Structural and dynamical properties of complexes between inhibitors and papain like protease enzyme of coronavirus 2019: Theoretical study

Chananya Rajchakom

A Thesis Submitted in Partial Fulfillment of Requirements for
degree of Master of Science in Chemistry
December 2024

Copyright of Mahasarakham University

สมบัติเชิงโครงสร้างและพลวัติของสารประกอบเชิงซ้อนระหว่างสารยับยั้งและเอนไซม์ปาเปนไลค์โปรตีเอส ของโคโรนาไวรัส 2019: การศึกษาเชิงทฤษฎี

วิทยานิพนธ์

ของ

ชนัญญา ราชคม

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเคมี

ธันวาคม 2567

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

Structural and dynamical properties of complexes between inhibitors and papain like protease enzyme of coronavirus 2019: Theoretical study

Chananya Rajchakom

A Thesis Submitted in Partial Fulfillment of Requirements
for Master of Science (Chemistry)
December 2024

Copyright of Mahasarakham University

The examining committee has unanimously approved this Thesis, submitted by Miss Chananya Rajchakom , as a partial fulfillment of the requirements for the Master of Science Chemistry at Mahasarakham University

Examining Committee

..... Chairman
(Assoc. Prof. Theerapong Puangmali , Ph.D.)

..... Advisor
(Assoc. Prof. Nadtanet Nunthaboot , Ph.D.)

..... Co-advisor
(Assoc. Prof. Thanyada Rungrotmongkol , Ph.D.)

..... Committee
(Asst. Prof. Banchob Wanno , Ph.D.)

..... Committee
(Asst. Prof. Khongvit Prasitnok , Ph.D.)

Mahasarakham University has granted approval to accept this Thesis as a partial fulfillment of the requirements for the Master of Science Chemistry

..... (Prof. Pairot Pramual , Ph.D.)
Dean of The Faculty of Science

..... (Assoc. Prof. Krit Chaimoon , Ph.D.)
Dean of Graduate School

TITLE Structural and dynamical properties of complexes between inhibitors and papain like protease enzyme of coronavirus 2019: Theoretical study

AUTHOR Chananya Rajchakom

ADVISORS Associate Professor Nadtanet Nunthaboot , Ph.D.
Associate Professor Thanyada Rungrotmongkol , Ph.D.

DEGREE Master of Science **MAJOR** Chemistry

UNIVERSITY Mahasarakham University **YEAR** 2024

ABSTRACT

COVID-19 has spread to numerous countries over five years, resulting in 774 million cases and 7 million deaths worldwide. Challenges, including ineffective medications, vaccine hesitancy, gene mutations, and high drug costs, emphasize the necessity for new inhibitors. SARS-CoV-2's papain-like protease (PLpro), which plays a crucial role in cleaving the replicase polyprotein at three essential sites vital for viral replication, has emerged as an attractive target for SAS-CoV-2 treatment. This study utilized structure-based virtual screening, pharmacokinetic analysis, molecular docking, and molecular dynamics (MD) simulation to identify phytochemical inhibitors targeting SARS-CoV-2 PLpro. Based on Lipinski's rule of five, a group of 45 phytochemicals sourced from Thai medicinal plants underwent initial screening. Among them, ADMET and molecular docking calculations further analyzed 27 compounds meeting the criteria. Out of these phytochemical molecules, (3R)-7,4'-dihydroxy-8-methoxyhomoisoflavane, bergenin, capparispine, and feruloyl tyramine representing the top 15 percent of docked compounds, underwent three cycles of 100 ns. MD simulations to assess stability and interactions at the PLpro binding pocket. All ligands except for feruloyl tyramine demonstrated extensive interactions with PLpro residues, including D164, R166, P247, P248, Y264, Y268, Q269, Y273, and D302, through both hydrogen bonds and hydrophobic interactions. MM/GBSA binding free energy, decomposition energy and hydrogen bond formation indicated that these three phytochemicals ((3R)-7,4'-dihydroxy-8-methoxyhomoisoflavane, bergenin, capparispine) showed strong and favorable binding energies towards PLpro of SARS-CoV-2, suggesting their potential as inhibitors. This detailed information could significantly aid in developing and optimizing effective SARS-CoV-2 inhibitors.

Keyword : Phytochemicals, SARS-CoV-2, PLpro, Molecular docking, MD simulation

ACKNOWLEDGEMENTS

This research project was financially supported by Mahasarakham University. I am sincerely grateful for the institution's support and encouragement throughout my academic journey. I would like to express my deepest gratitude to the Center of Excellence for Innovation in Chemistry (PERCH–CIC) and the Ministry of Higher Education, Science, Research, and Innovation for their generous financial support. I also extend my thanks to the Supramolecular Chemistry Research Unit (Mahasarakham University) and the Center of Excellence in Structural and Computational Biology (Chulalongkorn University) for providing the computational facilities essential to this research. I am particularly indebted to my advisor, Assoc. Prof. Nadtanet Nunthaboot, Ph.D., for her invaluable guidance, constructive feedback, and continuous encouragement. Her expertise and mentorship have been pivotal in shaping the success of this thesis. My heartfelt thanks also go to my co-advisor, Assoc. Prof. Thanyada Rungrotmongkol, Ph.D., whose insights and advice have been integral to the completion of this study. Special thanks to Dr. Nitchakan Darai of Walailak University for her exceptional guidance and expertise, as well as to Miss Thitiya Boonma, a Ph.D. candidate, for her unwavering support, mentorship, and insightful contributions throughout this journey. I am profoundly grateful to the members of my examining committee, Assoc. Prof. Theerapong Puangmali, Ph.D. (Chairman), Asst. Prof. Dr. Banchob Wanno, Ph.D., and Asst. Prof. Dr. Khongvit Prasitnok, Ph.D., for their invaluable feedback, rigorous evaluation, and constructive suggestions, all of which have greatly enhanced the quality of this thesis. Finally, I wish to express my heartfelt gratitude to my family and friends for their unwavering support and encouragement. Their belief in me has been a source of strength and inspiration throughout this challenging journey. This thesis represents the culmination of collective efforts, guidance, and contributions from all those mentioned above. I am deeply appreciative of their invaluable support.

Chananya Rajchakom

TABLE OF CONTENTS

Page

ABSTRACT ..... D
ACKNOWLEDGEMENTS ..... E
TABLE OF CONTENTS ..... F
LIST OF TABLES I
LIST OF FIGURES ..... J
CHAPTER I INTRODUCTION 13
1.1 Rationale and background 13
1.2 Purposes of the research 16
1.3 Scope of the research 16
1.4 Expected outcome 17
CHAPTER II COMPUTATIONAL THEORY AND LITERATURE REVIEWS 18
2.1 Molecular docking 18
2.2 Molecular dynamics simulation 21
2.3 Ensemble and trajectory 22
2.3.1 Canonical ensemble (NVT) 23
The canonical ensemble contains all possible states in thermal equilibrium with a heat bath. The system remains at the absolute temperature of T but may exchange energy with the heat bath. Three system parameters are fixed throughout the simulation: the absolute temperature (*T*), the number of atoms (*N*), and the volume (*V*). T is the most influential parameter among the system states. 23
2.3.2 Micro canonical ensemble (NVE) 23
2.3.3 Isothermal–isobaric ensemble (NPT) 23
2.4 Force-Field 23
2.4.1 Stretching energy ($E_{stretch}$) 23
2.4.2 Bending energy (E_{bend}) 23
2.4.3 Torsion energy ($E_{torsion}$) 24

2.4.4 Non-covalent interactions....24
2.5 Algorithms....24
2.6 Periodic boundary conditions....26
2.7 Cut-off and minimum image convention....26
2.8 Free energy calculations....27
2.8.1 Solvation interaction energy (SIE)....27
2.8.2 Molecular Mechanics Generalized Born Surface Area (MMGBSA)....29
2.8.3 Thermodynamic integration (TI)....29
2.8.4 Free energy landscape....30
2.8.5 Umbrella sampling....32
2.9 Literature reviews....34
2.9.1 Drug Repurposing....34
2.9.2 New compound searching....39
CHAPTER III COMPUTATIONAL DETAILS....53
3.1 Phytochemical compounds....53
3.2 Drug-likeness properties....53
3.3 Molecular docking....54
3.4 Molecular dynamics simulation....54
CHAPTER IV RESULTS AND DISCUSSION....56
4.1 Drug likeness and pharmacokinetic ADMET predictions....56
4.2 Molecular Docking Calculation....67
4.3 Molecular Dynamics (MD) simulation....69
4.3.1 System stability....69
4.3.2 Solvated interaction energy (SIE) Binding energy....70
4.3.3 Molecular Mechanics General Born surface area (MM/GBSA) Binding free energy....71
4.3.4 Understanding the Key Residues in Ligand Binding to SARS-CoV-2 PLpro....77
4.3.5 Protein-ligand interactions....80
CHAPTER V CONCLUSIONS....84

5.1 Future work 85
5.1.1 Protease Assay 85
5.1.2 Viral Replication Inhibition Assay 85
5.1.3 Cytotoxicity assays 86
REFERENCES 2
BIOGRAPHY 10

พหูนํ ปณฺฑิโต ชีเว

LIST OF TABLES

Page

Table 1 Decomposition of MM-GBSA binding energies for phenformin, quercetin and ritonavir in complex with COVID-19 PLpro through 50 ns MD simulations. 34

Table 2 Some repurposed drugs are predicted to be potential inhibitors for PLpro of SARS-CoV-2. 38

Table 4 Individual component contribution in binding free energy calculations for native PLpro and its ligand-bound complexes. 44

Table 3 MMGBSA scores for MD simulated complexes. 47

Table 5 Some new selected compounds predicted to be potential inhibitors for PLpro of SARS-CoV-2. 50

Table 6 Drug-likeness properties of phytocompounds predicted by SwissADME. ...57

Table 7 ADMET analysis predictions for selected phytocompounds using the pkCSM software 64

Table 8 Docking scores (binding energy in kcal/mol) of the phytocompounds against SARS-CoV-2 PLpro, through GOLD (as the values appear) for the best docked compounds. 68

Table 9 MM/GBSA binding energy (kcal/mol) components for a given MD run of protein-ligand complexes. 76

พหูนํ ปณฺฑิโต ชีเว

LIST OF FIGURES

Page

Figure 1 Structure of SARS-CoV-2 [5]. 13

Figure 2 Structure of papain-like protease. 15

Figure 3 shows lock-and-key model. 18

Figure 4 Scheme of the MD simulation [28]. 22

Figure 5 The periodic boundary conditions in two dimensions [32]. 26

Figure 6 The spherical cut-off and the minimum image convention [33]. 27

Figure 7 The thermodynamic cycle involves events A and B, which represent the binding of two distinct ligands to a protein, while events C and D signify the conversion of one ligand to the other in both the bound and solvated states, respectively. The free energy differences between processes A and C can be determined by calculating the free energy differences between processes B and D [40]. 30

Figure 8 A free energy landscape sketch shows two macrostates, each represented by a basin, reflecting distinct protein conformations. The collective variable x describes the slow transition between states, with the transition timescale determined by the energy barrier height between the macrostates [41]. 32

Figure 9 PLpro apo structure in khaki color (PDB ID: 6W9C) showing the binding site residues with the highly flexible loop (T265-H272) in red and the rest (L162-E167, P247-P248, Y264-Y273 and T301) in blue. The highly flexible amino acids (3-83, 182-200, 223-235 and 306-315) are shown in yellow [47]. 35

Figure 10 Binding interactions, RMSD, RMSF, Radius of gyration and SASA plots with CAP and PLpro 36

Figure 11 Illustrations represent the process of the study. 37

Figure 12 Predicted binding modes of four PLpro-inhibitor systems: PLpro-G805_0497 (a and b), PLpro-F403_0159 (c and d), PLpro-F112_0109 (e and f), and PLpro-D754_0006 (g and h) obtained from structure-based virtual screening. The protein PLpro is shown in cartoons and colored in blue. Hydrogen bond and π-π interactions are shown as dashed lines and colored in red and blue, respectively. 39

Figure 13 A) Molecular binding interactions and B) Binding mode of proposed food compounds inside the active site cavity of SARS-CoV-2 PLpro protein. 40

Figure 14 Donor-acceptor interactions obtained by docking of I-asarinin and receptor 4OVZ of PLpro protease of COVID-19. 41

Figure 15 Rational design of the SARS-CoV-2 PLpro inhibitors based on the peptidomimetic VIR250 inhibitor. (A) 3D and 2D structures of VIR250 with ligand modifications, (B) 2D structure of modified VIR250 and their binding free energy prediction comparison with VIR250 against SARS-CoV-2 PLpro derived from MM-PB(GB)SA calculations, (C) hydrogen bond occupation, and (D) per-residue decomposition free energy (ΔG_{bind} residue) of modified VIR250 and the A2/SARS-CoV-2 PLpro complex. Calculations are obtained from one snapshot of the complex after system minimization and solvation in the TIP3P model. 42

Figure 16 3D and 2D images of desacetylgedunin (DCG), with the highest binding affinity towards PLpro. 43

Figure 17 2D interaction diagram for the 100ns simulation trajectory of SN00334175/7JN2 (A) and SN00162745/7JN2 (B) complex. 44

Figure 18 Rational drug design of the SARS-CoV-2 PLpro inhibitors. 2D structure of lead 2 with possible modified fragments. 45

Figure 19 2D interaction diagrams of aspergillipeptide F with PLpro active site residues. 46

Figure 20 the 2D and 3D interactions of the compounds COMP1–COMP5. A superimposed view shows the comparison between the co-crystallized ligand (in green) in the 7CMD structure and its re-docked conformation (in purple), with a Root Mean Square Deviation (RMSD) of 0.388 Å, indicating high accuracy in docking. In the 2D format, hydrogen bonds are represented by blue and green dotted arrows, while in the 3D format, they are shown as black dotted lines. For color references and further details, readers are directed to the online version of the article. 48

Figure 21 2-dimensional analysis of non-bonded interaction between compound UKR1129266 and PLpro 49

Figure 22 RMSD, and R_g profiles for all protein-ligand complexes over time for each of the three replicates (Run 1–3).....70

Figure 23 The ΔGbindSIE (kcal/mol) results are based on the SIE method for all protein-ligand complexes over time for each of the three replicates (Run 1–3).....71

Figure 24 (A) Decomposition of binding energy per residue on SARS-CoV-2 PLpro utilizing the MM/GBSA approach. (B) The 3D structures with energy contributions of < 5 kcal/mol of ligands in association with SARS-CoV-2 PLpro are identified and colored appropriately. 79

Figure 25 Representative 3D structures (Top) and 2D srtructures (Bottom) displaying protein-ligand interactions of each simulated system. Hydrogen bond forming between protein and ligand is displayed by green dot lines.....83

พหูนํ ปณฺฑิโต ชีเว

CHAPTER I INTRODUCTION

1.1 Rationale and background

The global pandemic of coronavirus 2019 or COVID-19, caused by the severe acute respiratory syndrome 2 (SARS-CoV-2) virus emerged in December 2019 [1], has become a significant impact on the global economy and industries [2]. As of August 2024, there were reported to be approximately 775 million confirmed cases of COVID-19 and 7 million fatalities worldwide due to SARS-CoV-2 infection [3]. SARS-CoV-2 belongs to the family *Coronaviridae* and the genus *Betacoronavirus.* It is a large positive-sense single-stranded RNA virus[4]. The virus is composed of several structural and non-structural proteins. The structural proteins include the spike protein (S), envelope protein (E), membrane protein (M), and nucleocapsid protein (N). In addition to these, there are 16 species of non-structural proteins present in SARS-CoV-2 [5] (Figure 1).

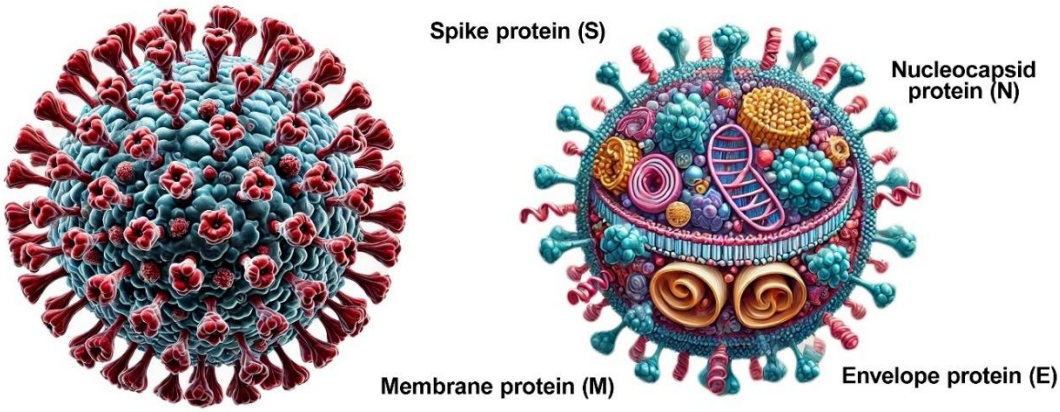

Figure 1 Structure of SARS-CoV-2 [5].

Currently, COVID-19 vaccinations have played a crucial role in preventing SARS-CoV-2 infection and reducing the severity of the disease. Additional treatment options are still needed to effectively address the current pandemic. While several existing drugs have been repurposed and tested for their efficacy against COVID-19, such as favipiravir [6], remdesivir [7], hydroxychloroquine, and chloroquine [8, 9], the results have been mixed, and none of them have shown consistently satisfactory levels of efficacy [10].

Papain-like protease (PLpro) is a non-structural protein 3 (nsp3) in SARS-CoV-2, consisting of multiple domains: a ubiquitin-like domain, a thumb domain, a zinc-binding domain, and a palm domain [11]. The catalytic triad, made up of C111, H272, and D286, forms the active site of PLpro, located at the interface between the thumb and palm domains (Figure 2). The BL2 loop of PLpro consists of a flexible sequence of amino acids (267–271) that are located in a key area of the enzyme. These amino acids help break down viral proteins and keep the virus from being recognized by the human immune system. This loop is essential for PLpro flexibility, enabling it to bind effectively with substrates or inhibitors. This flexibility is crucial for PLpro regulation because it lets it attach to viral proteins, break down parts that are needed for viral replication, and help put proteins together. Consequently, the BL2 loop enhances PLpro functionality and aids in immune evasion.

PLpro functions as a proteolytic enzyme, cleaving the polyprotein replicase at three specific sites to release key proteins, NSP1-3, which are vital for viral replication. PLpro recognizes specific LXGG motifs in its substrates for cleavage. In addition to its role in viral replication, PLpro is also pivotal in inhibiting the host's immune response. Upon infection, the innate immune system triggers an antiviral response by modifying host cell proteins through the addition of ubiquitin (Ub) and interferon-stimulated gene product 15 (ISG15) [12]. These modifications are part of the host’s defense against viral infection. However, PLpro efficiently removes ISG15 and ubiquitin modifications from proteins via its catalytic cysteine cleavage domain, which reduces inflammation and inhibits antiviral signaling, thereby enabling the virus to evade the host’s immune response effectively.

The development of PLpro inhibitors for SARS-CoV-2 is important because PLpro is a key enzyme involved in both viral replication and immune evasion. It is crucial to ensure that these inhibitors specifically target the viral PLpro without interfering with human proteases. Several factors indicate that PLpro inhibitors can be designed to be highly specific, which will reduce the risk of adverse side effects in the human body. The SARS-CoV-2 PLpro has a unique structure that differs from human proteases, particularly in the active site, which can recognize and cleave the molecular sequence “LXGG↓XX” found in the viral polyprotein. This sequence is rarely found in human proteins [13]. As a result, PLpro can be targeted without disrupting human

proteases such as cathepsins and caspases, which have different active sites and molecular sequences, reducing the chance that PLpro inhibitors will affect these enzymes in the human body.

Given the critical role of PLpro in SARS-CoV-2 replication and immune suppression, it has become a promising target for antiviral drug development. Researchers are studying ways to inhibit PLpro activity, either by directly targeting the catalytic site or by blocking the interaction between PLpro and its substrates. The hope is that inhibiting PLpro can halt viral replication and restore the host's immune response, leading to an effective antiviral effect. It should be noted that the development of drugs targeting PLpro remains an active area of study, and further research is needed to evaluate the potential of PLpro inhibitors in the treatment of COVID-19 [14].

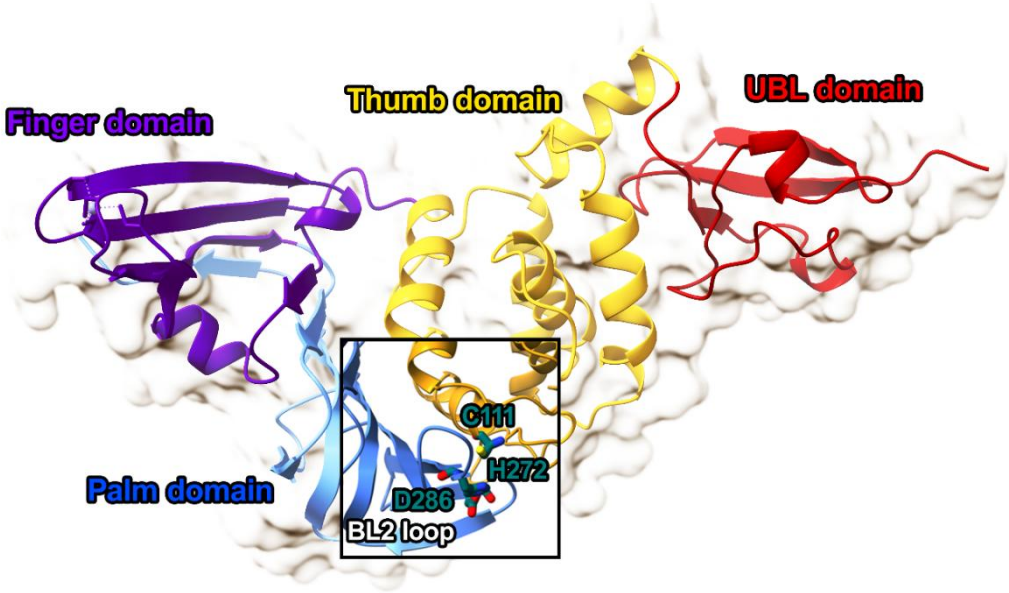

Figure 2 Structure of papain-like protease.

