Arab American University
Faculty of Graduate Studies
Department of Health Sciences
Master Program in Molecular and Cellular Biosciences

Molecular Mechanisms Controlling Insulin Signaling: Computational and Bioinformatic Approaches

Siwar Sati Abu Saiefen
202212123

Supervision Committee:
Prof. Hilal Zaid
Prof. Birgit Strodel
Dr. Siba Shanak

This Thesis Was Submitted in Partial Fulfillment of the Requirements for the Master Degree in Molecular and Cellular Biosciences.

Palestine, 7 / 2025

© Arab American University. All rights reserved.

Arab American University
Faculty of Graduate Studies
Department of Health Sciences
Master Program in Molecular and Cellular Biosciences

Thesis Approval

Molecular Mechanisms Controlling Insulin Signaling: Computational and Bioinformatic Approaches

Siwar Sati Abu Saiefen
202212123

This thesis was defended successfully on 31/07/2025 and approved by:

Thesis Committee Members:

Name	Title	Signature
1. Prof. Hilal Zaid	Main Supervisor	
2. Prof. Birgit Strodel	Members of Supervision Committee	
3. Dr. Siba Shanak	Members of Supervision Committee	Siba

Palestine, 7 / 2025

Declaration

I declare that, except where explicit reference is made to the contribution of others, this thesis is substantially my own work and has not been submitted for any other degree at the Arab American University or any other institution.

Student Name: Siwar Sati Abu Saiefen

Student ID: 202212123

Signature: Siwar Sati Abu Saiefen

Date of Submitting the Final Version of the Thesis: 12 / 8 / 2025

Acknowledgements

Always give thanks and praise to God. I want to thank those who were the lights and supporters in my journey, to my dear father, my first supporter and pride, who never failed to support me morally and financially, to my family and friends, who surrounded me with love and attention and encouraged me at every stage, to my supervisor Prof. Hilal Zaid, who opened wide horizons of knowledge for me, to whom I owe great thanks and gratitude, and to Prof. Birgit Strodel, for her careful supervision, patience, and continuous guidance during the preparation of this thesis.

To my dear university, Arab American University, which embraced me scientifically and academically. To the Research Center-Jülich (Forschungszentrum Jülich) for providing the ideal environment and the necessary support to complete this research.

I dedicate this work, in gratitude and thanks to everyone who contributed to its completion and supported me in my scientific journey.

Siwar Sati Abu Saiefen

Molecular Mechanisms Controlling Insulin Signaling: Computational and Bioinformatic Approaches

Siwar Sati Abu Saiefen
Prof. Hilal Zaid
Prof. Birgit Strodel,
Dr. Siba Shanak

Abstract

Background: Diabetes and insulin resistance are two major global public health issues, with a sever effect on the public welfare and general health. Novel targets for therapeutic interventions for these diseases can only be determined with better knowledge of the molecular mechanisms of insulin signaling. Phytochemicals, bioactive plant chemicals, have been found to alter insulin signaling pathways. The likelihood of these compounds interfering with critical inhibitory proteins in the insulin signaling pathway, namely Akt substrate of 160 kDa (AS160) and phosphatase and tensin homolog (PTEN) is an overly optimistic window of intervention. Computational and bioinformatics methods can shed light in understanding such kinds of molecular interaction.

Aim: This study aims to understand the molecular interactions between six selected phytochemicals and key proteins in insulin signaling pathway (AS160 and PTEN). The goal is to study the potential therapeutic targets for the selected phytochemicals in the insulin signaling pathway to overcome insulin resistance.

Methodology: A computational approach of molecular docking simulation and molecular dynamics (MD) simulations used to predict the binding affinity (binding free energy) and binding interaction of the selected six phytochemicals (3,4-dihydroxybenzoic acid, 4-methoxybenzoic acid, 3,4,5-trihydroxybenzoic acid, 2-hydroxy-4-methoxybenzoic acid, 4-hydroxybenzoic acid, and benzoic acid) with two key target proteins in insulin signaling pathway (AS160 and PTEN). Three-dimensional (3D) structures of the phytochemicals obtained from PubChem database. Crystal structures for PTEN and AS160 were retrieved from the Protein Data Bank with PDB IDs 1D5R and 3QYB, respectively. Docking using AutoDockFR (ADFR) software was done to identify the binding modes and binding affinities of protein-phytochemical complexes. For MD simulations GROMACS was used to assess the binding stability and flexibility of the protein-phytochemical complexes over the time course.

Results: Docking experiments revealed that the chosen phytochemicals exhibited greater binding affinities towards PTEN than AS160. Among the tested ligands, 3,4-dihydroxybenzoic acid and 2-hydroxy-4-methoxybenzoic acid both exhibited the highest affinities of -7.4 kcal/mol. MD simulations revealed that ligand-PTEN complexes were stable, especially with the above two phytochemicals, as their RMSD values were low. AS160-ligand complexes did not show stable interactions and possessed greater RMSD values. Hydrogen bonding analysis suggests that stability within the compound is due to methoxy and hydroxyl groups since stability in the complex declines dramatically in the absence of hydrogen bonding.

Conclusion: Molecular docking and MD simulation studies helped in understanding the binding interactions between AS160 and PTEN with the selected phytochemicals. The results revealed a strong and stable interaction between several of these ligands and PTEN. Conversely, the AS160-ligands complexes showed an unstable and weak interaction. These results indicated that further studies in PTEN as a therapeutic target and the promising phytochemicals as inhibitors to treat diabetes are promising and worthwhile. However, studies on AS160 and the current

phytochemicals need further editing and understanding options such as ligand improvement or alternative AS160 binding sites will be helpful to enhance the interactions.

Keywords: Diabetes, GROMACS, AutoDockFR, Molecular Docking, Molecular Dynamics.

Table of Contents

#	Title	Page

Declaration ... i

Acknowledgements ... ii

Abstract ... iii

List of Tables ... viii

List of Figures ... ix

List of Definitions of Abbreviations ... x

Chapter One: Introduction ... 1

1.1 Study Background ... 1

1.2 Study Objectives ... 2

1.3 Study Questions ... 2

1.4 Hypotheses ... 3

1.5 Study Limitations ... 3

Chapter Two: Literature Review ... 4

2.1 Diabetes ... 4

2.1.1 Introduction to Diabetes ... 4

2.1.2 Diabetes Classification ... 5

2.1.3 Normal Rang and Diagnosis ... 6

2.1.4 Treatments available ... 7

2.2 Insulin signaling pathway ... 8

2.2.1 Overview of the Insulin signaling pathway ... 8

2.2.2 Key components of Insulin signaling pathway ... 9

2.2.3 Glucose uptake mechanism ... 11

2.2.4 Dysregulation in Diabetes ... 12

2.2.5 Therapeutic targets in the Insulin signaling pathway ... 13

2.3 Medical plants and active compounds ... 14

2.3.1 Overview of Medicinal Plants in Diabetes Treatment ... 14

2.3.2 Antidiabetic phytochemicals 15
2.3.3 Challenges and Opportunities in drug discovery 15
2.4 Bioinformatics 16
2.4.1 Bioinformatics in Drug Discovery 17
2.4.2 Molecular Docking 17
2.4.3 Molecular Dynamics 18
2.4.4 Merits and Limitations 18
Chapter Three: Methodology 20
3.1 Molecular Docking 20
3.1.1 Selection and preparation of proteins 20
3.1.2 Selection and preparation of ligands 21
3.1.3 Molecular docking with AutoDockFR (ADFR) 21
3.1.4 Post-Docking Analysis 23
3.2 Molecular Dynamics (MD) simulation 23
3.2.1 System Setup 23
3.2.3 Simulation runs by GROMACS 24
3.2.4. Post-simulation Analysis 25
Chapter Four: Result 27
4.1 Molecular Docking 27
4.1.2 Binding Site Examination of PTEN and Ligands via PyMOL 32
4.2 Molecular Dynamics 36
4.2.1 MD of PTEN and AS160 in the Apo form 36
4.2.1.1 Simulation Setup 36
4.2.2 MD of PTEN with Phytochemicals 41
4.2.3 MD of AS160 with Phytochemicals 48
Chapter Five: Discussion 50
5.1 Molecular Docking 50
5.1.1 Results Overview 50
5.1.2 Interaction Patterns with AS160 51
5.1.3 Interaction Patterns with PTEN 51

5.2 Molecular Dynamic Simulation 51
5.2.1 Apo Proteins simulations 51
5.2.2 Insight into PTEN-phytochemical interactions 52
5.2.3 Insight into AS160-phytochemical interactions 52
5.2.4 Hydrogen bond analysis 53
5.3 Future Directions 53
5.4 Conclusion 53
References 54
ملخص 57

List of Tables

Table #	Title of Table	Page

Table 2.1: key components and their functions.10

Table 4.1: Binding energies of Phytochemicals to AS160 and PTEN. The binding energies of AS160 and PTEN are calculated by the respective docking scores.27

Table 4.2: PTEN Binding Site Residues with six Phytochemicals after MD run.46

List of Figures

Figure #	Title of Figure	Page

Figure 4.1: Binding sites of AS160 for specific phytochemicals.29

Figure 4.2: Ligand interaction binding sites of AS160:30

Figure 4.3: Binding Sites of selected phytochemicals with PTEN.33

Figure 4.4: Ligand Interaction Binding Sites of PTEN.34

Figure 4.5: Temperature fluctuations during the MD simulations of the apo forms of PTEN and AS160.37

Figure 4.6: Different forms of energy, namely potential, kinetic, and total, from the MD simulations of apo PTEN and AS160.38

Figure 4.7: PTEN's Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), and Radius of Gyration (Rg).40

Figure 4.8: AS160 Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), and Radius of Gyration (Rg).40

Figure 4.9: PTEN-Ligand Complexes RMSD Analysis.42

Figure 4.10: Hydrogen bond analysis in the absence of ligands carboxyl group (deprotonated state).43

Figure 4.11: Hydrogen bond analysis in the presence of a carboxyl group (Protonated).44

Figure 4.12: RMSD Analysis of AS160-Ligand Complexes.48

List of Definitions of Abbreviations

Abbreviations	Title
3D	Three-dimensional
ACE	Angiotensin-converting enzyme
ADA	American Diabetes Association
ADFR	AutoDockFR
AKT	Protein kinase B
ALA	Alanine
AMBER	Assisted Model Building with Energy Refinement
AMP	Adenosine monophosphate
AMPK	AMP-activated protein kinase
ARG	Arginine
ARBs	Angiotensin II receptor blockers
AS160	Akt substrate of 160 kDa
ASP	Aspartic acid
ATP	Adenosine triphosphate
BC	Before Christ
CaMK	Ca^{2+}/calmodulin-dependent protein kinase
CRISPR	Clustered regularly interspaced short palindromic repeats
CYS	Cysteine
DM	Diabetes mellitus
DPP-4	Dipeptidyl peptidase IV

FPG	Fasting plasma glucose
GDM	Gestational diabetes mellitus
GFR	Glomerular filtration rate
GLN	Glutamine
GLP1R	Glucagon-like peptide-1 receptor
GLUT4	Glucose transporter type 4
GLY	Glycine
GSK-3	Glycogen synthase kinase-3
HbA1c	Hemoglobin A1c
HIS	Histidine
IDF	International Diabetes Federation
ILE	Isoleucine
IR	Insulin receptors
IRS	Insulin receptor substrates
LEU	Leucine
LYS	Lysine
MAPK	Mitogen-activated protein kinase
MDA	Malondialdehydes
MD	Molecular dynamics
MODY	Maturity-onset diabetes of the young
mTOR	Mammalian target of rapamycin
NCBI	National Center for Biotechnology Information

