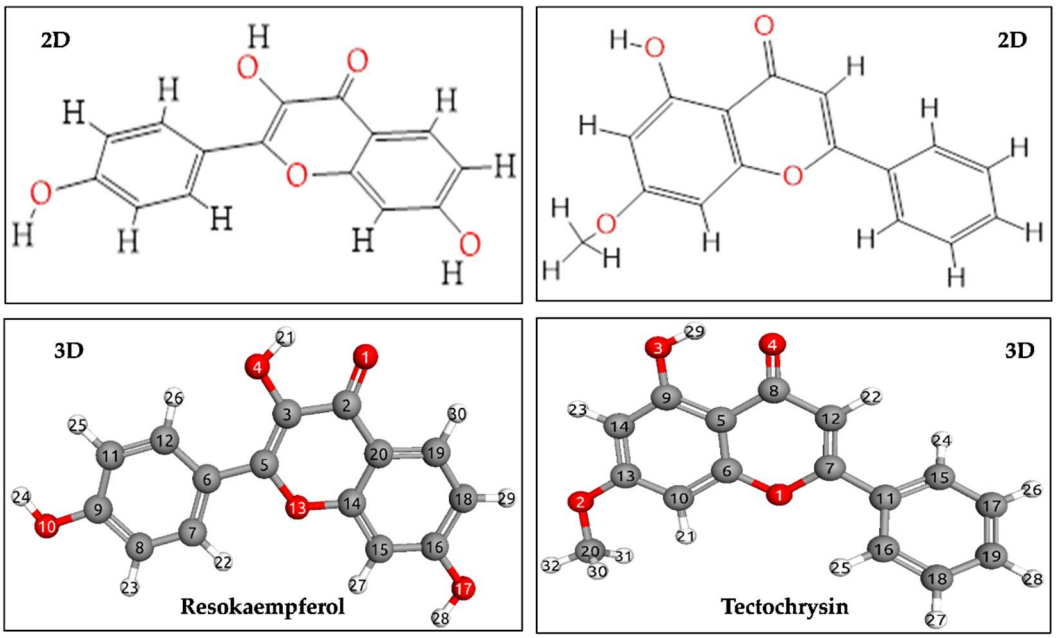

Figure 0.1. *2D and 3D molecular structures of the ligands resokaempferol and tectochrysin [84].*

The molecular configuration of resokaempferol, composed from 30 atoms, features 84 distinct normal modes of vibration. To effectively capture these vibrational dynamics, 84 linearly independent internal coordinates are utilized, including 21 C-H modes (numbered 4-10, 36-42, and 61-67). This ligand showcases a varied set of internal vibrational coordinates, consisting of 29 stretching modes (numbered 1-29), 28 bending modes (numbered 30-57), and 27 torsion modes (numbered 58-84).

The molecular structure of tectochrysin, comprising 32 atoms, exhibits 90 distinct normal modes of vibration. These vibrational motions are adequately represented by 90 linearly independent internal coordinates. Among these, 33 C-H modes (numbered 2 through 12, 36 through 46, and 63 through 73) are employed. The compound exhibits a diverse range of internal vibrational coordinates, including 31 distinct stretching modes (numbered 1-31), 30 bending modes (numbered 32-61), and 29 torsion modes (numbered 62-90).

For the two ligands, a comparative vibrational analysis of the functional groups was conducted, as presented in Figures 3.4 and 3.5. The upper panels (a) display the experimental IR spectra, obtained from laboratory measurements. The red line represents the observed absorbance, reflecting the actual vibrational behavior of the ligands as they interact with IR radiation. The green areas overlaid on the experimental data indicate the spectral windows in which the same functional group regions were identified in the theoretical spectra, in order to allow a direct visual comparison. Conversely, the lower panels (b) showcase the theoretical IR spectra, where the red lines indicate the calculated absorbance values across the wavenumber range. The green shaded regions denote the vibrational contributions from various functional groups within the ligands, as predicted by quantum chemical calculations. These theoretical spectra provide insights into the expected vibrational transitions based on molecular simulations.

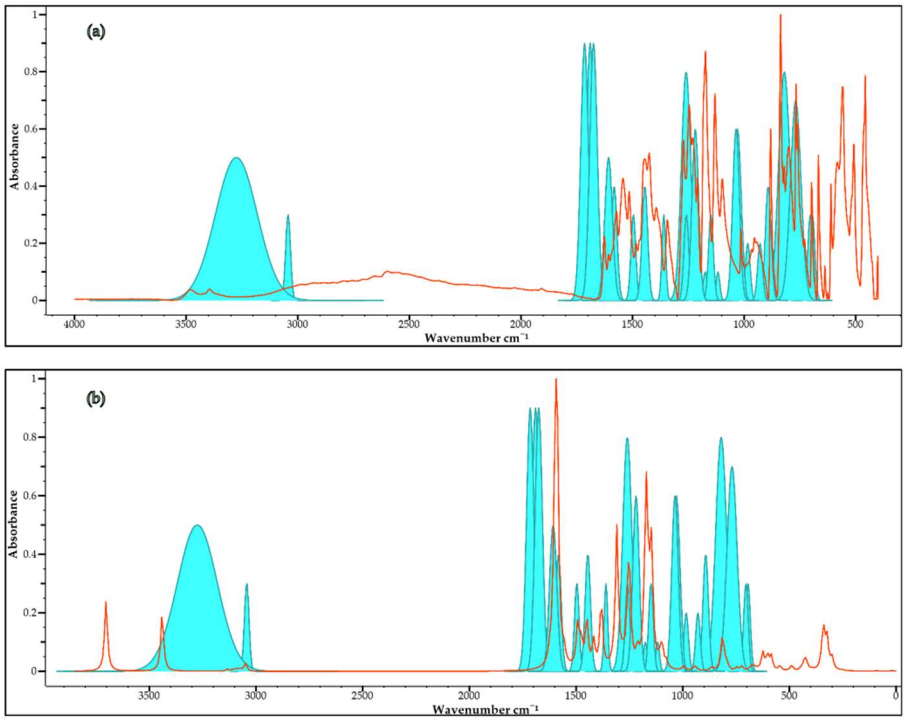

Figure 0.2. *Comparative vibrational analysis of functional groups (green) present in the spectra of resokaempferol (red): (a) experimental IR spectrum and (b) scaled theoretical IR spectrum.*

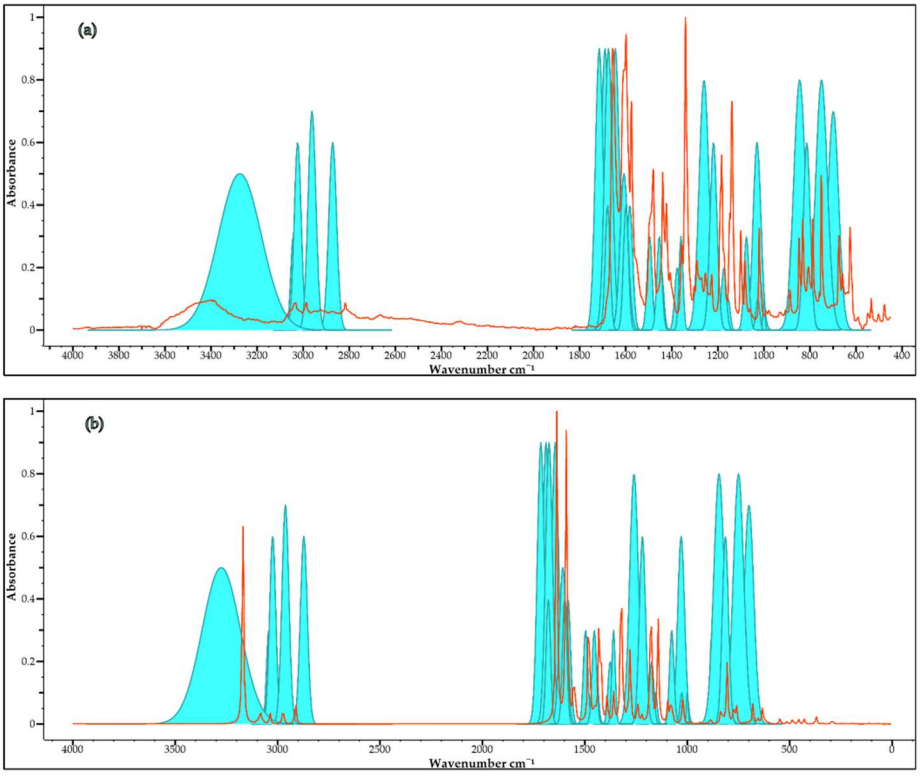

Figure 0.3. *Comparative vibrational analysis of functional groups (green) present in the spectra of tectochrysin (red): (a) experimental IR spectrum and (b) scaled theoretical IR spectrum.*

From the analysis of the vibrational modes of the different functional groups present in the molecular structure of the two studied compounds, we can infer observations regarding the relationship between specific functional groups and their potential as ligands for inhibiting cancer-related proteins.

Resokaempferol features ether groups (Ph-O-C), while tectochrysin displays unique alkane and alkene groups. These groups contribute different chemical properties, such as polarity and flexibility, which might affect the overall binding dynamics. Ethers present in resokaempferol might enhance molecular solubility, affecting the drug-like properties and possibly improving interaction with the hydrophilic regions of receptors. Conversely, the alkane and alkene groups present in tectochrysin could influence the ligand's ability to penetrate deeper into hydrophobic pockets of the proteins.

The potential of these two molecules as inhibitors rests significantly on these functional groups and their capacity to interact favorably with the protein structure. A deeper understanding of these interactions, ideally supplemented by molecular docking studies and experimental binding assays, would further elucidate their potential effectiveness in targeting PI3K alpha in cancer therapies.

3.1.2 Quantitative evaluation of quantum chemical parameters

Electronic parameters play an important role in predicting the biological properties of molecules. Flavonoids, such as resokaempferol and tectochrysin, have been evaluated through theoretical chemical analyses to assess their potential as drugs. By using the B3LYP/6-

311G++(d,p) method, we have obtained the values of the energies of the frontier molecular orbitals E_{HOMO} and $E_{LUMO,}$ as well as the dipole moment (DM) [127].

The results presented in Table 3.9 provide a quantification of the molecular orbital energies, specifically the energies of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), as well as the energy gaps (ΔE) between them, calculated for both ligands.

Table 0.1. *Molecular orbital energies and energy gaps for the resokaempferol and tectochrysin ligands [84].*

Ligand	E_{HOMO} (hartree/eV)	E_{LUMO} (hartree/eV)	E_{GAP} (hartree/eV)[1]
Resokaempferol	- 0,22/- 5,96	- 0,08/- 2,17	0,14/3,79
Tectochrysin	- 0,23/- 6,38	- 0,08/- 2,26	0,15/4,13

[1]1 Hartree = 27,21 eV/ 219474,63 cm^{-1}/ 627,51 kcal/mol

The graphical representation presented in Figure 3.6 is associated with Table 3.9 and provides a comparative analysis of the molecular energy range for resokaempferol and tectochrysin [84].

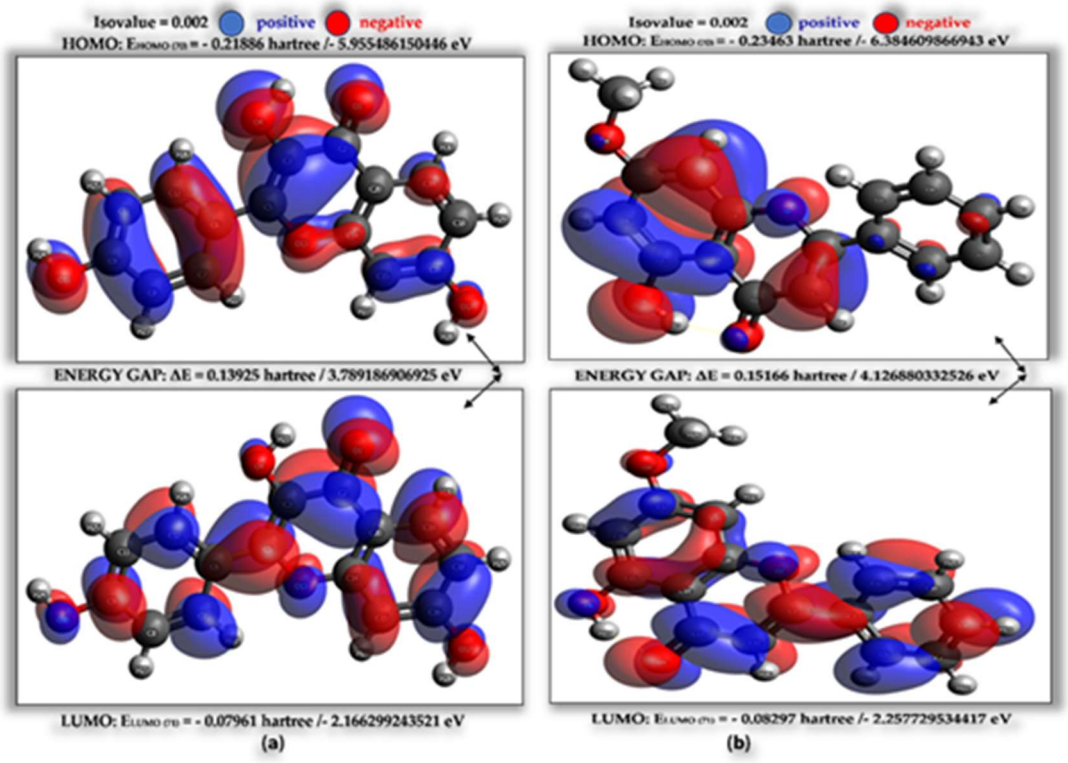

Figure 0.4 *Comparative analysis of the molecular energy gaps for: (a) resokaempferol and (b) tectochrysin [84].*

The energy gap between the HOMO and LUMO for tectochrysin (4,1269 eV) is larger than that for resokaempferol (3,7892 eV). This signifies that tectochrysin has a greater chemical

stability and lower reactivity than resokaempferol, as a larger amount of energy is required to excite an electron from the HOMO to LUMO.

Table 3.10 outlines the dipole moment and polarizability of each ligand. These properties are pivotal for understanding the ligands' interactions with their environment, be it within a solvent or a biological system. The dipole moment of tectochrysin (6,69 debye) is larger than the dipole moment of resokaempferol (3,47 debye). This suggests a more asymmetric distribution of electronic charge within the molecule of tectochrysin, which can enhance interactions with other polar substances. Both ligands exhibit similar values of polarizability, which indicates that they have comparable abilities to deform their electron clouds under external electric fields, suggesting similar susceptibilities to van der Waals interactions [84].

Table 0.2. *Dipole moment and the polarizability for the analyzed ligands [84].*

Ligand	Dipole moment (Debye)	Polarizability ($Å^3$)
Resokaempferol	3,47	32,44
Tectochrysin	6,69	32,21

The data obtained for the E_{HOMO}, E_{LUMO} and E_{gap} energies facilitated the determination of other chemical parameters (ionization potential (IP), electron affinity (EA), chemical hardness (η), chemical softness (σ), electronegativity (χ), electrophilicity index (ω)). The results are presented in Table 3.11. They indicate that tectochrysin generally displays higher values across most quantum chemical properties, indicative of a molecule with a better chemical stability and a stronger ability to attract and retain electrons. These attributes suggest that tectochrysin might exhibit lower reactivity but higher potential for interacting with other chemical entities in a controlled manner, making it potentially more suitable for applications where stability is crucial. Conversely, the relatively higher softness and lower stability of resokaempferol might make it more reactive, which could be advantageous in dynamic chemical environments. [84].

Table 0.3. *Molecular descriptors of resokaempferol and tectochrysin [84].*

Quantum Chemical Properties	Resokaempferol (eV)	Tectochrysin (eV)
IP	5,96	6,38
EA	2,17	2,26
η	1,89	2,06
σ	0,53	0,48
χ	4,06	4,32
ω	4,35	4,52

The molecular electrostatic potential (MEP) map obtained for each compound individually enhance these findings, in order to predict the interactions of these ligands with ions, proteins, and nucleic acids. In these maps, presented in figure 3.7, the color gradient illustrates the electrostatic potential from positive (blue) to negative (red). The red zones around the oxygen atoms denote areas of high electron density, indicative of negative electrostatic potential where

electrophilic attack is favorable. Conversely, the positive regions are concentrated around all the hydrogen atoms, which are the most likely sites for nucleophilic attack.

The maps indicate that both flavonoids exhibit regions with variable electrostatic potential, which may influence how they interact with other molecules, particularly in binding to target proteins in biological systems. In addition to understanding intermolecular interactions (ligand-receptor), the MEP map also provides a visual explanation of biological properties. For example, regions with negative potential (blue) are often associated with antioxidant properties due to their ability to donate electrons to free radicals [84].

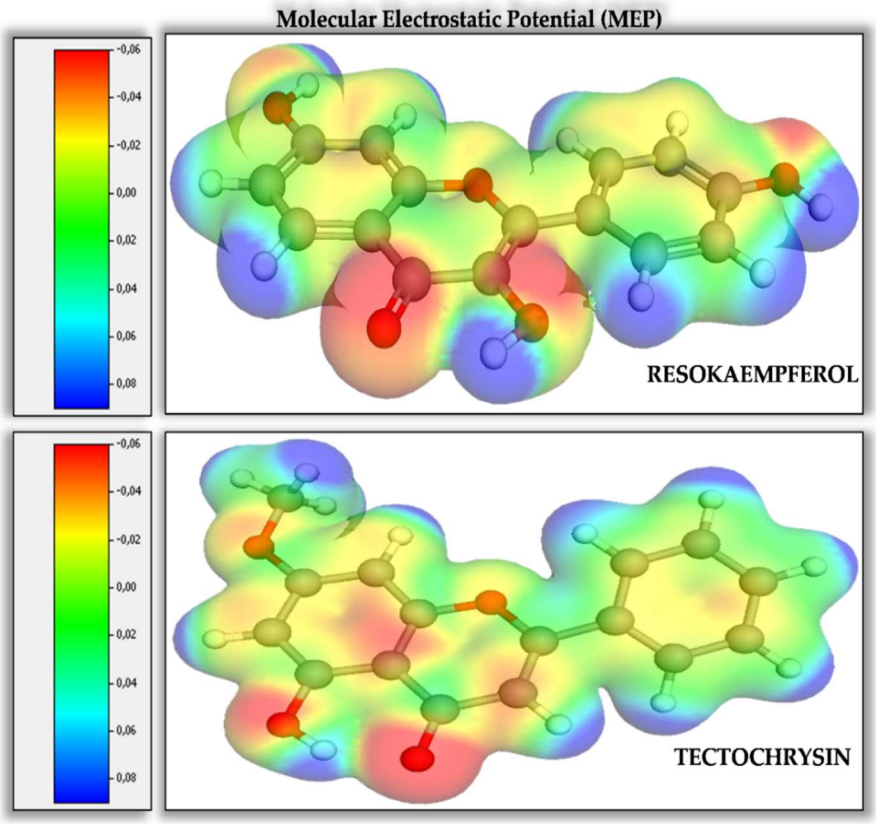

Figure 0.5. *Molecular electrostatic potential on electron density (MEP) of the studied ligands [84].*

3.1.3 The physico-chemical and pharmacokinetic profiles

The analysis of the physico-chemical and pharmacokinetic profiles of the two ligands (resokaempferol and tectochrysin) is another step in identifying them as potential inhibitors of the PI3Kα protein and its mutation, which are targets in cancer therapy. The radar charts (figure 2.13) presented in chapter 2, are supplemented by the data listed in table 3.13. The molecular weight values of the two compounds (resokaempferol: 270,240 g/mol and tectochrysin: 268.260 g/mol) show that the ligands meet one of the criteria for good absorption and permeability.