พหูนํ ปณฺฑิโต ชีเว

Medicinal plants mainly contain phytochemicals such as alkaloids, phenolic compounds, flavonoids, terpenoids and other organic molecules, have been a primary source of materials for treatment of many diseases. These phytochemicals show a wide range of biological activities such as antibacterial [15], anti-inflammatory [16] and anticancer [17] activities. The use of natural products derived from herbs is appealing due to their affordability and widespread availability. Hence, this research work focuses on exploring the potential of medicinal plants, specifically Thai herbs, as alternative treatments for common illnesses, including COVID-19. In Thai traditional medicine, recipes with fever-reducing properties, such as Harak, Khiaohom, and Chanthaleela, are used to treat moderate to high fevers. These recipes consist of various herbs, including *Tiliacora triandra* (Colebr.) Diels, *Harrisonia perforata* (Blanco) Merr, *Capparis micracantha* DC., *Dracaena cochinchinensis* (Lour.) S.C.Chen, *Azadirachta indica* A.Juss., *Ficus racemosa* L., *Tinospora crispa* (L.) Hook. f. & Thomson, and others. To seek for potential compounds inhibiting PLpro, a target of SARS-CoV-2, the phytochemicals found in these Thai traditional medicine recipes were collected and screened using various tools such as Lipinski's rule of five, ADMET analysis, molecular docking, and MD simulation. By examining the phytochemical constituents of selected herbs, the researchers aim to gain insight into the chemical compounds present in these plants that may exhibit strong inhibitory activity against the PLpro enzyme of COVID-19.

1.2 Purposes of the research

1.2.1 To screen phytochemical compounds that might show good inhibitory potencies against PLpro enzyme of coronavirus 2019

1.2.2 To study the interaction and dynamical behaviors of the PLpro of coronavirus 2019 with inhibitors

1.3 Scope of the research

1.3.1 To screen compounds/phytochemical compounds with the targeted PLpro of coronavirus 2019

1.3.2 To perform molecular dynamics simulation of the complexes between PLpro of coronavirus 2019 and inhibitors

1.4 Expected outcome

1.4.1 To obtain candidate compounds/phytochemical compounds showing high potency and selectivity to the PLpro of coronavirus 2019

1.4.2 In depth understanding of mechanism of action and interactions between PL-inhibitor complexes

พหูนํ ปณฺฑิโต ชีเว

CHAPTER II COMPUTATIONAL THEORY AND LITERATURE REVIEWS

Theoretical methods utilized in this study and the previously reported research works related to PLpro are described.

2.1 Molecular docking

Molecular docking is a computational method used to predict the binding modes and interactions between protein (host) and ligand (guest) [18]. Figure 3 shows lock-and-key model [19], which refers to the rigid docking of receptor and ligand to find the correct orientation for the "key" to open the "lock" [20].

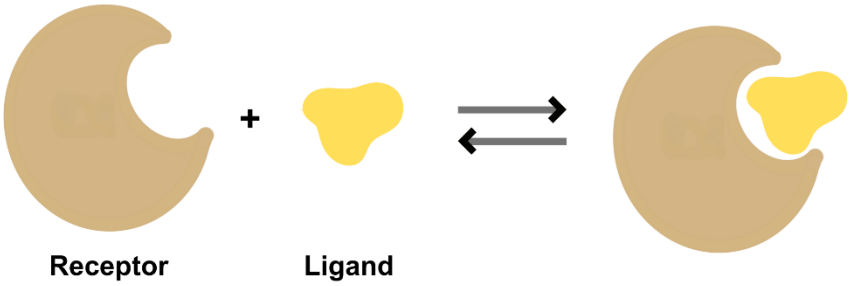

Figure 3 shows lock-and-key model.

The interaction between protein and ligand is a process of systematic thermodynamic equilibrium, and by which the complex structure formed should be the possible conformation with the lowest binding free energy. These simplified scoring functions commonly presume that the binding free energy may be represented as a sum of numerous additive components representing diverse contributions to the binding free energy [21]. $\Delta G_{binding}$ which is calculated as follows:

$$\Delta G_{binding} = \Delta G_{vdW} + \Delta G_{ele} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta G_{tors} \quad (1)$$

ΔG_{vdW} is The Lennard-Jones (LJ) 12-6 potential that describes the energy interaction between neutral atoms or molecules [22]. It is expressed as:

$$V(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \quad (2)$$

Where:

$V(r)$ is the potential energy between the two particles as a function of their separation distance (r).

ε is the depth of the potential well, representing the strength of the attractive interaction.

σ is the distance at which the potential energy is zero, related to the size of the particles.

The $\left(\frac{\sigma}{r}\right)^{12}$ term: Represents the repulsive forces arising from the electron clouds' overlap at close distances. Dominates at short distances.

The $\left(\frac{\sigma}{r}\right)^{6}$ term: Represents the attractive forces, primarily due to van der Waals dispersion forces. Dominates at intermediate distances.

ΔG_{elec} represents the change in free energy associated with the electrostatic interaction between two charged species in a solvent environment. It considers both the Coulombic interaction between the charges and the influence of the solvent's dielectric constant on this interaction [23]. ΔG_{elec}, in terms of a simplified Coulombic interaction within a dielectric medium:

$$\Delta G_{\text{elec}} = -\frac{1}{2}\frac{q_1 q_2}{4\pi\varepsilon_0\varepsilon_r}\frac{1}{\alpha} \tag{3}$$

Where:

q_1, q_2 is charges of the interacting species.

ε_0 is permittivity of free space.

ε_r is dielectric constant of the solvent.

α is distance between the charges (often approximated as the sum of their ionic radii)

ΔG_{hbond} is the 12-10 Potential with Goodford Directionality. The 12-10 potential is another mathematical model that describes intermolecular interactions, similar to the Lennard-Jones potential. It is beneficial for modeling hydrogen bonding interactions, and Goodford's Directionality is essential for the directional aspect of hydrogen bonds. Unlike purely distance-based interactions, the strength of a hydrogen bond is strongly influenced by the angle between the donor, hydrogen, and acceptor atoms. A linear alignment (close to 180°) is optimal, and deviations from this angle weaken the hydrogen bond. This directional dependence is essential for accurately modeling hydrogen bonds in molecular systems [24]. The potential can be expressed as:

$$\Delta G_{\text{hbond}} = \varepsilon\left[5\left(\frac{r_0}{r}\right)^{12} - 6\left(\frac{r_0}{r}\right)^{10}\right]\cdot \int(\theta) \tag{4}$$

Where:

ΔG_{hbond} is the hydrogen bond interaction energy.

ε is the well depth of the potential, representing the strength of the hydrogen bond.

r_0 is the equilibrium distance between the donor and acceptor atoms.

r is the actual distance between the donor and acceptor atoms.

$\int(\theta)$ is an angular function that modulates the interaction energy based on the hydrogen bond angle (θ) and incorporates Goodford's directionality concept, ensuring that the hydrogen bond is strongest when the donor, hydrogen, and acceptor atoms are aligned optimally.

12-10 Exponents are chosen to provide a steeper repulsive wall and a shallower attractive well compared to the 12-6 Lennard-Jones potential, which is more suitable for describing the stronger and more directional nature of hydrogen bonds.

ΔG_{desolv} is calculated form Stouten Pairwise Atomic Solvation Parameters are a set of empirical values used to estimate the desolvation free energy associated with the transfer of a molecule from a solvent environment to a protein binding site. GOLD program utilizes a simplified version of the Stouten model, which does not explicitly consider the effect of partial charges [25]. The equation is:

$$\Delta G_{desolv} = \sum\sum S_{ij} \quad (5)$$

Where:

S_{ij} is solvation parameter for the pairwise interaction between atom types i and j

The summation is over all pairs of atoms in the ligand that lose contact with the solvent upon binding.

ΔG_{tors} are often used to represent the change in free energy associated with the loss of conformational entropy due to the restriction of rotatable bonds upon ligand binding to a protein based on the number of rotatable bonds (N_{rot}) in the ligand [26, 27]:

$$\Delta G_{tors} = RT\ln(3^{N_{rot}}) \quad (6)$$

Where:

R is gas constant

T is Temperature

N_{rot} is Number of rotatable bonds in the ligand.

2.2 Molecular dynamics simulation

Molecular dynamics (MD) simulation is a statistical tool based on Newton's second law. This approach assumes that every particle in the system behaves like a Newtonian particle, and quantum behaviors are entirely disregarded. Electronic motions are ignored, and electrons are supposed to remain in their ground state and modify their dynamics quickly when atomic locations change (the Born-Oppenheimer approximation). Only classical mechanics is employed to describe the motion of the particles. As a result, the equation of motion, $F = ma$, applies to particles where F is the force, m represents mass, and a is the particle's acceleration. Once each atom's position and velocity are known, the system's state can be anticipated, and new positions and velocities can be determined. The method can be performed indefinitely to obtain a trajectory of atomic motions. The first step of the MD simulation algorithm (Figure 4) includes the assignment of coordinates for each atom in the system, initial velocities, and a simulation time step (Δt) is defined. After that, for each time step of the simulation, the forces between atoms, which are determined by an interatomic potential (force field), are computed, and equations of motion are integrated.

พหูนํ ปณฺฑิโต ชีเว

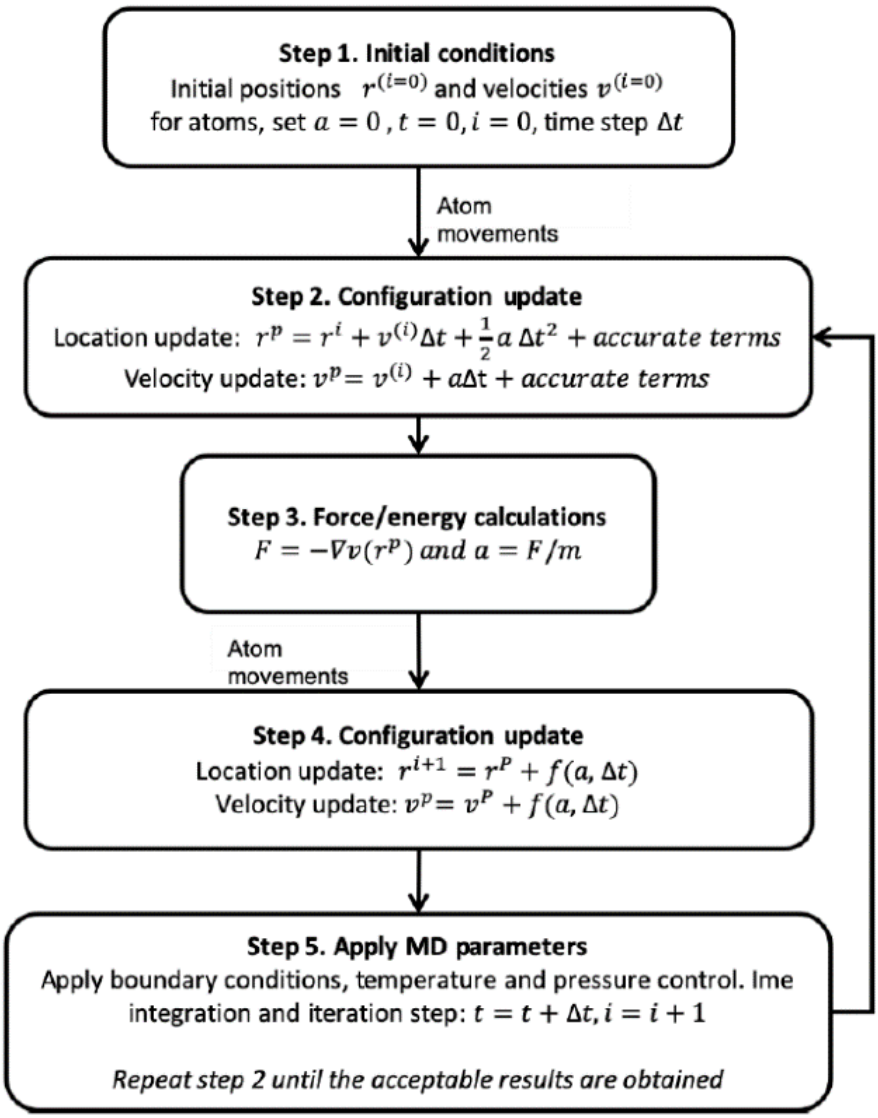

Figure 4 Scheme of the MD simulation [28].

2.3 Ensemble and trajectory

The output of the simulation comes in the form of ensemble of frames. All frames share the same macroscopic/thermodynamic state but may differ in the microscopic states. Each frame represents the system at a specific point in time (a specific microscopic state). If the ensemble is sequence (time) dependent, it is called a trajectory. In this case, the trajectory represents the time-dependent evolution of the system.

2.3.1 Canonical ensemble (NVT)

The canonical ensemble contains all possible states in thermal equilibrium with a heat bath. The system remains at the absolute temperature of T but may exchange energy with the heat bath. Three system parameters are fixed throughout the simulation: the absolute temperature (*T*), the number of atoms (*N*), and the volume (*V*). T is the most influential parameter among the system states.

2.3.2 Micro canonical ensemble (NVE)

The micro canonical ensemble represents an isolated system. No change in mass/number of atoms (*N*), Volume (*V*), nor exchange of Energy (*E*) is allowed.

2.3.3 Isothermal–isobaric ensemble (NPT)

The system has a fixed temperature (*T*), hence it is isothermal, and fixed pressure (*P*), hence it is isobaric [29].

2.4 Force-Field

The Force Field is the summation of bonded and non-bonded terms or covalent and non-covalent interactions among the atoms and molecules. A simple molecular mechanics energy equation is given by:

$$E_{\text{total}} = E_{\text{bonded}} + E_{\text{non-bonded}} \quad (7)$$

$$E_{\text{bonded}} = E_{\text{stretch}} + E_{\text{bend}} + E_{\text{torsion}} \quad (8)$$

$$E_{\text{non-bonded}} = E_{\text{electrostatic}} + E_{\text{van der Waals}} \quad (9)$$

2.4.1 Stretching energy (E_{stretch})

The stretching energy associated with vibration about the equilibrium bond length. The k_b parameter controls the stiffness of the bond spring, while r_0 defines its equilibrium length. Unique k_b and r_o parameters are assigned to each pair of bonded atoms based on their type such as C-C, C-H, O-C, etc.

$$E_{\text{stretch}} = \sum_{\text{bonded}} k_b (r - r_0)^2 \quad (10)$$

2.4.2 Bending energy (E_{bend})

The bending energy associated with vibration about the equilibrium bond length. The k_b parameter controls the stiffness of the angle, while θ defines its equilibrium angle. Unique parameters for angle binding are assigned to each bonded triplet of atoms based in their type such as C-C-C, C-O-C, C-C-H, etc.

$$E_{\text{bend}} = \sum_{\text{angles}} k_\theta (\theta - \theta)^2 \quad (11)$$

2.4.3 Torsion energy (E_{torsion})

Torsional energy fluctuates as covalent bonds (such as C-C, C-N, and C-O single bonds) rotate. The power reaches its maximum when the torsion angle (t) is 0˚ and when interactions occur between atoms separated by three sigma bonds. The A parameter governs the amplitude of the energy curve, while the n parameter controls its periodicity and Ø shifts the curve along the rotational angle axis (τ). These parameters are determined through curve fitting. Specific torsional rotation parameters are assigned to each set of four bonded atoms based on their types, such as C-C-C-C, C-O-C-N, H-C-C-H, etc.

$$E_{\text{torsion}} = \sum_{\text{torsion}} A[1 + cos(n\tau - \emptyset)] \quad (12)$$

2.4.4 Non-covalent interactions

The non-bonded energy represents the pairwise sum of the energies of all possible interacting non-bonded atom

$$E_{\text{non-bonded}} = \sum_{ij}^{\text{atompairs}} E_{\text{eletrostatics}} + \sum_{ij}^{\text{atompairs}} E_{\text{van der Waals}} \quad (13)$$

$$E_{\text{non-bonded}} = \sum_i \sum_j \frac{-A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} + \sum_i \sum_j \frac{q_i q_j}{r_{ij}} \quad (14)$$

Where A determines the degree of stickiness of the van der Waals attraction and B determines the degree of hardness of the atoms, r_{ij} is the distance between atoms i and j, and q_i is the partial atomic charge. With equation 8, it is possible to observe that the energy depends on non-bonded interactions for van der Waals attraction, repulsion, and electrostatic interaction [30].

2.5 Algorithms

MD simulation to view the dynamic evolution of biological systems on a temporal scale can be defined by employing Newton's Laws of motion.

$$F_i = m_i a_i \quad (15)$$

$$F_i(t) = m_i \frac{d^2 r_i(t)}{dt^2} \quad (16)$$

In equations 16 and 17, r_i and m_i represent the position and mass of atoms i, respectively, and $F_i(t)$ is the force acting on atom i at time t. These classical equations of motion are integrated using the finite difference method. finite difference methods are techniques used to generate MD trajectories with continuous potential models. The basic idea is that the total force of each particle is calculated on a time scale as

the vector sum of its interactions with other particles, rather than as the vector sum of its interactions with other particles.

Algorithms for integrating the equation of motion often rely on the finite difference method. A core assumption in these algorithms is that dynamic properties (such as position, velocity, and acceleration) can be approximated using the Taylor Series Expansion Method. The Taylor series expansion method provides a framework for predicting the position and other dynamic properties at a future time $t + \delta t$, using the current values at time t. This expansion includes position, velocity, and acceleration terms, including higher-order terms for improved accuracy.

Position Update Using Velocity and Acceleration as an equation (18):

$$r(t + \delta t) = r(t) + v(t)\delta(t) + \frac{1}{2}a(t)\delta t^2 + \cdots \quad (17)$$

Where:

$r(t)$ = is the position at time t, $v(t)$ is the velocity at time t, $a(t)$ is the acceleration at time t, δt is a small time step.

Position Update Using Alternative Acceleration Function as an equation (19):

$$r(t + \delta t) = r(t) + v(t)\delta(t) + \frac{1}{2}b(t)\delta t^2 + \cdots \quad (18)$$

Here, $b(t)$ represents a different rate of change in acceleration compared to $a(t)$, possibly accounting for external forces or other influences on the motion.

Equation (20) is Acceleration Update Using Jerk (Rate of Change of Acceleration)

$$(t + \delta t) = \mathrm{a}(t) + \mathrm{b}(t)\delta(t) + \cdots \quad (19)$$

Where:

$a(t)$ is the acceleration at time

$b(t)$ represents the jerk (rate of change of acceleration).

These equations are fundamental to many numerical integration algorithms in physics and engineering simulations, including Verlet Integration and Euler's Method. These techniques allow for the approximation of particle positions and velocities over small time intervals [31].

2.6 Periodic boundary conditions

The periodic boundary condition addresses surface interaction effects (Figure 5), particularly in simulations of small systems, where interactions between particles and the walls could lead to inaccurate system properties. In this method, the cubic simulation box is replicated in space, allowing particles to interact not only with others within the same box but also with particles in adjacent replicated boxes. As a result, when a particle exits the simulation box, its corresponding image particle enters from the opposite side with the same velocity, ensuring the conservation of overall mass and momentum within the system.

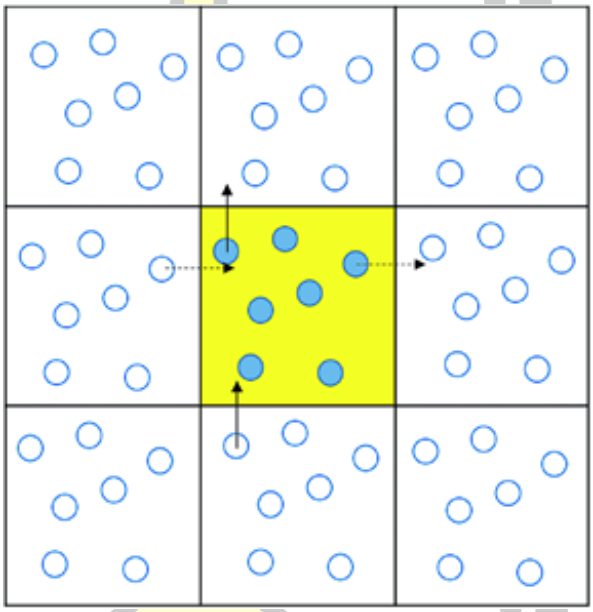

Figure 5 The periodic boundary conditions in two dimensions [32].

2.7 Cut-off and minimum image convention

When a cut-off is applied, the interactions between all pairs of atoms that are further apart than the cut-off value are set to zero. In the base image convention, each atom sees at most one image of every other atom in the system. This suggests that only the closest atoms interact with a fixed atom, as shown in Figure 6. Minimum image convection is a very explicit way to fix a cut-off length; in contrast, each system particle could see its image. On the contrary, minimum image convection could introduce adopted periodicity in the system. These periodicities could not influence the precision of properties derived as in the case of molecular dynamics simulations of well-ordered systems. Still, in the case of liquid or vitreous systems, the periodicity could induce the formation of unreal-ordered phases. Applying the

minimum image convention means cutting off non-bonded terms that are less than half the length of the shortest dimension.

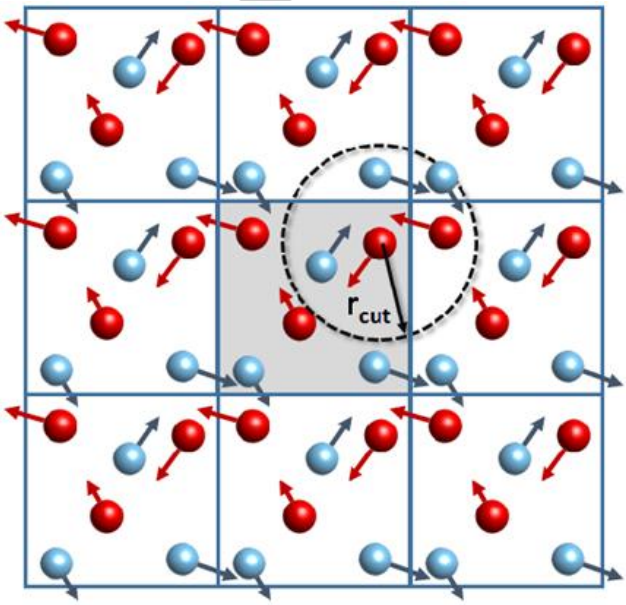

Figure 6 The spherical cut-off and the minimum image convention [33].