NF-κB	Nuclear factor kappa B
NPT	Isothermal–isobaric ensemble
NVT	Canonical ensemble
OGTT	Oral glucose tolerance test
PDB	Protein Data Bank
PI3K	Phosphoinositide 3-kinase
PIP3	Phosphatidylinositol-3,4,5-trisphosphate
PTEN	Phosphatase and tensin homolog
PTP1B	Protein tyrosine phosphatase 1B
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
ROS	Reactive oxygen species
Rg	Radius of gyration
SD	Standard deviation
SGLT-2	Sodium-glucose cotransporter 2
TAC	Total antioxidant capacity
T1DM	Type 1 diabetes mellitus
T2DM	Type 2 diabetes mellitus
THR	Threonine

Chapter One: Introduction

1.1 Study Background

Type 2 diabetes mellitus (T2DM) is one of the most prevalent metabolic disorders that results from insulin resistance that impairs the ability of the body to metabolize glucose properly. It is now one of the most important health issues worldwide. According to projections of the International Diabetes Federation (IDF), the number of people with diabetes worldwide is expected to reach 643 million in 2030 and up to 783 million by 2045 (International Diabetes Federation, 2021). T2DM is characterized by prolonged hyperglycemia that can lead to complication-associated diseases such as cardiovascular diseases, neuropathy, nephropathy, and retinopathy (Shank et al., 2022). It can be managed effectively through a combination of lifestyle changes, pharmacological agents, and searching for alternatives to therapy.

The insulin signaling pathway plays a key role in the regulation of glucose metabolism. One of the major events in this signaling is the enhancement of 'glucose transporter type 4' (GLUT4) translocation from intracellular vesicles to the plasma membrane in the presence of insulin, which increases glucose uptake into the cells (Shanak et al., 2022). Key regulators of this event include Akt substrate of 160 kDa (AS160) and phosphatase and tensin homolog (PTEN). AS160 promotes GLUT4 translocation, while PTEN inhibits the pathway by dephosphorylating PIP3 to PIP2 and thus stopping insulin signaling (Zaid et al., 2008). Such antagonist roles open new therapeutic avenues for both PTEN and AS160 for potentiating glucose uptake and insulin sensitivity.

Natural products, especially from plant sources, have also been studied for possible antidiabetic agents. Plants bioactive compounds have been the most interesting subject of research especially as they possess few side effects, and have much therapeutic potential (Zaid et al., 2018). Indeed, tens of potential antidiabetic medicinal plant extracts were examined at Zaid`s lab in vitro and in vivo to assess their antidiabetic activity. Moreover, the chemical composition of the active extract was examined. Upon the potential antidiabetic phytochemicals, benzoic acid, 4-methoxybenzoic acid, 3,4,5-trihydroxybenzoic acid, 2-hydroxy-4-methoxybenzoic acid, 4-hydroxybenzoic acid, and 2,4-dihydroxybenzoic acid are the most promising. They were tested as pure phytochemicals *in vivo* and *in vitro*. They enhanced GLUT4 translocation in muscle cells and reduced blood sugar in diabetic mice (data not published). Therefore, the above-mentioned compounds were chosen to

be tested here for their potential interaction with key proteins in the insulin signaling pathway, AS160 and PTEN, *in silico.*

Today, the modern methods in computational drug discovery have increasingly proved helpful through the use of in silico techniques such as molecular docking and molecular dynamics (MD) simulation as tools to predict protein-ligand interaction outcomes, the basis for therapeutic cardinal importance in evaluating binding affinities, stability at the molecular level, and mechanisms of interaction that, in turn, may provide important insight into finding potential drug candidates able to treat T2DM.

1.2 Study Objectives

The study investigates the interaction between selected phytochemicals (benzoic acid, 4-methoxybenzoic acid, 3,4,5-trihydroxybenzoic acid, 2-hydroxy-4-methoxybenzoic acid, 4-hydroxybenzoic acid, and 2,4-dihydroxybenzoic acid) and two proteins (PTEN and AS160) using computational approaches. The specific objectives of the study include:

(a) Characterize molecular interactions between benzoic acid derivatives with target proteins PTEN and AS160.
(b) Perform molecular docking analysis to evaluate the binding affinities of these phytochemicals to specific target proteins.
(c) Perform molecular dynamics (MD) simulations to ascertain the stability and dynamic behavior of these interactions.
(d) Help develop new and better plant-based therapeutic agents T2DM.

1.3 Study Questions

This study aims to answer the following questions as follows:

(a) Which phytochemicals exhibit the strongest binding affinities with PTEN and AS160?
(b) What are the key residues, binding modes, and mechanisms involved in these protein-ligand interactions?

These answers could make it possible to advance the new plant-created medicines aimed at curing T2DM.

1.4 Hypotheses

(a) Some phytochemicals interact with PTEN and AS160, thereby influencing insulin signaling.

(b) Computer modeling of those interactions will help provide novel plant-based therapeutics for diabetes.

1.5 Study Limitations

The study is confined to a computational approach regarding the interaction between benzoic acid derivatives and the two proteins of the insulin signaling pathway, PTEN and AS160. The study used *in silico* methods, including molecular docking and MD simulations, which provide binding affinity and interaction dynamics predictions but not wet lab biological validation. Though PTEN and AS160 play significant roles in insulin signaling, there are several other proteins critically involved in the pathology of diabetes that are beyond the aim of this research. Some of these limitations are part of ongoing research projects beyond the scope of this thesis, which includes *in vitro* and *in vivo* testing of these phytochemicals.

Chapter Two: Literature Review

2.1 Diabetes

2.1.1 Introduction to Diabetes

Diabetes mellitus (DM) is defined as a chronic metabolic disease, characterized by prolonged hyperglycemia (elevated blood glucose levels). This disorder arises from a defect in insulin action, secretion, or both (Diagnosis and classification of diabetes mellitus 2010). Affecting millions of people around the world, DM's high rate of morbidity and mortality is considered a significant health challenge and one of the most common hormonal-related disorders (Ibrahim, 2019). Insulin hormone is one of the key players in the glucose regulation mechanism in the body; it has a critical role in glucose hemostasis. Insulin is produced by beta islets of Langerhans in the pancreas, and it works synergically with glucagon to maintain normal glucose levels in the blood. Diabetes leads to an imbalance in insulin work by impairing the body's ability to secrete insulin (insulin absent), which is caused by an autoimmune disease that destroys pancreatic beta cells, or impaired insulin action (insulin resistance). As a result, the blood glucose levels will be elevated (hyperglycemia).

Several factors could be considered risk factors for developing Diabetes like genetic predisposition, obesity, physical inactivity, age, dietary habits, ethnicity, history of gestational diabetes, stress, and mental health (International Diabetes Federation, 2021). Uncontrolled diabetes could lead to various complications, that could be subdivided into microvascular and macrovascular complications. Microvascular complications include retinopathy (damage in retina blood vessels that lead to vision illnesses and blindness), nephropathy (kidney damage), and neuropathy (nerve damage). Macrovascular complications include cardiovascular diseases (heart diseases), diabetic foot ulcers, and infection (World Health Organization, 2016).

DM has many subclasses, including maternity-onset diabetes of young people (MODY), neonatal diabetes, type 1 diabetes (T1DM), type 2 diabetes (T2DM), gestational diabetes, and secondary cases (Yameny, 2024). The major subclasses of DM are type 1 diabetes, which results from an autoimmune disorder that attacks pancreatic cells which affects insulin secretion, and type 2 diabetes which is caused by insulin resistance (problem in insulin action).

Diabetes mellitus, derived from the Greek word "diabetes" meaning siphon, and the Latin word “mellitus” meaning sweet, was coined between 250 and 300 BC (Sapra & Bhandari, 2023).

Ancient civilizations found the sweet urine in diabetes hence the name "Diabetes Mellitus." Insulin was finally purified from cows' pancreas in the year 1922, enabling treatment with great effectiveness. The disease still prevails commonly in its chronic form worldwide, with the US having the seventh place in the death toll. As reported by the International Diabetes Federation (IDF), in 2021 there were ~ 537 million adult people who have diabetes which represents 10% of the adult population. This number is predicted to increase to 643 million by 2030 and 783 million in 2024 if the current situation continues (International Diabetes Federation, 2021).

2.1.2 Diabetes Classification

According to the World Health Organization different forms of diabetes exist with heterogeneous etiologies and clinical presentations (Reuters, 2024). Indeed, DM can be classified into 4 major types:

(a) **Type 1 Diabetes mellites (T1DM)** or also as was known insulin-dependent diabetes, is considered an autoimmune condition where the body's T-cells start to destroy the beta cells of Langerhans in the pancreas (insulin producer cells), this destruction leads to absent or extreme lowering insulin level in blood. According to the International Diabetes Federation, 2021 ~5-10% of diabetes cases belong to T1DM. This type of DM is usually diagnosed in children and adolescents (International Diabetes Federation, 2021); hence it is called juvenile-onset diabetes or young people's disease. Although it could be diagnosed at any age. People diagnosed with T1DM need insulin therapy throughout their lives and a lifelong diet.

(b) **Type 2 Diabetes mellites (T2DM)** is also known as non-insulin-dependent or adult-onset diabetes. In 2021, the International Diabetes Federation estimated that ~90-95% of diabetes cases belong to T2DM, usually found in middle-age and older adults have a chronic elevation in blood glucose levels and have an unhealthy lifestyle (International Diabetes Federation, 2021). T2DM is characterized by insulin resistance, unlike a normal or above-normal production of insulin. After prolonging period, the possibilities that the pancreas loses its competence to produce sufficient insulin for the body's needs is increases. This can happen due to aging, obesity, a sedentary life, and other unhealthy lifestyle habits related to poor eating.

(c) **Gestational Diabetes mellites (GDM)** develop during pregnancy usually in the second or third trimester, which normally disappears after giving birth (Sapra, Vaqar, & Bhandari, 2019). However, those women who had gestational diabetes are at higher risk of eventually getting type 2 diabetes. Moreover, there are many other risks of this disease the mother may carry, and for the child within her womb childbirth complications, and even probable health problems related to that child. To screen for their risk of future diabetes, women with gestational diabetes should follow up on their health during pregnancy and after delivery (American Diabetes Association, 2023).

1) **Other specific types** besides the common types such as Type 1 and Type 2, there exist a few more rare forms of diabetes that need to be considered. One such form is Maturity-Onset Diabetes of the Young, MODY. Unlike typical Type 2 diabetes that presents later in life, this is a genetic disorder affecting mostly insulin secretion, at a younger age group (Yameny, 2024). There is secondary diabetes, which is treated separately and is caused by hormonal imbalances or other medical conditions. For example, pancreatitis and Cushing's syndrome are disorders related to diabetes because they affect insulin production and regulation (Thayer et al., 2020).