Table 0.4. *Physico-chemical properties of resokaempferol and tectochrysin.*

Physico-chemical Property	Resokaempferol	Tectochrysin
Molecular Weight	270,240	268,260
Van der Waals volume	265,186	273,692

Density	1,018	0,979
nHA	5	4
nHD	3	1
nRot	1	2
nRing	3	3
MaxRing	10	10
nHet	5	4
fChar	0	0
nRig	18	18
Flexibility	0,056	0,111
Stereo Centers	0	0
TPSA	90,900	59,670
AMR	73,990	75,203
Hy	1,103	-0,320
LogS	-3,828	-4,390
LogP	2,073	3,850
XLogP	3,168	3,141
XLogP2	10,036	9,865
XLogP3	2,690	3,850
ALogP	2,038	2,852
ALogP2	4,155	8,137
LogD7.4	2,098	3,072
pka (Acid)	7,196	7,478
pka (Baza)	2,932	3,177
Melting point	288,744	184,312
Boiling point	378,783	331,085
Ui	3,906	3,906

In summary, resokaempferol seems to have more favorable properties for bioavailability, i.e., a higher number of hydrogen bond acceptors and donors, a higher TPSA, and a better hydrophilic factor. Tectochrysin, while slightly less favorable in these aspects, may have better membrane permeability due to its higher LogD7.4 value. In addition, its increased flexibility could allow a better adaptation to the protein target's active site. The physico-chemical profiles of both compounds suggest that they have the potential to be good PI3Kα inhibitors, but their efficacy would ultimately need to be validated through biological assays and clinical trials. A series of chemical ADMET properties and their associated molecular descriptors for resokaempferol and tectochrysin were evaluated and empirically interpreted based on specific assumptions, as detailed in Table 3.14 [84].

Table 0.5. *ADMET properties and molecular descriptors of resokaempferol and tectochrysin.[84].*

ADMET Property	Molecular Descriptor *	Resokaempferol		Tectochrysin	
		Predicted Value/Probability	Empirical Decision	Predicted Value/Probability	Empirical Decision
Absorption	Caco-2	-5,62	poor	-4,78	excellent
	Pgp-inhibitor	0,14 (--)	excellent	0,96 (+++)	poor
	Pgp-substrat	0,27 (--)	excellent	0,07 (---)	excellent
	HIA	0,03 (---)	excellent	0,01 (---)	excellent
	$F_{20\%}$	0,32 (-)	medium	0,05 (---)	excellent
	$F_{30\%}$	0,78 (++)	poor	0,45 (-)	medium
	MDCK	-4,85	poor	-4,68	excellent
Distribution	PPB	96,66	poor	98,51	poor
	VDss	0,20	excellent	0,77	excellent
	BBB Penetration	0,01 (---)	excellent	0,2 (--)	excellent
	Fu	3,33	poor	0,81	poor
Metabolism	CYP1A2 inhibitor	0,75 (++)	poor	1,00 (+++)	poor

ADMET Property	Molecular Descriptor *	Resokaempferol		Tectochrysin	
		Predicted Value/Probability	Empirical Decision	Predicted Value/Probability	Empirical Decision
	CYP1A2 substrate	0,12 (--)	excellent	0,79 (++)	poor
	CYP2C19 inhibitor	0,25 (--)	excellent	0,98 (+++)	poor
	CYP2C19 substrate	0,001 (---)	excellent	0,02 (---)	excellent
	CYP2C9 inhibitor	0,89 (++)	poor	0,038 (---)	excellent
	CYP2C9 substrate	0,32 (-)	medium	0,98 (+++)	poor
	CYP2D6 inhibitor	0,01 (---)	excellent	0,81 (++)	poor
	CYP2D6 substrate	0,75 (++)	poor	0,98 (+++)	poor
	CYP3A4 inhibitor	0,92 (+++)	poor	0,98 (+++)	poor
	CYP3A4 substrate	0,01 (---)	excellent	0,01 (---)	excellent
Excretion	CL_{plasma}	7,14	medium	5,28	medium
	$T_{1/2}$	1,49	medium	0,75	poor
Toxicity	hERG Blockers	0,11	excellent	0,13	excellent
	H-HT	0,40	medium	0,46	medium
	DILI	0,67	medium	0,94	poor
	AMES Mutagenicity	0,55	medium	0,64	medium
	FDAMDD	0,74	poor	0,73	poor
	Skin Sensitization	0,63	medium	0,42	medium
	Carcinogenicity	0,79	poor	0,81	poor
	Eye Corrosion	0,77	poor	0,32	medium
	Eye Irritation	0,99	poor	0,99	poor
	Respiratory Toxicity	0,64	medium	0,77	poor

The ADMET evaluation for resokaempferol shows that, although certain descriptors indicate poor absorption, others suggest favorable properties. Additionally, its probability of human intestinal absorption (HIA) is 0.03 (---), which classifies it as excellent, further suggesting a favorable absorption potential. The results presented in Table 3.14 underline the complexity of the interactions among the factors affecting absorption and highlight the need for a comprehensive assessment beyond individual molecular descriptors. Additional experimental validation will help clarify and resolve the evident discrepancies [84].

In silico prediction of the HERG channel blocking potential by the compounds resokaempferol and tectochrysin suggests that these flavonoids may exhibit cardiotoxic activity. This prediction requires detailed analysis and interpretation, considering several essential factors. Therefore, the next steps in evaluating these compounds should include in vitro electrophysiological tests on cardiac cells to verify their potential for blocking HERG channels under real physiological conditions [84].

Both compounds show mixed results in toxicity endpoints. Tectochrysin, however, tends to exhibit higher toxicity across several parameters including DILI (Drug-Induced Liver Injury),

carcinogenicity, and respiratory toxicity, which could limit its use due to safety concerns. Both compounds show potential eye irritation and are poor in eye corrosion assessments, indicating a risk in these specific toxicological areas. Overall, tectochrysin, despite its better absorption and distribution profiles, presents significant challenges in metabolism and toxicity that could impede its development as a safe therapeutic agent. Resokaempferol, with a more favorable metabolic profile and lower levels of certain toxicities, might offer a safer alternative, though its absorption characteristics are less ideal [84].

Another comparative analysis of the two compounds (resokaempferol and tectochrysin) pertains to the description of specific medicinal chemistry characteristics to highlight their potential as ligands. The results of this analysis have been consolidated in the table 3.15.

Table 0.6. *Comparative analysis of medicinal chemistry attributes and toxicophore rules for resokaempferol and tectochrysin [84].*

Medicinal Chemistry	Resokaempferol		Tectochrysin	
	Predicted Value	Empirical decision	Predicted Value	Empirical decision
Drug-likeness	0,63	poor	0,78	excellent
SAscore	3,08	excellent	3,01	excellent
Fsp3	0	poor	0,06	poor
MCE-18	17	poor	16	poor
NPscore	1,04	medium	0,95	medium
Lipinski Rule	0	excellent	0	excellent
Pfizer Rule	0	excellent	2	poor
GSK Rule	0	excellent	0	(2 conditions satisfied)[2]
Golden Triangle	0	excellent	0	excellent
PAINS	0	excellent	0	excellent
BMS	0	excellent	0	excellent
NonBiodegradable	0	excellent	0	excellent
SureChEMBL Rule	0	excellent	0	excellent

[1]: 2 fulfilled conditions: logP> 3; TPSA < 75;

Overall, while tectochrysin shows slightly better drug-likeness and potentially meets more medicinal chemistry criteria favorably, it may face challenges due to its pharmacokinetic profile, as indicated by Pfizer's rule. Resokaempferol, although having a lower drug-likeness score, does not exhibit these pharmacokinetic concerns, possibly making it more advantageous in early-phase drug development. Both compounds, however, would benefit from further optimization to enhance their profiles as potential therapeutic agents targeting cancer proteins and their mutants [84].

3.1.4 Molecular docking of native structures and ligands of interest on receptors PI3Kα and H1047R

For our research, we utilized the crystal structures sourced from the RCSB PDB online database) to conduct re-docking and cross-docking studies. *Re-docking*, also known as *self-docking* involved reproducing the co-crystallized binding geometry and orientation of the ligands within the context of rigid macromolecular frameworks. *Cross-do*cking was applied to examine

the interactions of the targeted flavonoid ligands, resokaempferol and tectochrysin, with the rigid structures of both the wild type and the H1047R mutant of the PI3Kα protein.

The *self-docking* study was conducted to delve into the interaction dynamics and binding affinities of native molecular pairs. Using the crystalline structures 7K71 and 8TS9 obtained from the RCSB PDB online database [129], we initially isolated the wild-type variant of the catalytic alpha subunit of phosphatidylinositol 4,5-bisphosphate 3-kinase and its H1047R mutation using the Chimera X software, as illustrated in figure 3.10 We further isolated the native inhibitors from these X-ray structures, specifically the VYP ligand 5-(2-morpholin-4-ylpyrimidin-4-yl)pyrimidin-2-amine, from the 7K71 complex and UE9 ligand 5-[3-fluoro-5-(trifluoromethyl)benzamido]-N-methyl-6-(2-methylanilino)pyridine-3-carboxamide, from the 8TS9 complex, by using their name, synonyms or sequence entry number [84].

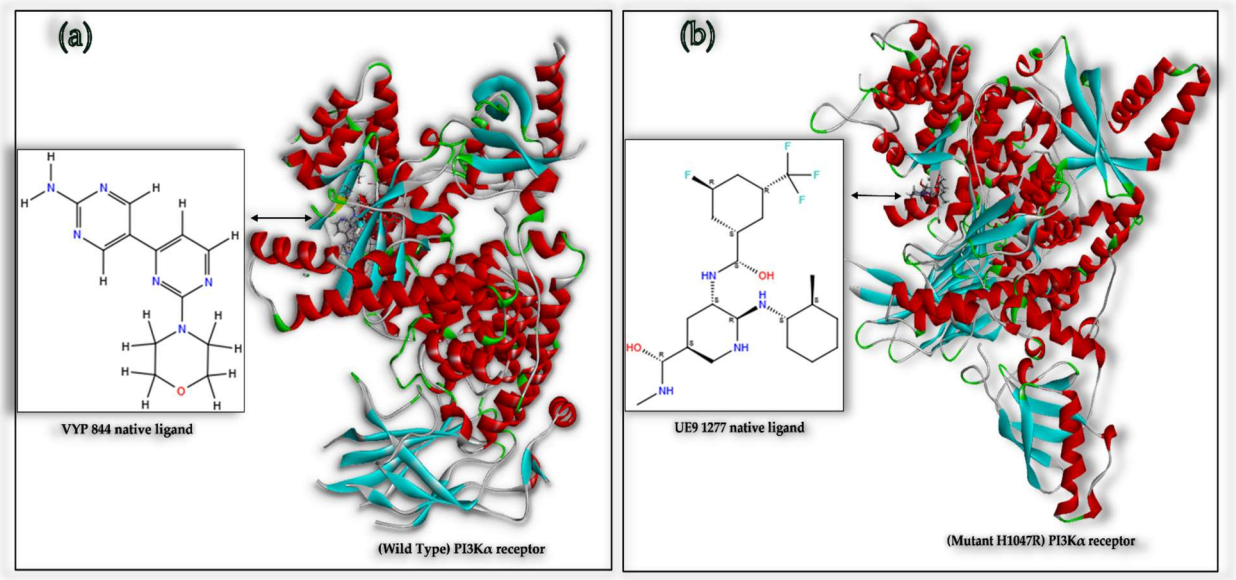

Figure 0.6. *Comparative visualization of PI3Kα receptor-native ligand pairs: (a) wild-type with 2-(morpholin-4-yl)[4,5'-bipyrimidin]-2'-amine and (b) H1047R mutant with 5-[3-fluoro-5-(trifluoromethyl)benzamido]-N-methyl-6-(2-methylanilino)pyridine-3-carboxamide [84].*

The first isolated ligand VYP from the 7K71 complex, classified as a non-polymer, comprises 33 atoms, 35 bonds, 12 of which are aromatic, with a formal charge of 0 and no chiral atoms, and characterized by the chemical formula $C_{12}H_{14}N_6O$, with a molecular weight of 258.279 g/mol. The second isolated inhibitor UE9, classified as a non-polymer the chemical formula: $C_{22}H_{18}F_4N_4O_2$ and a molecular weight of 446.398 g/mol. It comprises 50 atoms, 52 bonds, 18 of which are aromatic, with a formal charge of 0 and no chiral atoms [130].

The phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform, (7K71- PDB entry), is an essential component in cellular signaling and enzymatic processes, encapsulating key structural and functional attributes. This protein, in its wild-type form without any mutations, consists of 843 amino acid residues and encompasses a total of 5886 atoms. It is classified as a signaling protein and a transferase, originating from the *Homo sapiens* organism [84].

The other biological target, the H1047R mutation belonging to the same protein family, identified in the crystalline structure 8TS9, exhibits a functionally modified structure. This mutant contains 1004 amino acid residues and a total of 16435 atoms. Similar to PI3Kα – wild type, this

variant is derived from *Homo sapiens* and shares the same family and domain. However, as opposed to PI3Kα, it has a unique protein chain, designated as chain A, which highlights its distinct structure [84].

Files obtained from the PDB often require corrections before being used in molecular docking, due to missing atoms or the presence of water molecules—issues that can significantly affect the accuracy of docking simulations. The procedure for preparing receptors for docking was presented in Chapter 2. In the next step, the two ligands were prepared using the LigPrep tools along with Epik from the Maestro Suite, including corrections, conversions, and optimizations of the structures at pH 7.0±2.0. A structure-based drug design approach was used to identify binding sites on the target proteins, with the aim of effectively inhibiting cellular progression. These binding sites were obtained from the crystalline structures using manual correlation techniques through the PyMol software [100–102,131,132].

In the next step, the two ligands were prepared using the LigPrep tools along with Epik from the Maestro Suite, including corrections, conversions, and optimizations of the structures at pH 7.0±2.0. A structure-based drug design approach was used to identify binding sites on the target proteins, with the aim of effectively inhibiting cellular progression. These binding sites were obtained from the crystalline structures using manual correlation techniques through the PyMol software [84].

The two molecular structures of PI3Kα protein and their native and new potential inhibitors were subjected to virtual screening through re-docking and cross-docking using three computational programs: AutoDock, Vina, and Glide. A comparative analysis of the molecular re-docking, which included binding energy metrics for the PI3Kα receptor variants, was presented in tables 3.17 to 3.19. This analysis provided critical insights into the interaction dynamics and efficacy of the inhibitors across different receptor configurations.

The comparative analysis (table 3.17) emphasizes not only the differing efficacies of each ligand but also the distinct interactions within the protein environments. Starting with VYP, the analysis reveals that this ligand exhibits strong affinity for the wild type protein as indicated by a free energy of binding of -8,44 kcal/mol. This suggests a highly stable complex formation.

Resokaempferol presents an intriguing case with its evaluation against both protein variants. It shows a stronger binding to the mutant (-9.22 kcal/mol) than to the wild type (-8.73 kcal/mol), coupled with a substantially lower inhibition constant in the mutant (175.41 nM), highlighting its enhanced potency and stronger interactions with the mutant variant. The comparable ligand efficiency and intermolecular energy between the two forms suggest consistent interaction dynamics, despite the significant structural deviations as indicated by high RMSD values, particularly in the mutant form.

Tectochrysin, studied in both protein forms, exhibits the weakest binding energies among the analyzed ligands but presents the lowest inhibition constant for the wild type (18.60 nM), suggesting exceptional potency in this context. The energy metrics such as intermolecular, torsional free, and unbound system's energy remain consistent between the variants, reflecting stable binding characteristics, albeit with notable structural deviations as shown by RMSD measurements.

Table 0.7. *Comparative analysis of molecular docking obtained by using AutoDock: binding energy metrics for receptor variants [84].*

Ligand	Best Docking Conformation	Receptor	
		PI3Kα-wild type	H1047R
VYP	Free Energy of Binding (kcal/mol)	-8,44	-
	Inhibition Constant, Ki (nM)	653,08	-
	Ligand Efficiency (docking energy) (kcal/mol)	-0,44	-
	Intermolecular energy (kcal/mol)	-9,33	-
	Total Internal Energy (kcal/mol)	-0,77	-
	Electrostatic Energy (kcal/mol)	0,00	-
	van der Waals + Hydrogen bonds + Desolations. Energy (kcal/mol)	-9,33	-
	Torsional Free Energy (kcal/mol)	0,89	-
	Unbound System's Energy (kcal/mol)	-0,77	-
	RMSD from reference structure (Å)	0,631	-
UE9	Free Energy of Binding (kcal/mol)	-	-8,34
	Inhibition Constant, Ki (nM)	-	772,54
	Ligand Efficiency (docking energy) (kcal/mol)	-	-0,26
	Intermolecular energy (kcal/mol)	-	-10,13
	Total Internal Energy (kcal/mol)	-	-1,82
	Electrostatic Energy (kcal/mol)	-	-0,15
	van der Waals + Hydrogen bonds + DE solvation. Energy (kcal/mol)	-	-9,97
	Torsional Free Energy (kcal/mol)	-	1,79
	Unbound System's Energy (kcal/mol)	-	-1,82
	RMSD from reference structure (Å)	-	12,96
Resokaempferol	Free Energy of Binding (kcal/mol)	-8,73	-9,22
	Inhibition Constant, Ki (nM)	395,76	175,41
	Ligand Efficiency (docking energy) (kcal/mol)	-0,44	-0,46
	Intermolecular energy (kcal/mol)	-9,93	-10,41
	Total Internal Energy (kcal/mol)	-1,11	-1,09
	Electrostatic Energy (kcal/mol)	0,00	0,00
	van der Waals + Hydrogen bonds + desolations. Energy (kcal/mol)	-9,93	-10,41
	Torsional Free Energy (kcal/mol)	1,19	1,19
	Unbound System's Energy (kcal/mol)	-1,11	-1,09
	RMSD from reference structure (Å)	31,39	33,29
Tectochrysin	Free Energy of Binding (kcal/mol)	-6,45	-6,16
	Inhibition Constant, Ki (nM)	18,60	30,48
	Ligand Efficiency (docking energy) (kcal/mol)	-0,32	-0,31
	Intermolecular energy (kcal/mol)	-7,35	-7,06
	Total Internal Energy (kcal/mol)	-1,03	-1,03
	Electrostatic Energy (kcal/mol)	-0,1	-0,06
	van der Waals + Hydrogen bonds + DE solvation. Energy (kcal/mol)	-7,25	-6,99
	Torsional Free Energy (kcal/mol)	0,89	0,89
	Unbound System's Energy (kcal/mol)	-1,03	-1,03
	RMSD from reference structure (Å)	32,38	34,39

The comparative analysis of molecular docking results presented in Table 3.18, which were obtained with the Vina program [113,116], provides insightful details into the binding energy

metrics. VYP demonstrates a strong affinity for the wild-type receptor with a binding energy of -8,0 kcal/mol, suggesting a robust and stable interaction. In contrast, UE9 shows a remarkably strong affinity for the H1047R mutant, recorded at -11,05 kcal/mol. This notably lower energy value indicates an exceptionally stable binding conformation, highlighting the ligand's enhanced potency and specificity towards the mutant receptor, H1047R. Resokaempferol, evaluated against both types of receptors, shows a slightly stronger binding with the wild-type protein (-8,227 kcal/mol) compared to the mutant (-8,045 kcal/mol). The small difference in binding energy suggests that resokaempferol maintains a consistent interaction with the PI3Kα receptor, regardless of its mutational state, offering a versatile therapeutic profile that could be effective across a wider range of PI3Kα-related pathologies. Tectochrysin exhibits a similar tendency, with a marginally stronger interaction with the wild-type protein (-8,254 kcal/mol) than the mutant variant (-7,844 kcal/mol). The relatively close binding energy values between the two forms indicate that tectochrysin, like resokaempferol, could also be suitable for targeting both the normal forms and the mutations of the receptor, although with a slight preference for the wild type.