2.8 Free energy calculations

2.8.1 Solvation interaction energy (SIE)

The SIE approach utilizes empirical equations [34, 35] to predict the binding free energy (ΔG_{bind}) of protein-ligand interactions in an aqueous environment, as described in the following equation [36, 37].

$$\Delta G_{\text{bind}}(\rho,D_{\text{in}},\alpha,\gamma,C) = \alpha\left[E_{\text{vdW}}+E_{\text{coul}}(D_{\text{in}})+\Delta G^{R}_{\text{bind}}(\rho,D_{\text{in}})+\gamma\Delta\text{MSA}(\rho)\right]+C \quad (20)$$

Where:

E_{vdW} represents the energy associated with van der Waals interactions between the protein and ligand molecules when bound together.

E_{coul} represents the energy associated with electrostatic (Coulomb) interactions between the protein and ligand in the bound state. These interactions are a result of the attraction or repulsion between charged or partially charged atoms within the molecules, and they play a significant role in the binding process.

E_{vdW} and E_{coul} were calculated using the AMBER molecular mechanics (GAFF) force field.

ΔG^{R}_{bind} is the change in solvation reaction field energy between the bound and unbound states, which is influenced by the parameter known as the solute interior dielectric constant (D_{in}), similar to Coulomb energy.

ΔMSA represents the molecular surface area change upon binding, with the area coefficient (γ) determining its correlation with the non-polar solvation energy. The parameter ρ is the AMBER van der Waals radii linear scaling coefficient. Additionally, the coefficient α serves as a global scaling factor, while C acts as a translation constant, aligning the SIE score with the magnitude of experimental values from a training set.

The optimal standard parameters are as follows:

ρ is typically associated with scaling the van der Waals interaction energy. A value of 1.6 suggests that the van der Waals interactions are slightly amplified in the model compared to their raw calculated values.

D_{in} is likely related to the dielectric constant used in the calculation of electrostatic interactions. A value of 2.25 indicates that the model is considering a lower dielectric environment than that of bulk water (which has a dielectric constant of around 80). This lower value could reflect the less polar environment at the protein-ligand interface.

γ is often referred to as the surface tension coefficient or area coefficient. It connects the change in solvent-accessible surface area (ΔMSA) upon binding to the non-polar solvation-free energy. The units (kcal $mol^{-1}Å^{-2}$) indicate that it represents the energy penalty or gain per unit area of surface exposed to or buried from the solvent.

α scales the contribution of the electrostatic interaction energy to the overall binding free energy. A value of 0.105 suggests that electrostatic interactions play a role, but their contribution is moderated compared to other factors.

C is a constant term in the SIE equation with a −2.89 kcal/mol value. It is an overall correction factor that accounts for various energetic contributions not explicitly captured by the other terms in the model.

2.8.2 Molecular Mechanics Generalized Born Surface Area (MMGBSA)

The MMGBSA is a widely used computational method to estimate the binding free energy between two molecules, such as a ligand and a receptor. It combines molecular mechanics for intramolecular and intermolecular forces with solvation effects modeled by the Generalized Born (polar) and Surface Area (non-polar) terms. The method is commonly employed in drug discovery to assess and rank the binding affinities of ligands for their protein targets [38].

The binding free energy (ΔG_{bind}) is given by:

$$\Delta G_{bind} = E_{MM} + \Delta G_{solvation} - T\Delta S \quad (21)$$

Where:

E_{MM} is molecular mechanics energy (bonded, van der Waals, and electrostatic interactions).

$\Delta G_{solvation}$ is solvation-free energy that includes polar (Generalized Born) and non-polar (Surface Area) contributions.

$T\Delta S$ is an entropic contribution, often neglected due to computational complexity.

2.8.3 Thermodynamic integration (TI)

Thermodynamic integration (TI) calculations determine the free energy difference between two states, A and B, by coupling them through a parameter called λ, which acts as an extra, nonspatial variable. This λ framework enables the calculation of the free energy difference between the two states as follows:

$$\Delta G^{0}_{TI} = \int_0^1 \langle \frac{\partial V(\lambda)}{\lambda \partial} \rangle_\lambda \, d\lambda \quad (22)$$

The thermodynamic cycle presented in Figure 7 enables the comparison of results from a series of TI calculations with measurable physical quantities.

Processes A and B represent the binding of two different ligands to a protein, while processes C and D involve transformations between the two ligands. In process C, this transformation happens while the ligand is bound to the protein, and in process D, while the ligand is dissolved in water. Since $\Delta G_C - \Delta G_D = \Delta G_A - \Delta G_B$, TI calculations can be used to compute relative binding free energies, making them valuable for drug design or lead optimization. In equation 1, V(λ) represents the λ-coupled potential function, corresponding to V(A) when λ = 0 and V(B) when λ = 1. The integration is performed over the average λ derivative of the potential function at

different λ values. Since analytical integration is rarely possible, simulations are conducted at discrete λ points, and the integral is evaluated numerically through interpolation. One advantage of TI calculations is that independent molecular dynamics (MD) simulations at fixed λ values can be performed in parallel, enhancing efficiency. Additionally, extra λ points can be introduced at any time to improve the accuracy of the calculations [39].

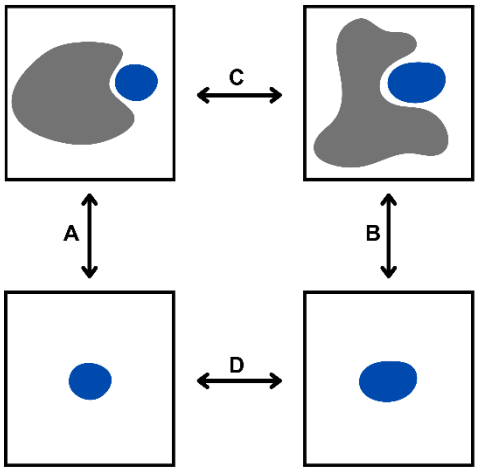

Figure 7 The thermodynamic cycle involves events A and B, which represent the binding of two distinct ligands to a protein, while events C and D signify the conversion of one ligand to the other in both the bound and solvated states, respectively. The free energy differences between processes A and C can be determined by calculating the free energy differences between processes B and D [40].

2.8.4 Free energy landscape

The function of a biological system is defined by its free energy landscape, which represents the probability of the system being in a particular state under equilibrium conditions [41]. This landscape serves as a map of biological function, with multiple pathways between states, some requiring less energy than others. The system moves through this map by balancing entropic and internal energies, influenced by dissipative forces, temperature, and the system's structure [42]. This concept is captured in the following equation.

$$G = H(U) - TS \quad (23)$$

Where:

G is Gibbs free energy

H is enthalpy

U is internal energy

T is temperature

S is entropy (S)

In biophysics, free energy usually refers to Gibbs free energy, but Helmholtz free energy (F) is also used. The choice depends on the ensemble: Gibbs free energy applies in the NPT ensemble (constant particle number, pressure, and temperature), while Helmholtz free energy is for the NVT ensemble (constant particle number, volume, and temperature). In practice, Gibbs free energy is commonly used in experiments, but both NVT and NPT ensembles are relevant in biophysics due to small volume fluctuations. The free energy landscape, representing a probability distribution, encapsulates a biological system's function and is related to the Boltzmann factor.

$$\rho(x) \propto e^{-G(x)/k_B T} \quad (24)$$

In this context, $\rho(x)$ is the probability density function along a reaction coordinate (x), G(x) is the Gibbs free energy, k_B is Boltzmann's constant, and T is the temperature. The reaction coordinate (or collective variable) describes transitions between states, but identifying these variables can be challenging, as they often involve complex motions in protein conformational changes. If multiple states exist in the free energy landscape, more than one collective variable may be required.

$$S = k_B \ln(\Omega) \quad (25)$$

State can refer to either microstates or macrostates. A microstate represents a specific configuration, while a macrostate is a collection of microstates, typically forming a basin in the free energy landscape. A macrostate has non-zero entropy, which is defined as:

In this context, k_B is Boltzmann's constant and Ω represents the number of microstates. Microstates have zero entropy, so their free energy is determined by enthalpy alone. For macrostates, entropy is proportional to the logarithm of the macrostate's hyper-volume, or the width of the basin in a 1D example. As temperature increases, macrostates widen and the free energy landscape flattens. Lower energy

states are more probable, meaning transitions between low-energy microstates are faster, while crossing energy barriers between macrostates is slower [43].

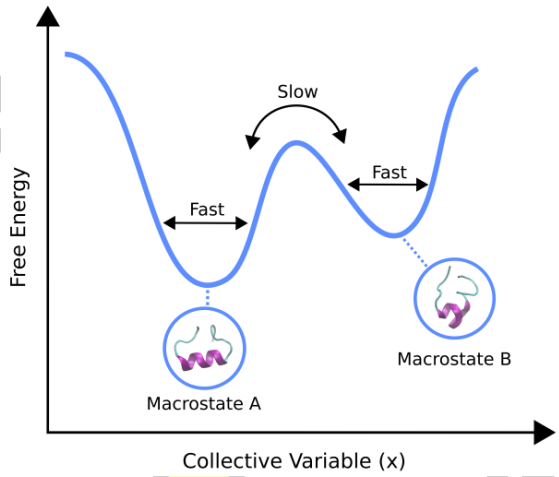

Figure 8 A free energy landscape sketch shows two macrostates, each represented by a basin, reflecting distinct protein conformations. The collective variable x describes the slow transition between states, with the transition timescale determined by the energy barrier height between the macrostates [41].

2.8.5 Umbrella sampling

Umbrella Sampling is a computational technique used to calculate the Potential of Mean Force (PMF), which represents the free energy profile along a reaction coordinate, such as the distance between two molecules or the rotation of a bond. It is commonly employed in molecular dynamics simulations to study rare events and processes where direct sampling is inefficient.

In Umbrella Sampling, the system is biased using a harmonic potential (or "umbrella") to sample different regions of the reaction coordinate more effectively. This method helps overcome energy barriers and ensures adequate sampling across the entire reaction coordinate space. After the simulations, these biased distributions are "unbiased" using statistical techniques such as the Weighted Histogram Analysis Method (WHAM) to compute the PMF [44].

The PMF, W(x), along a reaction coordinate x, can be described as:

$$W(x) = -k_B T \ln P(x) \quad (25)$$

Where:

W(x) is the Potential of Mean Force at point x along the reaction coordinate.

P(x) is the probability distribution of the system at x.

kB is Boltzmann's constant.

T is the temperature.

In umbrella sampling, biased simulations are conducted with a harmonic potential of the form:

$$U_{bias}(x) = \frac{1}{2}k(x-x_0)^2 \quad (26)$$

Where:

$U_{bias}(x)$ is the biasing potential.

k is the force constant of the harmonic potential.

x_0 is the center of the biasing window (the point along the reaction coordinate where the bias is applied).

The probability distribution obtained from each biased simulation is corrected to obtain the unbiased distribution, which is then used to calculate the PMF [45].

พหูนํ ปณฺฑิโต ชีเว

2.9 Literature reviews

The screening and searching for compounds against PLpro target can be divided into two main strategies i.e., drug repurposing and new compound searching/screening.

2.9.1 Drug Repurposing

Kandeel, M., et al., performed virtual screening of 1697 clinical FDA-approved drugs against SARS-CoV-2 PLpro [46]. The results showed the compounds with the highest docking scores, including 26 compounds with a docking score of –7 or higher. Only 10 drugs with the highest estimated docking scores and favorable pharmacokinetics were subjected to MD simulations, followed by molecular mechanics/generalized Born surface area (MM/GBSA) binding energy calculations. Three drugs, phenformin, quercetin, and ritonavir, showed favorable binding free energies (Table 1) of -56.6, -40.9, and -37.6 kcal/mol, respectively. Among the three drugs, energetic and structural analyses showed that phenformin was more stable than quercetin and ritonavir. The list of drugs provided herein constitutes a primer for clinical application in COVID-19 patients and guidance for further antiviral studies.

Table 1 Decomposition of MM-GBSA binding energies for phenformin, quercetin and ritonavir in complex with COVID-19 PLpro through 50 ns MD simulations.

Drug name	Calculated MM-GBSA binding energy (kcal/mol)						
	ΔE_{Vdw}	ΔE_{ele}	ΔE_{GB}	ΔE_{SUR}	ΔG_{gas}	ΔG_{Solv}	$\Delta G_{binding}$
Phenformin	-23.2	-124.1	94.8	-4.0	-147.3	90.7	-56.5
Quercetin	-33.5	-75.0	72.7	-5.2	-108.4	67.6	-40.9
Ritonavir	-45.0	-26.7	40.0	-5.9	-71.7	34.0	-37.6

Drug repurposing of DrugBank database was carried out by adopting ensemble docking approach for multiple conformations of PLpro SARS-CoV-2 molecular target [47]. The three drugs (Benserazide, Dobutamine and Masoprocol) showed the best docking scores and interaction with Y268 and Q269 key binding residues as illustrated Figure 7. Further MD simulations suggested the superior stability and binding of dobutamine and masoprocol inside the binding site compared to Benserazide. This approach can facilitate identifying drugs for repositioning via

targeting multiple conformations of a crucial target for the rapidly emerging COVID-19 pandemic.

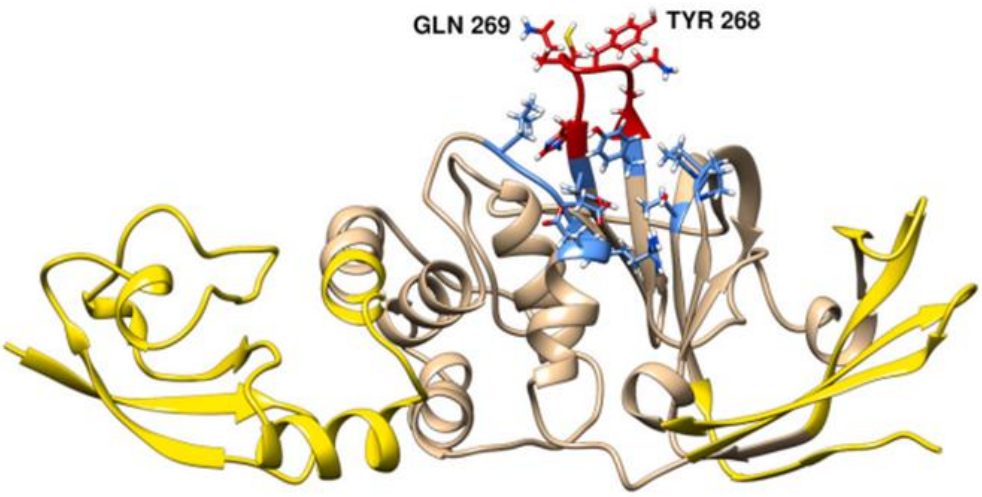

Figure 9 PLpro apo structure in khaki color (PDB ID: 6W9C) showing the binding site residues with the highly flexible loop (T265-H272) in red and the rest (L162-E167, P247-P248, Y264-Y273 and T301) in blue. The highly flexible amino acids (3-83, 182-200, 223-235 and 306-315) are shown in yellow [47].

Reza, R., et al., [48] study explores the potential of seventeen anti-lung cancer drugs to combat COVID-19 by screening them against essential SARS-CoV-2 proteins: Main Protease (Mpro), Papain-like Protease (PLpro), and Spike Glycoprotein. ADMET profiling was employed to assess the pharmacokinetics and toxicity of these drugs. In molecular docking studies, Capmatinib (CAP) emerged as the top candidate, showing the highest binding affinity and lowest inhibition constant (K_i) against SARS-CoV-2 proteins. Molecular Dynamics (MD) simulations confirmed that CAP induced significant conformational changes in the proteins, further supported by analyses of RMSD (Root Mean Square Deviation), RMSF (Root Mean Square Fluctuation), and binding energy. MMPBSA calculations revealed that CAP had the highest binding energy with PLpro compared to other proteins, indicating a strong inhibitory effect. ADMET profiling suggested that CAP has a favorable toxicity profile. The results reveal the intriguing potential of CAP as a therapeutic agent against COVID-19, sparking further interest and engagement in experimental validation and drug discovery efforts targeting SARS-CoV-2.

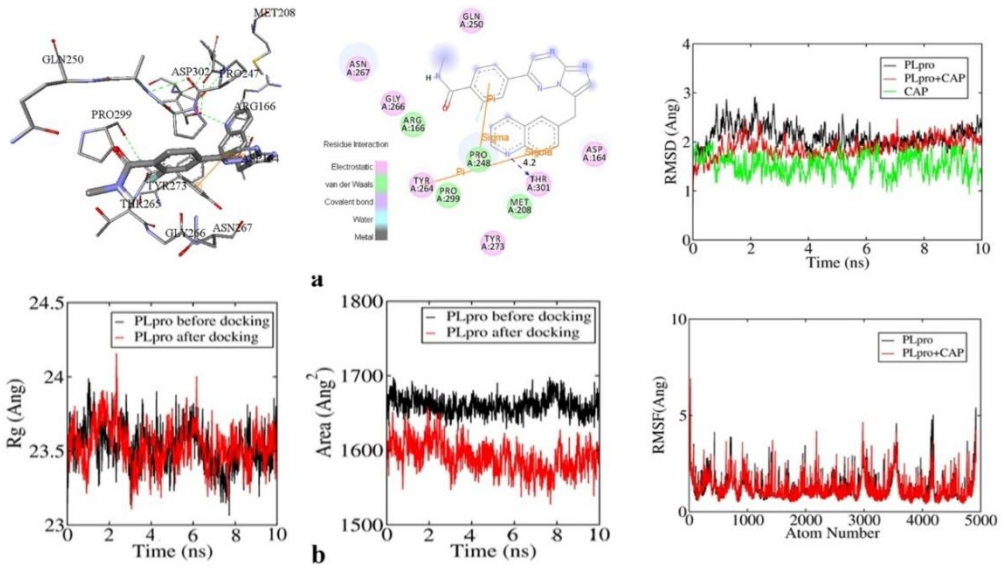

Figure 10 Binding interactions, RMSD, RMSF, Radius of gyration and SASA plots with CAP and PLpro

The researchers of Patel, R., et al., [49] focused on GRL0617, an established inhibitor of PLpro in SARS-CoV, which is also effective against SARS-CoV-2 because of the high similarity (over 80%) between the PLpro enzymes of both viruses. Notably, conserved amino acids such as Tyr268 in SARS-CoV-2 were identified for their crucial role in interacting with GRL0617, mainly through hydrophobic interactions that stabilize the binding. Then, Screening of Antibacterial Compounds. To identify new potential inhibitors of SARS-CoV-2 PLpro, an extensive database of 8581 antibacterial compounds was downloaded from the ZINC database. These compounds were filtered based on their molecular weight (between 250 and 400 Daltons) and structural features resembling GRL0617. Priority was given to compounds with aromatic rings capable of forming hydrophobic pi interactions and a central –CONH group that could form hydrogen bonds with the critical Y268 residue. The filtered compounds were docked against SARS-CoV-2 PLpro to evaluate their potential as inhibitors. This docking analysis identified five top hits that demonstrated strong interactions with the crucial Y268 residue. Among the five compounds, ZINC44459905 emerged as the lead compound, displaying the strongest interaction with Y268.

Further analysis through molecular dynamics (MD) simulations confirmed the stability of the ZINC44459905-PLpro complex, suggesting that this compound could effectively inhibit PLpro's function. Based on the docking and simulation results, the study proposes ZINC44459905 as a potential lead molecule for further in vitro and in vivo experiments. This compound is suggested as a promising candidate for the development of therapeutic treatments against COVID-19, targeting the critical PLpro enzyme to disrupt viral replication and immune evasion mechanisms.

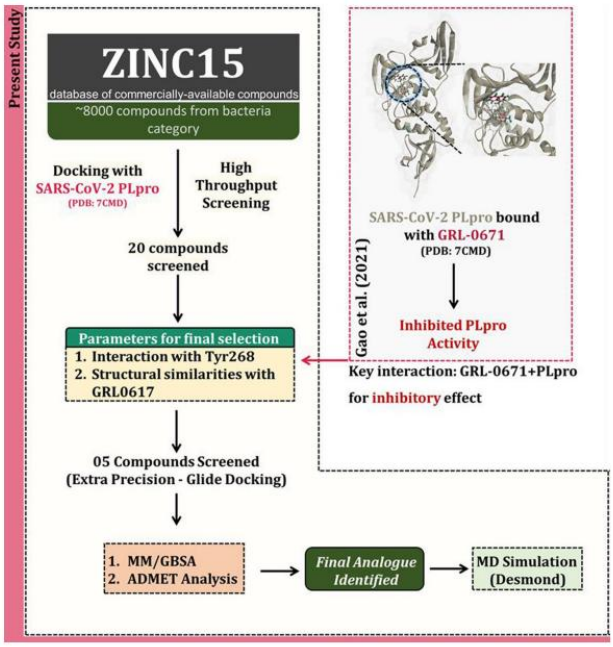

Figure 11 Illustrations represent the process of the study.

Table 2 Some repurposed drugs are predicted to be potential inhibitors for PLpro of SARS-CoV-2.

Chemical	Database		Results			REF
	Inhibitor	PDB ID	Molecular docking		MD simulation	
			Docking score (kcal/mol)	H-bond	Binding free energy (kcal/mol)	
Phenformin	FDA-approved drugs	6W9C	-7.23	-	-56.5 (MMGBSA)	[46]
Benserazide	DrugBank database	7JRN	-11.15	Q269 Y273 D164	-	[47]
Capmatinib	PubChem	6W9C	-7.70	-	-8.34 (MMGBSA)	[48]
2-(2-((benzofuran-2-carboxamido) methyl)-5-methoxy-1H-indol-1-yl)acetic acid (ZINC44459905)	ZINC15 database	7CMD	-5.70	Q269 D164	-50.96 (MMGBSA)	[49]

พหูนํ ปณฺฑิโต ชีเว

2.9.2 New compound searching

Pang, J., et al., discovered small molecules inhibited SARS-CoV-2 PLpro [50]. Data analysis of the simulations shows the top four compounds (F403_0159, F112_0109, G805_0497, and D754_0006) identified as potent SARS-CoV-2 PLpro inhibitors. MD simulations indicated that the contribution of van der Waals interaction dominated the binding free energies of these compounds, which may be attributed to the hydrophobicity of the active site of PLpro. Furthermore, as shown in Figure 8, all four compounds formed conservative hydrogen bonds with residues such as D164, Q269, and Y273.