2.1.3 Normal Rang and Diagnosis

Several laboratory tests are commonly used to diagnose diabetes, including fasting plasma glucose (FPG), oral glucose tolerant test (OGTT), hemoglobin A1c (HbA1c), random plasma glucose, and fructosamine. These tests evaluate the glucose levels in blood. A newer diagnostic approach involves a cytological method based on the qualitative analysis of oral epithelial cells. This method is founded on the theory that blood glucose levels influence the number of epithelial cells in the oral cavity. However, it still requires confirmation and support from standard laboratory tests to be considered reliable (Aneed et al., 2024). Another emerging approach involves measuring oxidative stress biomarkers such as total antioxidant capacity (TAC) and malondialdehyde (MDA). These are important indicators of cellular damage and oxidative stress and are increasingly being studied in the context of diabetes diagnosis (Pawar, Sontakke, & Pawar, 2023).

For early diagnosis, the best choice is fasting plasma glucose level and HbA1c, HbA1c represents the average of blood glucose levels for the past 2-3 months (Yameny, 2024). The cut-off value for identifying diabetes is 126 mg/dl or higher for FPG, 200 mg/dl or higher for random glucose level,

and HbA1c should be greater than 5.6% in this case the patient diagnoses DM. If the levels of FPG and HbA1c are at the borderline (gray zone) the test of choice to conform DM is the oral glucose tolerant test OGTT, this test requires fastening (no food or drinks intake for at least 8 hours, except the water) and starts when the patient drinks a glucose solution (glucose-rich drink) that contain 75g of glucose, then the glucose levels in the blood will be measure after 1 and 2 hours, a glucose level 200mg/dl or higher indicates DM. This test is also used to diagnose Gestational diabetes in pregnant women, it is done as a screening test for women at 25 weeks' gestation (Thayer et al., 2020). Patients with FPG levels between 100-125 mg/dl or OGTT of 140-200 mg/dl are identified as prediabetic, which might precede T2DM in the future.

Elevated blood sugar levels, such as HbA1c levels of 6.5% or higher, fasting plasma glucose levels of 126 mg/dL or higher, two-hour plasma glucose levels of 200 mg/dL or higher during a 75-gram oral glucose tolerance test, and random plasma glucose levels of 200 mg/dL or higher are the characteristics of diabetes, according to the American Diabetes Association (ADA). The United States Preventative Services Task Force recommends screening overweight adults between the ages of 40-70 years, whereas American Diabetes Association recommends screening all adults above the age of 45 years regardless of risk factors (Selph et al., 2015).

2.1.4 Treatments available

Until now DM has had no cure. Managing DM requires a lifelong complex treatment that includes: first, controlling lifestyle by losing weight for obese or overweight diabetic people, having an active life by doing regular exercises for at least 3 hours weekly, and have a balanced diet (diabetes-friendly diet) which mainly restrict overall calories intake and reduce carbohydrate intake and replace with vegetables, fruits, and whole grains.

Second, medication intake helps in controlling insulin levels in the blood. Insulin injections are essential in the case of T1DM because, no internal insulin (insulin is extremely low or absent), and insulin injections are also required to manage T2DM’s last stages due to beta cells in the pancreas will lose their function in the last stages. In T2DM use oral medication such as biguanides like metformin, sulfonylureas, meglitinides, alpha-glucosidase inhibitors, thiazolidinediones, glucagon-like peptide-1 agonists, inhibitors of dipeptidyl peptidase IV (DPP-4), selective amylinomimetics, and sodium-glucose transporter-2 (SGLT-2) alongside lifestyle modification are particularly important in managing glucose levels (Knowler et al., 2002). Metformin is the first

choice and preferred drug that doctors prescribe in T2DM management and works in lowering baseline and after-meal (postprandial) plasma glucose.

Finally, monitoring and evaluating glucose levels, is essential to carry out regular laboratory tests every six months, consisting of HbA1c and FPG tests. Other tests are available and can also help in monitoring general health because diabetes can cause so many complications. Tests like diabetic retinal examination and neurological tests help diabetes management. Moreover, healthcare providers may recommend that patients inspect their feet daily for lesions that neuropathy may render them oblivious to. It may be a case of prescribing low-dose tricyclic antidepressants, duloxetine, anticonvulsants, topical capsaicin, and analgesics for diabetic patients in the treatment of neuropathic pain. Urine microalbumin testing can measure glomerular filtration rate (GFR) and diagnose early kidney changes due to diabetes when albuminuria is more than 30 mg/g creatinine. As the first-line treatment to delay progression from microalbuminuria to macroalbuminuria in both Type 1 and Type 2 diabetes mellitus, ACE inhibitors, and ARBs are given due to their antiproteinuric effects.

Monitoring blood pressure is also critical as the ADA recommends Diabetic people's blood pressure must be within 130 mmHg systolic and 85 mmHg diastolic, regular testing is important and recommended (de Boer et al., 2017). Balance in diabetic management is important. However, over management could lead to hypoglycemia (low blood glucose level), which could be fatal in the case of a diabetic patient.

2.2 Insulin signaling pathway

2.2.1 Overview of the Insulin signaling pathway

The insulin signaling pathway plays a critical role in maintaining glucose levels and general energy balance. When insulin, the peptide hormone secreted by β-cells of the pancreas, binds to insulin receptors (IR) on the cell surfaces will initiate a metabolic response in its target cells (muscle, liver, and adipose tissue). After binding to the IR, a cascade of intracellular signaling events is initiated

by activating insulin receptor substrates (IRS), phosphoinositide 3-kinase (PI3K), and protein kinase B (Akt), which will increase glucose uptake and glycogen synthesis and inhibit fat metabolism. These events would define the cellular responses to insulin. In consequence, it must consider a tightly regulated cascade of events that allow glucose homeostasis to work with some metabolic flexibility, building variations in nutrient availability (Shanak et al., 2022).

Insulin signaling irregularities are now considered a hallmark of metabolic diseases, ranging from the induction of insulin resistance to type 2 diabetes mellitus. The normal metabolic condition allows insulin signaling to show a positive metabolic status for glucose storage and energy conservation. Other confounding factors such as a predisposing factor due to genetics, obesity, and chronic inflammation could be added on thereby affecting glucose intake to exacerbate hyperglycemia and increase levels of blood glucose. The available biochemical mechanisms by which insulin act takes greater relevance for the considerations that have a place in the pathophysiology connected with disorders in metabolism hence providing a rationale for the identifications of the targets for drug discovery to restore metabolism homeostasis (Shanak et al., 2022).

2.2.2 Key components of Insulin signaling pathway

Insulin signaling is an intracellular signaling pathway present in this complicated arrangement of proteins interacting within certain agreements to maintain glucose homeostasis and metabolic balance. Activation of insulin receptor (IR) via binding insulin hormone or a member of the insulin-like group leads to subsequent activation of various signaling cascades in receptor tyrosine kinases. IR has been implicated in insulin autophosphorylation on certain tyrosine residues such as TYR-960 because upon binding of insulin to IR conformational changes are observed in the receptor molecule. After phosphorylation and activation of the receptor, other proteins such as PI3K and IRSs are docked into the receptor. The IRS proteins can now recruit PI3K, which in turn synthesizes the second messenger phosphoinositide (3,4,5)-trisphosphate (PIP3), which activates the master regulator Akt (Petersen & Shulman, 2018).

Akt plays a vital role in many cellular functions, these functions involve lipid storage, glucose metabolism, and proliferation in insulin response. After activation, Akt utilizes phosphorylation by phosphorylating numerous downstream targets such as AS160. That will eventually stimulate the translocation of vesicles containing glucose transporter type 4 (GLUT4) to the cell membrane,

facilitating glucose uptake by the cells. In addition, Akt inhibits glycogen synthase kinase-3 (GSK-3), therefore promoting glycogen synthase activity and glycogen synthesis. Indeed, one of the main functions of Akt is the facilitation of cell survival through the inhibition of pro-apoptotic factors and/or the stimulation of the mammalian target of rapamycin (mTOR), an important growth-regulating pathway controlling lipid and protein synthesis. This perspective highlights the capacities of insulin signaling in determining the growth, proliferation, and differentiation of cells against the background of R as/ mitogen-activated protein kinase (MAPK) signaling pathways (Akhtar & Sah, 2020).

These proteins interact with other proteins that have been identified and create a variety of functional systems to regulate insulin signaling precisely and instantly. Proteins are under constant surveillance to ensure they are quickly inactivated upon entering pathways that precede insulin signaling, so they do not significantly alter the insulin responses. As characteristic negative modulators, protein tyrosine phosphatase 1B (PTP1B) and phosphatase and tensin homology (PTEN) inhibit signaling through the removal of phosphates (dephosphorylation) from key intermediates. Grb2 and Shc are adapter proteins that facilitate insulin signaling to a broader spectrum of cellular functions by linking a multiplicity of signals. The metabolic response to insulin is further amplified by crosstalk of the insulin signaling pathway with numerous others signaling pathways resulting in modulation and enhancements in glucose utilization (Latva-Rasku et al., 2017). Disruption of any of these critical components, through genetic mutations, chronic inflammation, or environmental triggers, will impair the insulin signaling cascade, leading to insulin resistance and eventually to metabolic disorders like type 2 diabetes.

Knowledge of these molecular components and their interactions has provided therapeutic targets to improve insulin sensitivity in restoring metabolic balance in such individuals. **Table 1** summarizes the key components and their functions.

Table 2.1: key components and their functions.

Component	Function
Insulin Receptor (IR)	Initiates signaling by binding insulin
Insulin Receptor Substrates (IRS)	Links IR to downstream pathways
Phosphoinositide 3-Kinase (PI3K)	Generates PIP3, a key second messenger

Protein Kinase B (Akt)	Mediates glucose uptake and metabolic effects
Glucose Transporter Type 4 (GLUT4)	Facilitates glucose entry into cells
Negative Regulators (PTEN, PTP1B)	Terminate or downregulate signaling
mTOR Pathway	Supports protein and lipid synthesis
Ras/MAPK Pathway	Regulates growth and differentiation

2.2.3 Glucose uptake mechanism

Glucose uptake from plasma by muscle and fat cells is a vital step. The most important glucose transporter is GLUT4, which is translocated to membrane surfaces by different pathways that are transducing the signals of both insulin and exercise in most mammals. Initiation of the process occurs once the actual hormone binds to a receptor on the cell surface to cascade numberless signals inside the cell. The binding of insulin to its receptor activates a series of intracellular signaling events, phosphorylating its receptors. The insulin signaling pathway starts with phosphoinositide-3-kinase (PI3K) producing phosphatidylinositol-3,4,5-bisphosphate (PIP3). PIP3 recruits and activates Akt, protein kinase B, a key player in the AS160 signaling pathway. Inactive AS160 prevents GLUT4 vesicles through inhibition of the Rab proteins involved in transport, thus inhibiting the glucose intake into a cell. By its phosphorylation, AS160 activity is toned down; this, in effect, allows active Rab proteins to move GLUT4 vesicles onto the plasma membrane (van Gerwen, Shun-Shion, & Fazakerley, 2023).