Table 0.8. *Comparative analysis of molecular docking using Vina: binding energy metrics for PI3Kα receptor variants [84].*

Best Docking Conformation	Ligand	Receptor	
		Wild Type	H1047R
Affinity (kcal/mol)	VYP	-8,0	-
	UE9	-	-11,05
	Resokaempferol	-8,23	-8,05
	Tectochrysin	-8,25	-7,84

The information summarized in table 3.19 was obtained from the molecular docking performed by using the Glide software [84,104]. For VYP, the XP Glide Score for the wild type receptor is -7.520, indicating a moderately strong interaction, which is consistent with a stable ligand-receptor complex formation. The corresponding Glide Ligand Efficiency of -0.396 further supports the effectiveness of VYP in terms of energy per heavy atom, optimizing its design for potent binding. UE9, on the other hand, demonstrates a significant interaction with the H1047R mutant, reflected by an XP Glide Score of -10,02. This suggests an exceptionally strong and stable binding, and the Glide Ligand Efficiency of -0,31 indicates a relatively efficient use of its molecular structure for binding to the mutant receptor H1047R.

Resokaempferol shows robust binding to both the wild type and mutant receptors, with an XP Glide Score of -9,63 for the wild type and -7,40 for the mutant. This indicates a stronger affinity for the wild type, complemented by a higher Glide Ligand Efficiency of -0,48 compared to -0,37 for the mutant. This data suggests that Resokaempferol is a versatile ligand capable of effectively targeting both receptor variants, although with a preference for the wild type. Tectochrysin similarly evaluated for both receptor types, exhibits an XP Glide Score of -8,54 for the wild type and -6,47 for the mutant The Glide Ligand Efficiency scores of -0,43 for the wild type and -0,32 for the mutant indicate a reasonable use of molecular structure for effective binding, although it shows a better affinity and efficiency for the wild type [84].

Table 0.9. *Comparative analysis of molecular docking using Glide: binding energy metrics for PI3Kα receptor variants [84].*

Ligand	Best Docking Conformation	Receptor	
		PI3Kα wild type	H1047R
VYP	XP Glide Score	-7,52	-
	Glide Ligand Efficiency	-0,39	-
UE9	XP Glide Score	-	-10,02
	Glide Ligand Efficiency	-	-0,31
Resokaempferol	XP Glide Score	-9,63	-7,40
	Glide Ligand Efficiency	-0,48	-0,37
Tectochrysin	XP Glide Score	-8,54	-6,47
	Glide Ligand Efficiency	-0,43	-0,32

The comprehensive analysis of molecular docking results highlights the differential binding affinities exhibited by each ligand across various computational platforms, illuminating their potential as therapeutic agents [84]:

- VYP exhibited its highest binding affinity towards the wild type receptor using AutoDock 4, closely followed by Vina, while its least effective interaction was observed with Glide.
- UE9 displayed a notable affinity for the H1047R mutant variant, achieving its strongest binding in the Vina program, with subsequent high affinity scores in Glide and AutoDock 4. This underlines its potential effectiveness against specific receptor mutations.
- Resokaempferol displays affinity for PI3Kα – wild type, the best score was obtained using Glide, followed by AutoDock 4. However, its interaction with the wild type was the least favorable in Vina. When considering the mutant variant, AutoDock 4 provided the most favorable binding, closely followed by Vina, while Glide showed the least efficacy.
- Tectochrysin performed best with the wild type receptor in Glide, followed by a strong showing in Vina, its least effective docking occurred with AutoDock 4. In the case of the mutant variant H1047R, the strongest binding was obtained with Vina. Its second-best results were obtained with Glide, while the weakest performance was obtained in AutoDock 4.

The variations in the inhibition potential indicators highlight the differences in how these programs assess the interactions between the ligand and receptor. Despite these variations, all three programs—AutoDock 4, AutoDock Vina, and Glide—indicate good results in re-docking and cross-docking, emphasizing the potential of the two ligands (resokaempferol and tectochrysin) to inhibit both the normal form and the H1047R mutation of the cancer-associated protein.

3.1.5 Structural alignment validation and MolProbity analysis

The structural alignment validation with the native structure is critically important in the process of *in silico* re-docking of native ligands and other proposed potential ligands for several fundamental scientific and technical reasons. Ensuring the accuracy of molecular modelling is paramount: aligning the target protein structure with its native conformation guarantees that the molecular model used in simulations is as close as possible to the natural state, thereby yielding precise and relevant results in molecular docking simulations. This validation also preserves the integrity of the binding site, ensuring that the identified binding sites in the modelled protein correspond accurately to those in the native structure. Such alignment is essential for the correct docking of ligands, reflecting the molecular interactions that would occur *in vivo* [84].

Moreover, structural alignment validation enhances the comparability of results. It allows for a direct comparison between results obtained for native and proposed ligands, eliminating systematic errors that could arise from conformational differences between the modelled and native structures. Additionally, this process helps identify and correct errors that may have been introduced during the modelling process, such as distortions of side chains, improper hydrogen bond assignments, or the presence of inappropriate residues in the active site. The methodological reliability is significantly improved through structural alignment validation, providing a reference standard for evaluating the robustness of *in silico* re-docking methodologies. This validation step ensures that the simulation and modelling methods used are both robust and reproducible. Furthermore, it secures the bio-logical relevance of the model; without structural alignment, the model might fail to appropriately reflect the biological context of the target protein. Ensuring that the modelled conformation is biologically relevant allows for a more realistic assessment of the therapeutic potential of the studied ligands [84].

Accurately modelling the dynamic nature of protein conformations in response to ligand binding represents a challenge in the field of protein-ligand docking. To ascertain the reliability of the docking outcomes, we have applied structural alignment methodologies and analyzed the configurations resulting from both self-docking and cross-docking experiments. These models were superimposed onto native co-crystallized complexes, providing a robust framework for validating the conformational integrity of the predicted interactions. This approach ensures that the docking simulations faithfully represent the structural dynamics observed in biological systems.

To this aim, we have used the Pairwise Structure Alignment by employing the online tools available on the RCSB PDB platform [133]. Structure alignment is a computational approach that seeks to establish correspondence between residues of two or more macromolecular structures by optimally super-posing their shapes and three-dimensional conformations. This technique includes options for pairwise structure alignment, where structures are compared in pairs to ascertain their spatial and conformational similarities.

Following the Pairwise Structure Alignment, measures describing the extent of structural similarity have been documented in tables 3.20 – 3.23. The results from aligning all referenced structures are quantitatively characterized by several measures [84,136]:

- *The root means square deviation* (RMSD), expressed in angstroms (Å), is calculated between the aligned pairs of backbone C-alpha atoms in the superposed structures. A lower RMSD indicates a better structural alignment between the pair of structures, making it a commonly reported metric in structural comparisons. However, it is sensitive to deviations in local

structure, and residues in loops that are not well-aligned are typically excluded from the RMSD calculation, which is then performed only by using the residues that can be effectively aligned;

- *The template modeling score (TM-score)* is a useful parameter for assessing how well the structures align globally and for effectively comparing structures of different sizes, avoiding sensitivity to local structural differences that may influence other measures such as RMSD. This score varies between 0 and 1, where 1 signifies a perfect match, and 0 indicates no match between the two structures. TM-scores below 0.2 usually suggest that the proteins are unrelated, whereas scores above 0.5 generally indicate that the proteins share the same fold, often classified according to databases like Scop or Cath.
- *The identity* represents the percentage of paired residues in the alignment that are identical in sequence.
- *The equivalent residues* refer to the number of residue pairs that are structurally equivalent in the alignment (the number of amino acids that are compared between structures to evaluate structural and functional similarity).
- *The sequence length* denotes the total number of polymeric residues in the deposited sequence for a given chain.
- The modeled residues indicate the number of residues with coordinates that were used for the structure alignment, providing insight into the extent of the model covered by the alignment.

Table 0.10. *Comparative structural alignment validation of the 7K71 crystal structure with Glide, AutoDock 4 (AD4), and Vina self-docking outputs for the Wild-type (WT) PI3Kα protein complexed with the native VYP inhibitor [84].*

Pairwise Structure Alignment	7K71 RCSB PDB	Glide Re-Docking	AD 4 Re-Docking	Vina Re-Docking
Chain	A	A	A	A
RMSD	-	0,17	0	0
TM-score	-	1	1	1
Identity	-	99%	100%	67%
Aligned Residues	-	843	843	843
Sequence Length	946	843	843	843
Modelled Residues	843	843	843	843

Table 0.11. *Comparative structural alignment validation of the 8TS9 crystal structure with Glide, AutoDock 4 (AD4), and Vina self-docking outputs for the mutant H1047R PI3Kα protein complexed with the native UE9 inhibitor [84].*

Pairwise Structure Alignment	8TS9 RCSB PDB	Glide Re-Docking	AD 4 Re-Docking	Vina Re-Docking
Chain	A	A	A	A
RMSD	-	0.17	0	0
TM-score	-	1	1	1
Identity	-	99%	100%	100%

Aligned Residues	-	1,004	1,004	1,004
Sequence Length	1,060	1,004	1,004	1,004
Modelled Residues	1,004	1,004	1,004	1,004

Table 0.12. *Comparative structural alignment validation of the 7K71 crystal structure with Glide, AutoDock 4 (AD4), and Vina cross-docking outputs for the Wild-type (WT) PI3Kα protein complexed with the potential ligands resokaempferol and tectochrysin [84].*

Pairwise Structure Alignment	7K71 RCSB PDB	Glide cross Docking	AD 4 cross Docking	Vina cross Docking
Chain	A	A	A	A
RMSD	-	0,29	0	0
TM-score	-	1	1	1
Identity	-	99%	100%	67%
Aligned Residues	-	843	843	843
Sequence Length	946	843	843	843
Modelled Residues	843	843	843	843

Table 0.13. *Comparative structural alignment validation of the 8TS9 crystal structure with Glide, AutoDock 4 (AD4), and Vina cross-docking outputs for the mutant H1047R PI3Kα protein complexed with the potential ligands resokaempferol and tectochrysin [84].*

Pairwise Structure Alignment	8TS9 RCSB PDB	Glide cross docking	AD 4 cross docking	Vina cross docking
Chain	A	A	A	A
RMSD	-	0.17	0	0
TM-score	-	1	1	1
Identity	-	99%	100%	100%
Aligned Residues	-	1,004	1,004	1,004
Sequence Length	1,060	1,004	1,004	1,004
Modelled Residues	1,004	1,004	1,004	1,004

All these findings highlight the successful alignment and identity match between the self-docking and cross-docking pairs with their corresponding native structures. Depending on the docking programs utilized, the identity percentages between the aligned pairs ranged impressively from 67% to 100%. Additionally, the RMSD values were very favorable (all below 2), indicating a high degree of structural congruence. Furthermore, the TM-scores were exceptionally high, all reaching the maximum value of 1, underscoring the precision and effectiveness of our

docking approaches in replicating native-like interactions. In conclusion, structural alignment validation with the native structure is an indispensable step in the *in silico re-docking* process. It ensures the accuracy, comparability, reliability, and biological relevance of the conducted simulations and evaluations. This significantly contributes to the confidence in the obtained results and underpins subsequent decisions in the development of new drug therapies [84].

MolProbity is a versatile web server designed for the quality validation of 3D structures, encompassing proteins, nucleic acids, and their complexes. It offers comprehensive all-atom contact analysis to identify and resolve any steric clashes within the molecules. Additionally, it provides advanced diagnostics on dihedral angles and can calculate and visually represent hydrogen bonds and van der Waals contacts across interfaces between components [137].

The analysis was conducted by evaluating several important parameters [137]:

- *Clash score* measures the number of steric clashes per 1000 atoms in the model, where a steric clash is defined as an instance where atoms are closer together than allowed by their van der Waals radii, indicating potential structural problems;
- *Poor rotamers*: are those amino acid side chains that do not conform to energetically favorable rotational positions of the side chains;
- *Favored rotamers:* represents those conformations that are in an energetically favorable position;
- *Ramachandran outliers:* are residues whose backbone dihedral angles (phi and psi) are in regions of the Ramachandran plot that are rarely occupied in reliable structures;
- *Ramachandran favored:* refers to residues that fall within the most favorable regions of the Ramachandran plot;
- *Ramachandran distribution Z score* is a statistical measure that represents the overall deviation of the set of phi-psi dihedral angles from those typically found in high-quality structures, providing a normalized assessment of backbone conformational quality;
- *MolProbity score* is an overall quality score derived from the analysis of all other metrics, summarizing the model's quality with a single number where lower scores indicate better structural integrity;
- *Cβ deviations greater than 0,25 Å* indicate where the beta carbon atoms deviate significantly from their expected positions, which can affect the overall structural stability;
- *Bad bonds*: are indicators of chemical bonds in the protein that deviate significantly from expected values, often resulting in structural distortions;
- Bad angles: are indicators of bond angles in the protein that deviate significantly from expected values, often resulting in structural distortions;
- *CaBLAM outliers* represent non-ideal conformations in the protein backbone detected by *C-alpha Based Low-resolution Annotation Method*, a measure used particularly for validating low-resolution structures;
- *CA Geometry outliers* involve deviations in the geometry of alpha carbon atoms, indicating potential inaccuracies in the protein backbone's modelling.

The results obtained through docking, after evaluating the parameters, were classified into three levels: *good, caution and warning*. These classifications are based on standardized references provided in the "Key to Table Cutoffs" made available by the MolProbity platform. This structured approach ensures that each aspect of the molecular structure is assessed according to established benchmarks, allowing for a clear and consistent interpretation of the structural quality and potential areas of concern within the analyzed models. The results of this complex analysis are presented in tables 3.24 – 3.29.

***Table 0.14**. MolProbity analysis of all-atom contacts and geometry for the best cognate docking output of (Wild Type) PI3Kα protein in complex with native inhibitor VYP*

Category	Metrics	Raw count	Percentage	Validation[1]
All-atom contacts	Clashscore, all atoms2	0		Good: 100th percentile* (N=1784, all resolutions)
Protein geometry	Poor rotamers	9	1.18%	Caution: 0.3% < Outliers ≤ 1.5%
	Favoured rotamers	738	96.72%	Caution: 98% > Favoured ≥ 95%
	Ramachandran outliers	2	0.25%	Caution: 0.05% < Outliers ≤ 0.5%
	Ramachandran favoured	760	96.32%	Caution: 98% > Favoured ≥ 95%
	Ramachandran distribution Z score	-1.46 ± 0.27		Good: abs(Z-score) ≤ 2
	MolProbity score3	0.80		Good: 100th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations > 0.25Å	1	0.12%	Caution: 0 < Outliers < 5%
	Bad bonds	0/6,928	0.00%	Good: Outlier bonds < 0.01%
	Bad angles	66/9,335	0.71%	Warning: Outlier angles ≥ 0.5%
Low-resolution criteria	CaBLAM outliers	6	0.7%	Good: Outliers ≤ 1%
	CA Geometry outliers	5	0.62%	Caution: 0.5% < Outliers < 1%

[1] Key to table cutoffs: http://molprobity.biochem.duke.edu/help/validation_options/summary_table_guide.html.

***Table 0.15**. MolProbity analysis of all-atom contacts and geometry for the best cognate docking output of (mutant H1047R) PI3Kα protein in complex with native inhibitor UE9.*

Category	Metrics	Raw count	Percentage	Validation[1]
All-atom contacts	Clashscore, all atoms	1.9		Good: 99th percentile* (N=1784, all resolutions)
Protein geometry	Poor rotamers	13	1.43%	Caution: 0.3% < Outliers ≤ 1.5%
	Favoured rotamers	858	94.49%	Warning: Favoured < 95%
	Ramachandran outliers	1	0.10%	Caution: 0.05% < Outliers ≤ 0.5%
	Ramachandran favoured	932	97.29%	Caution: 98% > Favoured ≥ 95%
	Ramachandran distribution Z score	-1.81 ± 0.25		Good: abs(Z-score) ≤ 2
	MolProbity score3	1.21		Good: 99th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations > 0.25Å	4	0.42%	Caution: 0 < Outliers < 5%
	Bad bonds	0/8,232	0.00%	Good: Outlier bonds < 0.01%
	Bad angles	80/11,100	0.72%	Warning: Outlier angles ≥ 0.5%
Low-resolution criteria	CaBLAM outliers	9	0.9%	Good: Outliers ≤ 1%
	CA Geometry outliers	4	0.41%	Good: Outliers ≤ 0.5%

[1] Key to table cutoffs: http://molprobity.biochem.duke.edu/help/validation_options/summary_table_guide.html.

***Table 0.16.** MolProbity analysis of all-atom contacts and geometry for the best cross-docking output of (Wild Type) PI3Kα protein in complex with potential inhibitor resokaempferol*

Category	Metrics	Raw count	Percentage	Validation[1]
All-atom contacts	Clashscore	2,33		Good: 99th percentile* (N=1784, all resolutions)
Protein geometry	Poor rotamers	7	0.92%	Caution: 0.3% < Outliers ≤ 1.5%
	Favored rotamers	732	95.81%	Caution: 98% > Favored ≥ 95%
	Ramachandran outliers	1	0.13%	Caution: Outliers > 0.5% and Outlier count = 1
	Ramachandran favored	761	96.09%	Caution: 98% > Favored ≥ 95%
	Ramachandran distribution Z score	-1.91 ± 0.27		Good: abs(Z-score) ≤ 2
	MolProbity score3	1.28		Good: 99th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations > 0.25Å	1	0.12%	Caution: 0 < Outliers < 5%
	Bad bonds	0/6,928	0.00%	Good: Outlier bonds < 0.01%
	Bad angles	67/9,339	0.72%	Warning: Outlier angles ≥ 0.5%
Low-resolution criteria	CaBLAM outliers	11	1.40%	Good: Outliers ≤ 1%
	CA Geometry outliers	4	0.50%	Caution: Outliers ≤ 0.5%

[1] Key to table cutoffs: http://molprobity.biochem.duke.edu/help/validation_options/summary_table_guide.html.

Table 0.17. *MolProbity analysis of all-atom contacts and geometry for the best cross-docking output of (mutant H1047R) PI3Kα protein in complex with potential inhibitor resokaempferol.*

Category	Metrics	Raw count	Percentage	Validation[1]
All-atom contacts	Clashscore, all atoms2	1.84		Good: 99th percentile* (N=1784, all resolutions)
Protein geometry	Poor rotamers	13	1.43%	Caution: 0.3% < Outliers ≤ 1.5%
	Favored rotamers	858	94.49%	Warning: Favored < 95%
	Ramachandran outliers	1	0.10%	Caution: Outliers > 0.5% and Outlier count = 1
	Ramachandran favored	932	97.29%	Caution: 98% > Favored ≥ 95%
	Ramachandran distribution Z score	-1.81 ± 0.25		Good: abs(Z-score) ≤ 2
	MolProbity score3	1.20		Good: 99th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations > 0.25Å	4	0.42%	Caution: 0 < Outliers < 5%
	Bad bonds	0/8,232	0.00%	Good: Outlier bonds < 0.01%
	Bad angles	80/11,100	0.72%	Warning: Outlier angles ≥ 0.5%
Low-resolution criteria	CaBLAM outliers	9	0.9%	Good: Outliers ≤ 1%
	CA Geometry outliers	4	0.41%	Good: Outliers ≤ 0.5%

[1] Key to table cutoffs: http://molprobity.biochem.duke.edu/help/validation_options/summary_table_guide.html.