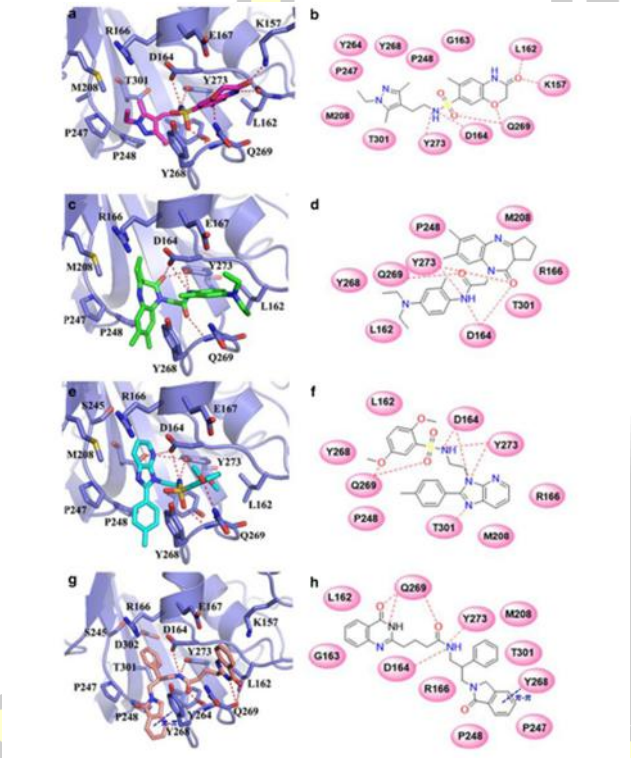

Figure 12 Predicted binding modes of four PLpro-inhibitor systems: PLpro-G805_0497 (a and b), PLpro-F403_0159 (c and d), PLpro-F112_0109 (e and f), and PLpro-D754_0006 (g and h) obtained from structure-based virtual screening. The protein PLpro is shown in cartoons and colored in blue. Hydrogen bond and π-π interactions are shown as dashed lines and colored in red and blue, respectively.

Bhowmick, S., et al., reported a few specific food compounds that can bind tightly with the SARS-CoV-2 PLpro protein identified through virtual screening, molecular docking, MD simulations, and MM-GBSA-based binding free energy calculations [51]. This research also provides a deeper understanding of the binding modes of the four proposed food compounds (FDB001395, FDB029219, FDB030757, and FDB031079) with SARS-CoV-2 PLpro protein. Molecular docking revealed several significant inter-molecular binding contacts between the functional groups of the identified food compounds and the catalytic amino acids of the SARS-CoV-2 PLpro protein (Figure 11), which were also confirmed by all-atom MD simulation studies in a dynamic state. The binding free energy showed the proposed compounds displayed better binding affinity for the SARS-CoV-2 PLpro than the standard compound VBY (control). Data analysis of MD simulations shows characteristics of both the protein backbone and the food compound, such as RMSD and RMSF, RoG, SASA, H-bond interaction profiles, etc., and it was discovered that the backbone of SARS-CoV-2 PLpro remained very stable even after binding with the suggested molecules in comparison to the standard inhibitor. Moreover, the binding free energy demonstrated strong ΔG values for all complexes, ranging from –15.56 to –28.59 kcal/mol. Overall, the extensive computational study explained that all proposed food compounds might be acting as crucial SARS-CoV-2 PLpro inhibitors or modulators for successful therapeutic application in COVID-19.

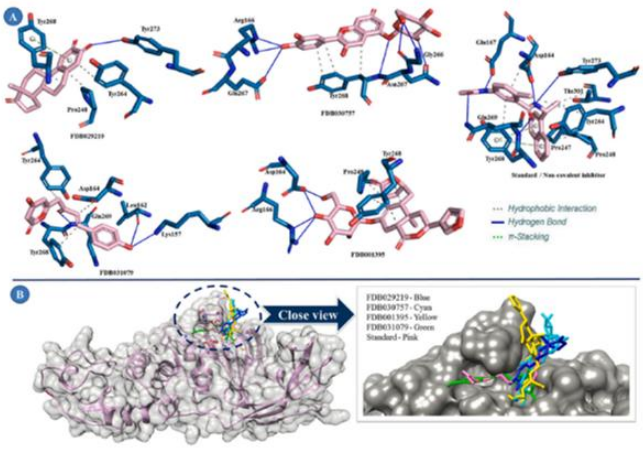

Figure 13 A) Molecular binding interactions and B) Binding mode of proposed food compounds inside the active site cavity of SARS-CoV-2 PLpro protein.

Lakhera, S., et al., investigated phytoconstituents of medicinal herb '*Piper Longum*' against SARS-CoV-2 PLpro using screen of the pharmacokinetic properties [52]. Molecular docking has revealed that I-asarinin, a component of Piper Longum with the best binding affinity score of −10.8 kcal/mol and following most of the ADMET properties, has forced us to assume it is a better component to be used as a drug against SARS-CoV-2 PLpro. Furthermore, the MD simulation results have justified the assumption by showing acceptable values of calculated binding energy. H-bond interactions and the radius of gyration, although, explained the better stability and compactness of proteins, all these properties uplift the motive to use I-asarinin as a potential drug against COVID-19 (Figure 12), and the authors believe that this *in silico* study will lead to drug development for the treatment of COVID-19.

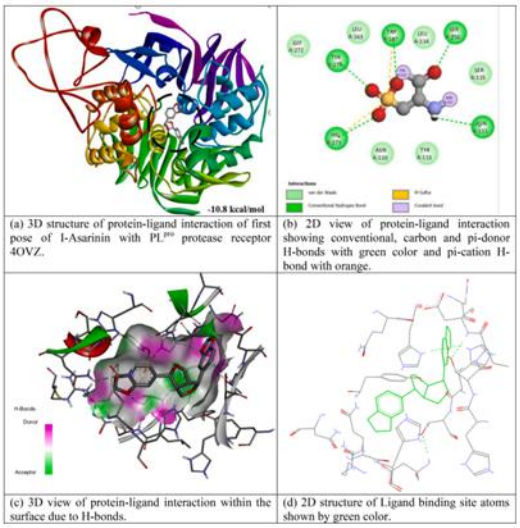

Figure 14 Donor-acceptor interactions obtained by docking of I-asarinin and receptor 4OVZ of PLpro protease of COVID-19.

Sanachai, K., et al., employed all-atom MD simulations and binding free energy calculations based on MM- PB(GB)SA and SIE methods to elucidate and compare the binding behaviors of five inhibitors derived from peptidomimetic inhibitors (VIR250 and VIR251) and naphthalene-based inhibitors (GRL-0617, compound 3, and compound Y96) against SARS-CoV-2 PLpro [53]. The results showed that all inhibitors interacted within the PLpro active site. The van der Waals interactions and hydrogen bond formation in residues G163 and G271 with peptidomimetics and the Q269 residue with naphthalene-based inhibitors were the main energy contributions that stabilized the protein-ligand complexes. The

simulation results showed that VIR250 had the highest binding efficiency with SARS-CoV-2 PLpro of the five inhibitors tested. Rational drug design to replace the aromatic rings, including heteroatoms (e.g., thiazolopyridine), at the P2 and P4 sites could help to improve the inhibitor-binding efficiency, and increased the nonpolar moiety (e.g., ethene) at the N-terminal of VIR250 to enhance hydrophobic interactions with residues at the P1 and P3 sites (Figure 13).

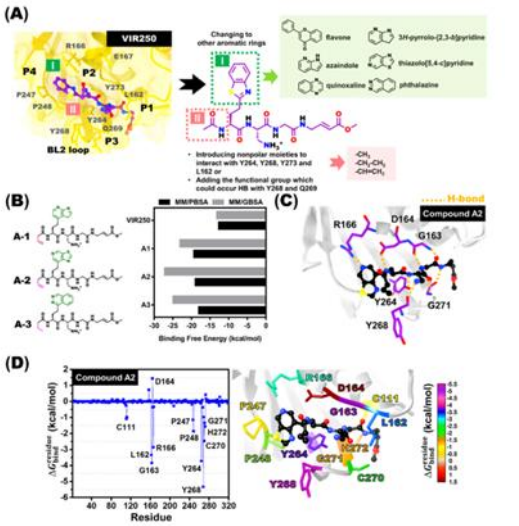

Figure 15 Rational design of the SARS-CoV-2 PLpro inhibitors based on the peptidomimetic VIR250 inhibitor. (A) 3D and 2D structures of VIR250 with ligand modifications, (B) 2D structure of modified VIR250 and their binding free energy prediction comparison with VIR250 against SARS-CoV-2 PLpro derived from MM-PB(GB)SA calculations, (C) hydrogen bond occupation, and (D) per-residue decomposition free energy (ΔG_{bind} residue) of modified VIR250 and the A2/SARS-CoV-2 PLpro complex. Calculations are obtained from one snapshot of the complex after system minimization and solvation in the TIP3P model.

Baildya, N., and co-workers screened potential drugs from Azadirachta Indica (Neem) for SARS-CoV-2 PLpro by molecular docking study along with MD simulation [54]. Desacetylgedunin (DCG) found in Neem seed showed the highest binding affinity towards PLpro as shown in Figure 14. Furthermore, MD simulation studies supported by standard analysis (e.g. root mean square deviation and fluctuation (RMSD, RMSF), radius of gyration, solvent accessible surface area (SASA)) showed large impact on the structure of PLpro by DCG.

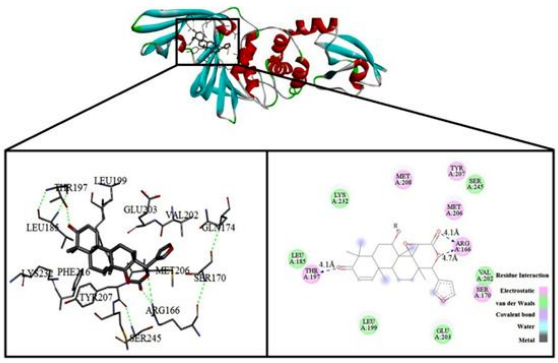

Figure 16 3D and 2D images of desacetylgedunin (DCG), with the highest binding affinity towards PLpro.

Singh, E., et al., investigated the structural influences of ligand binding on PLpro [55]. ZINC-in-Trials screening for SARS-CoV-2 PLpro Was performed. The resulting compounds obtained from molecular docking calculations, ZINC000000596945, ZINC000064033452, and VIR251 (control molecule), were subjected to MD simulation. The essential dynamics analyses utilize principal component analysis, a dynamic cross-correlation matrix, a free energy landscape, and time-dependent essential dynamics to predict the structural changes observed in PLpro upon ligand binding in a simulated environment. The binding free energy calculations based on MM/PBSA of the two selected molecules, ZINC000000596945 (-41.23 ± 3.70 kcal/mol) and ZINC000064033452 (-25.10 ± 2.65 kcal/mol), revealed significant values, indicating that they are potential inhibitors of PLpro from SARS-CoV-2 (Table 4).

พหูนํ ปณฺฑิโต ชีเว

Table 3 Individual component contribution in binding free energy calculations for native PLpro and its ligand-bound complexes.

System	Van der Waals energy (kcal/mol)	Electrostatic energy (kcal/mol)	Polar solvation energy (kcal/mol)	SASA energy (kcal/mol)	Binding free energy (kcal/mol)
ZINC000000596945	-44.70 ± 3.39	-9.86 ± 4.59	17.34 ± 3.52	-4.00 ± 0.21	-41.23 ± 3.70
ZINC000064033452	-39.25 ± 2.93	-5.78 ± 2.54	24.07 ± 3.43	-4.13 ± 0.27	-25.10 ± 2.65
VIR251	-17.38 ± 2.48	-22.68 ± 6.01	35.11 ± 7.70	-2.17 ± 0.26	-7.13 ± 5.69

Based on the research work done by Jupudi, S., et al., [56]. The docking scores of SN00334175 and SN00162745 are -10.58 kcal/mol and -9.93 kcal/mol, respectively. The MD simulation of SN00334175/7JN2 and SN00162745/7JN2 revealed that these complexes were stabilized with ligand binding interactions forming with G266, N267, Y268, Y273, T301 and D302, K157, L162, D164, R166, Q167, P248 and Y264. Moreover, van der Waal energy and hydrophobic energy terms are major contributors to total binding free energy (Figure 15).

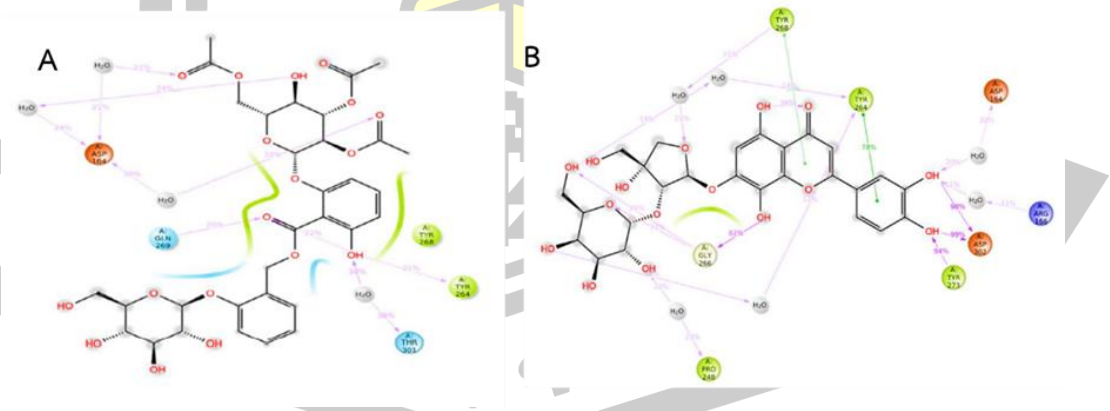

Figure 17 2D interaction diagram for the 100ns simulation trajectory of SN00334175/7JN2 (A) and SN00162745/7JN2 (B) complex.

Based on the research work by Elseginy, S.A. and M.M. Anwar, [57]. ZINC101291108 (lead 1) and ZINC16449029 (lead 2) were identified as potent SARS-CoV-2 PLpro inhibitors with IC_{50} values of 0.085 mM and 0.063 mM, respectively. MD simulations were performed for lead 1, 2 and several reported SARS-CoV-2 inhibitors. The simulations confirmed the stability of both lead compounds and showed that they adopted two confirmations during the simulation period. The per-residue decomposition results revealed that the key residues involved in inhibitor binding were E167, P247, P248, Y264, Y268 and Q269. H-bond analyses showed H-bonds with G266 and N267 and salt bridges with G209 and Y273, which are essential for strengthening the substrate-binding pocket. Both inhibitors showed hydrophobic interactions with the S4 site and BL2 loop residues. The RMSD of the BL2 loop with the two inhibitors was investigated, and the results showed that the Y268 and Q269 BL2 loop residues moved outward to accommodate the large size of lead 2. The van der Waals interaction was the main energy contribution that stabilized lead 2, while van der Waals and electrostatic interactions were the main energy contributions stabilizing lead 1. The rational design of lead 2 suggested that replacement of the 2-(2-hydroxybenzylidene) hydrazine moiety with naphthalene or nitrobenzene at the P4 position of lead 2 and introduce polar substituents as aniline and benzoate groups at position P1 to enhance hydrophobic interactions and H-bonds, respectively as illustrated in Figure 17.

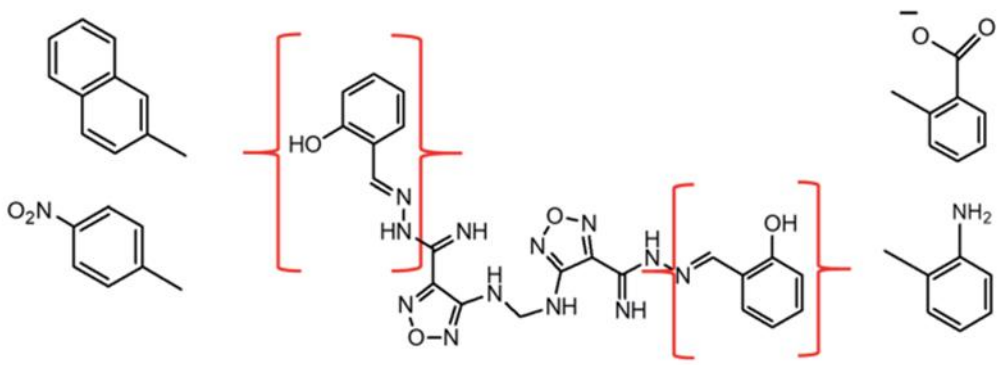

Figure 18 Rational drug design of the SARS-CoV-2 PLpro inhibitors. 2D structure of lead 2 with possible modified fragments.

Based on the research work by Thangavel, N. and M. Albratty, [58] predicted 4 olive secoiridoids as SARS-CoV-2 PLpro inhibitors. Docking results showed that demethyloleuropein aglycone was the top-ranked compound. MD simulation analysis predicted that demethyloleuropein aglycone was a potent non-covalent inhibitor of SARS-CoV-2 PLpro based on the nature and stability of its interactions with multiple binding sites. MM-GBSA energy of -94.54 ± 6.05 kcal/mol indicates good stability. In addition, for 98 % of the simulation time, two phenolic hydroxy groups of the demethyloleuropein aglycone maintained two hydrogen bonds with Asp302 of PLpro, specifying the significance of the groups in receptor binding (Figure 18).

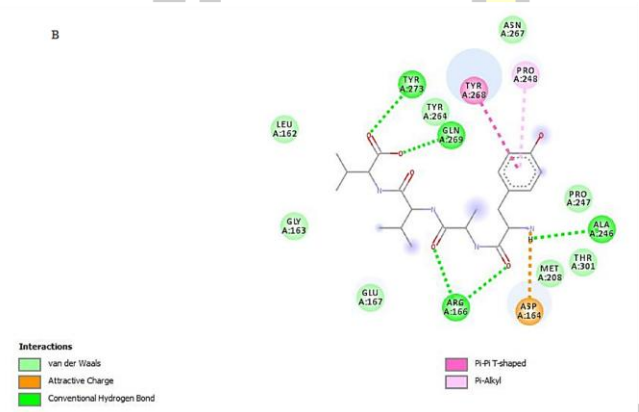

Figure 19 2D interaction diagrams of aspergillipeptide F with PLpro active site residues.

Based on the research work by Selvaraj, V., et al., [59] study explores the in silico inhibitory potential of 28 polyphenolic compounds against the SARS-CoV-2 papain-like protease (PLpro), which is crucial for viral replication. Among these compounds, amentoflavone, tiliroside, papyriflavanol A, and the antiviral drug indinavir exhibited strong binding affinities and were selected for further analysis. Amentoflavone, tiliroside, and papyriflavanol A demonstrated high stability and inhibitory potential, particularly by binding to the T158 and L162 dyad in PLpro. Further ADME (absorption, distribution, metabolism, and excretion) and DFT (density functional theory) analyses of amentoflavone and papyriflavanol A indicated excellent pharmacokinetic and molecular electrostatic properties. Molecular dynamics (MD) simulations and MM-GBSA analysis showed that amentoflavone, tiliroside, and indinavir effectively bind to the PLpro active site, with the amentoflavone-PLpro complex yielding a high MM-GBSA score of -106.56 kcal/mol. Although these

findings are promising, human clinical trials are necessary to confirm their therapeutic relevance.

Table 4 MMGBSA scores for MD simulated complexes.

Complexes	**MMGBSA score (-kcal/mol)**
PLpro (6W9C)-Amentaflavone	-106.56
PLpro (6W9C)-Tiliroside	-90.13
PLpro (6W9C)-Indinavir	-81.55

According to the analysis of Waqas, M., et al., [60] explore the discovery of potential drug candidates targeting the SARS-CoV-2 papain-like protease (PLpro), a critical enzyme for viral replication and immune evasion. Using computational and enzymatic methods, the study identified five natural compounds with strong interactions and high binding energies with PLpro, suggesting they could disrupt its activity. Molecular dynamics simulations showed that these compounds form stable complexes with the enzyme, potentially altering its normal functions. Among the identified compounds, COMP4 was the most effective in inhibiting protease and deubiquitinase activities, followed by COMP1–COMP3 and COMP5. In vitro tests demonstrated low cytotoxicity in human BJ cells at a 30 μM dosage. These findings suggest that the selected compounds have therapeutic potential for treating COVID-19 by inhibiting PLpro, and further optimization and in vivo testing are needed to improve their efficacy as drug candidates.