Not only the insulin, muscle contraction during exercise increases the uptake of glucose also. Exercise activates the important cellular energy sensor AMP-activated protein kinase (AMPK). It is high in AMP/ATP ratio during contraction that activates AMPK, leading to TBC1D1 phosphorylation, a highly related protein to AS160, which mediates the translocation of GLUT4 independent of insulin. Calcium signaling increases the translocation of GLUT4 to the membrane during muscle contraction and activates pathways like Ca2+/calmodulin-dependent protein kinase (CaMK). Exercise-induced reactive oxygen species (ROS) activate AMPK and mitogen-activated protein kinase (MAPK or MAP kinase) activation, increasing cellular glucose availability and finally promoting glucose uptake (Richter & Hargreaves, 2013). Such dual regulation, mediated by insulin for postprandial glucose uptake and by exercise for energy demands, ensures tight control over glucose homeostasis.

The regulation of GLUT4 includes not just its movement to the cell surface but also its uptake and recycling within the cell. Upon glucose level decline, GLUT4 internalizes through clathrin-mediated vesicles and is stored intracellularly, this provides another balance mechanism for intracellular exocytosis and endocytosis. Dysregulation of these processes or impairment in PI3K-Akt or AMPK pathways, will develop insulin resistance and associated metabolic disorders, such as type 2 diabetes. For instance, AMPK deficiency during insulin resistance results in the downregulation of exercise-induced glucose uptake while blockade of AS160 phosphorylation suppresses GLUT4 translocation (Sakamoto & Holman, 2008). Thus, interventions targeting AS160, TBC1D1, and pathways leading to GLUT4 translocation at the stage of insulin resistance are currently targeted at attempting to increase glucose uptake.

Understanding the transport of glucose into the cells is essential for understanding metabolic disorders. Therefore, it shows how the body regulates insulin signaling, anthracycline-related signaling pathways, and GLUT4 by depending on the body glucose needs. This information not only sheds light on energy regulation but also gives a guideline of promise therapeutic of the diseases such as diabetes and obesity.

2.2.4 Dysregulation in Diabetes

Disturbance in glucose homeostasis resulting from impaired insulin secretion and/or insulin resistance defines diabetes in the long run, with specific reference to T2D. Muscle and adipose tissues are the insulin-responsive tissues that mediate glucose disposal during insulin stimulation and are logically affected owing to a form of disequilibrium in the insulin signaling pathway, hence causing insulin resistance. One of the most prominent abnormalities noted in type 2 diabetes is impaired activation of the IRS and its downstream signaling molecules. The overall effect is a decrease in glucose uptake and GLUT4 translocation to the plasma membrane through decreased activation of Akt and PI3K. These two crucial components of the pathway are chronic hyperglycemia or its secondary effects and cause inflammation and oxidative stress. These induce serine phosphorylation of IRS proteins leading to insulation of the effectiveness of signaling by insulin (Ramasubbu & Devi Rajeswari, 2023).

Another major area is the excessive activation of protein tyrosine phosphatases such as PTP1B, which inhibit insulin receptor signal effectors like IRS through their actions. Rab GTPases and their regulators, such as AS160 and TBC1D1, have been implicated in the phenomenon of insulin

resistance with reduced GLUT4 translocation. AMP-activated protein kinase (AMPK)-mediated pathways activated during muscle contraction are preserved in diabetes and could act as a compensatory mechanism for glucose transport. However, continuing insulin stress may mediate this compensatory response, further exacerbating metabolic damage (Frontiers in Physiology, 2022).

2.2.5 Therapeutic targets in the Insulin signaling pathway

Due to the many effects of insulin signaling, diabetes therapies to enhance glucose uptake target multiple sites in this pathway. PI3K-Akt signaling pathway, which is important for insulin-mediated glucose uptake. Compounds that directly or indirectly regulate Akt, particularly those focused on elevating Akt activity, have emerged. PTEN can also be a target within the signaling pathway since it negatively regulates the action of Akt by dephosphorylating PIP3 (Seong, 2024). PTEN overexpression figures as a member of the mapped signal-transduction pathway regulating insulin resistance and pharmacologic inhibition of this phosphatase are currently tested as a potential opportunity to improve insulin signaling and sugar absorption.

Metformin, a drug formerly used to treat patients with type 2 diabetes, activates AMPK, the latter acts by inducing GLUT4 translocation into the plasma membrane and hence stimulating glucose uptake, independent of insulin. Other proposed molecular targets in the newest therapeutic approaches are AS160 or rab GTPase TBC1D1, the latter being the principal regulator of GLUT4 trafficking (Zaid et al., 2008). Enhancement of such phosphorylation results in restoration of GLUT4 translocation even under insulin-resistant circumstances. There is much promising research into small molecules that will mimic AS160 phosphorylation or Rab GTPases activation.

Inflammation and oxidative stress, important contributory factors in insulin resistance, also offer therapeutic opportunities. Anti-inflammatory agents, which inhibit nuclear factor kappa B (NF-κB) signaling or employ scavenger methods of ROS, are under clinical investigation. Such agents seem promising in ameliorating systemic and local inflamed conditions. Besides their well-defined main mechanisms directed against incretin or sodium-glucose transporter 2 pathways of renal glucose reabsorption, GLP-1 receptor agonists and SGLT2 inhibitors, however, also carry secondary health benefits in systemic inflammation reduction and insulin sensitivity enhancement.

Furthermore, CRISPR-Cas9, and other gene-modifying technologies, are now being examined to fix genetic defects in primary signaling components such as PTEN. Biologics also represents another novel approach to personalized therapies in metabolic disorders aimed at inhibiting adipokines, particularly adiponectin, and leptin, restoring systemic metabolic balance and improving insulin sensitivity. Lastly, lifestyle modifications, exercises, and dietary changes are considered wonderful therapies for their roles in enhancing AMPK activation and GLUT4 translocation, adding to pharmacological therapies (Li, Ma, & Hao, 2024).

2.3 Medical plants and active compounds

2.3.1 Overview of Medicinal Plants in Diabetes Treatment

Medicinal plants are utilized in traditional medicines all over the world for many conditions including diabetes mellitus, and as their inherent value was more elaborately described, further potential uses have served to maintain it as a valued choice in treating diabetes. More than 80% of people get their primary health care through traditional medicine, according to WHO, especially in the developing world (Tran, Pham, & Le, 2020). These plants are thought to be a much richer source of bioactive compounds with therapeutic properties and low toxicity in comparison to synthetic drugs. With over 400 plant species documented to exert hypoglycemic activity, plant-based treatments for the disease are well established the efficacy of these plants-extracted, decocted, or powdered-has been validated through both tradition and research.

Bioactive agents in charge of medicines are a complex group of phytochemicals, consisting of flavonoids, alkaloids, terpenoids, and saponins. These molecules encompass a wide variety of biological features like hypoglycemic, anti-inflammatory, and antioxidant activities (Anshika et al., 2022).

Many bioactive compounds invoked reductions in blood glucose concentration through different mechanisms: enhancing the uptake of glucose in muscle and adipose tissues, delaying glucose absorption in the intestines, and enhancing insulin sensitivity and secretion (Patel et al., 2012). Furthermore, they also show some abilities to ameliorate oxidative stress and inflammation, which are the foremost pathophysiology-related events of diabetes. Notably, flavonoids have led to diverse signal pathways modulation for glucose metabolism. The complementary interaction of

these avenues implies the usefulness of phytochemicals for adjunctive therapy in diabetic pathology (Hanhineva et al., 2010).

2.3.2 Antidiabetic phytochemicals

Phytochemicals of medicinal plants with antidiabetic activity are used in diabetes management. Flavonoids, phenolics, alkaloids, and terpenoids are a few natural compounds that influence blood glucose levels by virtue of multiple mechanisms.

Coffee and some fruits are helpful in reducing glucose levels, due to the presence of chlorogenic acid, so inhibiting postprandial rise in blood glucose. Another example is bitter melon, a fruit utilized by numerous various cultures, containing Charantin acting for effective control of the blood sugar level as an insulin substitute. Also, Berberine, a compound in plants such as Berberis vulgaris, inhibits the production of glucose and acts to increase receptiveness to insulin. Curcumin is a yellow pigment present in turmeric, and aside from flavoring, it inhibits inflammatory activities and preserves the insulin-secreting cells of the pancreas. The other which cannot be disregarded is ginseng, particularly Panax, rich in ginsenosides, which enhance insulin action and glucose metabolic utilization (Zaid et al., 2019).

These phytochemicals will be capable of controlling glucose levels in the body and acting as antioxidants and anti-inflammatory substances; thus, inclusion of these in overall wellness is an added benefit. These phytochemicals are researched every now and then and are recommended-safe formulations and alternatives for diabetes management, as an adjunct to a person's journey towards good health. It is comforting that nature, after all, has some facets that can contribute to the life of a diabetic patient.

2.3.3 Challenges and Opportunities in drug discovery

Although the medical plant field has many promising potentials in diabetes treatment, it is still facing several challenges in the drug discovery process. One of the main problems faced is the standardization of the extraction from the plants, as the concentration of the phytochemicals can differ from one plant to another due to growing conditions, plant species, and the method of extraction (Zaid et al., 2008). Also, one plant can contain a complex of phytochemicals not just one and that will raise another problem which is the isolation and identification of the wanted phytochemical. Another problem is the absence of well-designed and thoroughly conducted

studies that involve humans studying the safety and efficiency of bioactive compounds. Even if the results that were obtained from the animal model are promising and demonstrate the antidiabetic effect of the phytochemical, a clinical trial on humans is still needed to ensure the effect and the side effects of the phytochemical. Furthermore, the interaction of the herb with another drug intake should be conceded, and studying phytochemical-drug interaction is needed.

Another bright possibility for the medicinal plant industry does exist, despite these obstacles. Phytochemistry and bioinformatics are moving toward better facilitation of the research of plant chemicals, and methods have been developed to separate and identify bioactive compounds with ease. Also, the rising interest in natural products helps bring financial support and researcher's interest in the medicinal plant field. Drug delivery systems like nanoparticle development enhance bioavailability and efficacy regarding these bioactive compounds.

Consequently, medicinal plants hold hope since several bioactive substances possess the capability against diabetes. The other side of the coin is that phytomedicine may serve as the basis for drug synthesis.

2.4 Bioinformatics

Bioinformatics is an interdisciplinary area allowing the analysis and interpretation of biological data using knowledge and methods from biology, computer sciences, and mathematics. It is one of the areas where this discipline has already found its footing with respect to drug discovery to become one of the important tools for making predictions in drug efficacy and toxicity, deducing molecular interactions, and identifying potential targets for drug discovery. Under its application, the researchers, by applying computational tools, can curtail the timeline of developing a drug and cut down on the cost of experimental research. Bioinformatics offers methods to effectively structure, analyze, and extract meaningful insights from the vast amounts of biological data emerging in the wake of biomedical questions such as those surrounding the multitopic big data from the perceived cellular functions of proteomics and next-generation sequencing. Bioinformatics modeling of structure-function relationships in complex biological systems allows for greater accuracy in a traditional drug discovery process (Anshika et al., 2022).

2.4.1 Bioinformatics in Drug Discovery

Various high-throughput genomics and proteomic datasets are processed to assist researchers in identifying novel potential targets for chemotherapeutic development. Virtual screening refers to the ability to evaluate multiple high-throughput libraries simultaneously in the search for promising drug candidates, often in concert with structure-based drug design where the design of the inhibitor is predicated upon geometrical considerations of how the three-dimensional structure of the molecules fits together. Computational technologies such as homology modeling and molecular docking allow for providing predictions of small molecules with macromolecular interactions. A mixture of various forms of data such as clinical data, gene expression profiles, and protein-protein interactions render it an efficient drug discovery pipeline. In early-phase drug development, machine learning algorithms and AI techniques can p redict the likelihood of toxicity and thus have been introduced to boost pharmacodynamics (Malathi & Ramaiah, 2018).