Table 0.18. *MolProbity analysis of all-atom contacts and geometry for the best cross-docking output of (Wild Type) PI3Kα protein in complex with potential inhibitor tectochrysin.*

Category	Metrics	Raw count	Percentage	Validation[1]
All-atom contacts	Clashscore, all atoms2	2.33		Good: 99th percentile* (N=1784, all resolutions)
Protein geometry	Poor rotamers	7	0.92%	Caution: 0.3% < Outliers ≤ 1.5%
	Favored rotamers	732	95.81%	Warning: Favored < 95%
	Ramachandran outliers	1	0.13%	Caution: Outliers > 0.5% and Outlier count = 1
	Ramachandran favored	761	96.09%	Caution: 98% > Favored ≥ 95%
	Ramachandran distribution Z score	-1.91 ± 0.27		Good: abs(Z-score) ≤ 2
	MolProbity score3	1.28		Good: 99th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations > 0.25Å	1	0.12%	Caution: 0 < Outliers < 5%
	Bad bonds	0/6,928	0.00%	Good: Outlier bonds < 0.01%
	Bad angles	67/9,339	0.72%	Warning: Outlier angles ≥ 0.5%
Low-resolution criteria	CaBLAM outliers	11	1.40%	Good: Outliers ≤ 1%
	CA Geometry outliers	4	0.50%	Good: Outliers ≤ 0.5%

1 Key to table cutoffs: http://molprobity.biochem.duke.edu/help/validation_options/summary_table_guide.html.

***Table 0.19**. MolProbity analysis of all-atom contacts and geometry for the best cross-docking output of (mutant H1047R) PI3Kα protein in complex with potential inhibitor tectochrysin*

Category	Metrics	Raw count	Percentage	Validation[1]
All-atom contacts	Clashscore, all atoms2	1.9		Good: 99th percentile* (N=1784, all resolutions)
Protein geometry	Poor rotamers	13	1.43%	Caution: 0.3% < Outliers ≤ 1.5%
	Favored rotamers	858	94.49%	Warning: Favored < 95%
	Ramachandran outliers	1	0.10%	Caution: Outliers > 0.5% and Outlier count = 1
	Ramachandran favored	932	97.29%	Caution: 98% > Favored ≥ 95%
	Ramachandran distribution Z score	-1.81 ± 0.25		Good: abs(Z-score) ≤ 2
	MolProbity score3	1.21		Good: 99th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations > 0.25Å	4	0.42%	Caution: 0 < Outliers < 5%
	Bad bonds	0/8,232	0.00%	Good: Outlier bonds < 0.01%
	Bad angles	80/11,100	0.72%	Warning: Outlier angles ≥ 0.5%
Low-resolution criteria	CaBLAM outliers	9	0.90%	Good: Outliers ≤ 1%
	CA Geometry outliers	4	0.41%	Good: Outliers ≤ 0.5%

1 Key to table cutoffs: http://molprobity.biochem.duke.edu/help/validation_options/summary_table_guide.html.

The analysis of the best molecular pairs, performed with the MolProbity platform, reveals that the majority of the metrics showed good results, highlighting the structural integrity and accuracy of the targeted modelled complexes. The MolProbity score, which reflects the overall quality of the model, was outstanding. However, there were minor noted discrepancies in the positioning of some beta carbons, suggesting areas for potential refinement. Regarding bond and angle integrity, the absence of bad bonds indicated perfect covalent connectivity. However, a small proportion of angles were flagged as outliers, necessitating a closer inspection of angle configurations in specific areas of the protein structure to ensure structural correctness. Overall,

the analysis validates the robustness of the structural models, with most metrics showing excellent conformity to expected standards. The few areas flagged for caution provide valuable insights for targeted refinements, aiming to enhance the precision of our molecular simulations.

3.1.6 Analysis of molecular interactions

Discovery Studio and PyMOL software packages were employed to analyze the interactions between native inhibitors and potential ligands with variants of the PI3Kα protein (wild type and mutant) [84,132,138]. These tools facilitated the detailed examination of two-dimensional, three-dimensional, and surface annotations, providing a multifaceted view of how these molecules interact at the molecular level. This approach not only enhanced our understanding of ligand binding, but also highlighted key interaction points and structural features within the protein complexes.

These molecular docking studies indicate that the docking of the ligands at the inhibition active sites of the receptor proteins exhibit a variety of intricate interactions. These interactions include conventional hydrogen bonds, carbon hydrogen bonds, unfavorable acceptor-acceptor bonds, Pi-Pi stacked bonds, one Pi-Pi T-shaped bond, alkyl bonds, Pi-alkyl bonds, Pi-sulfur interactions, halogen (fluorine) bonds, and Pi-anion interactions. Each type of interaction plays a critical role in the molecular docking process and can influence the inhibitory activity of the ligands either positively or negatively [84].

Graphical representations of the interactions between the PI3Kα protein (the normal form and the mutation) and the native inhibitors are presented in figure 3.11.

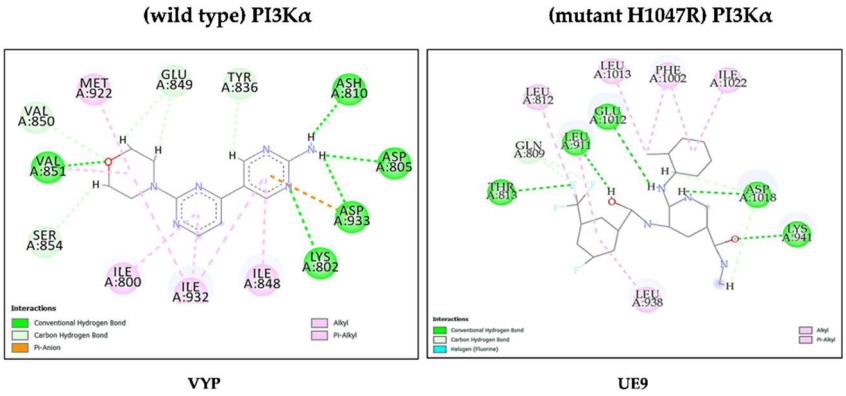

Figure 0.7. *Two-dimensional interaction diagram of the native inhibitor and PI3Kα protein: VYP with the (wild-type) PI3Kα and UE9 with the (mutant H1047R) PI3Kα [84].*

The three-dimensional representation of hydrogen bonds (acceptor-donor) and hydrophobic surfaces, with values ranging from -3 to 3, was created to assist in visualizing these molecular interactions. These maps provide a detailed representation of the spatial distribution and intensity of the interactions, highlighting the critical areas within the binding site (see figure 3.12).

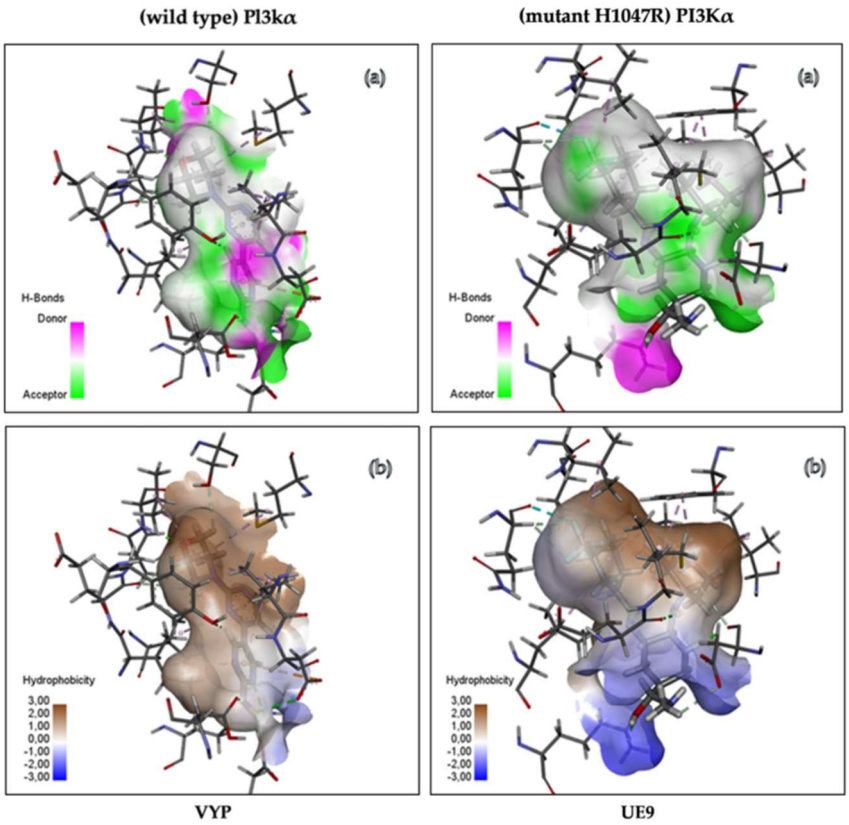

***Figure 0.8**. Three-dimensional visualization of the interactions between the native inhibitors and PI3Kα protein: VYP with the wild-type PI3Kα and UE9 with the H1047R mutant PI3Kα, highlighting (a) hydrogen bonding and (b) hydrophobic surfaces [84].*

In the interaction between resokaempferol and the wild-type PI3Kα protein (see figure 3.13), the ligand demonstrates a variety of binding interactions at the inhibition active sites the ligand demonstrates a variety of binding interactions at the inhibition active sites VAL 851, GLU 849 and LYS 802, along with one carbon-hydrogen bond involving the residue VAL 850. There is also a Pi-Sulfur bond with the residue MET 922, defined as a type of interaction where the π-electron cloud of an aromatic ring interacts with the sulfur atom of a thiol group, enhancing the stability and binding affinity of the ligand. Furthermore, a Pi-Pi T-shaped interaction is observed with the residue TRP 780, characterized by the perpendicular orientation of two aromatic rings facilitating strong aromatic interactions. Additionally, there are five Pi-alkyl bonds, two of which are with the residue ILE 800, two with the residue ILE 932, and one with the residue ILE 848, contributing to the hydrophobic interactions that stabilize the ligand within the binding pocket [84].

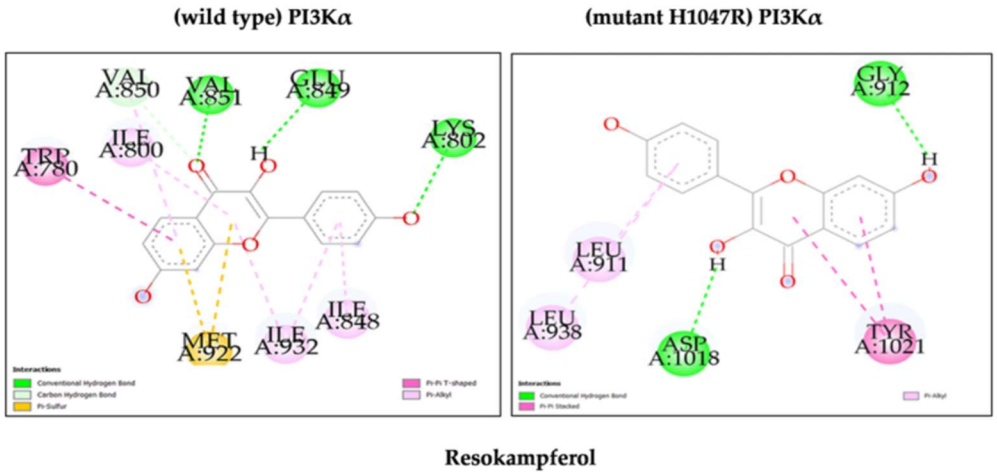

Figure 0.9. *Two-dimensional interaction diagram of the potential ligand resokaempferol with the (wild-type) PI3Kα and the (mutant H1047R) PI3Kα protein [84].*

In the interaction between resokaempferol and the mutant H1047R PI3Kα protein, the ligand forms several distinct binding interactions at the receptor's inhibition active sites. These include two conventional hydrogen bonds with the residues ASP 1018 and GLY 912, which are critical for stabilizing the ligand within the binding pocket. Furthermore, the ligand engages in two Pi-Pi stacked bonds with the residue TYR 1021, defined as interactions where two aromatic rings are parallel to each other, enhancing the aromatic interactions that are pivotal for its binding affinity. Additionally, two Pi-alkyl bonds with the residues LEU 911 and LEU 938 contribute to the hydrophobic interactions, further securing the ligand in place (see figure 3.13). The associated 3D representation illustrating the hydrogen bonds and hydrophobic surfaces is shown in figure 3.14 [84].

In the molecular interaction between tectochrysin and the wild-type PI3Kα protein, the ligand demonstrates a complex array of binding interactions at the active inhibition sites (see figures 3.15 and 3.16). These include three conventional hydrogen bonds with the residue VAL 851, enhancing the stability of the ligand within the binding pocket. Additionally, there is one carbon-hydrogen bond with the residue VAL 850, contributing to the overall binding conformation. The ligand also forms two Pi-Sulphur bonds with the residue MET 922, which are interactions that occur between the π-electrons of an aromatic ring and a Sulphur atom. This type of bond is significant in molecular recognition and can influence the binding affinity and specificity of molecules, especially in biological systems [84].

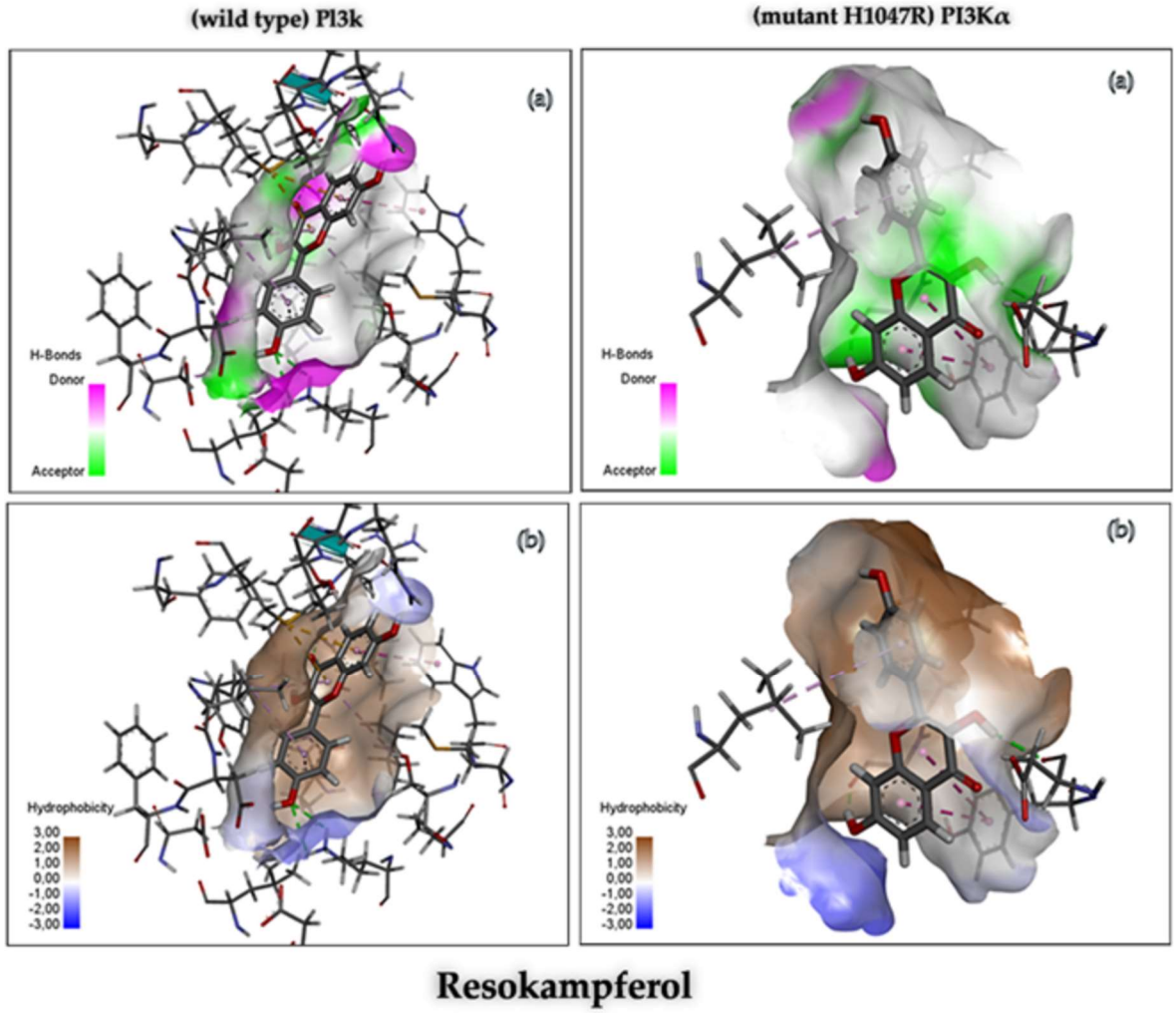

Figure 0.10. *Three-dimensional visualization of the potential ligand resokaempferol interacting with the (wild-type) PI3Kα and (mutant H1047R) PI3Kα protein highlighting (a) hydrogen bonding and (b) hydrophobic surfaces. [84].*

There is also one Pi-Pi T-shaped bond with the residue TRP 780, a type of non-covalent interaction that occurs between the π-electrons of two aromatic rings, where one ring is positioned perpendicular to the other, forming a “T” shape. This orientation allows for interactions between the electron clouds of the rings, contributing to the overall stability and specificity of the molecular complex. Moreover, there are six Pi-alkyl bonds, out of which two involve the residue ILE 932 and one each with the residues ILE 800, ILE 848, TRP 780, and VAL 850, further stabilizing the ligand through hydrophobic interactions [84].

In the interaction between tectochrysin and the mutant H1047R PI3Kα protein, the ligand engages in various binding interactions at the active inhibition sites. This includes one conventional hydrogen bond with the residue ASP 1018, which is important for stabilizing the ligand within the binding pocket. Additionally, three carbon-hydrogen bonds are formed, two with the residue GLY 912 and one with LYS 941, contributing to the structural integrity of the ligand-receptor complex. The ligand also forms one unfavorable acceptor-acceptor bond with the residue GLU 1012, which requires careful consideration due to its potential to destabilize the interaction. Furthermore, there is one Pi-Pi stacked bond with the residue PHE 1002, which is a non-covalent interaction that occurs between the π-electrons of adjacent aromatic rings. These interactions are crucial in stabilizing the three-dimensional structures of biological molecules, such as DNA and proteins, and play a significant role in molecular recognition processes. Additionally, there is one

Pi-Pi T-shaped bond with the residue TYR 1021, which enhances the aromatic interactions essential for the ligand affinity. The binding profile is further complemented by four alkyl and Pi-alkyl bonds with the residues LEU 812, LEU 938, PHE 937, and LYS 941, enhancing the hydrophobic interactions that contribute to the overall binding efficacy (see figure 3.15). Figure 3.16 shows the hydrogen bonding and the hydrophobicity surfaces [84].