พหูนํ ปณฺฑิโต ชีเว

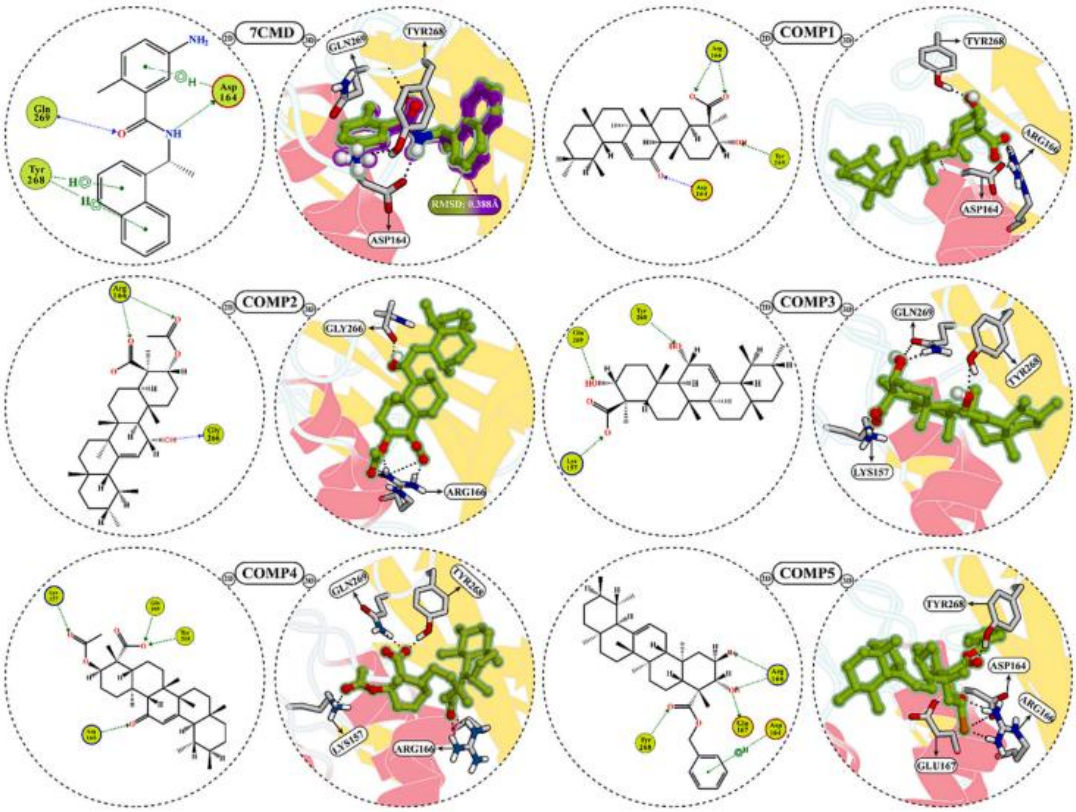

Figure 20 the 2D and 3D interactions of the compounds COMP1–COMP5. A superimposed view shows the comparison between the co-crystallized ligand (in green) in the 7CMD structure and its re-docked conformation (in purple), with a Root Mean Square Deviation (RMSD) of 0.388 Å, indicating high accuracy in docking. In the 2D format, hydrogen bonds are represented by blue and green dotted arrows, while in the 3D format, they are shown as black dotted lines. For color references and further details, readers are directed to the online version of the article.

According to the analysis of Gao, H., R. Dai, and R. Su, [61] discuss the severe global impact of the SARS-CoV-2 virus, emphasizing that managing its long-term effects on public health, the economy, and mental health will be a significant challenge. It identifies the papain-like protease (PLpro) as a promising target for antiviral drug development. Using pharmacophore-based drug design, the study screened and identified ten potential inhibitors, with compound UKR1129266 showing the best binding to PLpro through molecular docking analysis. Inhibiting PLpro is essential for blocking viral replication and preventing immune system disruptions, such as inflammatory flares observed in COVID-19 patients. Despite challenges in developing PLpro inhibitors, they are seen as essential for fighting

current and future coronavirus pandemics. The findings offer crucial insights for the design of antiviral drugs targeting SARS-CoV-2.

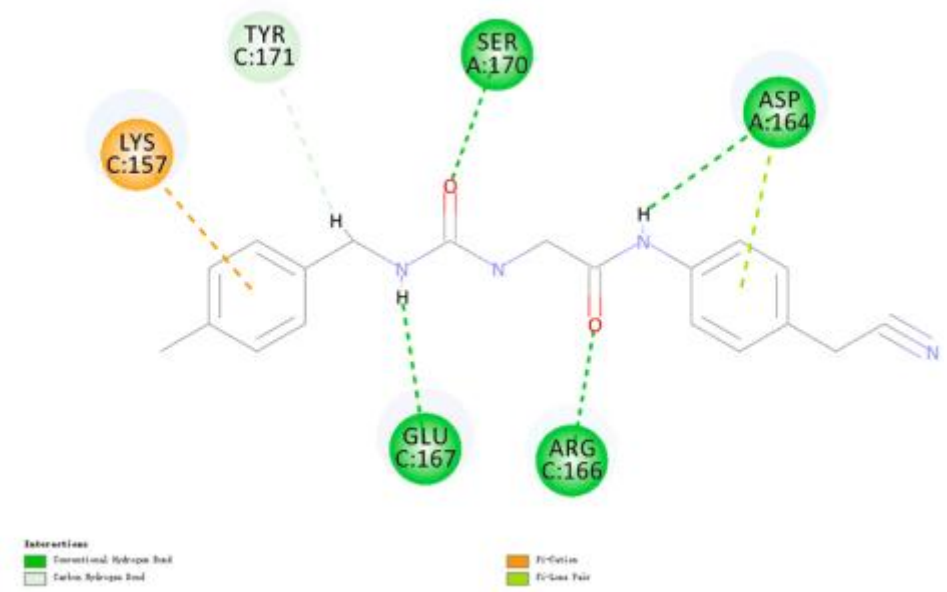

Figure 21 2-dimensional analysis of non-bonded interaction between compound UKR1129266 and PLpro

พหูนํ ปณฺฑิโต ชีเว

Table 5 Some new selected compounds predicted to be potential inhibitors for PLpro of SARS-CoV-2.

Chemical Name	Database		Results			REF
	Inhibitor	PDB ID	Molecular docking		MD simulation	
			Docking score (kcal/mol)	H-bond	Binding free energy (kcal/mol)	
N-[(2S)-3-(1-oxo-1,3-dihydro-2H-isoindol-2-yl)-2-phenylpropyl]-4-(4-oxo-3,4-dihydroquinazolin-2-yl) butanamide (ID: D754_0006)	ChemDiv database	6WX4	-	Q269 Y273 D164	− 36.60 (MMGBSA)	[50]
p-coumaroyl triacetic acid lactone (ID: FDB031079)	FooDB database	7JIW	-7.00	A240 Y258 Q263 T295	-28.597 (MMGBSA)	[51]
I-asarinin	Phytochemical (*Piper Longum*)	4OVZ	-10.80	N111 S116 W107 Y274 H273	− 11.9 (MMGBSA)	[52]
VIR250	VIR250 (6WUU) VIR251 (6WX4) GRL-	6W9C 6WX4 7JRN 7JIW 7KOL	-	-	−13.68 (MMPBSA) −15.02 (MMGBSA)	[53]

	0617 (7JRN) compound 3 (7JIW) compound Y96 (7KOL)					
Desacetylgedunin	Phytochemical (*Azadirachta Indica*)	6W9C	−7.3	-	-	[54]
(2S)-2-(4-(6-Phenylpyridin-3-ylmethoxy)phenyl) propanoic acid (ID: ZINC000000596945)	DrugDiscovery@ TACC portal	6WX4	-8.4 (Autodock)	H272 Y273 D286	-41.23 (MMPBSA)	[55]
Paclitaxel (ID: SN00334175)	SuperNatural Database	7JN2	-10.58	R166 Y264 G266 Y268 Q269	-58.344 (MMGBSA)	[56]
Lead 2 (ID: ZINC16449029)	ZINC database	6WZU	-79.58 (BUDE)	Q269 D164	-55.6 (MMGBSA)	[57]
Aspergillipeptide F	DrugBank PubChem	7LBS	-9.9 (Vina) -9.7	R166 Y273 Q269	-58.08 (MMGBSA)	[62]

	ChEMBL ZINC Natural Product Database		(Autodock)	A246		
Amentaflavone	polyphenolic compound	6W9C	-10.8 (AutoDock vina)	D108 T158 Q161 L162	-106.56 (MMGBSA)	[59]
2-bromo-3-11-dioxo-urs-12-en-24-oate (COM4)	natural compounds Boswellic acid derivatives	7CMD	-8.38 MOE 2022.02	K157 R166 Y268 Q269	-49.16	[60]
Ukrain (ID: UKR1129266)	Chelidonium majus	6WUU	-35.98 (CDOCKER)	D164 Q167	-	[61]

พหูนํ ปณฺฑิโต ชีเว

CHAPTER III COMPUTATIONAL DETAILS

3.1 Phytochemical compounds

The selection of Thai medicinal plants for this study was based on the Thai medical plant database, focusing on plants with antipyretic (fever-reducing) properties (https://ttdkl.dtam.moph.go.th) such as Harak, Khiaohom, and Chanthaleela, which are used in traditional Thai medicine recipes. These herbs are traditionally used to treat moderate to high fevers. From these plants, 45 phytocompounds were identified and assessed for their drug-like potential using Lipinski's rule [63] through the SwissADME web server (http://www.swissadme.ch) [64]. Phytocompounds meeting these criteria were further evaluated for their ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) properties using the pkCSM (http://biosig.unimelb.edu.au/pkcsm/) web platform. [65].

3.2 Drug-likeness properties

3D structures of the collected Thai medicinal plants were obtained from PubChem database, www.pubchem.ncbi.nlm.nih.gov. The selection of the Thai medicinal plants was premised on the Thai medical plant database with antipyretic properties (https://ttdkl.dtam.moph.go.th). The search provided fifty isolated phytocompounds from these plants were screened for their drug-likeness based on Lipinski's rule with additional parameters such as molar solubility (logS), molar volume (v), and number of rotatable bonds using SwissADME (http://www.swissadme.ch/) [64]. Further, those phytocompounds which fulfilled the criteria of drug-likeness were checked for their Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) properties. The pharmacokinetics properties evaluation were done with the help SwissADME online server [64]. After the ADMET analysis, the phytocompounds showed efficient pharmacokinetic parameters were selected for the molecular docking analysis.

3.3 Molecular docking

Our detailed molecular docking process was crucial in our study of SARS-CoV-2 PLpro. We discovered that the size of an inhibitor affects the conformation of the enzyme's BL2 loop in both its closed and open forms [66-69]. Specifically, binding of the small molecule GRL0617 led to a more closed conformation, narrowing the substrate cleft between the BL2 loop and the α3-to-α4 loop [57, 70]. In contrast, the larger molecule VIR251 caused this cleft to widen. Based on these findings, we chose to model the PLpro enzyme in its open BL2 loop conformation for our further analysis. To validate our docking method, we first docked GRL0617, which we obtained from a ligand-bound structure (PDB: 7CMD), into the active site of the unbound PLpro enzyme (PDB: 6WZU) [71]. This docking was performed using the GOLD CSDS 2023.3 software with default genetic algorithm settings [72], conducting 100 independent docking runs per molecule. We carefully considered the ligand binding area around SARS-CoV-2 PLpro, defined by residues within a 10 Å radius of the co-crystallized ligand GRL-0617 (X = –27.43 Å, Y = 30.00 Å, Z = 27.53 Å) [53, 57]. The best docking conformations for each compound were selected based on the GOLD score and were subsequently used for further analysis.

3.4 Molecular dynamics simulation

The best poses of protein-ligand complexes were obtained from molecular docking calculation and were used to further perform MD simulations using AMBER16 [73]. Force fields using FF14SB [74] and the second generation of GAFF (GAFF2) [75] were used as parameters for protein and ligands, respectively. The electrostatic potential (ESP) charges of each ligand were computed the HF/6-31g(d) level of theory and then the restrained electrostatic potential (RESP) charges were created by a charge fitting calculation using antechamber module implemented in AMBER16 [73]. This module was also used to neutralize the complexes by introducing Cl^- counterions, as well as to solve the complexes using a TIP3P 12.0 Å water box. The systems generated were then partially minimized for 500 steps with a 500 kcal/mol restraint potential and then fully minimized without energy restraint for 500 steps. Gradual heating of the systems was then performed from 0 K to 300 K for 50 ps under the NVT ensemble using Langevin thermostat. Equilibration of the systems at 300 K for 100 ns without energy restraint at a constant pressure of 1 atm

was performed under the NPT ensemble using Berendsen barostat [76]. The Particle-Mesh Ewald (PME) algorithm [77] was used for dialing with long-range interactions, whereas short-range cutoff of 12 Å for nonbonded interaction. The final MD simulation was performed at 100 ns for all the systems while the SHAKE algorithm [78] was employed to constrict all atomic hydrogen bonds. The MD trajectories were collected every 20 ps, and the calculation was extracted from the last 20 ns. While, The MM/GBSA method was also employed to predict the binding free energy of protein-ligand complexes [40]. The binding free energy (ΔG) is calculated using the following equation:

$$\Delta G = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad (22)$$

$$\Delta E_{MM} = \Delta E_{bonded} + \Delta E_{vdW} - \Delta E_{ele} \quad (23)$$

$$\Delta G_{sol} = \Delta G_{psolv} + \Delta G_{nsolv} \quad (24)$$

Where, ΔE_{MM} is the solvation free energy, and ΔG_{sol} is the conformational entropy. The entropy contribution from changes in translational, rotational, and vibrational degrees of freedom $-T\Delta S$ was calculated using normal mode analysis. ΔE_{MM} is further divided into bonded energy (ΔE_{bonded}), van der Waals energy (ΔE_{vdw}), electrostatic energy (ΔE_{ele}). The total solvation free energy (ΔG_{sol}) is the sum of the polar solvation (ΔG_{psol}) and non-polar energy (ΔG_{npol})

พหูนํ ปณฺฑิโต ชีเว

CHAPTER IV RESULTS AND DISCUSSION

4.1 Drug likeness and pharmacokinetic ADMET predictions

To enhance the design of molecules with favorable bioavailability, Chris Lipinski of Pfizer introduced a set of guidelines known as Lipinski's Rules of Five, commonly referred to as Lipinski's or the Rule of 5 [79]. These rules serve as a benchmark for evaluating the drug-likeness of compounds, particularly their potential to be orally active drugs in humans. Lipinski's rules are defined by four key criteria which the details are provided below:

(i) The compound must have a molecular weight not exceeding 500 g/mol. A lower molecular weight generally facilitates better cell membrane permeability, critical for oral bioavailability. Molecules that exceed this limit are often too large to easily traverse cellular membranes, which can impede their effectiveness as drugs

(ii) The compound should have no more than five hydrogen bond donors, typically comprising hydroxyl (OH) and amine (NH) groups. A higher number of hydrogen bond donors can reduce a compound's ability to cross lipid membranes, as excessive hydrogen bonding increases water solubility at the expense of membrane permeability.

(iii) The molecule should contain no more than ten hydrogen bond acceptors, including nitrogen and oxygen atoms. Similar to hydrogen bond donors, a higher count of acceptors can hinder a compound's ability to penetrate lipid membranes, negatively impacting its absorption.

(iv) The octanol-water partition coefficient, denoted by ClogP, should not exceed five. ClogP measures a compound's lipophilicity, indicating its distribution between hydrophobic (lipid) and hydrophilic (aqueous) environments. A ClogP value above five suggests that a compound is overly lipophilic, which may lead to poor solubility in water, thus affecting absorption and distribution.

The compounds that comply with Lipinski's Rule of Five demonstrate favorable drug-like properties, indicating strong potential for effective oral activity. These compounds not only align with the specified physicochemical parameters but also exceed them, making them highly promising candidates for further drug development. On the other hand, compounds that do not meet Lipinski's criteria show deviations from one or more guidelines, suggesting potential challenges in achieving

oral efficacy. As shown in Table 6, out of the 47 screened compounds, only 27 adhere to Lipinski's Rule of Five.

Table 6 Drug-likeness properties of phytocompounds predicted by SwissADME.

No.	Compounds	MW	ClogP	HBDs	HBAs	Follow Lipinski's rule
1	(3R)–6–4–dihydroxy–8–methoxyhomoisoflavane	286.32	2.9	2	4	Yes
2	(3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane	284.35	3.45	2	3	Yes
3	(3R)–7–4–dihydroxy–5–methoxyhomoisoflavane	286.32	2.9	2	4	Yes
4	(Z)–2ß–hydroxy–14–hydro–ß–santalol	238.37	2.89	2	2	Yes
5	(Z)–lanceol	220.35	4.01	1	1	Yes
6	(Z)–α–santalol	220.35	3.39	1	1	Yes
8	15-N-acetylcapparisine	507.58	2.12	6	3	No
9	28–Deoxonimbolide	452.54	4.22	0	6	Yes
10	2R–(Z)–campherene–2,13–diol	238.37	2.89	2	2	Yes
11	3,29-O-Dibenzoyloxykarounidiol	648.91	10.79	0	4	Yes
12	3-Deacetyl-3-cinnamoylazadirachtin	808.82	1.38	16	3	No
13	Aoibaclyin	642.82	3.61	5	9	No
14	Artemisinin	282.33	2.39	0	5	Yes
15	Astilbin	450.39	-0.29	7	11	No
16	Atractylodin	182.22	2.93	0	1	Yes
17	Bergenin	328.27	–1.53	5	9	Yes
18	Borapetoside B	552.57	-0.36	5	12	No
19	Borapetoside C	536.57	0.67	4	11	No
20	Bourjotinolone A	470.73	6.71	2	3	No

21	Capparispine	436.52	0.48	4	5	Yes
22	Daucosterol	576.85	5.85	4	6	No
23	Engeletin	434.39	0.01	10	6	No
24	Eurycomanol	410.42	-2.04	6	9	No
25	Eurycomanone	408.4	–1.83	5	9	Yes
26	Feruloyl tyramine	313.35	2.37	3	4	Yes
27	Gluanol acetate	468.75	9.05	0	2	No
28	Hanisonin	514.56	2.32	2	9	No
29	Hinesol	222.37	3.92	1	1	Yes
30	Hispidulin	300.26	2.59	3	6	Yes
31	Imperatorin	270.28	3.88	0	4	Yes
32	Isoastilbin	450.39	-0.29	7	11	No
33	Isocodonocarpine–noniso	466.55	0.84	4	5	Yes
34	Isoimperatorin	270.28	3.88	0	4	Yes
35	Lupeol acetate	468.75	8.45	0	2	No
36	Nimbandiol	456.53	2.92	2	7	Yes
37	Obacunone	454.51	3.6	0	7	Yes
38	Oxypeucedanin	286.28	3.1	0	5	Yes
39	Pectolinarigenin	314.29	2.89	2	6	Yes
40	Perforaquassin B	360.44	2.72	0	5	Yes
41	Resveratrol	228.24	2.76	3	3	Yes
42	Scopoletin	192.17	1.51	1	4	Yes
43	ß–eudesmol	222.37	3.92	1	1	Yes
44	Taraxerol	426.72	8.17	1	1	No
45	Tiliacorine	576.68	5.42	1	7	No
46	Tiliacorinine	576.68	5.42	1	7	No
47	**GRL0617**	**304.39**	**3.91**	**2**	**1**	**Yes**

Compounds that adhere to all four of Lipinski's criteria are likelier to possess desirable pharmacokinetic properties, including effective absorption, distribution, metabolism, excretion, and toxicity (ADMET) profiles shown in Table 7. These properties are crucial for ensuring the efficacy and safety of drugs administered orally. By meeting these standards, the compounds demonstrate potential not only for effective therapeutic action but also for minimized risks related to poor absorption or undesirable metabolic pathways. Thus, compliance with Lipinski's Rule of 5 serves as a key indicator of a compound's suitability for further development as an oral drug.

Aqueous solubility is a critical factor in the absorption of drug compounds, directly influencing their bioavailability and overall pharmacokinetic profile. Solubility is often measured using log S values, which represent the logarithm of the compound's solubility in water. Compounds with log S values below -2 are generally considered to have adequate solubility, which enhances their likelihood of being effectively absorbed in the gastrointestinal tract [80]. For oral drugs, sufficient aqueous solubility is essential as it ensures that the compound remains dissolved in gastrointestinal fluids, facilitating its passage across the intestinal membrane and into systemic circulation. Beyond solubility, intestinal absorption is another key factor in determining a compound's oral bioavailability. The percentage of intestinal absorption reflects a compound's ability to permeate the intestinal barrier a critical step for any orally administered drug to reach its target site within the body. Compounds with intestinal absorption rates exceeding 30% are typically viewed as having the potential to effectively cross the intestinal barrier, making them more promising candidates for oral drug development [81-83]. This threshold serves as a valuable benchmark in early drug discovery, helping to filter out molecules that may face absorption challenges. All compounds demonstrated calculated human intestinal absorption values greater than 60%, far exceeding the recommended criterion of 30%. This suggests that these compounds have a high absorption potential, enhancing their suitability as orally administered drugs. The high absorption rates observed for our collects compounds are particularly advantageous because they correlate with improved oral bioavailability, a critical parameter for drug effectiveness.

These findings underscore the importance of considering both solubility and absorption during the early stages of drug development. While solubility ensures that a compound remains dissolved, absorption determines its ability to cross the intestinal barrier and reach systemic circulation. The high intestinal absorption (>60%) observed in this study highlights the compounds' potential as drug candidates with promising oral bioavailability profiles.

Distribution is a vital component of pharmacokinetics, defining how a drug moves and functions within the body once it enters the bloodstream. After absorption, the drug is transported to various tissues and organs, influencing its therapeutic effectiveness, duration of action, and potential side effects. This stage is crucial in determining the overall success of the drug as a treatment.

The blood-brain barrier (BBB) plays a critical role in protecting the central nervous system (CNS) by tightly regulating the movement of substances between the bloodstream and the brain. It selectively permits only certain molecules, such as water and lipid-soluble substances, to pass into brain tissue while blocking others. This selective permeability is essential for maintaining the brain's microenvironment and shielding it from toxins, pathogens, and other harmful substances. Evaluating a drug's ability to cross the BBB is crucial for understanding its potential therapeutic effects on the CNS and assessing the risk of adverse side effects. A key measure of a drug's ability to penetrate the BBB is the log BB value, which indicates the drug's distribution between the blood and the brain. Drugs with log BB values greater than 0.3 are considered effective at crossing the BBB, making them suitable candidates for targeting CNS disorders. [84-86]. In this study, most compounds, except for compounds 1, 5, 6, 10, and 13, had log BB values below 0.3, indicating that they would be poorly distributed within the brain. This suggests limited CNS accessibility, which can be beneficial for non-CNS-targeting drugs but a limitation for drugs intended to treat brain disorders. Another parameter used to assess CNS penetration is the log PS, which measures the permeability-surface area coefficient a critical factor in determining how well a compound can penetrate brain tissue. Compounds with log PS values less than -2 are considered poorly distributed to the brain, indicating minimal CNS penetration [87, 88]. Most of the substances in this investigation showed log PS values below -2, suggesting a restricted ability to enter the brain.