2.4.2 Molecular Docking

Molecular docking is a procedure whereby computer-generated binding modes of the drug candidate and its biological target are simulated. The process consists of several steps, which undergo tight coupling to prepare the ligand out of small molecules in a suitable form in theoretical studies of interaction and the target structure's preparation, based on either the actual crystallographic data or derived modeled protein structures. The flow of the operation is to evaluate the binding energies of different ligand-receptor pairs and poses of the most probable binding conformations using a scoring function in the process of docking and hence predicting the interaction of the ligand with the receptor.

Docking is a key principle of structure-based drug design of the key and precious methodologies for researchers to shorten the list of drug candidates down to the most probable course of experimental validation. Binding affinity predicts several considerations like the flexibility of the protein, solvation effects, and scoring function accuracy. It can be attributed to the recent improvements in the performance of the docking tools, the new developments that have taken place in docking algorithms, and the inclusion of molecular dynamics. This also helps in identifying the candidate binding hotspots and gives mechanistic insight into drug-receptor interactions that are extremely important in lead optimization (Scripps Research, 2019).

2.4.3 Molecular Dynamics

The principle revolves around studying the time-dependent dynamic behavior of biomolecules and how they interact with drug candidates. Molecular dynamics (MD) simulations are excellently placed to yield stability assessments for protein-ligand complexes, binding free energies for gauging the strength of molecular interactions, and provide a basis for conformational changes that are relevant to drug potential. When conducting MD simulations, a drug molecule is treated as being inside a solvent infused with ions inside the binding site; thus, it furnishes an overall picture of the molecular interactions. MD simulation itself is a conveyor of information on molecular flexibility and the actual dynamics that occur between the completely static computational docking results and the actual biological systems. The continually increasing availability of high-performance computing resources is making MD simulations an indispensable tool in drug discovery, giving scientists the capacity to characterize drug-target interactions while making rational drug design interventions very accurately (Samantray et al., 2020).

2.4.4 Merits and Limitations

Bioinformatics in drug discovery has many merits. Primarily, it reduces costs, and expensive wet-lab experiments are avoided, while the process is accelerated. Data-driven methods give better precision to target identification and lead optimization. Predictive models can catch adverse side effects and toxicity long before the candidates reach clinical development. Moreover, it identifies drug repositioning through interaction analyses between available drugs and new targets in ways that hasten the time at which therapeutics make it into the market. The application of artificial intelligence to bioinformatics further expands the role of drug discovery in automating complicated analyses to produce new candidates for more effective drugs (Santa Maria, Wang, & Camargo, 2023).

However, there are also some limitations in the case of bioinformatics. Computational analyses require a big processing capacity and resources, and the prediction sometimes does not go with the experimental results. Accuracy in computational models relies on the quality and completeness of available biological databases; moreover, simplifying assumptions of the model may not represent all the complexities of the biological system. Furthermore, though bioinformatics tools give particularly useful insights, experimental verification becomes imperative to confirm computational predictions. Thus, though challenges persist, bioinformatics is still a revolution that

ushers in the development of drugs by combining computational efficiency with biological knowledge to develop more effective and targeted therapeutic solutions. The future of bioinformatics-driven drug discovery depends on the integration of multi-omics data, artificial intelligence, and high-level computation methods for an in-depth understanding of disease mechanisms and to facilitate the design of precision medicine strategies (Ahmed, Maldonado, & Durrant, 2023).

Chapter Three: Methodology

Diabetes mellitus is a metabolic disease manifested by continuous hyperglycemia or elevated blood glucose levels due to anomalies in insulin secretion, action, or both. Despite advancements in diabetes care, alternative therapies are still in demand. Computational techniques like, molecular docking and molecular dynamics (MD) simulations are applicable methods of computational drug design; both can be employed for the atomic level prediction of the binding mode of the putative small molecule inhibitors against target proteins. These models give atomistic-level details, which provide insight into the process of creating better medications.

Diabetes influences the expression of numerous essential proteins, including PTEN, AKT, AS160, insulin receptors, and other enzymes within the insulin signaling pathway, which play crucial roles in glucose homeostasis. In this study, a group of six phytochemicals were checked as probable therapeutic agents by using bioinformatics methodologies. Further, the phytochemicals were validated for their binding with a diabetes target protein by molecular docking and MD simulations.

3.1 Molecular Docking

3.1.1 Selection and preparation of proteins

After the literature review and analysis of insulin signaling, PTEN (phosphatase and tensin homolog) in the phosphatase domain and AS160 (Akt substrate of 160 kDa) in the Rab-GAP domain were chosen as target proteins. Their 3D crystal structures were retrieved from the Protein Data Bank with PDB IDs 1D5R and 3QYB, respectively (RCSB PDB, 2024; Lee et al., 1999; Sano et al., 2003). Structures were first refined before docking by adding hydrogen atoms, removing water molecules, filling in missing residues and adjusting bond orders with the protein preparation wizard of Maestro Schrödinger. Dimensions of the active site box were defined from previous studies and known binding sites to be 64 Å × 62 Å × 64 Å for PTEN and 72 Å × 68 Å × 50 Å for AS160.

The binding site of AS160 within the Rab-GAP domain was determined by superimposing its crystal structure with 2G77 (Crystal Structure of the Gyp1 TBC domain in complex with Rab33 GTPase bound to GDP and AlF3) and 5S2J (PanDDA analysis group deposition - Crystal Structure of SARS-CoV-2 Nsp3 macrodomain in complex with Z509756472). After alignment, the key active site residues identified were GLY972, ARG973, THR974, PHE975, and THR977, where

ARG973 is the catalytic residue. In PTEN in the phosphatase domain, from previous studies, the catalytic residue identified was CYS124 and the active site residues were HIS96, CYS124, LYS128, ARG130, and GLN171.

3.1.2 Selection and preparation of ligands

Six phytochemicals, according to their established biological activities and promising outcomes in lab experiment findings, were selected for the purpose of the present investigation as potential modifiers of glucose metabolism: benzoic acid (protein data bank (pdb) ID:243), 4-methoxybenzoic acid (pdb ID:7478), 3,4,5-trihydroxybenzoic acid (pdb ID:370), 2-hydroxy-4-methoxybenzoic acid (pdb ID:7477), 4-hydroxybenzoic acid (pdb ID:135), and 2,4-dihydroxybenzoic acid (pdb ID:72). The 3D structures of the ligands were retrieved from the PubChem Database.

The molecular properties of all ligands were analyzed, based on LogP, hydrogen bond donors/acceptors, and rotatable bonds, to check compatibility with the binding sites of AS160 and PTEN. To make sure that correct docking would take place, PyMOL and AutoDock Tools were used to optimize each ligand in terms of appropriate geometry, partial charges, and protonation states. Each ligand was represented in its active form; superfluous groups were deleted, and partial charges were carefully readjusted. Torsional bonds captured ligand flexibility during docking, while hydrogens were added to stabilize each structure where necessary.

These preparations had allowed for the fine-tuning of each ligand for more precise docking interactions with AS160 and PTEN; hence, it allowed for a more accurate analysis of the possible effects each ligand had on glucose metabolism. This preparatory step is a prerequisite for predictions at the level of molecular interactions and binding affinities within the context of insulin signaling.

3.1.3 Molecular docking with AutoDockFR (ADFR)

Auto Dock flexible receptor-ligand docking, called AutodockFR (ADFR), is a major software available for molecular docking in drug design. It is a component of the AutoDock suite, which has long been used for predicting binding modes and energies between small molecules and proteins. ADFR treats not only the ligand as flexible but also allows for flexibility of specific receptor regions in a simulation which provides better approximation than rigid receptor docking

to how binding will take place (Scripps Research, 2019). The software adopts a rule-based hierarchical approach to represent ligand flexibility and utilizes multiple search algorithms for predicting binding poses of small molecules in the protein active site. We selected ADFR for this study, as it is highly accurate in describing complicated receptor-ligand interactions and can efficiently handle big datasets.

ADFR uses a systematic docking methodology to determine the optimal orientation of ligands within the defined receptor binding site. In this work, the docking was performed by means of the Lamarckian Genetic Algorithm, which is efficient in exploring the conformational space by combining evolutionary algorithms with local search methods. To exhaustively explore the conformational space and take into consideration multiple ligand rotations and translations within the binding site, several independent docking experiments were carried out. This would increase the likelihood of finding the most energetically favorable pose. ADFR attempts to generate a conformation with a lower binding free energy by changing the position of the ligand during docking through energy minimization and to increase the likelihood that the scoring function will converge toward an optimal estimate of the ligand's binding affinity to the target protein.

ADFR has a complex scoring function to calculate the affinity of each ligand for the protein. The function calculates the binding free energy considering a variety of critical factors such as van der Waals forces, electrostatic interactions, hydrogen bonding, and desolvation effects. From these interactions, a score is generated for every ligand conformation, and then the poses are ranked by their binding energies. The lower the energy score, the stronger and more constant is the bond between the ligand and the receptor. To explore a wide conformational space and efficiently search for the best ligand orientations, the Lamarckian Genetic Algorithm (LGA) was employed in ADFR as the docking engine. This hybrid algorithm combines global search strategies with local optimization to enhance the accuracy of docking results.

The scoring function was vital for identifying the most probable ligand conformations in this study, which were then thoroughly analyzed to determine their potential to inhibit the target protein. So, we focused on the complexes with the strongest binding energies (as the energies are negative and the more negative, is the strongest binding), because we were only interested in the best interactions and better interactions were indicated by more negative binding energies. We applied an energy cutoff so that all poses with binding energies greater than some arbitrary values are

eliminated and thus we only look at the most favorable interactions between ligands and proteins for further scrutiny.

3.1.4 Post-Docking Analysis

The post-docking analysis was designed to assess the binding modes and affinities of ligands to AS160 and PTEN proteins correspondingly using ADFR for docking. Several important validations had been performed to ensure that the predicted interactions were of high reliability but at the same time informative.

Further, the docking poses were visually inspected in PyMOL, whereby full details could be obtained on the orientation of the ligands within the binding pockets of the proteins. The qualitative assessment hereafter provided views on the predicted interactions that helped in identifying the key residues involved in ligand binding, important for elucidating the mechanism of action.

Binding affinities were evaluated using the scoring functions of ADFR, enabling a comparative analysis of how each ligand interacts with the target proteins. These scores have been analyzed for the ranking of ligands in order of their predicted potencies, underlining compounds likely to be further investigated in experimental assays. Besides affirmation of the binding sites' potential and ligands' affinities, the results of post-docking analysis have identified the interactions that, with a high degree of likelihood, would have to be there for the biological activity of such compounds.

This comprehensive approach provides a platform for further experimental validation and detailed exploration in drug development.