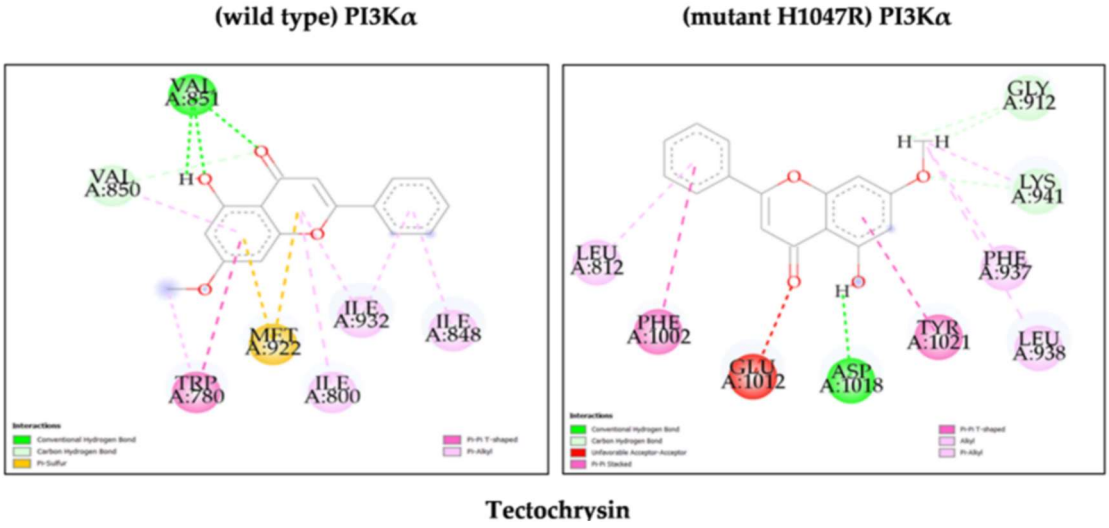

Figure 0.11. *Two-dimensional interaction diagram of the potential ligand tectochrysin with the (wild-type) PI3Kα and the (mutant H1047R) PI3Kα protein [84].*

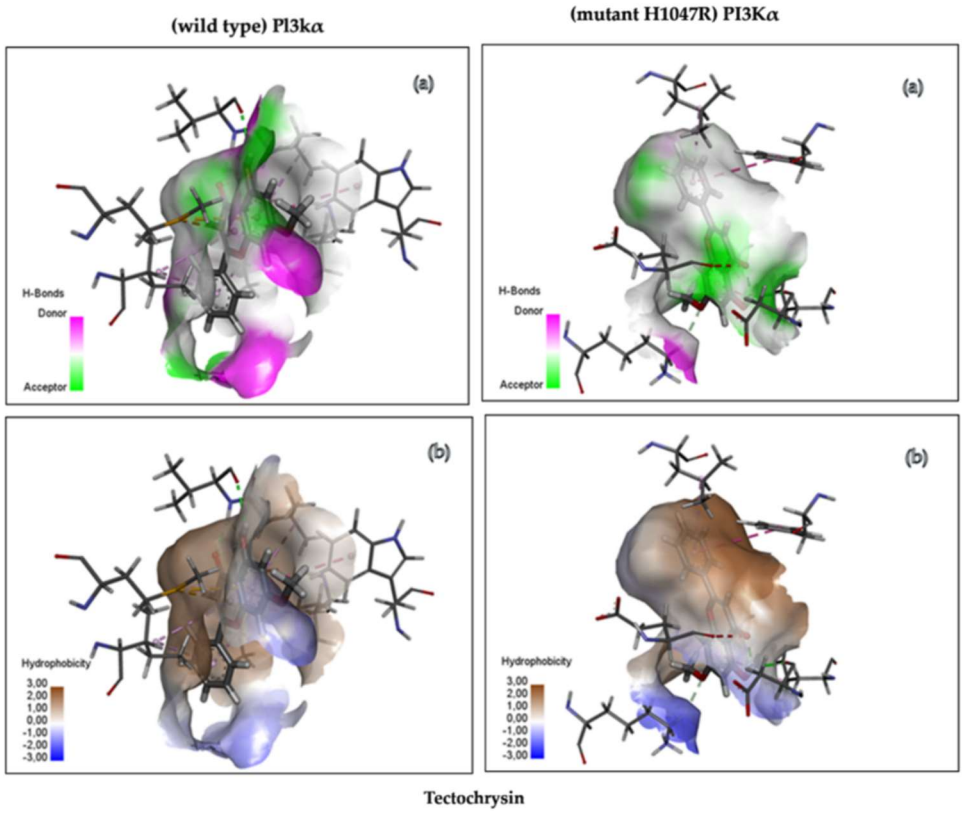

Figure 0.12. *Three-dimensional visualization of the potential ligand tectochrysin interacting with the (wild-type) PI3Kα and (mutant H1047R) PI3Kα protein highlighting (**a**) hydrogen bonding and (**b**) hydrophobic surfaces [84].*

3.2 Chemical potential evaluation of kaempferol based on molecular modeling [139]

The flavonoid kaempferol (3,5,7-trihydroxy-2-(4-hydroxyphenyl)chroman-4-one), whose description was provided in Chapters 1 and 2, was analyzed from the point of view of its inhibitory potential on the PI3Kα protein (wild type). Its molecular structure, shown in Figure 3.17, was optimized by using the *Gaussian 09* program [139,140].

The spatial orientation of kaempferol has the following characteristics: both the chroman backbone and phenol B ring are planar; the phenol group of the B ring linked to the basic chroman structure at the C2 position forms a torsion angle, α (c19-C8-C7-C2) that has a value of 0.04°; between the H29 hydrogen of the –OH group (O19-H29) from C3 position, as well as between the H31 hydrogen of the -OH group (O30-H31) from the C5 position that belong to the benzopyran nucleus and the O18 oxygen of the C=O carbonyl group (C9=O18) of the pyran C ring, two intramolecular hydrogen bonds are formed. The intramolecular bond formed between the H29 hydrogen atom and the O18 oxygen atom has a length of 1.99 Å, while the intramolecular hydrogen bond formed between the H31 hydrogen atom and the O18 oxygen atom has a length of 1.77 Å [139].

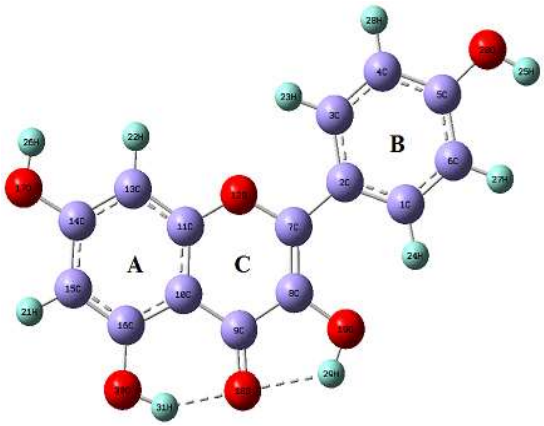

Figure 0.13. *Optimized molecular structure of 3,5,7-trihydroxy-2-(4-hydroxyphenyl)chromen-4-on (kaempferol) at B3LYP/6-311G(d,p) level [139].*

Since the crystal form of the studied compound could not be obtained, we have compared the theoretical parameters, i.e. bond lengths and angles, with the XRD parameters of the 5-fluoropyrimidine-2,4(1H,3H)-dione 3,5,7-trihydroxy-2-(4-hydroxyphenyl)-4H-1-benzopyran- 4 -one[141]. By using the Electronic Supplementary Publication from the Cambridge Crystallographic Data Centre (CCDC-897918), we obtained the geometrical parameters needed to compute the root mean square error (RMSE), which was computed as

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(y_{th}^{i} - y_{e}^{i})^2}{N}} \quad (1)$$

where N is the total number of theoretical geometries, y_{th} represents the theoretical values obtained by the DFT method and y_e are the geometrical parameters from XRD data [142].

The theoretical and experimental values of the geometric parameters are presented in table 3.30. To validate the method used for modelling, the theoretical values of the bond lengths graphically represented in figure 3.18 and those of the bond angles in figure 3.19 were compared with the corresponding experimental values obtained from X-ray structural analysis. Figure 3.18 shows that the differences between the theoretical and experimental values are very small, with the largest difference (0.01133) recorded for the C8-C9 bond and the smallest (-0.01118) for the C15-O6 interaction. The root mean square error is 0.00618 [139].

Table 0.20. *Centralization of theoretical and experimental values for kaempferol*

Bond	Experimental values XRD	Theoretical values B3LYP/6-311G(d,p)	Bond angle	Experimental values XRD	Theoretical values B3LYP/6-311G(d,p)
C2-C1	1,4	1,40665	C3-C2-C1	118	117,8033
C3-C2	1,4	1,40925	C2-C1-C6	121,4	120,8331
C4-C3	1,39	1,38368	C1-C2-C7	122,7	122,0805
C5-C4	1,39	1,39741	C3-C2-C7	119,3	120,1163
C6-C5	1,39	1,39643	C2-C3-C4	120,9	121,4345
C6-C1	1,39	1,38786	C3-C4-C5	120,1	119,9914
C5-O5	1,37	1,36103	C4-C5-C6	120,1	119,4572
C2-C7	1,47	1,46284	C1-C6-C5	119,5	120,4806
C8-C7	1,37	1,36706	C2-C7-C8	129	128,5738
C8-C9	1,44	1,45133	C7-C8-C9	120,9	121,5685
C9-C10	1,43	1,43295	C8-C9-C10	117	116,8491
C10-C11	1,4	1,40255	C9-C10-C11	119,5	119,1801
C10-C15	1,42	1,42539	C9-C10-C15	122,7	122,014
C11-C12	1,39	1,39044	C11-C10-C15	117,7	118,8059
C15-C14	1,38	1,38559	C10-C11-C12	122,8	121,8649
C12-C13	1,39	1,39655	C11-C12-C13	117,4	117,9445
C14-C13	1,4	1,40065	C12-C13-C14	122,3	121,9957
C7-O1	1,37	1,37618	C13-C14-C15	119	119,559
C13-O2	1,36	1,35767	C10-C15-C14	120,8	119,83
C8-O4	1,36	1,35535	C6-C5-O5	122	122,9768
C9-O3	1,26	1,25507	C4-C5-O5	117,9	117,566
C11-O1	1,36	1,35751	C2-C7-O1	111,1	112,3772
C15-O6	1,35	1,33882	C8-C7-O1	119,9	119,0489
			C7-C8-O4	121,4	123,8621
			C9-C8-O4	117,7	114,5694
			C8-C9-O3	120,8	118,5407
			C10-C9-O3	122,3	124,6103
			C10-C11-O1	120,4	120,6879
			C12-C13-O2	121,5	121,6981

			C14-C13-O2	116,2	116,3062
			C10-C15-O6	119,9	120,2527
			C14-C15-O6	119,2	119,9173
			C7-O1-C11	122,1	122,6654
			O1-C11-C12	116,8	117,4471

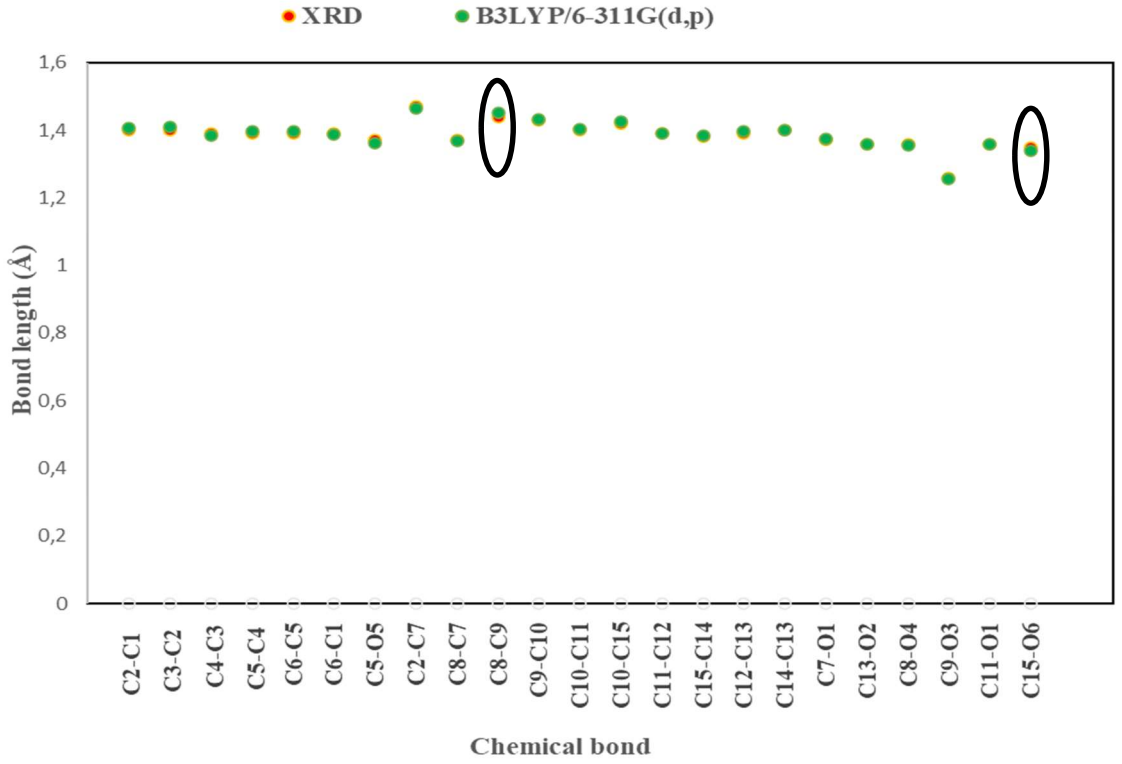

Figure 0.14. *Theoretical and experimental bond lengths of kaempferol. [139].*

The results presented in Fig. 3 indicate that in the case of bond angles, with the exception of four angles, i.e. C7-C8-O4 (2.4°), C9-C8-O4 (3.13°), C8-C9-O3 (2.25°) and C10-C9-O3 (2.31°), the differences between theoretical and experimental values are also quite small. The RMSE =1.05942, which corresponds to the differences between theoretical and experimental values of bond angles, is also acceptable [139].

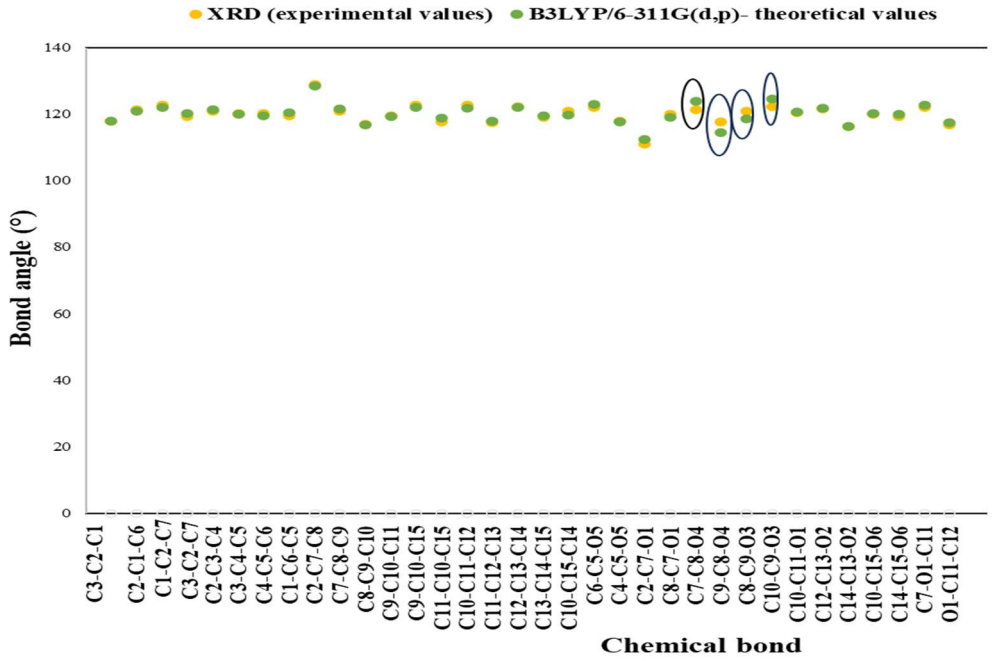

Figure 0.15. *Theoretical and experimental bond angles of kaempferol [139]*

After optimizing the molecular structure of kaempferol, the Gaussian 09 program proved to be a valuable tool in calculating the electronic parameters (dipole moment (DM) and the frontier orbital energies E_{HOMO} and E_{LUMO}) generated in a file with the .log extension. Based on the obtained values, other electronic parameters were computed, such as ionization potential (IE), electron affinity (EA), chemical hardness (η), chemical softness (σ), electronegativity (χ), and electrophilicity index (ω)[139,140,142]. The electronic parameters obtained for kaempferol are presented in table 3.31. They indicate the distribution of electrons in the molecule and estimate the chemical reactivity, assessing the ability of a molecule to give or accept electrons to the frontier molecular orbitals The electrophilicity index ω of kaempferol is 4.1 eV, i.e. larger than 2.0 eV suggesting the strong electrophilic character of the ligand [139].

Table 0.21. *Electronic parameters of kaempferol [139].*

Electronic parameter	Kaempferol
G° (Hartree)	-1029,06
Dipole moment (debye)	4,246
E_{LUMO} (eV)	-2,04
E_{HOMO} (eV)	-5,84
E_{gap} (eV)	3,8
η (eV)	1,9
σ (eV^{-1})	0,26
χ (eV)	3,94
μ (eV)	-3,94
ω (eV)	4,1
IE (eV)	5,84
EA (eV)	2,04

The highest occupied molecular orbital (HOMO) indicates the ability as a donor of electrons, while the lowest unoccupied molecular orbital (LUMO) shows the capacity of accepting electrons [143]. For kaempferol, the difference between the corresponding energies E_{HOMO} and E_{LUMO} is characterized by the parameter E_{gap} = 3,8 eV (see figure 3.20), which indicates a good chemical reactivity.

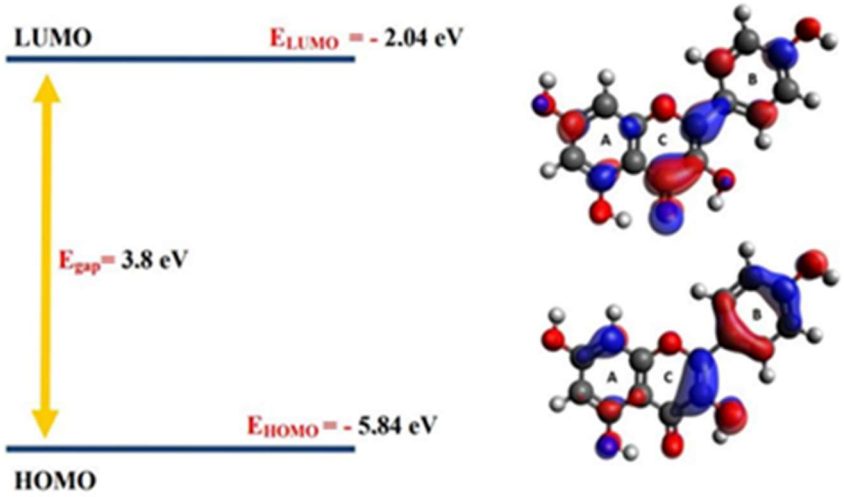

Figure 0.16. *Charge distribution of HOMO and LUMO molecular orbitals in the optimized structure of kaempferol [139].*

The Avogadro software was used to generate the molecular electrostatic potential (MEP) diagram [144,145]. The 3D MEP diagram, presented in figure 3.21, illustrates the relative polarity of kaempferol and its active sites prone to electrophilic respectively nucleophilic attacks. More specifically, it shows the relative polarity of kaempferol and its active sites prone to electrophilic/nucleophilic attacks. The regions with positive potential (represented in red) are located around the hydrogen atoms, which are the most probable active sites for nucleophilic attack [139].