Compounds 2, 6, 7, 10, 14, 19, and 25, on the other hand, indicated a higher chance of CNS penetration with log PS values exceeding -2.

Comprehending the log BB and log PS values of a substance offers valuable understanding regarding its dispersion across the central nervous system. When creating medications to treat neurological disorders, these measurements are especially crucial since they forecast how well the medication will work to target specific brain regions while reducing systemic adverse effects. Lower log BB and log PS values, Conversely, for drugs intended to act outside the CNS, lower log BB and log PS values are desirable to avoid unwanted neurological effects.

Metabolism is a pivotal stage in a drug's lifecycle within the body and a key component of pharmacokinetics, which explores how the body processes drugs. This metabolic process significantly influences the drug's effectiveness, duration of action, safety, and elimination and is crucial for achieving optimal therapeutic effects. The metabolism of drugs and their detoxification primarily rely on cytochrome P450 (CYP) enzymes, predominantly located in the liver [89]. These enzymes facilitate the conversion of drugs into metabolites, simplifying their excretion from the body. CYP2D6 and CYP3A4 are two primary CYP variants that significantly metabolize a wide range of medications [90]. Compounds 2, 3, 7, 16, 17, 18, and 26 were found to be possible CYP3A4 substrates and inhibitors in this investigation. This indicates that CYP3A4 plays a major role in the liver, where these chemicals are anticipated to undergo considerable metabolism [82, 91-93]. These substances are processed by CYP3A4 as they act as substrates, and their inhibitory potential suggests it might decrease the activity of the enzyme. This dual function is essential for the development of new drugs since it affects the drug's general safety profile, possible interactions with other drugs, and clearance rate. Predicting the behavior of these compounds in the body, including their metabolic pathways, potential drug-drug interactions, and potential adverse effects, is rendered easier by knowing what drugs interact with CYP3A4.

Excretion, the final stage in the pharmacokinetic journey of drugs, is a vital process that helps preserve internal balance. During this phase, metabolized compounds are expelled from the body, preventing the accumulation of drugs to potentially toxic levels. Understanding the excretion mechanisms and pathways is

critical for formulating safe and effective pharmaceutical treatments. Key indicators such as renal OCT2 (organic cation transporter 2) substrates and total clearance are used to gauge the efficiency with which medicines are cleared from the body relative to their concentrations internally. Total clearance measures the rate at which a drug is eliminated through metabolic and excretory processes. Drugs with lower total clearance values are cleared more slowly, potentially enhancing their stability and extending their therapeutic effects over a longer duration [94]. Eight compounds (4, 5, 6, 8, 14, 16, 18, and 27) with log(CL_{tot}) values larger than 0.7 were found to have high total clearance in this analysis. This implies that these substances are swiftly excreted from the body, which may restrict their efficiency if eliminated before they reach therapeutic levels. It may also be advantageous for lowering the danger of buildup and toxicity.

On the other hand, five compounds had poor clearance rates; their log(CLtot) values were less than 0.3, meaning they stayed in the body longer. These compounds were 1, 2, 20, 21, and 26. Maintaining therapeutic concentrations may benefit from this prolonged retention, which gives the medication enough time to reach its target before being eliminated. The remaining compounds had balanced excretion that supports efficient therapeutic action without quick elimination or excessive accumulation, indicating moderate clearance rates, with log(CL_{tot}) values ranging from 0.3 to 0.7 [18]. Except for 2R-(Z)-camphene-2,13-diol, none of the compounds under investigation were expected to be OCT2 substrates regarding renal excretion. Renal transporter OCT2 is essential for the renal absorption of many medications. Substances not OCT2 substrates are less likely to interact with this transporter, which lowers the possibility of OCT2 activity-related medication interactions and renal adverse effects.

In pharmacology, toxicity describes the extent to which a substance can harm humans or animals. For many reasons, managing and understanding toxicity is critical, establishing it as a fundamental issue in drug development, environmental health, and clinical environments. Hepatotoxicity and the AMES test are commonly employed as markers in toxicity assessments. Five chemicals under investigation, numbers 6, 9, 16, 18, and 26, showed signs of probable carcinogenicity and mutagenicity in the AMES test.

Moreover, feruloyl tyramine, capparispine, and isocodonocarpine-noniso demonstrated positive hepatotoxicity test findings, indicating a possibility for impairing normal liver function. These substances are cytotoxic, so they should be used sparingly and in low quantities. The molecules under investigation display advantageous pharmacokinetic characteristics, indicating a higher probability of exhibiting beneficial drug-like behavior.

พหูนํ ปณฺฑิโต ชีเว

Table 7 ADMET analysis predictions for selected phytocompounds using the pkCSM software

No.	Cosmpounds	Adsorption		Distribution		Metabolism				Excretion		Toxicity	
		Water solubility	Intestinal absorption (human)	BBB permeability	CNS permeability	CYP				Renal OCT2 substrate	Total Clear anc	AMES toxicity	Hepatotoxicity
						2D6	3A4	2D6	3A4				
						Substate		Inhibitor					
		Numeric (log mol/L)	Numeric (% Absorbed)	Numeric (log BB)	Numeric (log PS)	Categorical (Yes/No)				Categorical (Yes/No)	Numeric (log ml/min/kg)	Categorical (Yes/No)	
1	(3R)-6-4-dihydroxy-8-methoxyhomoisoflavane	-3.908	90.557	0.376	-2.143	No	No	No	Yes	No	0.229	No	No
2	(3R)-7,4'-dihydroxy-8-methoxyhomoisoflavane	-4.139	89.742	0.133	-1.950	No	Yes	No	Yes	No	0.197	Yes	No
3	(3R)-7-4-dihydroxy-5-methoxy-homoisoflavane	-3.740	89.578	0.127	-2.140	No	Yes	No	Yes	No	0.308	No	No
4	(Z)–2ß–Hydroxy–14–hydro–ß–santalol	-3.461	91.040	-0.091	-2.268	No	Yes	No	No	No	1.210	No	No
5	(Z)-lanceol	-4.411	93.990	0.648	-2.451	No	No	No	No	No	1.528	No	No
6	(Z)-α-santalol	-4.192	95.955	0.734	-1.418	No	Yes	No	No	No	1.058	No	No

7	28-Deoxonimbolide	-5.206	92.815	0.055	-1.604	No	Yes	No	Yes	No	0.317	No	No
8	2R-(Z)-campherene-2,13-diol	-3.215	94.118	-0.102	-2.133	No	No	No	No	Yes	1.217	No	No
9	Artemisinin	-3.678	97.543	0.235	-2.909	No	Yes	No	No	No	0.065	Yes	No
10	Atractylodin	-3.935	97.414	0.509	-1.689	No	No	No	No	No	0.422	No	No
11	Bergenin	-1.853	81.257	-0.430	-2.283	No	No	No	No	No	0.427	No	No
12	Capparispine	-3.156	84.654	-1.120	-2.389	No	Yes	No	No	No	0.564	No	Yes
13	Eurycomanone	-3.534	93.260	0.591	-2.309	No	No	No	No	No	0.703	No	No
14	Feruloyl tyramine	-3.292	98.571	0.070	-2.967	No	Yes	No	No	No	0.270	No	Yes
15	Hinesol	-4.429	97.755	0.176	-1.905	No	Yes	No	No	No	1.094	No	No
16	Hispidulin	-3.407	78.816	-0.692	-2.397	No	Yes	No	No	No	0.531	No	No
17	Imperatorin	-3.831	97.745	0.149	-2.788	No	Yes	Yes	Yes	No	0.994	Yes	No
18	Isocodonocarpine-noniso	-3.615	100	-0.650	-2.976	No	Yes	No	Yes	No	0.583	No	Yes
19	Isoimperatorin	-3.754	90.230	-0.715	-2.547	No	Yes	No	Yes	No	1.027	Yes	No
20	Nimbandiol	-4.882	100	-0.451	-1.934	No	Yes	No	No	No	0.462	No	No
21	Obacunone	-4.415	94.523	-0.580	-2.314	No	No	No	No	No	0.159	No	No
22	Oxypeucedanin	-3.046	95.277	-0.299	-2.230	No	No	No	No	No	0.716	No	No
23	Pectolinarigenin	-3.245	100	-0.258	-2.847	No	Yes	No	No	No	0.637	No	No
24	Perforaquassin B	-4.081	90.935	-0.048	-2.067	No	Yes	No	No	No	0.614	No	No

25	Scopoletin	-2.504	94.296	0.634	-1.858	No	Yes	No	No	No	0.73	No	No
26	Resveratrol	-3.178	93.361	-0.643	-2.854	No	Yes	No	Yes	No	0.076	Yes	No
27	ß-eudesmol	-4.900	63.774	-1.091	-3.903	No	No	No	No	No	1.032	No	No
28	**GRL0617**	-4.678	91.557	0.101	-1.645	No	Yes	Yes	No	No	0.221	Yes	No

พหูนํ ปณฺฑิโต ชีเว

4.2 Molecular Docking Calculation

To validate the docking protocol, the re-docking processes of GRL0617 to the active site of PLpro of SARS-COV-2 were preliminary performed as described in the computational method section. As depicted in Figure S1, the re-docked conformations of GRL0617 showed significant overlap with the native co-crystallized ligand and interacted with the same critical amino acids characterized by H-bonds (Y273 and D302), hydrophobic (G163, V165, R166, S245, A246, Y264, N267 and Y268) and π-π (L162, D164 and T301) interactions. These findings underscore the efficiency and validity of the docking protocol.

After the validation of docking procedures, molecular docking of the collected phytochemicals was performed similarly to that of GRL0617, and the resulting docking scores are summarized in Table 8. These phytochemicals were assessed based on their maximal docking scores, which served as criteria to evaluate their potential as drugs. The docking scores of the complexes formed between the studied molecules and the PLpro enzyme ranged from 35.0 to 55.0 Notably, feruloyl tyramine) scored the highest at 55.31, followed by capparispine (52.8), bergenin (51.3), and (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane (50.5). These scores were notably higher than that of GRL0617 (52.5). Consequently, the docking poses of these top-scoring ligands within 15 percent were selected for further investigation through MD simulations.

พหูนํ ปณฺฑิโต ชีเว

Table 8 Docking scores (binding energy in kcal/mol) of the phytocompounds against SARS-CoV-2 PLpro, through GOLD (as the values appear) for the best docked compounds.

No.	Compounds	GOLD score
1	(3R)–6–4–dihydroxy–8–methoxyhomoisoflavane	49.4
2	(3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane	50.5
3	(3R)–7–4–dihydroxy–5–methoxyhomoisoflavane	47.5
4	(Z)–2ß–Hydroxy–14–hydro–ß–santalol	47.1
5	(Z)–lanceol	49.4
6	(Z)–α–santalol	47.7
7	28–Deoxonimbolide	38.1
8	2R–(Z)–campherene–2,13–diol	46.0
9	Artemisinin	38.5
10	Atractylodin	38.2
11	Bergenin	51.3
12	Capparispine	52.8
13	Eurycomanone	38.5
14	Feruloyl tyramine	55.3
15	Hinesol	35.8
16	Hispidulin	46.1
17	Imperatorin	49.1
18	Isocodonocarpine–noniso	47.9
19	Isoimperatorin	47.3
20	Nimbandiol	42.0
21	Obacunone	37.4
22	Oxypeucedanin	47.7
23	Pectolinarigenin	45.3
24	Perforaquassin B	41.9
25	Resveratrol	49.1
26	Scopoletin	36.5
27	ß–eudesmol	34.7
28	**GRL0617**	**52.5**

4.3 Molecular Dynamics (MD) simulation

4.3.1 System stability

Root mean square deviation (RMSD) analysis was used to evaluate the binding stabilities of the top four protein-ligand complexes throughout a 100 ns simulation period. The results are shown in Figure 22. The ligands continuously stayed inside the binding pocket, exhibiting outstanding stability throughout the three simulation runs, as seen by the RMSD figure. Low RMSD values, typically between 2.00 and 3.00 Å, show little volatility and indicate that the complexes maintained relatively stable conformations over time.

To verify the stability and compactness of the protein-ligand complexes, the simulations involved the observation of the radius of gyration or R_g. R_g values measure the total compactness of the protein structure, giving information on the protein's ability to hold its shape during the simulation. R_g values varied by less than 1 Å during the simulations, with all complexes falling between 23.50 Å and 24.00 Å. This slight variation suggests that the protein-ligand complexes maintained their structural stability and compactness for the whole simulation time.

Both RMSD and R_g suggested the complexes in their most stable and equilibrated condition, snapshots taken during the final 20 ns of the simulation (between 80 and 100 ns) were chosen for additional examination. The information supports that the most representative configurations of the protein-ligand interactions serve as the foundation for the final structural evaluations and analyses. The selected complexes maintain a strong and stable binding, as confirmed by the stability in the RMSD and R_g analyses. This indicates that they are potential candidates for further study.

พหูนํ ปณฺฑิโต ชีเว

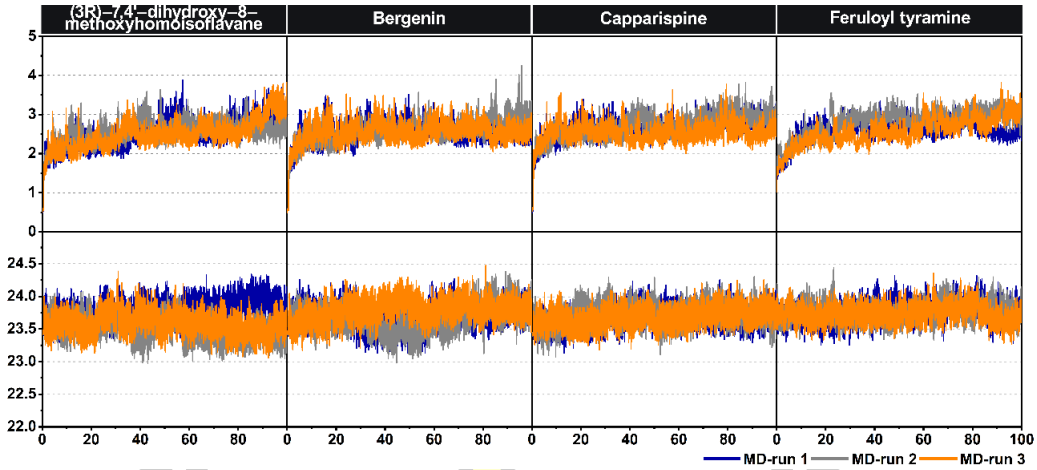

Figure 22 RMSD, and R_g profiles for all protein-ligand complexes over time for each of the three replicates (Run 1–3).

4.3.2 Solvated interaction energy (SIE) Binding energy

The binding strength of all protein-ligand complexes (ΔG_{bind}^{SIE}) was initially evaluated using the SIE (Solvated Interaction Energy) approach, which utilized snapshots from the last 20 nanoseconds of simulation from three independent runs. The SIE method estimates binding free energy in solvation by combining Coulomb and non-polar interaction components with the desolvation-free energy contribution. The results showed that the protein-ligand binding energies were energetically favorable across all three independent runs, with energy differences of less than 1.50 kcal/mol (Figure 23). The most stable binding free energies observed from the three runs were approximately -7.13, -7.57, -6.69, and -6.27 kcal/mol for (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane, bergenin, capparispine, and feruloyl tyramine, respectively. It's important to note that these binding stabilities differ from docking energy results, which do not consider structural dynamics and solvation effects. However, the differences in binding free energies among the four complexes were similar and not easily distinguishable. Therefore, an additional approach is necessary to differentiate their potential further. As previously mentioned, the binding free energies obtained from the three independent runs for each system showed no significant variation. Consequently, the simulation that yielded the most stable binding energy for each system was selected as the representative protein-ligand complex for further analysis, ensuring the reliability and consistency of the research

findings. The results indicated that the binding energies were consistently favorable across different simulations, with minimal energy variation, suggesting reliable interactions. However, since the energy differences between the ligands were minimal, it was challenging to distinguish their relative binding strengths. The meticulous process of choosing the simulation with the most stable energy for each ligand as the best representation for further analysis was crucial to ensure the robustness of our conclusions.

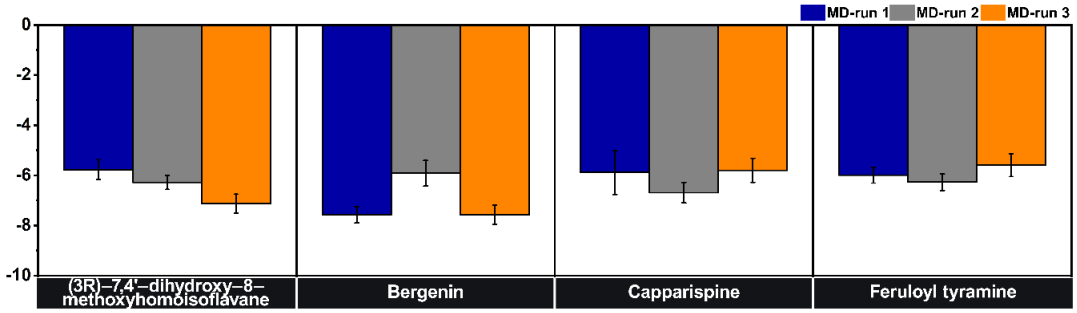

Figure 23 The $\Delta G_{\text{bind}}^{\text{SIE}}$ (kcal/mol) results are based on the SIE method for all protein-ligand complexes over time for each of the three replicates (Run 1–3).

4.3.3 Molecular Mechanics General Born surface area (MM/GBSA) Binding free energy

Previous studies have demonstrated the significant role of the MM/GBSA (Molecular Mechanics/Generalized Born Surface Area) method in enhancing the accuracy of docking calculations and effectively distinguishing between active and inactive molecules [65-67]. In this study, we conducted MM/GBSA calculations to estimate the binding free energy for each simulated system, using 100 snapshots from the last 20 nanoseconds of the best simulation run identified by the SIE approach. While MM/GBSA may not perfectly replicate experimental values, its ability to closely approximate these results underscores its usefulness in predicting relative binding affinities and understanding molecular interactions, thereby enlightening the scientific community.

Table 9 presents the binding free energy and components calculated by the MM/GBSA method for all the complexes studied. MM/GBSA offers a more apparent distinction between active and inactive compounds than the SIE results. The binding free energies for (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane, bergenin, capparispine, and feruloyl tyramine were -14.70, -12.48, -13.11, and -1.02 kcal/mol, respectively, with (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane showing the strongest binding affinity to PLpro. All systems demonstrated favorable binding free energies except for feruloyl tyramine. Although the ranking of the first three compounds differs from that obtained using the SIE method, both methods consistently indicate that feruloyl tyramine forms the least stable complex with PLpro among the four compounds. It is also noteworthy that the ranking of ligand potential differs from that of docking calculations, which typically rely on a single static protein conformation and may only be ideal for evaluating some ligands. Utilizing the MM/GBSA method to assess the binding free energy of various ligands with the PLpro protein provides valuable insights into the stability of these protein-ligand interactions, a crucial aspect in drug development.

By analyzing snapshots from the best-performing simulation run identified by the SIE approach, this study provided valuable predictions about how strongly different molecules bind to the target protein, aiding in distinguishing active compounds from inactive ones. The MM/GBSA method's contribution to this understanding is significant and should be appreciated by the scientific community. The findings revealed that (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane had the strongest binding affinity, while feruloyl tyramine exhibited the weakest. Despite some differences in compound rankings compared to other methods like SIE and traditional docking, the consistent conclusion across methods was that feruloyl tyramine was the least stable, suggesting its likely inactivity. This comparison underscores the importance of using multiple methods to understand molecular interactions comprehensively, providing reassurance about the robustness of our study.

A detailed analysis of the energy components revealed that van der Waals interactions, electrostatic interactions, and non-polar solvation energy contributed negatively to the overall binding energy, enhancing the stability of the protein-ligand complexes. In contrast, polar solvation energy contributed positively, working against binding stability. The van der Waals energy terms were relatively consistent across the four systems, ranging from approximately -30.00 to -43.00 kcal/mol. However, the contributions from electrostatic interactions varied significantly, ranging from around -6.00 to -156.00 kcal/mol. Electrostatic interactions significantly stabilized the bergenin and capparispine complexes, contributing most significantly to their negative binding energies. In contrast, this stabilizing effect was less pronounced in the (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane and feruloyl tyramine complexes. These variations in electrostatic contributions likely explain the differences in total binding free energies observed among the four complexes.

Bergenin and capparispine exhibited stronger binding affinity to PLpro, primarily through hydrogen bonding interactions with surrounding polar amino acids. On the other hand, van der Waals interactions from non-polar residues were a significant factor in stabilizing the feruloyl tyramine complex, contributing significantly to the total binding free energy. The notable increase in electrostatic energy observed in the capparispine complex can be attributed to its positively charged nature, allowing it to establish stronger electrostatic interactions than the other systems. The focus is dissecting the various energy components contributing to the binding stability of different protein-ligand complexes. The negative contributions (which lower binding free energy and thus enhance binding stability) primarily come from van der Waals, electrostatic interactions, and non-polar solvation energy. However, polar solvation energy works against binding stability by contributing positively. The electrostatic interactions vary greatly between the ligands, significantly influencing their overall binding stability. For example, bergenin and capparispine benefit most from electrostatic solid interactions, leading to more stable binding.

Conversely, (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane and feruloyl tyramine have less electrostatic stabilization, resulting in different overall binding energies. The analysis also highlights the importance of hydrogen bonding and van der Waals interactions in stabilizing these complexes, particularly for bergenin, capparispine, and feruloyl tyramine. The distinct energy profiles of these interactions explain why some compounds have stronger binding affinities than others (as discussed in the following section).