3.2 Molecular Dynamics (MD) simulation

3.2.1 System Setup

Because of its versatility in facilitating research and development in the field, GROMACS, one of the most widely used software packages for modeling biomolecular systems, was used to perform the MD simulations. The protein-ligand complex was first built using the proper force fields, such as the Amber14sb force field, to accurately model the protein-ligand interactions. Building the ligand topology was made possible by the ACPYPE and AMBER tools, which facilitated the process of converting ligand structures into GROMACS-compatible formats and accurately setting all simulation parameters. To preserve overall charge equilibrium after solvation with a TIP3P

water model which is frequently used to simulate physiological conditions the system was then balanced with counterions (0.15 M of Na^+ and Cl^-), as this concentration mimics the physiological ionic strength of human plasma and helps replicate the electrostatic environment found in vivo (Machado & Pantano, 2020).

3.2.2 Energy Minimization and Equilibration

Before producing the main Molecular Dynamics (MD) simulation, the molecular system must be stabilized first. In preparing the simulation setup, there are two key processes that should be performed first, energy minimization and equilibration.

The energy minimization phase seeks to resolve initial structural defects between the molecules. These may often be in the form of high energy contacts or steric interferences, which are regions with poorly positioned or tightly packed atoms. This step is significant because it is needed to begin the MD simulations, which is not possible when the forces acting in the system are remarkably high. During MD, the atomic positions are updated using GROMACS and the Amber14sb force field.

The energy minimization was completed, and the system was stabilized for the main simulation through a two-stage equilibration process. The first stage was to keep the volume constant while raising the system to the desired temperature, which was 310 K, or so-called constant volume-NVT. The NVT equilibration phase was carried out for 100 ps to allow the temperature to stabilize. Changes in temperature at controlled levels without changing pressure were made using GROMACS. After equilibration in the NVT phase, the system underwent equilibration in constant pressure-NPT, whereby volume could slightly change and arrive at the correct value for the required density at 1 atm of pressure. The NPT equilibration was then run for an additional 100 ps to ensure pressure and density stability, this reduced the variation in the pressure and volume further for complete stabilization of the system. Subsequently, the MD production phase was carried out.

3.2.3 Simulation runs by GROMACS

The production MD simulations were performed with the GROMACS 2021.3 package. These covered the proteins AS160 and PTEN bound to six ligands each and the simulation of the proteins

in isolation forms. With such a setup, one can carry out a deep analysis of the behavior of the proteins when isolated and in a ligand-bound state.

All the unrestrained simulations of the proteins were set to run for 100 nanoseconds. The simulation times for proteins where the ligand remained tightly bound was extended to 200 nanoseconds. This extended period was necessary to precisely capture the complex relationships between proteins and ligands, as well as to identify any potential structural changes. Moreover, as shown by Abraham et al. (2019), longer simulations are needed for an accurate explanation of molecular interactions, and these should be closely observed to fully perceive the dynamics and stability of a protein-ligand complex.

The MD simulations produced volumes of data that represent the movement and interaction of atoms in such systems. GROMACS utilized computational resources in efficiently tracking protein dynamics on the fly during the runs. The results obtained from such simulations are envisaged to enhance the understanding of how ligands affect the behavior of AS160 and PTEN and could inform future research or treatment strategies.

3.2.4. Post-simulation Analysis

The main analyses involved the Root Mean Square Deviation (RMSD) of the trajectory data obtained from the various sets of MD simulations carried out for both forms of the AS160 and PTEN proteins alone and in complex with six different ligands. This is a measure that tracks the structure of the protein for stability during simulation. Such a study clearly showed major structural transitions occurring during the simulations. The backbone atoms (N, Cα, and C) were aligned to the reference structure for proteins only to account for general structural shifts. The same set of backbone atoms was then used in calculating the RMSD. In protein-ligand complexes, the alignment was based on the atoms of the protein's backbone, and the RMSD of the ligand group was calculated to track the conformational stability and flexibility of the ligand within the binding site.

The Root Mean Square Fluctuation (RMSF) analyses have provided the regions of proteins showing higher mobility; further flexibility for each residue was investigated using RMSF plot calculations. The radius of gyration (Rg) calculation was done to probe the compactness of protein structures during simulation.

The hydrogen bond analysis was performed regarding the interaction of proteins and ligands with the number of hydrogen bonds formed and their stability. This nature of interaction is a key determinant of the actual binding profile of the ligands towards the proteins. Further, minimum distance analysis gave information about the proximity of the ligands to important protein residues that can be used to interpret interactions related to binding.

Chapter Four: Result

4.1 Molecular Docking

Docking studies were performed to investigate the interactions between the ligands and the target proteins, AS160 and PTEN. This sort of simulation allows the prediction of specific interactions and binding strength of the ligand-protein interaction. These analyses can give a glimpse into the molecular relationships that could influence therapeutic effects. Such types of analysis allow looking into the therapeutic potential of the compounds. Binding energies and key interactions in **Table 4.1** reveal how well these ligands are modulating protein activity. Additionally, it is important to note that the results show the best pose per complex, highlighting the most favorable binding scenarios for each ligand.

Docking scores included Gibbs free energy changes (ΔG), where lower values indicate better binding. In **Table 4.2**, the binding energies, or absolute values of the docking scores, will be shown.

Table 4.1: Binding energies of Phytochemicals to AS160 and PTEN. The binding energies of AS160 and PTEN are calculated by the respective docking scores.

Phytochemicals/Ligand	AS160 binding free energies (kcal/mol)	PTEN binding free energies (kcal/mol)
3,4 Dihydroxybenzoic Acid	-4.3	-7.4
4-Methoxybenzoic Acid	-3.8	-6.8
3,4,5-Trihydroxybenzoic Acid	-4.1	-7.2
2-Hydroxy-4-methoxy benzoic Acid	-4.0	-7.4
4-Hydroxybenzoic Acid	-3.9	-7.0
Benzoic Acid	-3.6	-6.5

The binding affinity of all tested ligands to AS160 and PTEN varies. Of importance, 3,4-Dihydroxybenzoic Acid and 4-Methoxybenzoic Acid showed moderate binding affinities toward AS160 with binding energies of -4.3 kcal/mol and -3.8 kcal/mol, respectively. However, when compared to PTEN, their binding affinities are highly increased, with 3,4-Dihydroxybenzoic Acid having a binding energy as low as -7.4 kcal/mol.

The other promising binding energies to both proteins are found for 3,4,5-Trihydroxybenzoic Acid and 2-Hydroxy-4-methoxybenzoic Acid, with binding energies of -7.2 kcal/mol and -7.4 kcal/mol for PTEN, respectively. This infers that these two ligands might serve as effective inhibitors for AS160 and PTEN proteins and therefore could be subjected to further biological investigation and elaboration of the mechanisms of their action.

Comparing the interactions of AS160 and PTEN while binding with the six different ligands, PTEN forms more substantial interactions, as reflected by its consistently lower binding energies. For example, 3,4-dihydroxybenzoic acid and 2-hydroxy-4-methoxybenzoic acid have the lowest binding energies with PTEN at -7.4 kcal/mol for both ligands. Benzoic acid has the lowest binding affinity against both proteins with binding energies of -3.6 kcal/mol with AS160 and -6.5 kcal/mol with PTEN, which stems from its rigidity and lower intensity of interaction. This analysis underlines the crucial aspect of ligand choice regarding the optimization of the binding efficacy and points out the different interaction profiles between the AS160 and PTEN.

4.1.1 Binding Site Examination of AS160 and Ligands via PyMOL

The binding sites of the six phytochemicals with the AS160 protein are analyzed in detail to enhance the understanding of their interactions. Understanding these interactions is particularly important for explaining the functional roles of AS160 and its importance in therapeutic applications. The unique binding characteristics of each phytochemical can provide critical clues about its biological functions in cellular contexts. By detailed analysis of such interactions, scientists may be able to find novel ways of creating targeted therapies by exploiting the features of these compounds.

The binding pocket of AS160 used for docking was not randomly selected; rather, it was identified through structural alignment with functionally related Rab-GAP domain proteins, specifically using PDB entries 2G77 and 5SZJ. Conserved active site residues between AS160 and these

homologs were identified through PyMOL-based alignment, guiding the definition of a likely functional binding region for ligand docking.

Figure 4.1 below represents the docking results, indicating that each phytochemical binds at similar locations within AS160; thus, a different binding pocket is occupied by each of the distinct ligands. It is interesting to note that 2-hydroxy-4-methoxybenzoic acid and 3,4-Dihydroxybenzoic acid bind in deeper pockets, explaining their comparatively high binding affinities **(Table 4.2)**. Benzoic acid, on the other hand, has the lowest binding affinity of all the ligands tested since its binding occurs in a more superficial pocket. Such findings indicate that deeper, more embedded binding sites form more stable interactions, which may increase the inhibitory potential of the corresponding ligands.

Further, the different binding poses and affinities create a possibility that each ligand might induce a different activity in AS160. It would therefore appear that the higher affinity ligands, for example, 3,4-Dihydroxybenzoic Acid, might stabilize some conformational states of AS160, thereby leading to possible divergence in activity changes of this protein in insulin signaling. Specific interactions between the ligands and amino acid residues at the binding sites of AS160 cause variations in the binding affinities that require further investigation, which will be detailed in the following sections.

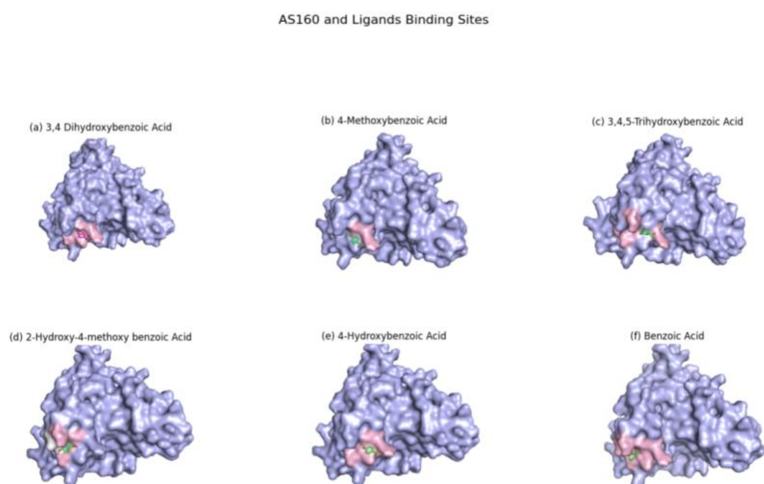

Figure 4.1: Binding sites of AS160 for specific phytochemicals.

The figure shows the docked poses of six different phytochemicals in the binding pocket of the AS160 protein. **The surface of AS160 is lavender**. The compounds represented by the panels are **(a)** 3,4 Dihydroxybenzoic Acid, **(b)** 4-Methoxybenzoic Acid, **(c)** 3,4,5-Trihydroxybenzoic Acid, **(d)** 2-Hydroxy-4-methoxybenzoic Acid, **(e)** 4-Hydroxybenzoic Acid, and **(f)** Benzoic Acid. The binding sites are similar given the differences in the ligands, which suggests they bind to the same region, most likely the active site of the protein. The fact that each ligand's binding site takes a different pose in this region, together with the similarity of the interacting amino acids, gives further support to the idea of a common binding site with slight adjustments in the orientation of the ligands. This suggests similar binding mechanisms, which may lead to the same interactions with AS160. To emphasize the **binding site, it is colored light pink**.