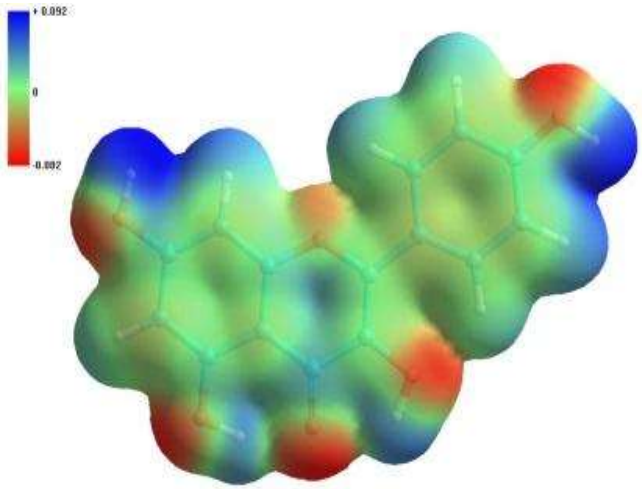

Figure 0.17. *Molecular electrostatic potential (MEP) map of kaempferol [139]*

Molecular docking was performed to establish the interactions between the potential inhibitor (kaempferol) and the PI3Kα protein. By using the AutoDock Vina software, a flexible docking was performed, in which the ligand undergoes translational, rotational, and intramolecular

torsional movements [113,114]. This approach allows the identification of the most favorable conformation in the active site of the PI3Kα protein, while the receptor is considered fixed [139].

Preparation of the receptor macromolecule and the ligand involved several steps using the Gaussian 09 and AutoDock Tools programs: generating a .pdbqt file for the protein; searching for the optimal position for the grid box, optimizing the ligand structure using the computational method DFT/B3LYP/6-311G(d,p); creating a .pdbqt file for the ligand; generating the conf.txt file, which contains all the necessary information for docking; running the AutoDock Vina program, and visualizing the results. The protein structure was obtained from the online database „Protein Data Bank" [88,113,139,146,147].

Following the docking of kaempferol into the active sites of PI3Kα kinase, a stable kinase-ligand complex was formed, with a binding energy of -8.1 kcal/mol (figure 3.22). The interactions of the ligand with neighboring amino acids in the active site of the protein are presented in figure 3.23. It can be observed that the ligand forms polar bonds (hydrogen atom with amino acids SER854A and ASP810A), nonpolar bonds (hydrophobic interactions with ILE932A and TRP780A), and a pi-cation interaction with LYS802.

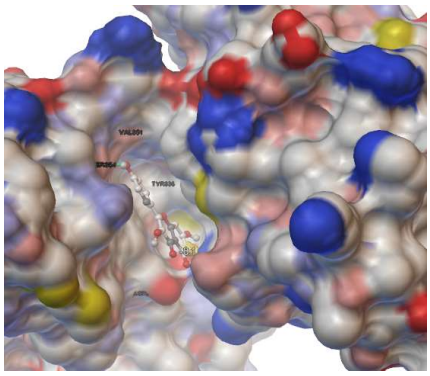

Figure 0.18. *The representation based on the polarity of the kaempferol docked in the active site of the PI3Kα protein.*

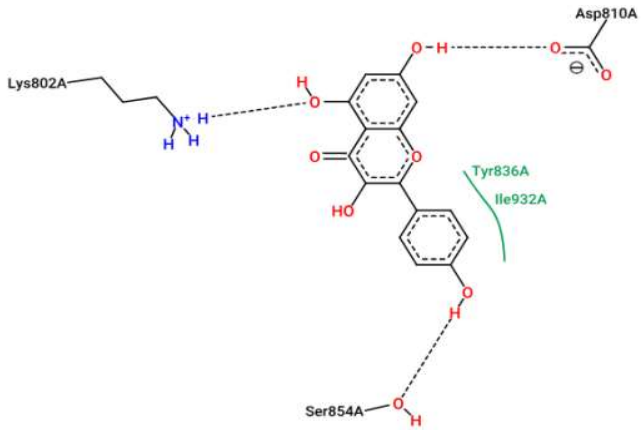

Figure 0.19*. Kaempferol in interaction with the neighboring amino-acids [139].*

To evaluate the potential inhibition of the PI3Kα kinase by kaempferol, we chose (2S)-1-N-[4-methyl-5-[2-(1,1,1-trifluoro-2-methylpropan-2-yl) pyridin-4-yl]-1,3-thiazol-2-yl]pyrrolidine -1,2-dicarboxamide as the standard. The later, also known as alpelisib, has an IC50 = 5.0 ± 0.01 nM [148].

The docking representation of the standard ligand alpelisib and the potential ligand kaempferol in the active site of the PI3Kα protein is shown in figure 3.24. The results indicate that the ligand and the standard almost overlap in the active site of the kinase [139].

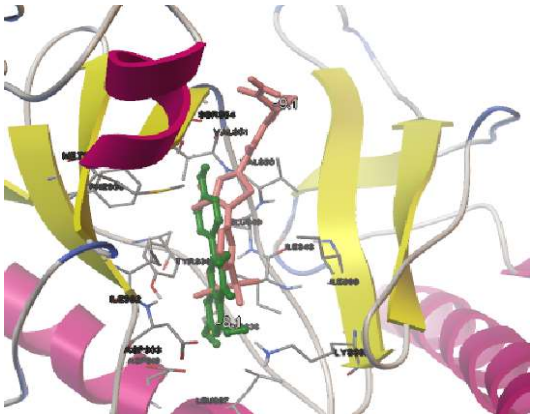

Figure 0.20. *Kaempferol and alpelisib are docked in the active site of PI3Kα kinase surface (green color – kaempferol and pink color - alpelisib)[139].*

Since kaempferol has a higher binding energy (-8.1 kcal/mol) than the standard ligand (-9.1 kcal/mol), we may conclude that the studied ligand forms a less stable complex with the protein macromolecule [139]. Alpelisib is a drug used in cancer therapy, with inhibitory activity [139,149]. Therefore, the investigated ligand can be considered a potential inhibitor of the PI3Kα kinase, but with moderate activity.

3.3 Experimental and computational studies of 3-(Nmorpholinylcarbodithioate)-2-(4-methoxyphenyl) chroman-4-one, a flavanone with potential therapeutical properties [59]

The synthetic flavanone, 3-(Nmorpholinylcarbodithioate)-2-(4-methoxyphenyl)chroman-4-one (figure 3.25), was analyzed in terms of its chemical potential and oral bioavailability to demonstrate its strong electrophilic character, which could thus be utilized in other studies in the field of drug design.

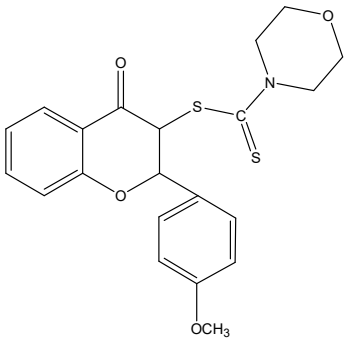

Figure 0.21. *Molecular structure of 3-(N-morpholinylcarbodithioate)-2-(4-methoxyphenyl)chroman-4-one [59].*

The structural and vibrational characterization of this synthetic flavanone was carried out by performing quantum mechanical calculations in the ground state and by using several hybrid functionals (B3LYP, PBE1PBE, CAM-B3LYP and ωB97X-D) in conjunction with the same basis set, namely 6-311G(d,p) [59]. The investigated flavanone has been obtained as a mixture of *anti* (figure 3.26) and *sin* (figure 3.27), he major component being the anti isomer [59].

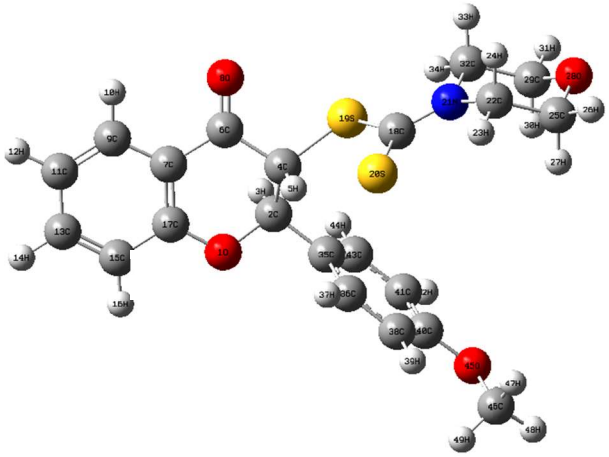

Figure 0.22. *Optimized molecular structure and atom numbering of the anti isomer of 3-(N-morpholinylcarbodithioate)-2-(4-methoxyphenyl)chroman-4-one, at B3LYP/6-311G(d,p) level [59].*

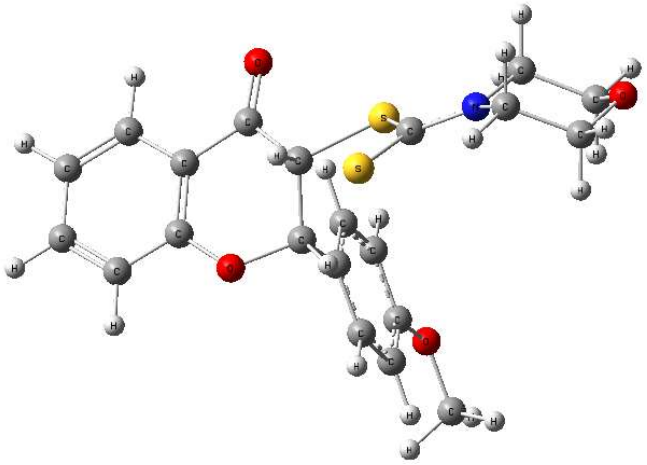

Figure 0.23. *Optimized molecular structure of the syn isomer of 3-(N-morpholinylcarbodithioate)-2-(4-methoxyphenyl)chroman-4-one, at the B3LYP/6-311G(d,p) level of theory [59].*

Following the testing of the four hybrid functionals: B3LYP, PBE1PBE, CAM-B3LYP, and ωB97X-D in correlation with 6-311G(d,p), we were able to conclude that the geometric parameters obtained with B3LYP/6-311G(d,p) are in agreement with those obtained experimentally from the XRD structure. Also, the theoretical wavenumbers are well correlated with the experimental ones. A detailed PED assignment of all the IR absorption bands has been achieved. The theoretical ^{13}C and ^{1}H chemical shifts (with respect to TMS) were compared with the experimental ones. They show a very good agreement in the case of the WP04/6-311++G(d,p) and PBE1PBE/6-311++G(d,p) methods for ^{13}C and in the case of theWP04/6-311++G(d,p) for ^{1}H.

Additionally, several molecular descriptors were investigated by analyzing the descriptors that characterize the steric, hydrophobic/hydrophilic and electronic interactions between a compound and its biological receptor. The quantum chemical parameters determined for the investigated flavanone indicated that it has a good stability and a low chemical reactivity. The value of dipole moment (3,15 debye for *anti* isomer and 3,80 debye for *syn* isomer) is between 3 and 5 debye, which suggests that the studied flavanone has the ability to form non-bonded interactions with a receptor. The electrophilicity index (3,61eV for anti isomer and 3,49 eV for syn

isomer) underscores the strong electrophilic character of the two isomers. The charge distribution of the frontier orbitals HOMO and LUMO for the anti isomer indicates a separation likely due to the presence of a single bond between the carbon atoms (C2-C3) (figure 3.28) [59].

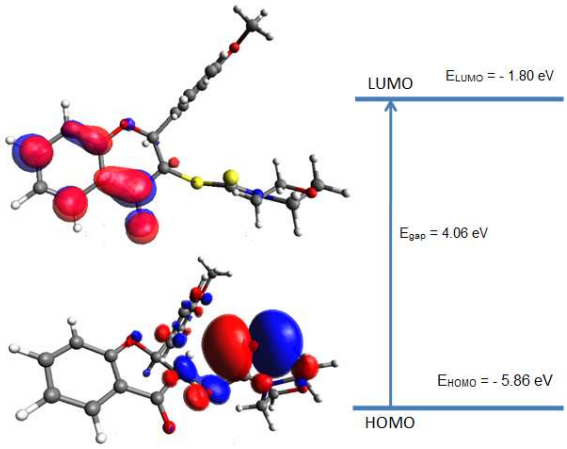

Figure 0.24. *Charge distribution of the frontier orbitals HOMO and LUMO for the optimized structure of the anti isomer of 3-(N-morpholinylcarbodithioate)-2-(4-methoxyphenyl)chroman-4-one [59].*

The bioavailability of the studied flavanone was indicated by a radar diagram, which represents six physico-chemical parameters. Their values (XLogP3 = 3.68, MW = 415.53 g/mol, TPSA = 105.39 $Å^2$, log S(ESOL) = - 4.72, fraction Csp^3 = 0.33 and RotB = 5) indicate that the studied flavanone has the oral bioavailability of drugs of interest, as well as suitable drug-like properties [59].

3.4 Methods for automating cannabinoids detection [150]

The objective of this research was to develop deep convolutional neural networks (DCNN) for the detection and identification of substances of abuse, with a focus on synthetic cannabinoids. In this study, two artificial intelligence models were used: a pre-trained *Convolutional Autoencoder Convolutional (CAE)* and a *Vision Transformer (ViT-B/32).* The optimization algorithms were tested in order to improve the model performance, thereby highlighting the advantages of innovative transfer learning methods that combine unsupervised and supervised techniques with the preprocessing of spectroscopic data [150,151].

The ATR-FTIR spectra used in this research were obtained from the pyDETECT-FTIR database, which contains data from public spectral libraries. Then, they were validated through independent analytical methods. The spectra were converted into monochromatic images (in .bmp format) to train and refine deep the DCNN. Two main spectral libraries were created: one non-forensic, with 11,000 images, that was used for pre-training the CAE, and a forensic one, containing 10,425 images, that was used for training the DCNN on substances of abuse. The spectral images were divided into three classes: class 1 contains 125 images of synthetic cannabinoids of the JWH type; class 2 contains images of other types of cannabinoids; class 3 consists of various other substances, including psychedelics, piperazines, dissociative, empathogens, stimulants, sedatives, and prescription medications [150].

The evaluation and optimization of the models involved the analysis of loss functions, such as cross-entropy and focal loss, as well as testing optimization techniques such as *Adaptive Moment Estimation* (ADAM), *Stochastic Gradient Descent* (SGD), *Sign Stochastic Gradient*

Descent (Sign SGD) and *Root Mean Square Propagation* (RMS Prop) [150]. The quantitative assessment of the artificial intelligence architectures was carried out by using a tool offered by the Wolfram Mathematica v.13.2 software, which was essential for calculating and visualizing the classification indices of success. Thus, information regarding the training dynamics could be recorded [150,152,153]. The results obtained for the developed DCNN are of particular importance in the identification, classification, and recognition of synthetic cannabinoids, having an important relevance in forensics.

Chapter 4. General conclusions and future directions for research and development

The research conducted as part of this doctoral thesis aimed primarily at identifying bioavailable substances, in order to contribute to the optimization and efficiency of production processes. Thus, the research activity was focused on two main directions:

- Utilization of advanced computational tools for processing IR spectra and the molecular descriptors of four bioavailable substances from the class flavonoids, with the purpose of optimizing the techniques for identifying compounds with high therapeutic potential, thereby enhancing the drug design process;
- Identification and classification of bioavailable substances from the class of JWH synthetic cannabinoids through the development of DCNNs.

The first research direction for identifying bioavailable substances from the flavonoid class involved:

- Optimization of molecular structures obtained from public databases by comparing the spectra of theoretical models with experimental spectra. Four computational techniques used for optimizing molecular structures were evaluated, and the best agreements between theoretical and experimental geometric parameters were found by using the hybrid functional B3LYP/6-311G(d,p);
- Characterization of these substances by creating physico-chemical profiles and determining molecular descriptors that highlight their bioavailability, classifying them as substances that can be used as drugs. The results obtained for the four flavonoids (two natural compounds and two synthetic compounds) support their bioavailable nature;
- Their identification as potential inhibitors of the PI3Kα protein by using molecular docking. The studied compounds, resokaempferol and tectochrysin, yielded good results in molecular docking, as indicated by three different programs (AutoDock, AutoDock Vina, and Glide) for two proteins from the PI3Kα kinase family (the normal form and the H1047R mutation). The originality and importance of the research stem from exploring these substances as potential inhibitors of the two proteins, which can provide essential insights for developing new targeted therapies for cancer treatment. The possible interactions of kaempferol with the PI3Kα kinase receptor demonstrated its antioxidant potential (i.e., the ability of kaempferol to inhibit the PI3Kα kinase, using alpelisib as a standard with an IC50 of 5.0±0.01 nM), and thus its inhibitory potential in cancer treatment.

The second research direction focused on:

- creating the physico-chemical profiles of five synthetic cannabinoids and comparing them with THC, the main active compound of cannabis, to establish their bioavailability. The results obtained indicate their oral bioavailability, as well as an increased tendency to accumulate in lipid environments rather than in aqueous ones.
- Developing neural networks for the automated identification and characterization of synthetic cannabinoids. The results obtained demonstrate that innovative transfer learning methods combined with spectroscopic data preprocessing provide significant benefits in the rapid and efficient detection of synthetic cannabinoids as prohibited substances.

Future directions for research and development

Regarding future research and development directions, we consider the following:

- identifying other compounds from the flavonoid class, as they represent a significant dietary resource for humans. Their diverse molecular structures, unique functionality, and rich pharmacological profiles have made them valuable targets for drug design;
- utilizing molecular dynamics for in-depth studies of the inhibitory potential on proteins responsible for cancer development;
- evaluating the inhibitory potential of flavonoids on other proteins, such as COX (cyclooxygenase), which is involved in inflammatory processes;
- studying cannabinoids as potential therapeutic agents in cancer treatments;
- developing additional classification models that allow for the inclusion of other categories of cannabinoids.

List of publications

List of ISI publications

- **Articles in process of being ISI indexed and/or ranked (Web of Science)**

1. **Paraschiv C.,** Gosav S., Burlacu C. M., Praisler M., Exploring the Inhibitory Efficacy of Resokaempferol and Tectochrysin on PI3Kα Protein by Combining Spectroscopy, DFT, and Molecular Docking Against Wild-Type and H1047R Mutant Forms, *Inventions* 9 (2024) 96. DOI: 10.3390/inventions9050096 ISI Impact Factor: 2.1 (2023) https://www.mdpi.com/2411-5134/9/5/96

2. **Paraschiv, C**., Gosav, S., Praisler, M., Chemical Potential Evaluation of Kaempferol Based on Molecular Modeling, *Advances in Digital Health and Medical Bioengineering. Proceedings of the 11th International Conference on E-Health and Bioengineering, EHB-2023*, November 9–10, 2023, Bucharest, Romania.