The non-polar solvation-free energies, which relate to the burial of solvent-accessible surface area (SASA) as the ligand binds to the protein, had a slightly favorable effect on the binding stability. This suggests that the non-polar components of the energy specifically, the sum of non-polar solvation energy ($\Delta G_{\text{solv-np}}$) and van der Waals interactions (ΔE_{vdw}) played a significant role in stabilizing the complexes between PLpro and the ligands. These non-polar forces help the ligand fit snugly into the protein's binding site, enhancing overall stability. While the electrostatic interactions (ΔE_{ele}) between the protein and ligands were generally favorable, they were counterbalanced by the polar solvation energy ($\Delta G_{\text{solv-ele}}$), which measures the energy cost of dissolving the charged groups in water. When these two terms are combined ($\Delta G_{\text{solv-ele}} + \Delta E_{\text{ele}}$), the net electrostatic contribution becomes positive, detracting from the binding free energy, making it less favorable. This effect was most pronounced in the PLpro/feruloyl tyramine complex, where the highest value of this combined electrostatic term resulted in a reduced overall binding free energy compared to other PLpro/ligand complexes. This reduction in binding free energy indicates a less stable complex, which could have implications for the effectiveness of feruloyl tyramine as a PLpro inhibitor.

Although electrostatic and polar interactions were critical factors in the energy differences observed between the complexes, the configurational entropy term ($-T\Delta S$) also played a role. This term, which reflects the loss of entropy or freedom of movement when the ligand binds to the protein, was relatively consistent across all systems. The loss of entropy is a measure of the reduction in the ligand's freedom of movement when it binds to the protein, and it contributed unfavorably to the binding

energy, indicating that binding reduced the ligands' flexibility in an energetically costly way.

Overall, the results suggest that most of the selected phytochemicals, except for feruloyl tyramine, show promising potential as inhibitors of PLpro. This finding instills hope and optimism in the potential of these phytochemicals in molecular biology. However, feruloyl tyramine's less favorable binding characteristics, likely due to the unfavorable net electrostatic contribution, underscore the need for further investigation to fully understand the molecular interactions that influence ligand binding to PLpro. The study reaffirms the potential of most of the phytochemicals tested as inhibitors of PLpro, except for feruloyl tyramine, which showed less favorable binding due to its unfavorable electrostatic contribution. This finding underscores the need for further research to fully understand molecular interactions, particularly for ligands that do not exhibit binding solid characteristics. The study's findings are significant in molecular biology, as they point to the potential of these phytochemicals as PLpro inhibitors and the importance of understanding their binding characteristics.

พหูนํ ปณฺฑิโต ชีเว

Table 9 MM/GBSA binding energy (kcal/mol) components for a given MD run of protein-ligand complexes.

Energy conponents (kcal/mol)	(3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane	Bergenin	Capparispine	Feruloyl tyramine
ΔE_{ele}	-26.26 ± 1.13	-63.38 ± 0.79	-156.05 ± 2.43	-5.97 ± 0.65
ΔE_{vdw}	-35.72 ± 0.44	-30.79 ± 0.42	-42.54 ± 0.48	-30.79 ± 0.50
ΔE_{MM}	-61.98 ± 1.05	-94.17 ± 0.66	-198.59 ± 2.25	-36.77 ± 0.81
$\Delta G_{Sol}^{GB/ele}$	33.58 ± 0.90	65.85 ± 0.40	169.38 ± 2.10	19.96 ± 0.59
$\Delta G_{Sol}^{GB/np}$	-4.51 ± 0.02	-4.53 ± 0.03	-6.00 ± 0.04	-4.21 ± 0.03
ΔG_{Sol}^{GB}	29.07 ± 0.90	61.32 ± 0.40	163.39 ± 2.08	15.75 ± 0.59
$\Delta G_{Sol}^{GB/ele} + \Delta E_{ele}$	7.33 ± 2.03	2.03 ± 1.20	13.33 ± 4.53	13.99 ± 1.25
$\Delta G_{Sol}^{GB/np} + \Delta E_{vdW}$	-40.24 ± 0.46	-35.32 ± 0.44	-48.54 ± 0.52	-35.01 ± 0.53
$-T\Delta S$	18.21 ± 1.09	20.36 ± 2.01	22.09 ± 0.92	20.02 ± 0.39
$\Delta G_{Total}^{MM/GBSA}$	-32.91 ± 0.44	-32.85 ± 0.47	-35.21 ± 0.44	-21.02 ± 0.39
$\Delta G_{Bind}^{MM/GBSA}$	-14.70 ± 0.66	-12.48 ± 1.54	-13.11 ± 0.49	-1.02 ± 0.21

4.3.4 Understanding the Key Residues in Ligand Binding to SARS-CoV-2 PLpro

To identify the critical residues involved in ligand binding to the SARS-CoV-2 papain-like protease (PLpro), researchers decomposed the MM/GBSA binding free energy to assess the interaction energy of each protein residue. This analysis allowed for a detailed understanding of how specific amino acids within PLpro contribute to the stability of the protein-ligand complexes.

The interaction spectra and binding modes between the protein and the ligands are illustrated in Figure 23 Residues that contributed more than 0.5 kcal/mol to the stabilization of the complex are highlighted, with negative energy values indicating stabilization (favorable interactions) and positive values indicating destabilization (unfavorable interactions). For the ligand (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane, significant stabilization was provided by residues L162, D164, V165, R166, Y264, Y268, and Q269, with energy contributions ranging from approximately -1.50 to -2.50 kcal/mol. These interactions involve hydrophobic contacts, where nonpolar residues interact with the ligand, π-π stacking, where aromatic rings of the protein and the ligand interact, and hydrogen bonding, where hydrogen atoms of the protein interact with electronegative atoms of the ligand.

Additionally, residues P248, Y273, T301, and D302 made smaller contributions of around -0.50 kcal/mol. The flexible β-hairpin BL2 loop is fascinating, especially residues Y268 and Q269, which play a crucial role in regulating the binding of viral protein substrates to PLpro [95]. In the case of the bergenin complex, the primary stabilizing residues were R166, M208, A246, P247, P248, Y273, T301, and D302, with energy contributions ranging from approximately -1.00 to -4.00 kcal/mol. The strongest stabilizing interactions were attributed to residues R166 and D302, which form hydrogen bonds with bergenin, enhancing the stability of the complex. For the capparispine complex, residues D164, P247, P248, Y264, Y268, and D302 were identified as critical stabilizers within the PLpro binding pocket, with energy contributions between -2.00 and -4.50 kcal/mol. D302 was particularly significant due to its role in forming a salt-bridge interaction between its carboxylate anion and the ammonium ion of capparispine. This strong interaction greatly enhances binding stability.

The feruloyl tyramine complex was stabilized primarily by residues P247, P248, Y264, and Y268. These residues played crucial roles in forming favorable interactions, which helped maintain the ligand's stability within the binding pocket of the PLpro enzyme. The analysis reveals that specific residues stabilize different ligands within the PLpro binding pocket. These findings align with previous studies, confirming the importance of these critical residues across various ligands in stabilizing protein-ligand complexes. Understanding these interactions is vital for designing effective inhibitors targeting SARS-CoV-2 PLpro, as these residues are crucial for strong and stable binding [53].

พหูนํ ปณฺฑิโต ชีเว

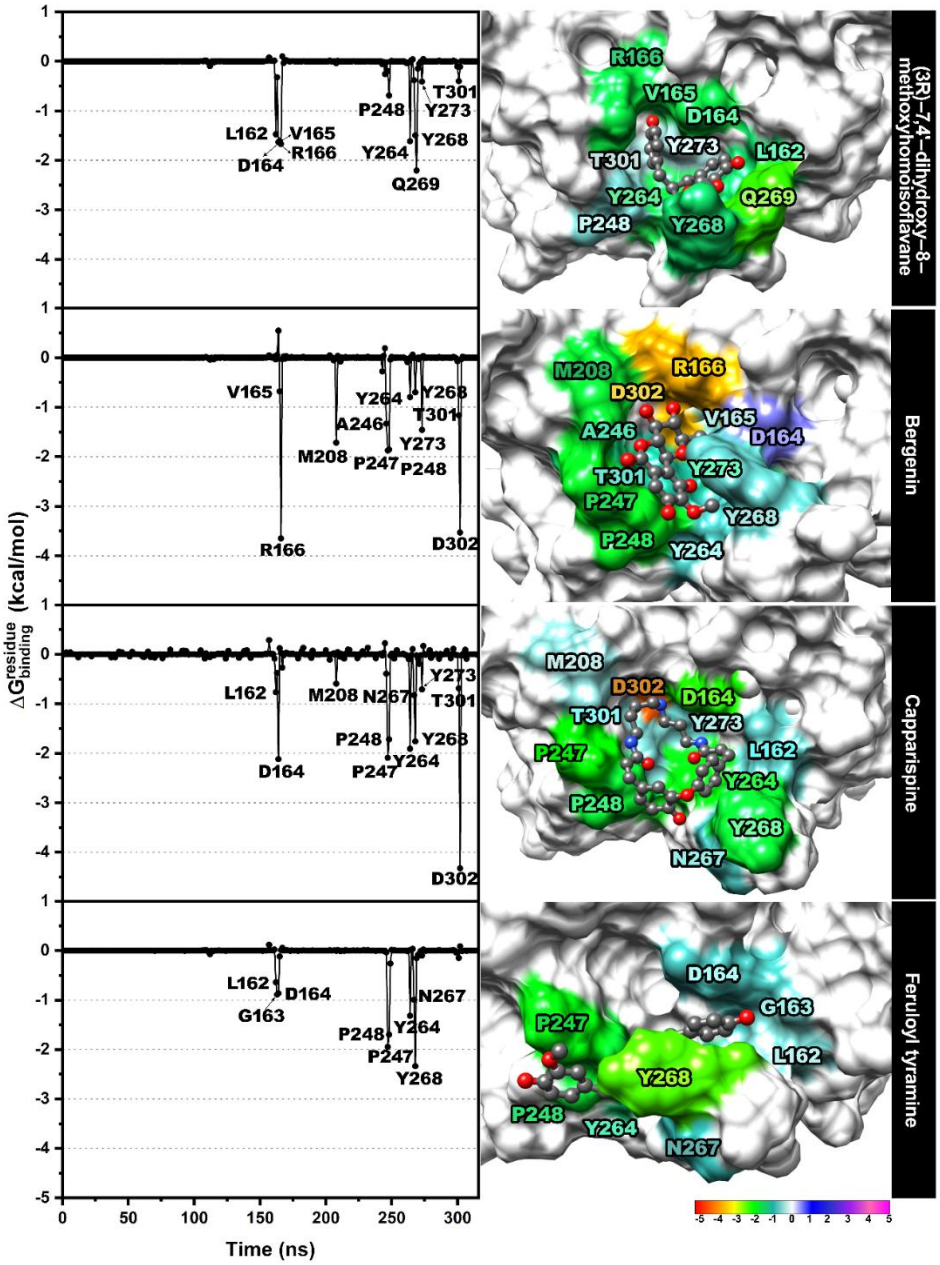

Figure 24 (A) Decomposition of binding energy per residue on SARS-CoV-2 PLpro utilizing the MM/GBSA approach. (B) The 3D structures with energy contributions of < 5 kcal/mol of ligands in association with SARS-CoV-2 PLpro are identified and colored appropriately.

4.3.5 Protein-ligand interactions

Hydrogen bonds, pivotal in bimolecular systems, were the subject of our research. We analyzed their evolution over the last 20 ns of the simulation using the cpptraj module. Applying a cutoff distance of 3.5 Å and an angle of 120º, we identified these bonds. The binding mode between the protein and ligand complexes was then visualized using Biovia Discovery Studio, providing a clear and comprehensive understanding. Figure 24, a visual representation of our findings, illustrates the percentage of hydrogen bond occupancy along with representative snapshots.

Throughout the simulation, strong and stable hydrogen bonds were observed between the hydroxyl (OH) groups of (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane and the key residues R166, Q269, and D302 within the PLpro enzyme. These hydrogen bonds persisted consistently, indicating their crucial role in maintaining the stability of the ligand within the binding pocket. Residue Q269, situated in the BL2 loop of the palm subdomain, was particularly notable for its involvement in hydrogen bond formation. This residue has been previously recognized for its interaction with other inhibitors, reinforcing its importance in ligand binding [53, 96, 97]. Beyond hydrogen bonding, the ligand's aromatic portions engaged in π-π stacking interactions with the aromatic residues Y264 and Y273. These π-π stacking interactions involve overlapping electron clouds between the aromatic rings of the ligand and protein, adding another layer of stability to the complex.

Moreover, (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane formed hydrophobic interactions with nearby non-polar residues such as L162, G163, V165, and P248. These hydrophobic interactions occur as the non-polar regions of the ligand and protein come into proximity, reducing their exposure to the aqueous environment and contributing to the overall binding stability. In addition to these interactions, π-alkyl interactions were also observed. These occurred between the aromatic ring of (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane and the side chain of L162, as well as between the aromatic moiety of Y268 and the methoxy group of the ligand. These π-alkyl interactions enhance the ligand's binding affinity by stabilizing its orientation within the binding site. The various types of molecular interactions contribute to the

binding stability of (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane within the PLpro enzyme of SARS-CoV-2. Hydrogen Bonds are strong interactions formed between the hydroxyl groups of the ligand and specific amino acid residues (R166, Q269, and D302) within PLpro. These bonds are critical for the stability of the ligand within the protein's binding pocket. The persistence of these hydrogen bonds throughout the simulation underscores their importance. Residue Q269 is in a flexible loop region of the protein. This residue is essential for binding and forming hydrogen bonds with the ligand. In previous studies, its involvement with other inhibitors highlighted its role as a critical interaction site within PLpro. These interactions collectively contribute to the strong binding affinity and stability of (3R)–7,4'–dihydroxy–8–methoxyhomoisoflavane within the PLpro enzyme, making it a potentially effective inhibitor.

Bergenin, a potential SARS-CoV-2 PLpro enzyme inhibitor, forms several strong hydrogen bonds with critical residues within the PLpro binding pocket, particularly R166 and D302. The hydroxyl group (OH) on the aromatic side chain of Y273 further stabilizes the interaction with bergenin. A unique π-π interaction, a type of hydrophobic interaction, is observed between the benzene ring of bergenin and the aromatic residue Y268. Bergenin also establishes non-polar contacts with the residues A246, P247, and P248. Additionally, a π-alkyl interaction occurs between the methoxy group of bergenin and the aromatic moiety of Y264. These collective interactions contribute significantly to the stable binding of bergenin within the PLpro binding pocket.

Capparispine, a potential SARS-CoV-2 PLpro enzyme inhibitor, forms a network of salt bridges with the carboxylate groups of residues D164 and D302, creating electrostatic solid interactions. One of the NH groups of capparispine also participates in a hydrogen bond with the backbone carbonyl of G163. The large cyclic ring of capparispine engages in hydrophobic contacts with multiple residues, including L162, G163, V165, A246, P247, P248, Y264, and Y273, further stabilizing its binding within the pocket.

In contrast, the feruloyl tyramine complex showed relatively fewer hydrogen bonding interactions with PLpro residues, with no hydrogen bonds detected between the protein and this ligand. The stabilization of feruloyl tyramine primarily relied on

non-polar interactions, including π-π interactions with residues Y264 and Y268, hydrophobic interactions with P247 and P248, and alkyl-π interactions with L162. This reduced level of interaction likely accounts for the diminished potency of feruloyl tyramine against PLpro compared to the other compounds.

The three selected phytochemicals, including (3R)-7,4 '-dihydroxy-8-methoxyhomoisoflavane, demonstrated strong interactions with PLpro residues, particularly involving the Y268 and Q269 residues of the BL2 loop. However, feruloyl tyramine exhibited fewer interactions, likely explaining its reduced efficacy against the SARS-CoV-2 PLpro enzyme. The Biological Significance of the Phytochemicals of (3R)–7,4'–Dihydroxy–8–Methoxyhomoisoflavane is flavonoid, isolated from *Dracaena cochinchinensis*, also known as dragon's blood (DB), has shown potential as an anti-osteoporosis agent [98], Bergenin is isocoumarin found in various plants like *Ardisia japonica, A. creanata, Bergenia crassifolia, B. purpurascens, Rodgersia sambucifolia* [99], bergenin exhibits a wide range of biological activities, including hepatoprotective effects [100], antifungal properties [101], anti-HIV activity [102], antiarrhythmic potential [103], hypolipidemic effects [104], and anticancer properties[105]. Capparispine belongs to the class of spermidine alkaloids; capparispine is known for its anti-inflammatory and antioxidant activities [106, 107] and feruloyl Tyramine is a phenylpropanoid compound isolated from various plants. Feruloyl tyramine exhibits several biological activities, including antimicrobial, antioxidant, anti-melanogenesis, and anticancer effects [108].

พหูนํ ปณฺฑิโต ชีเว

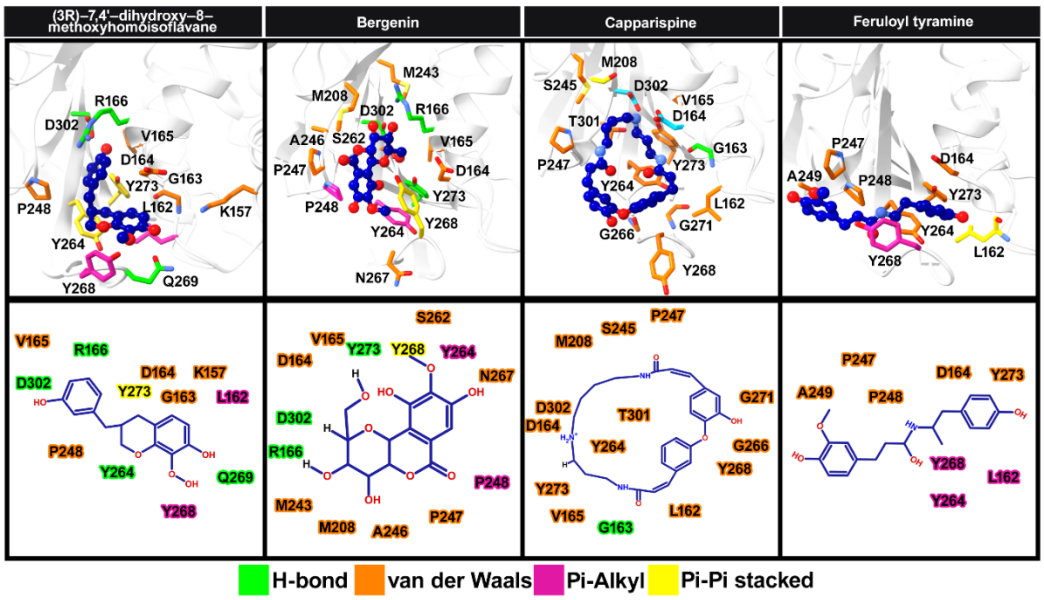

Figure 25 Representative 3D structures (Top) and 2D srtructures (Bottom) displaying protein-ligand interactions of each simulated system. Hydrogen bond forming between protein and ligand is displayed by green dot lines.

พหูน ปณฺฑิโต ชีเว

CHAPTER V CONCLUSIONS

In conclusion, Thailand is renowned for its rich biodiversity and abundant natural resources, particularly its extensive use of medicinal herbs in traditional medical treatments. The abundance of medicinal plants has significantly promoted research into the therapeutic application of plant extracts. Thai herbs have been recognized as a potent source of phytochemicals with the potential to treat viral diseases such as SARS-CoV-2.

The PLpro enzyme of SARS-CoV-2 plays a critical role in the viral replication process, making it a key target for developing treatments against SARS-CoV-2 or COVID-19. This study employed virtual screening techniques, including Lipinski's rule of five, ADMET prediction, molecular docking, and MD simulations, to investigate phytochemical compounds from Thai medicinal herbs. Of the 47 isolated phytochemical compounds, 27 adhered to Lipinski's rule of five criteria. These compounds were further analyzed through ADMET prediction and molecular docking. The four compounds with the Top15% highest docking scores were (3R)-7,4'-dihydroxy-8-methoxyhomoisoflavane, bergenin, capparispine, and feruloyl tyramine, which were selected for detailed MD simulations.

The findings revealed that, except for feruloyl tyramine, all other compounds established significant interactions with the PLpro enzyme via hydrogen bonds and hydrophobic interactions. SIE and MM/GBSA analyses confirmed that (3R)-7,4'-dihydroxy-8-methoxyhomoisoflavane, capparispine, and bergenin exhibited strong binding affinities to SARS-CoV-2 PLpro. Therefore, (3R)-7,4'-dihydroxy-8-methoxyhomoisoflavane from Dracaena cochinchinensis, bergenin from Ficus racemosa L., and capparispine from Capparis spinosa L. show potential as antiviral agents derived from Thai medicinal herbs, potentially inhibiting the PLpro enzyme of SARS-CoV-2. Although further validation is required, these compounds demonstrate promise as effective PLpro inhibitors for treating COVID-19.

In the future, the results of these simulations could lead to further laboratory experiments, such as toxicity studies and in vivo testing, to verify the bioactivity of these compounds. Additionally, advanced research into the structural properties of the compounds and enhancing their inhibitory efficiency could pave the way for developing more effective drugs. Integrating Thai herbal medicine with modern

COVID-19 treatments opens new opportunities for improving clinical therapy efficacy and advancing the fields of virology and drug development.

5.1 Future work

To confirm that the lead compounds can be potential inhibitors targeting PLpro of SARS-CoV-2, the following methods and procedures are suggested.

5.1.1 Protease Assay

The PLpro activity assay is a biochemical method used to measure the activity of the SARS-CoV-2 PLpro and evaluate the inhibitory effects of potential lead compounds. It involves preparing a purified PLpro enzyme and a specific substrate, typically a fluorogenic or chromogenic peptide that mimics the viral polyprotein cleavage site. The assay is set up by combining the enzyme, substrate, and varying concentrations of lead compounds in a reaction buffer, followed by incubation at an optimal temperature. The increase in fluorescence (or absorbance) is measured, indicating the extent of substrate cleavage, which allows for the quantification of PLpro activity and the determination of IC_{50} values for the inhibitors. Controls and replicates are essential for validating the results, making this assay critical in the drug discovery process targeting SARS-CoV-2.

5.1.2 Viral Replication Inhibition Assay

Viral replication inhibition assays are cell-based experiments designed to evaluate the efficacy of lead compounds against SARS-CoV-2 by measuring their ability to inhibit viral replication in infected cells. In this assay, permissive cell lines, such as Vero E6 or HEK293T, are inoculated with the virus and treated with varying concentrations of the test compounds. The viral load is assessed using techniques like qRT-PCR to quantify viral RNA, plaque assays to count infectious viral particles, or ELISA to detect viral proteins. This approach can determine the inhibitory potency of lead compounds and their potential therapeutic effects in a biologically relevant setting.