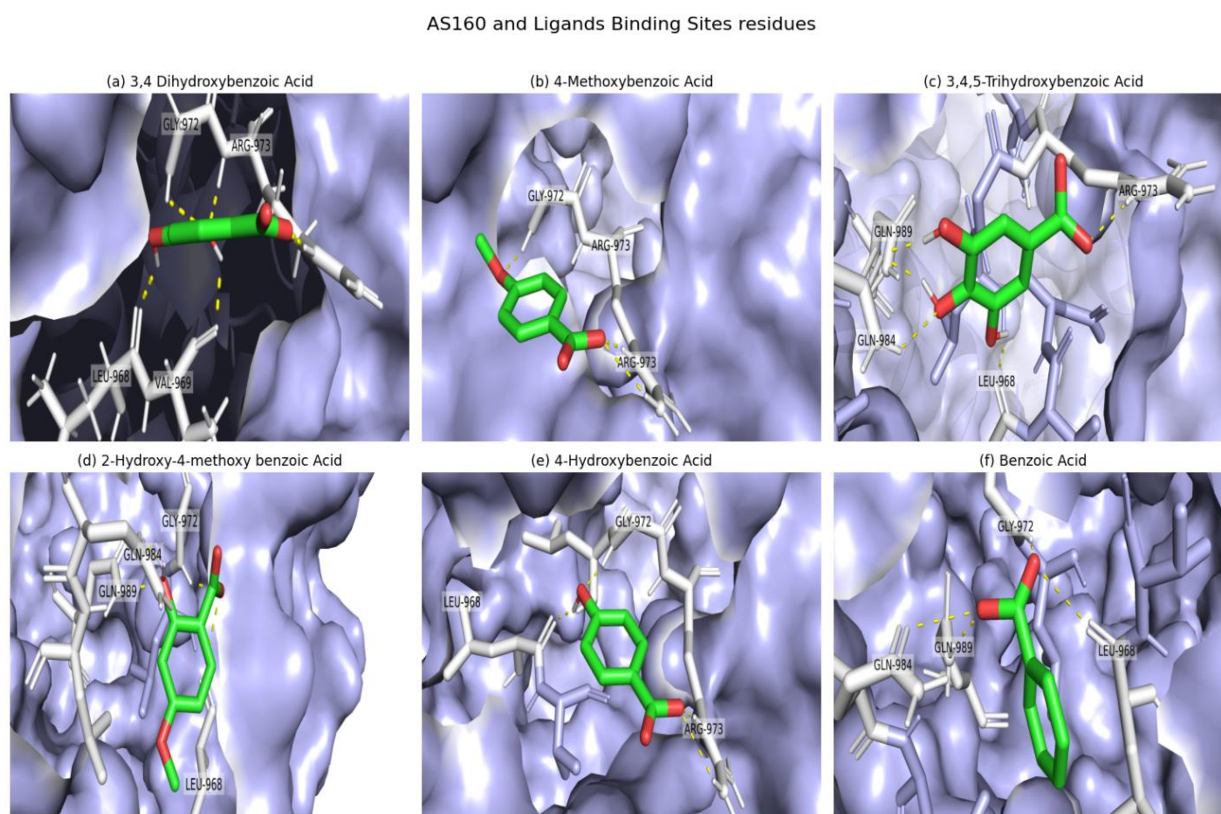

Figure 4.2: Ligand interaction binding sites of AS160:

(a) 3,4 Dihydroxybenzoic Acid bound to AS160 with key interacting residues VAL-959, LEU-968, GLY-972, and ARG-973. **(b)** 4-Methoxybenzoic Acid interacting with GLY-972 and ARG-

973. **(c)** 3,4,5-Trihydroxybenzoic Acid bound to LEU-968, ARG-973 and GLN-984. **(d)** 2-Hydroxy-4-methoxybenzoic Acid interacting with LEU-968, GLY-972, GLN-984 and GLN-989. **(e)** 4-Hydroxybenzoic Acid interacting with LEU-968, GLY-972, and ARG-973. **(f)** Benzoic Acid bound to LEU-968, GLY-972, GLN-984 and GLN-989. Each panel shows the ligand interaction site, with hydrogen bonds as dashed lines.

To interact with AS160, 3,4 Dihydroxybenzoic Acid formed hydrogen bonds with four essential residues: VAL-959, LEU-968, GLY-972, and ARG-973. Overall, five hydrogen bonds were created, which suggest robustness and durability of the interaction. This binding site suggests that GLY-972 and ARG-973 are the major contributors to ligand binding and, thus, important for ligand recognition and for stabilizing complexes. This is further supported by the presence of hydrophilic and hydrophobic residues at the binding site, which favors the mix of polar and non-polar interactions promoting the ligand's binding.

With 4-Methoxybenzoic Acid, AS160 formed three hydrogen bonds with GLY-972 and ARG-973. Even though the number of hydrogen bonds obtained was less in comparison with 3,4 Dihydroxybenzoic Acid, it still points toward a moderate binding strength between AS160 and the ligand. From the binding site analysis, it can be noted that ARG-973 would contribute to the general stability of the binding, while GLY-972 is crucial for the stabilization of the ligand.

In the case of 2-Hydroxy-4-methoxybenzoic Acid, four hydrogen bonds were formed between the ligand and LEU-968, GLY-972, GLN-984 and GLN-989. Its stabilization within the binding pocket was by the key residues LEU-968 and GLY-972. The interaction is of moderate strength according to the binding free energy (Table 4.1). Even though it formed fewer hydrogen bonds than 3,4 Dihydroxybenzoic Acid and 3,4,5-Trihydroxybenzoic Acid, a combination of hydrophilic and hydrophobic interactions suggests a stable binding complex.

4-Hydroxybenzoic Acid had made four hydrogen bonds with AS160 through interaction with LEU-968, GLY-972, and ARG-973. The GLY-972 and ARG-973 seemed to be the key players in the interaction which, besides holding the ligand tightly inside the binding pocket, further stabilized it. This interaction was as strong as that of 2-Hydroxy-4-methoxybenzoic Acid; hence, it can be said that the binding affinity of 4-Hydroxybenzoic Acid was moderate.

Finally, AS160 and benzoic acid interacted by forming four hydrogen bonds through ligand binding to LEU-968, GLY-972, GLN-984 and GLN-989. Based on the analysis of the binding site, the major residues in the stabilization of the ligand are GLY-972 and GLN-984. The nature of interaction between benzoic acid and AS160 is like that of 4-hydroxybenzoic acid. However, the involvement of GLN-984 and GLN-989 makes the binding pose dissimilar. These additional interaction residues may stabilize the ligand through a network of additional hydrogen bonds.

In summary, the number of hydrogen bonds and the residues involved suggest that the interaction of AS160 with the six ligands exhibits varying strengths and stability in binding. The maximum number of hydrogen bonds between AS160 and ligand, which is five, was reached by 3,4 Dihydroxybenzoic Acid and 3,4,5-Trihydroxybenzoic Acid, showing stronger or more stable interactions with AS160. 4-methoxybenzoic acid formed only three hydrogen bonds and thus has weaker interaction. The binding strength of the other three ligands, i.e., benzoic acid, 4-hydroxybenzoic acid, and 2-hydroxy-4-methoxybenzoic acid, falls in between, each forming four hydrogen bonds.

The frequently involved binding residues LEU-968, GLY-972 and ARG-973 demonstrate the importance of these amino acids in the ligand recognition and binding pattern of AS160. Other key interacting residues are GLN-984 and GLN-989, which thus also play a key role in the stabilization of the ligands inside the protein binding pocket.

4.1.2 Binding Site Examination of PTEN and Ligands via PyMOL

Understanding the binding sites of phytochemicals with the PTEN protein is, therefore, very important in exploring its role in biological processes and therapeutic potential. Such unique binding traits of each phytochemical have been found to reveal important insights into their cellular functions. **Figure 4.3** shows these interactions.

PTEN and Ligands Binding Sites

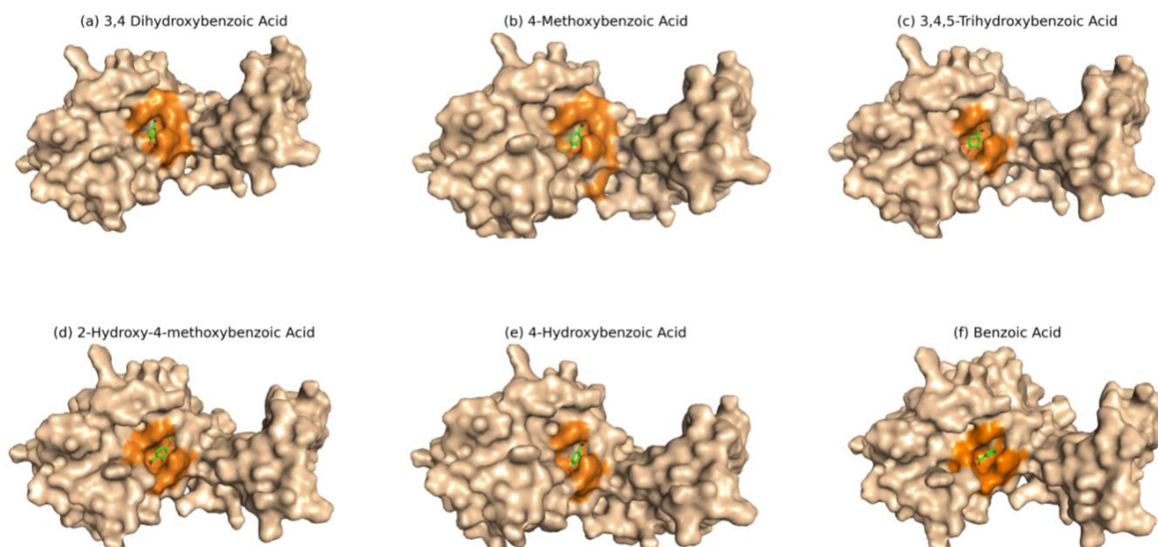

Figure 4.3: Binding Sites of selected phytochemicals with PTEN.

The figure shows the binding positions of six different ligands colored in red within the binding site of PTEN protein colored in light green. The panels correspond to: **(a)** 3,4 Dihydroxybenzoic Acid, **(b)** 4-Methoxybenzoic Acid, **(c)** 3,4,5-Trihydroxybenzoic Acid, **(d)** 2-Hydroxy-4-methoxybenzoic Acid, **(e)** 4-Hydroxybenzoic Acid, and **(f)** Benzoic Acid.

As found from the docking results, **Figure 4.3** indicates that each of these benzoic acid derivatives binds in common regions of the PTEN binding site. The known functional role of the binding site and the accessibility of the ligand is why it was selected for docking. Previous studies have shown this region of PTEN to be critical for enzymatic activity, which makes it an excellent target to study the interaction of the ligands. This infers that these ligands might interact via a variety of mechanisms. The binding affinities of 2-hydroxy-4-methoxybenzoic acid and 3,4-dihydroxybenzoic acid are higher (**see Table 4.1**) because they bind within deeper pockets of PTEN. With deeper binding pockets, more extensive, this would enable more extensive contacts, such as hydrogen bonding and hydrophobic contacts, to be formed with the resultant complex being more stable. On the other hand, the lowest binding affinity among the ligands studied here

is that of benzoic acid whose binding site is located in a more exposed and superficial area of the site. This suggests that, when bound in deeper and more enclosed pockets, ligands might be better able to modulate PTEN.