- **ISI indexed and/or ranked articles (Web of Science)**

3. Gosav, S., Praisler, M., Păduraru Cristea, N., Gosav, T., **Paraschiv, C.**, Experimental and computational studies of 3-(N-morpholinylcarbodithioate)-2-(4-methoxyphenyl)chroman-4-one, a flavanone with potential therapeutical properties, *Journal of Molecular Structure 1295 Part 1* (2024) 136627, DOI: 10.1016/j.molstruc.2023.136627, **WOS:001084143900001,** ISI Impact Factor: **4.0** (2023)
https://www.sciencedirect.com/science/article/abs/pii/S0022286023017179

ARTICLE CITED IN

1	Ali I., Mimouni F.Z., Belboukhari N., Sekkoum K., Locatelli M., Demir E., Yusuf K., Enantiomeric separation of flavanone on Chiralpak® IA column and determination of the chiral mechanism, *Biomedical Chromatography,* First published: 05 September 2024, DOI: 10.1002/bmc.6004 ISI Impact Factor: **1.8** (2023) https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/bmc.6004
2	Zhang Y., Shang X., Zhang Y., Zheng C., Physiological Function of Plant Flavonoids and Its Application in Poultry Production, *Chinese Journal of Animal Nutrition*, 36-9 (2024) 5468-5478. DOI: 10.12418/CJAN2024.466 **Scopus indexed**

4. Burlacu, C. M., Burlacu, A. C., Praisler, M., **Paraschiv C.**, Harnessing Deep Convolutional Neural Networks Detecting Synthetic Cannabinoids: A Hybrid Learning Strategy for Handling Class Imbalances in Limited Datasets, *Inventions* 8 (2023) 129. DOI: 10.3390/inventions8050129 **WOS:001089546600001** ISI Impact Factor: **2.1** (2023) https://www.mdpi.com/2411-5134/8/5/129

	ARTICLE CITED IN
1	Annaki I., Rahmoune M., Bourhaleb M., Overview of Data Augmentation Techniques in Time Series Analysis, *International Journal of Advanced Computer Science and Applications*, 15-1 (2024) 1201-1211. ISI Impact Factor: **0.7** (2023) https://thesai.org/Downloads/Volume15No1/Paper_118-Overview_of_Data_Augmentation_Techniques.pdf

Articles published in international data bases (BDI) journals

1. **Paraschiv (Nedelcu) C.**, Praisler M., Spectral Identification of Cannabinoids, *Annals of "Dunarea de Jos" University of Galati, Mathematics, Physics, Theoretical Mechanics*, Fascicle II, Year XV (XLVI) 2023, No. 1, pp. 5 - 10. DOI: 10.35219/ann-ugal-math-phys-mec.2023.1.02. ISSN 2067-2071
https://www.gup.ugal.ro/ugaljournals/index.php/math/article/view/6233

2. **Nedelcu (Paraschiv) C.**, M. Praisler, Cannabinoids: an exploratory review, *Annals of the "Dunarea de Jos" University of Galati, Mathematics, Physics, Theoretical Mechanics*, Fascicle II, Year XIV (XLV) 2022, No. 1, pp. 11–18. DOI: 10.35219/ann-ugal-math-phys-mec.2022.1.03.
https://www.gup.ugal.ro/ugaljournals/index.php/math/article/view/5509

Articles in process of publication in BDI journals

1. Paraschiv, C., Gosav, S., Praisler, M., Characterization of flavonoids: an explorative review, *Annals of "Dunarea de Jos" University of Galati, Mathematics, Physics, Theoretical Mechanics*, Fascicle II, Year XVI (XLVII) 2024, No. 1, DOI: 10.35219/ann-ugal-math-phys-mec.2024

Articles presented at international conferences

1. **Paraschiv, C**., Gosav, S., Praisler, M., Chemical Potential Evaluation of Kaempferol Based on Molecular Modeling, *Advances in Digital Health and Medical Bioengineering, Proceedings of the 11th International Conference on E-Health and Bioengineering, EHB-2023*, November 9–10, 2023, Bucharest, Romania. http://www.ehbconference.ro/

2. Gosav S., **Paraschiv C.,** Praisler M., Evaluation the chemical potential of some flavones using the DFT theory and molecular docking, *International Conference and Workshop "Interdisciplinary applications of advanced analytical and control techniques in environment, health and materials science - INTERVENT"*, 19-20 October 2023, Galați, Romania.
https://www.researchgate.net/publication/375548506_Abstract_Book_-_International_Conference_and_Workshop_Interdisciplinary_applications_of_advanced_analytical_and_control_techniques_in_environment_health_and_materials_science_-_INTERVENT_19-20_October

https://www.researchgate.net/publication/373217117_Interdisciplinary_applications_of_advanced_analytical_and_control_techniques_in_environment_health_and_materials_science_-_INTERVENT

Articles presented at national conferences

1. **Nedelcu (Paraschiv) C**., Gosav S., Burlacu C.M., Praisler M., Assessing the chemical potential of certain flavones based on DFT theory and molecular docking, *Book of abstracts SCDS-UDJG 2024, 12th Edition*, 6 - 7 June 2024, Galați, Romania https://cssd-udjg.ugal.ro/media/attachments/2024/05/31/book-of-abstract-2024-final_sdimi.pdf

2. **Nedelcu (Paraschiv) C**., Gosav S., Burlacu C.M., Praisler M., Characterization of flavonoids: an explorative review, *Book of abstracts SCDS-UDJG 2024, 12th Edition*, 6 - 7 June 2024, Galați, Romania https://cssd-udjg.ugal.ro/media/attachments/2024/05/31/book-of-abstract-2024-final_sdimi.pdf

3. **Nedelcu (Paraschiv) C.,** Negoiță C., Praisler M., KNN system screening for the main illicit drugs, *Book of abstracts SCDS-UDJG 2023, 11th Edition*, 8 - 9 June 2023, Galați, Romania.

4. **Nedelcu (Paraschiv) C.,** Praisler M., Spectral identification of cannabinoids – a review, *Book of abstracts SCDS-UDJG 2023, 11th Edition*, 8 - 9 June 2023, Galați, Romania.

5. **Nedelcu (Paraschiv) C.,** Praisler M., Cannabinoids: an exploratory review, *Book of abstracts SCDS-UDJG 2022, 10th Edition*, 9 - 10 June 2022, Galați, Romania.

Awards received

1. Second prize award in recognition of the scientific contribution to the Twelfth Edition of the Scientific Conference of the Doctoral Schools of „Dunărea de Jos", **Nedelcu (Paraschiv) C**., Gosav S., Burlacu C.M., Praisler M., Assessing the chemical potential of certain flavones based on DFT theory and molecular docking, *SCDS-UDJG 2024, 12th Edition*, 6 – 7 June 2024, Galați, Romania.

References

[1] P. Bawiec, J. Sawicki, P. Łasińska-Pracuta, M. Czop, I. Sowa, K. Iłowiecka, W. Koch, In Vitro Evaluation of Bioavailability of Se from Daily Food Rations and Dietary Supplements, Nutrients 15 (2023) 1511.

[2] N. Sharma, D. Saini, R.K. Kesharwani, P.C. Gupta, R.K. Keservani, Advances in Flavonoids for Human Health and Prevention of Diseases, 1st ed., Apple Academic Press, New York, 2023.

[3] World Health Organization, Cancer, (n.d.). https://www.who.int/news-room/fact-sheets/detail/cancer (accessed July 23, 2024).

[4] OECD, Profil de țară privind cancerul: România 2023, Organisation for Economic Co-operation and Development, Paris, 2023. https://www.oecd-ilibrary.org/social-issues-migration-health/profil-de-tara-privind-cancerul-romania-2023_20726aa6-ro (accessed July 23, 2024).

[5] V. Ghicavîi, N. Bacinschi, Gh. Gușuilă, Farmacologie, Ediția a II a revizuită și completată, Tipografia Centrală, Chișinău, 2019.

[6] A.N. Cristea, Farmacologie generală, Ediția a II-a, Editura Didactică și Pedagogică, București, 2018.

[7] Food Food and Drug Administration, Waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics

classification system: guidance for industry, Food Drug Adm. Cent. Drug Eval. Res. CDER (2015).

[8] V. P Chavda, Biological Classification System (BCS); with a New Perspective, MOJ Bioequivalence Bioavailab. 3 (2017).

[9] BDDCS-Biopharmaceutical Drug Disposition Classification System, (n.d.). https://www.solvobiotech.com/knowledge-center/bddcs.

[10] M.-A. Crocq, History of cannabis and the endocannabinoid system, Dialogues Clin. Neurosci. 22 (2020) 223–228.

[11] F.A. Javid, R.M. Phillips, S. Afshinjavid, R. Verde, A. Ligresti, Cannabinoid pharmacology in cancer research: A new hope for cancer patients?, Eur. J. Pharmacol. 775 (2016) 1–14.

[12] A. Pugazhendhi, N. Suganthy, T.P. Chau, A. Sharma, Y. Unpaprom, R. Ramaraj, I. Karuppusamy, K. Brindhadevi, Cannabinoids as anticancer and neuroprotective drugs: Structural insights and pharmacological interactions—A review, Process Biochem. 111 (2021) 9–31.

[13] N. Mangal, S. Erridge, N. Habib, A. Sadanandam, V. Reebye, M.H. Sodergren, Cannabinoids in the landscape of cancer, J. Cancer Res. Clin. Oncol. 147 (2021) 2507–2534.

[14] C. (Paraschiv) Nedelcu, M. Praisler, Cannabinoids: an exploratory review, Analele Univ. ”Dunărea Jos” Din Galați Fasc. II Mat. Fiz. Mec. Teor. 45 (2022) 11–18.

[15] Endocannabinoids: What are they and what do they do?, (2021). https://www.medicalnewstoday.com/articles/endocannabinoid (accessed July 30, 2024).

[16] U.N.O.O.D.A. Labor, world drug report 2021 (set of 5 booklets)., united nations, S.l., 2022.

[17] F. Zapata Arráez, J.M. Matey Cabañas, G. Montalvo García, C. García Ruiz, Chemical classification of new psychoactive substances (NPS), Microchem. J. 163 (2021) p.105877.

[18] A. Barbulescu, L. Barbes, C.-S. Dumitriu, Computer-Aided Classification of New Psychoactive Substances, J. Chem. 2021 (2021) 1–11.

[19] C.M. Burlacu, A.C. Burlacu, M. Praisler, Physico-chemical analysis, systematic benchmarking, and toxicological aspects of the JWH aminoalkylindole class-derived synthetic JWH cannabinoids, Ann. ”Dunarea Jos” Univ. Galati Fascicle II Math. Phys. Theor. Mech. 44 (2021) 34–45.

[20] C. Lidia, M. Eugenia, P. Doina, Cercetarea fitochimică a substanţelor biologic active din strugurii Vitis vinifera, in: n.d. https://ibn.idsi.md/vizualizare_articol/107906 (accessed July 30, 2024).

[21] A.N. Panche, A.D. Diwan, S.R. Chandra, Flavonoids: an overview, J. Nutr. Sci. 5 (2016) e47.

[22] N. Uludağ, G. Serdaroğlu, An improved synthesis, spectroscopic (FT-IR, NMR) study and DFT computational analysis (IR, NMR, UV–Vis, MEP diagrams, NBO, NLO, FMO) of the 1,5-methanoazocino[4,3-b]indole core structure, J. Mol. Struct. 1155 (2018) 548-560.

[23] M.C. Dias, D.C.G.A. Pinto, A.M.S. Silva, Plant Flavonoids: Chemical Characteristics and Biological Activity, Molecules 26 (2021) 5377.

[24] R. Shukla, V. Pandey, G.P. Vadnere, S. Lodhi, Role of Flavonoids in Management of Inflammatory Disorders, (2019) 293–322.

[25] L. Calmuțchi, E. Melentiev, G. Climenti, Study Of Same Flavonoides In Research Activities Focused On Student, Rev. Științifică 10 (2017) 125–131.

[26] T. Pinto, A. Vilela, Healthy Drinks with Lovely Colors: Phenolic Compounds as Constituents of Functional Beverages, Beverages 7 (2021) 12.

[27] M. Guasch-Ferré, J. Merino, Q. Sun, M. Fitó, J. Salas-Salvadó, Dietary Polyphenols, Mediterranean Diet, Prediabetes, and Type 2 Diabetes: A Narrative Review of the Evidence, Oxid. Med. Cell. Longev. 2017 (2017) 6723931.

[28] F. Nazzaro, F. Fratianni, V. De Feo, A. Battistelli, A.G. Da Cruz, R. Coppola, Polyphenols, the new frontiers of prebiotics, Adv. Food Nutr. Res. 94 (2020) 35–89.

[29] A.B. Das, V.V. Goud, C. Das, 9 - Phenolic Compounds as Functional Ingredients in Beverages, in: A.M. Grumezescu, A.M. Holban (Eds.), Value-Added Ingred. Enrich. Beverages, Academic Press, 2019: pp. 285–323.

[30] J. Bae, N. Kim, Y. Shin, S.-Y. Kim, Y.-J. Kim, Activity of catechins and their applications, Biomed. Dermatol. 4 (2020) 8.

[31] H. Speer, N. D'Cunha, N. Alexopoulos, A. Mckune, N. Naumovski, Anthocyanins and Human Health—A Focus on Oxidative Stress, Inflammation and Disease, Antioxidants 9 (2020) 366.

[32] A.H. Abari, M. Tayebi, Bioconversion of Genistein to Orobol by Bacillus subtilis Spore Displayed Tyrosinase and Monitoring the Anticancer Effects of Orobol on MCF-7 Breast Cancer Cells, Biotechnol. Bioprocess Eng. 24 (2019) 507–512.

[33] H. Liu, Y. Tian, Y. Zhou, Y. Kan, T. Wu, W. Xiao, Y. Luo, Multi-modular engineering of Saccharomyces cerevisiae for high-titre production of tyrosol and salidroside, Microb. Biotechnol. 14 (2021) 2605–2616.

[34] Human Metabolome Database: Showing metabocard for 3,4',7-Trihydroxyflavone (HMDB0034004), (n.d.). https://hmdb.ca/metabolites/HMDB0034004 (accessed July 31, 2024).

[35] PubChem, 3,7,4'-Trihydroxyflavone, (n.d.). https://pubchem.ncbi.nlm.nih.gov/compound/5281611 (accessed July 31, 2024).

[36] J. Zhang, J. Wu, X. Shi, D. Li, S. Yang, R. Zhang, B. Xia, G. Yang, A Propolis-Derived Small Molecule Tectochrysin Ameliorates Type 2 Diabetes in Mice by Activating Insulin Receptor β, Mol. Nutr. Food Res. 68 (2024) e2300283.

[37] M. Boufadi, S. Keddari, F. Moulai-Hacene, S. Chaa, Chemical Composition, Antioxidant and Anti-Inflammatory Properties of Salvia Officinalis Extract from Algeria, Pharmacogn. J. 13 (2021) 506–515.

[38] N. Kharazian, Flavonoid Constituents in Some Species of Salvia L. (Lamiaceae) in Iran, 25 (2014).

[39] B. He, F. Xu, T. Yan, F. Xiao, B. Wu, Y. Wang, K. Bi, Y. Jia, Tectochrysin from Alpinia Oxyphylla Miq. alleviates Aβ1-42 induced learning and memory impairments in mice, Eur. J. Pharmacol. 842 (2019) 365–372.

[40] PubChem, Tectochrysin, (n.d.). https://pubchem.ncbi.nlm.nih.gov/compound/5281954 (accessed July 31, 2024).

[41] M. Herrero, M. Plaza, A. Cifuentes, E. Ibáñez, Extraction techniques for the determination of phenolic compounds in food, Compr. Sampl. Sample Prep. 4 (2012) 159–180.

[42] M. Imran, B. Salehi, J. Sharifi-Rad, T. Aslam Gondal, F. Saeed, A. Imran, M. Shahbaz, P.V. Tsouh Fokou, M. Umair Arshad, H. Khan, S.G. Guerreiro, N. Martins, L.M. Estevinho, Kaempferol: A Key Emphasis to Its Anticancer Potential, Molecules 24 (2019) 2277.

[43] S. Parveen, I.U.H. Bhat, R. Bhat, Kaempferol and its derivatives: Biological activities and therapeutic potential, Asian Pac. J. Trop. Biomed. 13 (2023) 411.

[44] M.R. Khazdair, A. Anaeigoudari, G.A. Agbor, Anti-viral and anti-inflammatory effects of kaempferol and quercetin and COVID-2019: A scoping review, Asian Pac. J. Trop. Biomed. 11 (2021) 327.

[45] PubChem, Kaempferol, (n.d.). https://pubchem.ncbi.nlm.nih.gov/compound/5280863 (accessed July 31, 2024).

[46] M. Pourseyed Lazarjani, S. Torres, T. Hooker, C. Fowlie, O. Young, A. Seyfoddin, Methods for quantification of cannabinoids: a narrative review, J. Cannabis Res. 2 (2020) 35.

[47] ***, Mai mult decât văd ochii: spectrul electromagnetic, Sci. Sch. (n.d.). https://www.scienceinschool.org/ro/article/2011/em-3-ro/ (accessed August 1, 2024).

[48] C. Mignone, R. Barnes, More than meets the eye: the electromagnetic spectrum, (n.d.).

[49] ***, Spectrul Electromagnetic, ISA - Învățare Prin Sunet Si Atingere (n.d.). https://www.idea-isa.ro/lesson/spectrul-electromagnetic/ (accessed August 1, 2024).

[50] ***, The electromagnetic spectrum, (n.d.). https://www.esa.int/Science_Exploration/Space_Science/Integral/The_electromagnetic_spectrum (accessed August 1, 2024).

[51] ***, What Is Infrared Spectroscopy? Fundamentals & Applications, (n.d.). https://www.excedr.com/blog/what-is-infrared-spectroscopy (accessed August 1, 2024).

[52] ***, Infrared: Interpretation, Chem. Libr. (2013). https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Spectroscopy/Vibrational_Spectroscopy/Infrared_Spectroscopy/Infrared%3A_Interpretation (accessed August 1, 2024).

[53] ***, https://www.academia.edu/17645916/UNIVERSITATEA_DE_STIINTE_AGRONOMICE_SI_MEDICINA_VETERINARA_BUCURESTI (accessed August 1, 2024).

[54] M. Dorobanțu, Analiza structurala organica prin metode spectrale. Curs 5, (n.d.). https://chem.ubbcluj.ro/~darab/courses/Organic_Chemistry_I/Curs-5.pdf (accessed August 2, 2024).

[55] ***, Spectroscopia de vibrație (IR), (n.d.). http://chim.upt.ro/_old/comunicate-cadre/84604spectro_IR1.pdf (accessed August 2, 2024).

[56] M. Dorobanțu, Analiza structurala organica prin metode spectrale. Curs 4, (n.d.). https://chem.ubbcluj.ro/~darab/courses/Organic_Chemistry_I/Curs-4.pdf (accessed August 2, 2024).

[57] C. Paraschiv, M. Praisler, Spectral identification of cannabinoids, Analele Univ. ”Dunărea Jos” Din Galați Fasc. II Mat. Fiz. Mec. Teor. 46 (2023) 5–10.

[58] ***, SpectraBase, (n.d.). https://spectrabase.com/ (accessed August 2, 2024).

[59] S. Gosav, M. Praisler, N.P. Cristea, T. Gosav, C. Paraschiv, Experimental and computational studies of 3-(N-morpholinylcarbodithioate)-2-(4-methoxyphenyl)chroman-4-one, a flavanone with potential therapeutical properties, J. Mol. Struct. Part 1 1295 (2023) 136627.

[60] A. Mauri, M. Bertola, M. Alvascience, A New Software Suite for the QSAR Workflow Applied to the Blood–Brain Barrier Permeability, Int. J. Mol. Sci 23 (2022).

[61] D. Lagorce, L. Bouslama, J. Becot, M.A. Miteva, B.O. Villoutreix, FAF-Drugs4: free ADME-tox filtering computations for chemical biology and early stages drug discovery, Bioinformatics 33 (2017) 3658–3660.

[62] ***, https://chem.libretexts.org/Courses/Howard_University/General_Chemistry%3A_An_Atoms_First_Approach/Unit_1%3A__Atomic_Structure/Chapter_1%3A_Introduction/Chapter_1.7%3A__The_Mole_and_Molar_Mass (accessed August 3, 2024).

[63] ***, Calculating Molecular Weight, (n.d.). https://chemcollective.org/activities/tutorials/stoich/calculating_molecular_weight (accessed August 3, 2024).