5.1.3 Cytotoxicity assays

Cytotoxicity assays are experimental techniques used to evaluate the toxic effects of lead compounds on host cells, essential for determining the safety profile of potential therapeutics. These assays measure cell viability and membrane integrity following treatment with varying concentrations of compounds. Common methods include MTT or XTT assays, which assess metabolic activity by detecting the reduction of tetrazolium salts to form a colored product in living cells, and LDH release assays that quantify lactate dehydrogenase released from damaged cells. By analyzing the extent of cytotoxicity, we can ensure that observed antiviral effects are not due to harmful impacts on cell health, aiding in the selection of safe and effective compounds for further development.

พหูนํ ปณฺฑิโต ชีเว

พหูนํ ปณฺฑิโต ชีเว

REFERENCES

พหูนํ ปณฺฑิโต ชีเว

1. Pandey, A. and S. Yadav, 20 - Essentials of COVID-19 and treatment approaches, in Data Science for COVID-19, U. Kose, et al., Editors. 2022, Academic Press. p. 397-422.
2. Rehman, S.U., et al., Evolutionary Trajectory for the Emergence of Novel Coronavirus SARS-CoV-2. Pathogens, 2020. 9(3).
3. data.who.int, W.C.-d. COVID-19 Global circulation. Available from: https://data.who.int/dashboards/covid19/data?n=c.
4. Naqvi, A.A.T., et al., Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2020. 1866(10): p. 165878.
5. Satarker, S. and M. Nampoothiri, Structural Proteins in Severe Acute Respiratory Syndrome Coronavirus-2. Arch Med Res, 2020. 51(6): p. 482-491.
6. Naydenova, K., et al., Structure of the SARS-CoV-2 RNA-dependent RNA polymerase in the presence of favipiravir-RTP. Proceedings of the National Academy of Sciences, 2021. 118(7): p. e2021946118.
7. Al-Tawfiq, J.A., A.H. Al-Homoud, and Z.A. Memish, Remdesivir as a possible therapeutic option for the COVID-19. Travel Med Infect Dis, 2020. 34: p. 101615.
8. Baildya, N., N.N. Ghosh, and A.P. Chattopadhyay, Inhibitory capacity of chloroquine against SARS-COV-2 by effective binding with angiotensin converting enzyme-2 receptor: An insight from molecular docking and MD-simulation studies. Journal of Molecular Structure, 2021. 1230: p. 129891.
9. Roy, A., et al., Encapsulated hydroxychloroquine and chloroquine into cyclic oligosaccharides are the potential therapeutics for COVID-19: insights from first-principles calculations. Journal of Molecular Structure, 2022. 1247: p. 131371.
10. Ng, T.I., et al. Antiviral Drug Discovery for the Treatment of COVID-19 Infections. Viruses, 2022. 14, DOI: 10.3390/v14050961.
11. Shin, D., et al., Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. Nature, 2020. 587(7835): p. 657-662.
12. Faheem, et al., Druggable targets of SARS-CoV-2 and treatment opportunities for COVID-19. Bioorg Chem, 2020. 104: p. 104269.
13. Mielech, A.M., et al., Nidovirus papain-like proteases: multifunctional enzymes with protease, deubiquitinating and deISGylating activities. Virus Res, 2014. 194: p. 184-90.
14. Abd El-Aziz, T.M. and J.D. Stockand, Recent progress and challenges in drug development against COVID-19 coronavirus (SARS-CoV-2) - an update on the status. Infection, Genetics and Evolution, 2020. 83: p. 104327.
15. Turabi, K.S., et al., Chapter 10 - Computational screening of phytochemicals for anti-bacterial drug discovery, in Phytochemistry, Computational Tools and Databases in Drug Discovery, C. Egbuna, M. Rudrapal, and H. Tijjani, Editors. 2023, Elsevier. p. 213-243.
16. Howes, M.-J.R., Chapter 28 - Phytochemicals as Anti-inflammatory Nutraceuticals and Phytopharmaceuticals, in Immunity and Inflammation in Health and Disease, S. Chatterjee, W. Jungraithmayr, and D. Bagchi, Editors. 2018, Academic Press. p. 363-388.

17. Adedokun, K.A., et al., Chapter 18 - Therapeutic potentials of medicinal plants and significance of computational tools in anti-cancer drug discovery, in Phytochemistry, Computational Tools and Databases in Drug Discovery, C. Egbuna, M. Rudrapal, and H. Tijjani, Editors. 2023, Elsevier. p. 393-455.
18. Prieto-Martínez, F., M. Arciniega, and J. Medina-Franco, Molecular docking: current advances and challenges. TIP Revista Especializada en Ciencias Químico-Biológicas, 2018. 21.
19. Jiyu Fan, A.F., Le Zhang, Progress in molecular docking. Quant. Biol., 2019. 7(2): p. 83-89.
20. Yang, C., E.A. Chen, and Y. Zhang, Protein–Ligand Docking in the Machine-Learning Era. Molecules, 2022. 27(14): p. 4568.
21. Huey, R. and G. Morris, Using AutoDock with AutoDockTools: A Tutorial. 2008.
22. Jones, J.E. and S. Chapman, On the determination of molecular fields. —II. From the equation of state of a gas. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 1924. 106(738): p. 463-477.
23. Born, M., Volumen und Hydratationswärme der Ionen. Zeitschrift für Physik, 1920. 1: p. 45-48.
24. Goodford, P.J., A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. Journal of Medicinal Chemistry, 1985. 28(7): p. 849-857.
25. Stouten, P.F.W., et al., An Effective Solvation Term Based on Atomic Occupancies for Use in Protein Simulations. Molecular Simulation, 1993. 10(2-6): p. 97-120.
26. Gilson, M.K., et al., The statistical-thermodynamic basis for computation of binding affinities: a critical review. Biophys J, 1997. 72(3): p. 1047-69.
27. Gohlke, H. and G. Klebe, Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. Angew Chem Int Ed Engl, 2002. 41(15): p. 2644-76.
28. Grigera, J.R., Molecular dynamics simulation for ligand-receptor studies. Carbohydrates interactions in aqueous solutions. Curr Pharm Des, 2002. 8(17): p. 1579-604.
29. Justino, G.C., C.P. Nascimento, and M.C. Justino, Molecular dynamics simulations and analysis for bioinformatics undergraduate students. Biochem Mol Biol Educ, 2021. 49(4): p. 570-582.
30. Hehre, W.J., A guide to molecular mechanics and quantum chemical calculations. 2003: Irvine, CA : Wavefunction, [2003] ©2003.
31. Badar, M., et al., Molecular Dynamics Simulations: Concept, Methods, and Applications. 2020.
32. Katiyar, R.S. and P.K. Jha, Molecular simulations in drug delivery: Opportunities and challenges. WIREs Computational Molecular Science, 2018. 8(4): p. e1358.
33. Yang, B., Mechanical properties of graphene-borophene heterostructures grain boundaries Molecular dynamic modeling. 2020.

34. Wang, Y.-T., et al., Predictions of Binding for Dopamine D2 Receptor Antagonists by the SIE Method. Journal of Chemical Information and Modeling, 2009. 49(10): p. 2369-2375.
35. Wang, Y., et al. Identification Mechanism of BACE1 on Inhibitors Probed by Using Multiple Separate Molecular Dynamics Simulations and Comparative Calculations of Binding Free Energies. Molecules, 2023. 28, DOI: 10.3390/molecules28124773.
36. Naïm, M., et al., Solvated interaction energy (SIE) for scoring protein-ligand binding affinities. 1. Exploring the parameter space. J Chem Inf Model, 2007. 47(1): p. 122-33.
37. Perdih, A., U. Bren, and T. Solmajer, Binding free energy calculations of N-sulphonyl-glutamic acid inhibitors of MurD ligase. Journal of molecular modeling, 2009. 15: p. 983-96.
38. Genheden, S. and U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opinion on Drug Discovery, 2015. 10(5): p. 449-461.
39. Giangreco, I., et al., Insights into the complex formed by matrix metalloproteinase-2 and alloxan inhibitors: molecular dynamics simulations and free energy calculations. PLoS One, 2011. 6(10): p. e25597.
40. Giangreco, I., et al., Insights into the Complex Formed by Matrix Metalloproteinase-2 and Alloxan Inhibitors: Molecular Dynamics Simulations and Free Energy Calculations. PloS one, 2011. 6: p. e25597.
41. Bergh, C., From Static Structures to Free Energy Landscapes: Characterizing Conformational Transitions in Biological Macromolecules, in TRITA-SCI-FOU. 2023, KTH Royal Institute of Technology. p. 257.
42. Anfinsen, C.B., Principles that govern the folding of protein chains. Science, 1973. 181(4096): p. 223-30.
43. Zuckerman, D., Statistical Physics of Biomolecules: An Introduction. 2010. 1-325.
44. Kumar, S., et al., THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. Journal of Computational Chemistry, 1992. 13.
45. Torrie, G.M. and J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. Journal of Computational Physics, 1977. 23(2): p. 187-199.
46. Kandeel, M., et al., Repurposing of FDA-approved antivirals, antibiotics, anthelmintics, antioxidants, and cell protectives against SARS-CoV-2 papain-like protease. Journal of Biomolecular Structure and Dynamics, 2021. 39(14): p. 5129-5136.
47. Ismail, M.I., et al., Targeting multiple conformations of SARS-CoV2 Papain-Like Protease for drug repositioning: An in-silico study. Computers in Biology and Medicine, 2021. 131: p. 104295.
48. Reza, R., et al., Repurposing of anti-lung cancer drugs as multi-target inhibitors of SARS-CoV-2 proteins: An insight from molecular docking and MD-simulation study. Microbial Pathogenesis, 2022. 169: p. 105615.

49. Patel, R., et al., Repurposing the antibacterial drugs for inhibition of SARS-CoV2-PLpro using molecular docking, MD simulation and binding energy calculation. Molecular Diversity, 2022. 26(4): p. 2189-2209.
50. Pang, J., et al., Discovery of small molecule PLpro inhibitor against COVID-19 using structure-based virtual screening, molecular dynamics simulation, and molecular mechanics/Generalized Born surface area (MM/GBSA) calculation. Structural Chemistry, 2021. 32(2): p. 879-886.
51. Bhowmick, S., et al., Identification of potent food constituents as SARS-CoV-2 papain-like protease modulators through advanced pharmacoinformatics approaches. Journal of Molecular Graphics and Modelling, 2022. 111: p. 108113.
52. Lakhera, S., et al., In Silico Investigation of Phytoconstituents of Medicinal Herb 'Piper Longum' Against SARS-CoV-2 by Molecular Docking and Molecular Dynamics Analysis. Results Chem, 2021: p. 100199.
53. Sanachai, K., et al., In Silico Elucidation of Potent Inhibitors and Rational Drug Design against SARS-CoV-2 Papain-like Protease. The Journal of Physical Chemistry B, 2021. 125(50): p. 13644-13656.
54. Baildya, N., et al., Screening of potential drug from Azadirachta Indica (Neem) extracts for SARS-CoV-2: An insight from molecular docking and MD-simulation studies. Journal of Molecular Structure, 2021. 1227: p. 129390.
55. Singh, E., et al., A computational essential dynamics approach to investigate structural influences of ligand binding on Papain like protease from SARS-CoV-2. Computational Biology and Chemistry, 2022. 99: p. 107721.
56. Jupudi, S., et al., Identification of Papain-Like Protease inhibitors of SARS CoV-2 through HTVS, Molecular docking, MMGBSA and Molecular dynamics approach. South African Journal of Botany, 2022. 151: p. 82-91.
57. Elseginy, S.A. and M.M. Anwar, In silico analysis of SARS-CoV-2 papain-like protease potential inhibitors. RSC Advances, 2021. 11(61): p. 38616-38631.
58. Thangavel, N. and M. Albratty, Benchmarked molecular docking integrated molecular dynamics stability analysis for prediction of SARS-CoV-2 papain-like protease inhibition by olive secoiridoids. Journal of King Saud University - Science, 2023. 35(1): p. 102402.
59. Selvaraj, V., et al., Polyphenolic Phytochemicals Exhibit Promising SARS-COV-2 Papain Like Protease (PLpro) Inhibition Validated through a Computational Approach. Polycyclic Aromatic Compounds, 2023. 43(6): p. 5545-5566.
60. Waqas, M., et al., Targeting papain-like protease by natural products as novel therapeutic potential SARS-CoV-2. International Journal of Biological Macromolecules, 2024. 258: p. 128812.
61. Gao, H., R. Dai, and R. Su, Computer-aided drug design for the pain-like protease (PLpro) inhibitors against SARS-CoV-2. Biomedicine & Pharmacotherapy, 2023. 159: p. 114247.
62. Thangavel, N. and M. Albratty, Pharmacophore model-aided virtual screening combined with comparative molecular docking and molecular dynamics for

identification of marine natural products as SARS-CoV-2 papain-like protease inhibitors. Arabian Journal of Chemistry, 2022. 15(12): p. 104334.

63. Lipinski, C.A., Lead- and drug-like compounds: the rule-of-five revolution. Drug Discov Today Technol, 2004. 1(4): p. 337-41.
64. Daina, A., O. Michielin, and V. Zoete, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports, 2017. 7(1): p. 42717.
65. Pires, D.E.V., T.L. Blundell, and D.B. Ascher, pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. Journal of Medicinal Chemistry, 2015. 58(9): p. 4066-4072.
66. Ratia, K., et al., A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. Proc Natl Acad Sci U S A, 2008. 105(42): p. 16119-24.
67. Báez-Santos, Y.M., et al., X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases. J Med Chem, 2014. 57(6): p. 2393-412.
68. Ratia, K., et al., Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme. Proc Natl Acad Sci U S A, 2006. 103(15): p. 5717-22.
69. Ghosh, A.K., et al., Severe acute respiratory syndrome coronavirus papain-like novel protease inhibitors: design, synthesis, protein-ligand X-ray structure and biological evaluation. J Med Chem, 2010. 53(13): p. 4968-79.
70. Gao, X., et al., Crystal structure of SARS-CoV-2 papain-like protease. Acta Pharm Sin B, 2021. 11(1): p. 237-245.
71. Osipiuk, J., et al., Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. Nature Communications, 2021. 12(1): p. 743.
72. Verdonk, M.L., et al., Improved protein-ligand docking using GOLD. Proteins, 2003. 52(4): p. 609-23.
73. Götz, A.W., et al., Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. J Chem Theory Comput, 2012. 8(5): p. 1542-1555.
74. Tian, C., et al., ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. Journal of Chemical Theory and Computation, 2020. 16(1): p. 528-552.
75. He, X., et al., A fast and high-quality charge model for the next generation general AMBER force field. The Journal of Chemical Physics, 2020. 153(11): p. 114502.
76. Berendsen, H.J.C., et al., Molecular dynamics with coupling to an external bath. The Journal of Chemical Physics, 1984. 81(8): p. 3684-3690.
77. Darden, T., D. York, and L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. The Journal of Chemical Physics, 1993. 98(12): p. 10089-10092.
78. Ryckaert, J.-P., G. Ciccotti, and H. Berendsen, Numerical-Integration of Cartesian Equations of Motion of a System with Constraints – Molecular-Dynamics of N-Alkanes. Journal of Computational Physics, 1977. 23: p. 327-341.

79. Lipinski, C.A., et al., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews, 1997. 23(1): p. 3-25.
80. Touhtouh, J., et al., Determination of the antioxidant and enzyme-inhibiting activities and evaluation of selected terpenes' ADMET properties: in vitro and in silico approaches. Biochemical Systematics and Ecology, 2023. 111: p. 104733.
81. Kalantzi, L., et al., Characterization of the human upper gastrointestinal contents under conditions simulating bioavailability/bioequivalence studies. Pharm Res, 2006. 23(1): p. 165-76.
82. Norinder, U. and C.A. Bergström, Prediction of ADMET Properties. ChemMedChem, 2006. 1(9): p. 920-37.
83. Zafar, F., et al., Physicochemical and Pharmacokinetic Analysis of Anacardic Acid Derivatives. ACS Omega, 2020. 5(11): p. 6021-6030.
84. Pratama, M.R.F., H. Poerwono, and S. Siswodiharjo, ADMET properties of novel 5-O-benzoylpinostrobin derivatives. J Basic Clin Physiol Pharmacol, 2019. 30(6).
85. Matondo, A., et al., Oleanolic Acid, Ursolic Acid and Apigenin from Ocimum basilicum as Potential Inhibitors of the SARS-CoV-2 Main Protease: A Molecular Docking Study. International Journal of Pathogen Research, 2021. 6(2): p. 1-16.
86. Vardhan, S. and S.K. Sahoo, In silico ADMET and molecular docking study on searching potential inhibitors from limonoids and triterpenoids for COVID-19. Comput Biol Med, 2020. 124: p. 103936.
87. Han, Y., et al., In silico ADME and Toxicity Prediction of Ceftazidime and Its Impurities. Front Pharmacol, 2019. 10: p. 434.
88. Daoui, O., et al., QSAR, molecular docking and ADMET properties in silico studies of novel 4,5,6,7-tetrahydrobenzo[D]-thiazol-2-Y1 derivatives derived from dimedone as potent anti-tumor agents through inhibition of C-Met receptor tyrosine kinase. Heliyon, 2021. 7(7): p. e07463.
89. Kinirons, M.T. and M.S. O'Mahony, Drug metabolism and ageing. Br J Clin Pharmacol, 2004. 57(5): p. 540-4.
90. Rodrigues-Junior, V.S., et al., Nonclinical evaluation of IQG-607, an anti-tuberculosis candidate with potential use in combination drug therapy. Regul Toxicol Pharmacol, 2020. 111: p. 104553.
91. Hajji, H., et al., In Silico Investigation on the Beneficial Effects of Medicinal Plants on Diabetes and Obesity: Molecular Docking, Molecular Dynamic Simulations, and ADMET Studies. Biointerface Research in Applied Chemistry, 2021. 11: p. 6933-6949.
92. Hajji, H., et al., Computational approach investigation bioactive molecules from Saussurea Costus plant as SARS-CoV-2 main protease inhibitors using reverse docking, molecular dynamics simulation, and pharmacokinetic ADMET parameters. Comput Biol Med, 2022. 150: p. 106209.
93. Edache, E.I., et al., Design of some potent non-toxic autoimmune disorder inhibitors based on 2D-QSAR, CoMFA, molecular docking, and molecular dynamics investigations. Intelligent Pharmacy, 2024.

94. Winiwarter, S., et al., In silico ADME in drug design – enhancing the impact. ADMET and DMPK, 2018. 6(1): p. 15-33.
95. Shen, Z., et al., Potent, Novel SARS-CoV-2 PLpro Inhibitors Block Viral Replication in Monkey and Human Cell Cultures. bioRxiv, 2021.
96. Osipiuk, J., et al., Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. Nat Commun, 2021. 12(1): p. 743.
97. Pitsillou, E., et al., Identification of Small Molecule Inhibitors of the Deubiquitinating Activity of the SARS-CoV-2 Papain-Like Protease: in silico Molecular Docking Studies and in vitro Enzymatic Activity Assay. Front Chem, 2020. 8: p. 623971.
98. Xu, X., et al., Isolation and chatacterization of homoisoflavonoids from Dracaena cochinchinensis and their osteogenic activities in mouse mesenchymal stem cells. Journal of Pharmaceutical and Biomedical Analysis, 2016. 129: p. 466-472.
99. Yan, R., et al., Unravelling and reconstructing the biosynthetic pathway of bergenin. Nat Commun, 2024. 15(1): p. 3539.
100. Lim, H.K., et al., Hepatoprotective effects of bergenin, a major constituent of Mallotus japonicus, on carbon tetrachloride-intoxicated rats. J Ethnopharmacol, 2000. 72(3): p. 469-74.
101. Prithiviraj, B., et al., Antifungal activity of bergenin, a constituent of Flueggea microcarpa. Plant Pathology, 2003. 46: p. 224-228.
102. Piacente, S., et al., Constituents of Ardisia japonica and their in vitro anti-HIV activity. J Nat Prod, 1996. 59(6): p. 565-9.
103. Pu, H.L., et al., Bergenin is the antiarrhythmic principle of Fluggea virosa. Planta Med, 2002. 68(4): p. 372-4.
104. Jahromi, M.A.F., J.P.N. Chansouria, and A.B. Ray, Hypolipidaemic Activity in Rats of Bergenin, the Major Constituent of Flueggea microcarpa. Phytotherapy Research, 1992. 6.
105. Shi, X., et al., Anticancer activity of bergenin against cervical cancer cells involves apoptosis, cell cycle arrest, inhibition of cell migration and the STAT3 signalling pathway. Exp Ther Med, 2019. 17(5): p. 3525-3529.
106. Maresca, M., et al., Acute effect of Capparis spinosa root extracts on rat articular pain. Journal of Ethnopharmacology, 2016. 193: p. 456-465.
107. Khatib, M., et al., An insight on the alkaloid content of Capparis spinosa L. root by HPLC-DAD-MS, MS/MS and 1H qNMR. Journal of Pharmaceutical and Biomedical Analysis, 2016. 123: p. 53-62.
108. Jiang, Y., L. Yu, and M.H. Wang, N-trans-feruloyltyramine inhibits LPS-induced NO and PGE2 production in RAW 264.7 macrophages: Involvement of AP-1 and MAP kinase signalling pathways. Chem Biol Interact, 2015. 235: p. 56-62.

BIOGRAPHY

NAME	Miss Chananya Rajchakom
DATE OF BIRTH	6 June 2000
PLACE OF BIRTH	Chiangrai
ADDRESS	93 Moo 8, Bankruat district, Buriram 31180
EDUCATION	2017 Bankruatwittayakarn School 2022 Bachelor of Science degree (B.Sc.) in chemistry Mahasarakham University 2024 Master of Science degree (M.Sc.) in chemistry Mahasarakham University
Research grants & awards	Center of Excellence for Innovation in Chemistry (PERCH-CIC)

พหูนํ ปณฺฑิโต ชีเว