[64] ***, Definition of molecular weight - NCI Dictionary of Cancer Terms - NCI, (2011). https://www.cancer.gov/publications/dictionaries/cancer-terms/def/molecular-weight (accessed August 3, 2024).

[65] P.D. Leeson, Molecular inflation, attrition and the rule of five, Adv. Drug Deliv. Rev. 101 (2016) 22–33.

[66] ***, Molecular weight - MolModa Documentation, (n.d.). https://durrantlab.pitt.edu/molmoda/docs/structures/molprop/weight/ (accessed August 3, 2024).

[67] Edited by V. Poongavanam, V. Ramaswamy, Computational Drug Discovery. Methods and Applications, Wiley-VCH, 2024. https://www.wiley.com/en-us/Computational+Drug+Discovery%3A+Methods+and+Applications%2C+2+Volumes-p-9783527840731.

[68] S. Kralj, M. Jukič, U. Bren, Comparative Analyses of Medicinal Chemistry and Cheminformatics Filters with Accessible Implementation in Konstanz In-formation Miner (KNIME), Int. J. Mol. Sci 23 (2022).

[69] ***, van der Waals radius, Wikipedia (2024). https://en.wikipedia.org/w/index.php?title=Van_der_Waals_radius&oldid=1224084649 (accessed August 4, 2024).

[70] T.M. Martin, User's Guide for T.E.S.T. (Toxicity Estimation Software Tool), (n.d.).

[71] B. Chandrasekaran, S.N. Abed, O. Al-Attraqchi, K. Kuche, R.K. Tekade, Chapter 21 - Computer-Aided Prediction of Pharmacokinetic (ADMET) Properties, in: R.K. Tekade (Ed.), Dos. Form Des. Parameters, Academic Press, 2018: pp. 731–755.

[72] V. Baerle, An Overview Of Bioavailability In Silico Prediction, Biostudent 3 (2) (2020) 157–164.

[73] J. Plante, S. Werner, JPlogP: an improved logP predictor trained using predicted data, J. Cheminformatics 10 (2018) 61.

[74] ***, logP - MolModa Documentation, (n.d.). https://durrantlab.pitt.edu/molmoda/docs/structures/molprop/logp/ (accessed August 3, 2024).

[75] A. Avdeef, E. Fuguet, A. Llinàs, C. Ràfols, E. Bosch, G. Völgyi, T. Verbić, E. Boldyreva, K. Takács-Novák, Equilibrium solubility measurement of ionizable drugs – consensus recommendations for improving data quality, ADMET DMPK 4 (2016) 117–178.

[76] ***, Theory of aqueous solubility prediction | Chemaxon Docs, (n.d.). https://docs.chemaxon.com/display/docs/calculators_theory-of-aqueous-solubility-prediction.md#src-1806933-theoryofaqueoussolubilityprediction-examplesofpredictedph-dependentlogscurves (accessed August 20, 2024).

[77] M.L. Landry, J.J. Crawford, LogD Contributions of Substituents Commonly Used in Medicinal Chemistry, ACS Med. Chem. Lett. 11 (2019) 72–76.

[78] S. Bhal, Lipophilicity Descriptors: Understanding When to Use LogP & LogD, (n.d.).

[79] ***, Polar surface area, Wikipedia (2024). https://en.wikipedia.org/w/index.php?title=Polar_surface_area&oldid=1218207188 (accessed August 20, 2024).

[80] O.B. Akawa, F.O. Okunlola, M.I. Alahmdi, N.E. Abo-Dya, P.A. Sidhom, M.A.A. Ibrahim, M.F. Shibl, S. Khan, M.E.S. Soliman, Multi-cavity molecular descriptor interconnections: Enhanced protocol for prediction of serum albumin drug binding, Eur. J. Pharm. Biopharm. 194 (2024) 9–19.

[81] L. Fu, S. Shi, J. Yi, N. Wang, Y. He, Z. Wu, J. Peng, Y. Deng, W. Wang, C. Wu, A. Lyu, X. Zeng, W. Zhao, T. Hou, D. Cao, ADMETlab 3.0: an updated comprehensive online ADMET prediction platform enhanced with broader coverage, improved performance, API functionality and decision support, Nucleic Acids Res. 52 (2024) W422–W431.

[82] A. Daina, O. Michielin, V. Zoete, iLOGP: A Simple, Robust, and Efficient Description of *n* -Octanol/Water Partition Coefficient for Drug Design Using the GB/SA Approach, J. Chem. Inf. Model. 54 (2014) 3284–3301.

[83] A. Daina, O. Michielin, V. Zoete, SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci. Rep. 7 (2017) 42717.

[84] C. Paraschiv, S. Gosav, C.M. Burlacu, M. Praisler, Exploring the Inhibitory Efficacy of Resokaempferol and Tec-tochrysin on PI3Kα Protein by Combining DFT and Molecular Docking Against Wild-Type and H1047R Mutant Forms, Inventions 9 (2024) 96.

[85] B. Kieron, The ABC of DFT, (n.d.). https://dft.uci.edu/doc/g1.pdf (accessed August 24, 2024).

[86] S. Liu, ed., Front Matter, in: Concept. Density Funct. Theory, 1st ed., Wiley, 2022.

[87] P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, Phys. Rev. 136 (1964) B864–B871.

[88] ***,Density Functional (DFT) Methods | Gaussian.com, (n.d.). https://gaussian.com/dft/ (accessed August 24, 2024).

[89] F. Stanzione, I. Giangreco, J.C. Cole, Use of molecular docking computational tools in drug discovery, in: Prog. Med. Chem., Elsevier, 2021: pp. 273–343.

[90] S. Dastmalchi, M. Hamzeh-Mivehroud, B. Sokouti, eds., Methods and Algorithms for Molecular Docking-Based Drug Design and Discovery:, IGI Global, 2016.

[91] P. Castel, E. Toska, J.A. Engelman, M. Scaltriti, The present and future of PI3K inhibitors for cancer therapy, Nat. Cancer 2 (2021) 587–597.

[92] X. Yang, X. Zhang, M. Huang, K. Song, X. Li, M. Huang, L. Meng, J. Zhang, New Insights into PI3K Inhibitor Design using X-ray Structures of PI3Kα Complexed with a Potent Lead Compound, Sci. Rep. 7 (2017) 14572.

[93] Y. He, M.M. Sun, G.G. Zhang, J. Yang, K.S. Chen, W.W. Xu, B. Li, Targeting PI3K/Akt signal transduction for cancer therapy, Signal Transduct. Target. Ther. 6 (2021) 425.

[94] T.A. Zughaibi, M. Suhail, M. Tarique, S. Tabrez, Targeting PI3K/Akt/mTOR Pathway by Different Flavonoids: A Cancer Chemopreventive Approach, Int. J. Mol. Sci. 22 (2021) 12455.

[95] F. Janku, J.J. Wheler, A. Naing, G.S. Falchook, D.S. Hong, V.M. Stepanek, S. Fu, S.A. Piha-Paul, J.J. Lee, R. Luthra, A.M. Tsimberidou, R. Kurzrock, PIK3CA Mutation H1047R Is Associated with Response to PI3K/AKT/mTOR Signaling Pathway Inhibitors in Early-Phase Clinical Trials, Cancer Res. 73 (2013) 276–284.

[96] L. Nisa, P. Häfliger, M. Poliaková, R. Giger, P. Francica, D.M. Aebersold, R.-P. Charles, Y. Zimmer, M. Medová, PIK3CA hotspot mutations differentially impact responses to MET targeting in MET-driven and non-driven preclinical cancer models, Mol. Cancer 16 (2017) 93. https://doi.org/10.1186/s12943-017-0660-5.

[97] S. Guo, S. Loibl, G. von Minckwitz, S. Darb-Esfahani, B. Lederer, C. Denkert, PIK3CA H1047R Mutation Associated with a Lower Pathological Complete Response Rate in Triple-Negative Breast Cancer Patients Treated with Anthracycline-Taxane–Based Neoadjuvant Chemotherapy, Cancer Res. Treat. Off. J. Korean Cancer Assoc. 52 (2020) 689–696.

[98] H. Cheng, S.T.M Orr, S. Bailey, A. Brooun, P. Chen, J.G. Deal, Y.L. Deng, M.P. Edwards, G.M. Gallego, N. Grodsky, et al. Structure-Based Drug Design and Synthesis of PI3Kα-Selective Inhibitor (PF-06843195), J. Med. Chem. 64 (2021) 644-661.

[99] B.V. Bhaskar, A. Rammohan, T.M. Babu, G.Y. Zheng, W. Chen, W. Rajendra, G.V. Zyryanov, W. Gu, Molecular insight into isoform specific inhibition of PI3K-α and PKC-η with dietary agents through an ensemble pharmacophore and docking studies, Sci. Rep. 11 (2021) 12150.

[100] Schrödinger Release 2024-2: Epik, in: Schrödinger, LLC, New York, NY, 2024.

[101] Schrödinger Release 2024-2: LigPrep, in: Schrödinger, LLC, New York, NY, 2024.

[102] Schrödinger Release 2024-2: Maestro, in: Schrödinger, LLC, New York, NY, 2024.

[103] Schrödinger Release 2024-2: Prime, in: Schrödinger, LLC, New York, NY, 2024.

[104] Schrödinger Release 2024-2: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2024; Impact, Schrödinger, LLC, New York, NY; Prime, in: Schrödinger, LLC, New York, NY, 2024.

[105] M.S. Edited by Coumar, Fundamentals, Techniques, Resources and Applications, 1st ed., Academic Press, Elsevier, 2021.

[106] M. Mukhopadhyay, A Brief Survey On Bio Inspired Optimization Algorithms For Molecular Docking, Int. J. Adv. Eng. Technol. 7 (2014) 868–878.

[107] K.K. Chaudhary, N. Mishra, A Review on Molecular Docking: Novel Tool for Drug Discovery, JSM Chem 4 (2016) 1029.

[108] S. Tang, R. Chen, M. Lin, Q. Lin, Y. Zhu, J. Ding, H. Hu, M. Ling, J. Wu, Accelerating AutoDock Vina with GPUs, Molecules 27 (2022) 3041.

[109] T. Chen, X. Shu, H. Zhou, F.A. Beckford, M. Misir, Algorithm selection for protein–ligand docking: strategies and analysis on ACE, Sci. Rep. 13 (2023) 8219.

[110] AutoDock3.0.5_UserGuide.pdf, (n.d.). https://autodock.scripps.edu/wp-content/uploads/sites/56/2022/04/AutoDock3.0.5_UserGuide.pdf (accessed August 26, 2024).

[111] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, J. Comput. Chem. 19 (1998) 1639–1662.

[112] Center of Computational Structural Biology, AutoDock Vina Documentation, (n.d.). https://autodock-vina.readthedocs.io/en/latest/introduction.html (accessed August 26, 2024).

[113] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, J. Comput. Chem. 31 (2010) 455–461.

[114] ***, Manual, AutoDock Vina (n.d.). https://vina.scripps.edu/manual/ (accessed August 26, 2024).

[115] D. Santos-Martins, L. Solis-Vasquez, A.F. Tillack, M.F. Sanner, A. Koch, S. Forli, Accelerating AutoDock4 with GPUs and Gradient-Based Local Search, J. Chem. Theory Comput. 17 (2021) 1060–1073.

[116] J. Eberhardt, D. Santos-Martins, A.F. Tillack, S. Forli, AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings, J. Chem. Inf. Model. (2021).

[117] ***, Docking and scoring, Schrödinger (2021). https://www.schrodinger.com/life-science/learn/white-papers/docking-and-scoring/ (accessed August 26, 2024).

[118] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: a new

approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, J. Med. Chem. 47 (2004) 1739–1749.

[119] R. Meli, G.M. Morris, P.C. Biggin, Scoring Functions for Protein-Ligand Binding Affinity Prediction Using Structure-based Deep Learning: A Review, Front. Bioinforma. 2 (2022) 885983.

[120] Improved Molecular Docking of MAO-B Inhibitors with Glide, Biointerface Res. Appl. Chem. 13 (2022) 159.

[121] D.S. Wishart, A. Guo, E. Oler, F. Wang, A. Anjum, H. Peters, R. Dizon, Z. Sayeeda, S. Tian, B.L. Lee, M. Berjanskii, R. Mah, M. Yamamoto, J. Jovel, C. Torres-Calzada, M. Hiebert-Giesbrecht, V.W. Lui, D. Varshavi, D. Varshavi, D. Allen, D. Arndt, N. Khetarpal, A. Sivakumaran, K. Harford, S. Sanford, K. Yee, X. Cao, Z. Budinski, J. Liigand, L. Zhang, J. Zheng, R. Mandal, N. Karu, M. Dambrova, H.B. Schiöth, R. Greiner, V. Gautam, HMDB 5.0: the Human Metabolome Database for 2022, Nucleic Acids Res. 50 (2021) D622–D631.

[122] PCIDB, (n.d.). https://www.genome.jp/db/pcidb (accessed August 27, 2024).

[123] Z. Wang, R. Song, W. Chen, J. Wang, P. Wang, Z. Zhang, X. Zhang, F. Wan, Vibrational Spectra and Molecular Vibrational Behaviors of Dibenzyl Disulfide, Dibenzyl Sulphide and Bibenzyl, Int. J. Mol. Sci. 23 (2022) 1958.

[124] H. K.M., K. S. Madan, M. B.C., N. R., Vibrational analysis and physical property studies of 6-Methoxy-2-[(E)-phenyliminomethyl]-phenol in the THz, IR and UV–visible spectral regions, Spectrochim. Acta. A. Mol. Biomol. Spectrosc. 222 (2019) 117227.

[125] M.H. Jamróz, Vibrational Energy Distribution Analysis (VEDA): Scopes and limitations, Spectrochim. Acta. A. Mol. Biomol. Spectrosc. 114 (2013) 220–230.

[126] B. Sathya, M. Prasath, M. Selvapandiyan, K. Prabha, Vibrational Analysis (FT-IR and FT-Raman Spectra) and Molecular Docking Evaluation of MPTB in GABA Receptor, J. Clust. Sci. 30 (2019) 1025–1035.

[127] N. Paduraru, S. Gosav, M. Praisler, Chemometric Characterization of Some Flavonoids Active Against HT-29 Human Cancer Cells, E-Health Bioeng. Conf. EHB 2015 (2015).

[128] A. Sykuła, A. Kowalska-Baron, K. Gałęcki, P. Błazińska, E. Łodyga-Chruścińska, Structural and Spectral Investigation of a Series of Flavanone Derivatives, Molecules 26 (2021) 1298.

[129] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242.

[130] Pubchem., (n.d.). https://pubchem.ncbi.nlm.nih.gov (accessed May 15, 2024).

[131] R.B. Murphy, M.P. Repasky, J.R. Greenwood, I. Tubert-Brohman, S. Steven Jerome, R. Annabhimoju, N.A. Boyles, C.D. Schmitz, R. Abel, R. Farid, R.A. Friesner, WScore: A flexible and accurate treatment of explicit water molecules in ligand–receptor docking, Med. Chem 59 (2016) 4364-4384.

[132] The PyMOL Molecular Graphics System, Version 2.5.7, in: Schrödinger, LLC., n.d.

[133] RCSB PDB - Pairwise Structure Alignment Tool, (n.d.). https://www.rcsb.org/alignment (accessed September 2, 2024).

[134] Z. Li, L. Lukasz Jaroszewski, M. Iyer, M. Sedova, A. Godzik, FATCAT 2.0: towards a better understanding of the structural diversity of proteins, Nucleic Acids Res. 48 (2020) W60–W64.

[135] J. Ma, S. Wang, Algorithms, Applications, and Challenges of Protein Structure Alignment, Adv Protein Chem Struct Biol 94 (2014) 121–175.

[136] R.P.D. Bank, Pairwise Structure Alignment, (n.d.). https://www.rcsb.org/docs/tools/pairwise-structure-alignment (accessed September 2, 2024).

[137] C.J. Williams, J.J. Headd, N.W. Moriarty, M.G. Prisant, L.L. Videau, L.N. Deis, V. Verma, D.A. Keedy, B.J. Hintze, V.B. Chen, S. Jain, S.M. Lewis, W.B. 3rd Arendall, J. Snoeyink, P.D. Adams, S.C. Lovell, J.S. Richardson, D.C. Richardson, MolProbity: More and better reference data for improved all-atom structure validation, Protein Sci. 27 (2018) 293–315.

[138] BIOVIA, Dassault Systèmes, Discovery Studio, 24.1.0.23298, in: San Diego: Dassault Systèmes, 2024.

[139] C. Paraschiv, S. Gosav, M. Praisler, Chemical Potential Evaluation of Kaempferol Based on Molecular Modeling, in: Int. Conf. E-Health Bioeng. EHB 2023 11th Ed., 2023.

[140] Gaussian 09 W Software, (2003).

[141] Search results, (n.d.). https://research.manchester.ac.uk/en/datasets/ccdc-1429470-experimental-crystal-structure-determination (accessed July 27, 2023).

[142] S. Gosav, N. Paduraru, D. Maftei, M.L. Birsa, M. Praisler, Quantum chemical study of a derivative of 3-substituted dithiocarbamic flavanone, Spectrochim. Acta. A. Mol. Biomol. Spectrosc. 172 (2017) 115–125.

[143] S. Gosav, A. Ion, M. Praisler, DFT characterization of MDMA methylene homologue, a chemical compound with psychoactive properties, in: Sofia, Bulgaria, 2019: p. 170027.

[144] ***, Avogadro: an open-source molecular builder and visualization tool, (n.d.). http://avogadro.cc/.

[145] M.E. Hanwell, D.E. Curtis, D.C. Lonie, T. Vandermeersch, E. Zurek, G.R. Hutchison, Avogadro: An advanced semantic chemical editor, visualization, and analysis platform, J. Cheminformatics (2012).

[146] R.P.D. Bank, RCSB PDB - 4JPS: Co-crystal Structures of the Lipid Kinase PI3K alpha with Pan and Isoform Selective Inhibitors, (n.d.). https://www.rcsb.org/structure/4JPS (accessed September 8, 2023).

[147] J. Kapil, P. Shukla, A. Pathak, Review Article on Density Functional Theory, in: V.K. Jain, S. Rattan, A. Verma (Eds.), Recent Trends Mater. Devices, Springer, Singapore, 2020: pp. 211–220.

[148] W. Qin, N. Du, L. Zhang, X. Wu, Y. Hu, X. Li, N. Shen, Y. Li, B. Yang, C. Xu, Genistein alleviates pressure overload-induced cardiac dysfunction and interstitial fibrosis in mice, Br J Pharmacol 172 (2015) 5559–5572.

[149] M. Zhang, H. Jang, R. Nussinov, PI3K inhibitors: review and new strategies, Chem. Sci. 11 (2020) 5855–5865.

[150] C.M. Burlacu, A.C. Burlacu, M. Praisler, C. Paraschiv, Harnessing Deep Convolutional Neural Networks Detecting Synthetic Cannabinoids: A Hybrid Learning Strategy for Handling Class Imbalances in Limited Datasets, Inventions 8 (2023) 129.

[151] C.M. Burlacu, S. Gosav, B.A. Burlacu, M. Praisler, Convolutional Neural Network Detecting Synthetic Cannabinoids, in: In Proceedings of the International Conference on e-Health and Bioengineering (EHB), Iasi, Romania, 2021.

[152] D. Hellmann, The Python 3 Standard Library by Example, 2nd ed., Addison-Wesley Professional:, Boston, NA, USA, 2017.

[153] Wolfram Research, I. Mathematica Version 13.3, in: Wolfram Research, Inc.:, Champaign, IL, USA, 2023.