The binding sites and affinities have indicated that the variation may have variable effects on PTEN functional states depending on a ligand. For example, conformation changes of PTEN may be induced by high-affinity ligands like 3,4-Dihydroxybenzoic Acid. This may hence alter how PTEN regulates the signaling pathways. It provides a platform for further investigation, most importantly regarding how the activity of phosphatase from PTEN will be specifically modified by these compounds, considering different modes through which each ligand binds to selected amino acid residues present within the PTEN binding pocket. The present study therefore provides useful information on the binding behavior of these ligands, thus setting the stage for future experimental validation of how each compound influences PTEN's biological role and implications in signaling pathways.

PTEN and Ligands Binding Sites residues

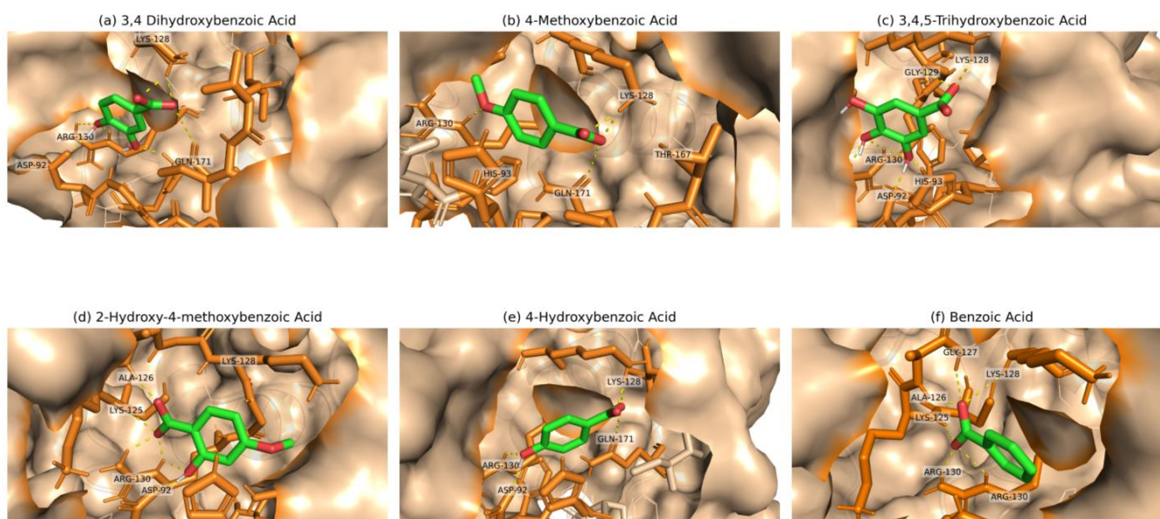

Figure 4.4: Ligand Interaction Binding Sites of PTEN.

(a) 3,4 Dihydroxybenzoic Acid bound to PTEN with key interacting residues ASP-92, LYS-128, ARG-130 and GLN-171. **(b)** 4-Methoxybenzoic Acid interacting with residues LYS-128, ARG-130 and GLN-171. **(c)** 3,4,5-Trihydroxybenzoic Acid bound to residues ASP-92, HIS-93, LYS-128, GLY- 129, and ARG-130. **(d)** 2-Hydroxy-4-methoxybenzoic Acid interacting with residues ASP-92, LYS-125, ALA-126, and ARG-130. **(e)** 4-Hydroxybenzoic Acid interacting with ASP-92, LYS-128, ARG-130 and GLN-171. **(f)** Benzoic Acid bound to residues LYS-125, ALA-126, GLY-127, LYS-128, and ARG-130. Each panel shows the ligand interaction site, with hydrogen bonds indicated by dashed lines.

Important interactions between 3,4-Dihydroxybenzoic Acid and PTEN are seen at ASP-92, LYS-128, ARG-130, and GLN-171. This ligand forms a stable binding environment by creating hydrogen bonds, especially with ASP-92 and ARG-130. The involvement of both positively charged (ARG-130) and polar (ASP-92) residues in the pocket enhances interaction, indicating high affinity binding. This stable interaction is further strengthened by extra hydrogen bonding provided by GLN-171.

4-Methoxybenzoic Acid interacts with ARG-130, LYS-128 and GLN-171. While fewer hydrogen bonds are formed compared to 3,4-Dihydroxybenzoic Acid, the interaction with ARG-130, an important binding pocket residue in PTEN results in a moderately stable binding environment. Although much weaker than hydroxylated compounds, the methoxy substitution on the ligand enabled a balanced mix of hydrophilic and hydrophobic interactions that allow a stable interaction.

ASP-92, LYS-128, ARG-130, and GLN-171 all exhibit binding interactions with 3,4,5-trihydroxybenzoic acid. This ligand forms a stable ligand-protein complex like 3,4-Dihydroxybenzoic Acid through multiple hydrogen bonds, especially with ARG-130 and ASP-92. The extra hydroxyl presents in 3,4,5-Trihydroxybenzoic Acid increases its hydrogen-bonding activity, which increases the stability of its binding and infers that it has a high affinity for PTEN.

2-Hydroxy-4-methoxybenzoic Acid interacts with the binding pocket of PTEN through interactions with Lys-128, Gly-127, Ala-126, and Arg-130. Gly-127 and Arg-130 provide stabilizing contacts, although this ligand forms fewer hydrogen bonds than the dihydroxybenzoic derivatives. The hydroxyl and methoxy groups provide a balance of interaction types, while the hydrophilic contributions from Gly-127 and Arg-130 and the hydrophobic involving Ala-126 create a moderately stable complex.

4-hydroxybenzoic acid binds PTEN through hydrogen bonds with Gly-129, Arg-130, His-93, and Asp-92. Interactions with Asp-92 confer additional stability on it, while the more stable interactions with Arg-130 and His-93 suggest a moderate binding affinity. This compound is, in general, weaker than trihydroxy benzoic derivatives in terms of overall binding strength but still shows a considerable level of stability inside the pocket of PTEN.

Benzoic Acid interacts with Ala-126, Lys-125, Arg-130, and Asp-92. This ligand, having no other functional groups besides hydroxyl or methoxy groups, is shallower in the binding profile inside PTEN's pocket; hence, fewer hydrogen bonds can form, reflecting poorer binding. The Arg-130 and Asp-92 can provide some stability, but compared to other ligands, the binding is low due to the lack of other interaction sites.

In summary, Arg-130 is the major stabilizing residue in the PTEN binding pocket, electrostatically or via hydrogen bonds interacting with each of the six ligands. The multi-hydroxyl group-containing ligands such as 3,4-Dihydroxybenzoic Acid and 3,4,5-Trihydroxybenzoic Acid can form extensive hydrogen bonds and thus have the largest binding affinities. Benzoic acid has the weakest interaction due to a lack of extra functional groups. That indicates that the presence and the relative position of hydroxyl or methoxy groups are crucial for the stabilization of the ligand increment and affinity in the PTEN active site. These facts could make it possible to prepare a benzoic acid derivative with an optimized PTEN binding profile.

4.2 Molecular Dynamics

4.2.1 MD of PTEN and AS160 in the Apo form

4.2.1.1 Simulation Setup

MD simulations were performed using the software GROMACS, version 2021.3 and the AMBER 14SB force field, which works well for protein simulations. Each protein system was solved in a water box, maintaining a 10 Å buffer between the protein and the box boundaries. Na^+ and Cl^- ions were then added to neutralize the system and provide a physiological ionic strength of 0.15 M. The temperature was kept at 310 K, while the pressure was kept at 1 bar using the Parinello-Rahman algorithm. The simulation time for each protein was 100 ns, involving 50,000,000 production MD steps of 2 fs.

The temperature fluctuations during the 100 ns MD runs of PTEN and AS160 in their apo states are shown in **figure 4.5 the "Temperature over Time" plot**. Early equilibration of the system

still causes larger spikes in the temperature immediately at 0 ps, indicating how the system acquires its target temperature of 310 K. After this transient period, the temperature starts to stabilize in a range between 295 and 325 K. At this stage, the system has reached steady-state conditions and thermal stability from the initial non-equilibrium state.

System Temperature vs Time

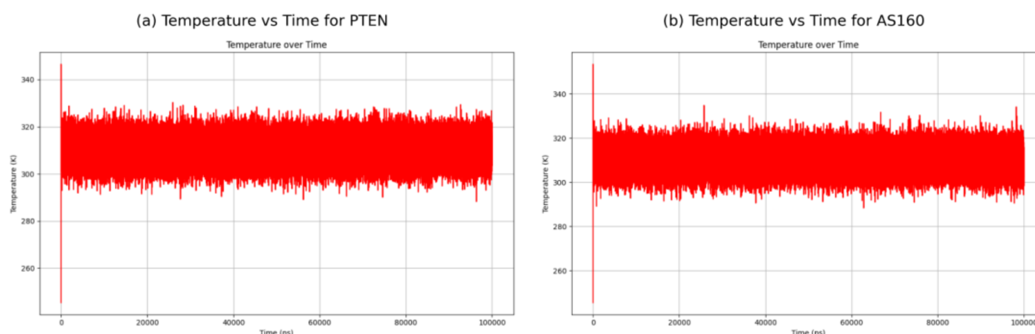

Figure 4.5: Temperature fluctuations during the MD simulations of the apo forms of PTEN and AS160.

(a) Temperature vs Time for PTEN, **(b)** Temperature vs Time for AS160. This plot confirms the thermal stability and equilibration of the system during the simulation. **X-axis:** Simulation time 100 ns based on ps, ranging from 0 to 100000 ps and **Y-Axis:** The temperature in Kelvin [K] which reflects the thermal state of this system.

The temperature remains constant at around 310 K after the initial fluctuation with very minor oscillations. This clearly shows that thermal equilibrium was reached by the system and hence regulated the system temperature as such. The fluctuations seen, as indeed expected in molecular dynamics simulations, demonstrate natural thermal motion and energy exchange between molecules. Such fluctuations indicate continuing system dynamics, not instability. The temperature stabilizes within a narrow range, which secures appropriate equilibration. Thus, validation of the simulation setup is obtained, ensuring that the data will be reliable for further structural and dynamic analysis.

The plot, "System Energy Components vs Time" in figure 4.6, shows convergence in the behavior of the potential, kinetic, and total energies for the 100 ns MD simulations of apo PTEN

(left) and AS160 (right). Overall, the potential energy has minor fluctuation tendencies during simulation time (blue line). This means that no major changes in the protein structure occur and that all the molecular interactions, such as van der Waals and electrostatic forces, between the proteins and the solvent in which they are embedded are stable. Besides this, the kinetic energy-red line-remains constant; this further confirms steady molecular motion, showing that the temperature of the system was controlled during the run.

The green line, representing the sum of kinetic and potential energies, refers to the total energy and remains negative with only a few fluctuations throughout the simulation. It is proof of the energy conservation of the system. Indeed, such a flat energy profile confirms that the system is stable and at equilibrium by showing no major external forces or errors influencing the system. The minor oscillations in all energy components serve as further evidence that the system reached equilibrium early on during the simulation and remained there throughout the 100 ns run. This stable energy performance makes the resulting data reliable for further study, reinforcing the correctness of the setup of the simulation, constituting a sound framework for further studies focused on dynamics and interactions of the protein.

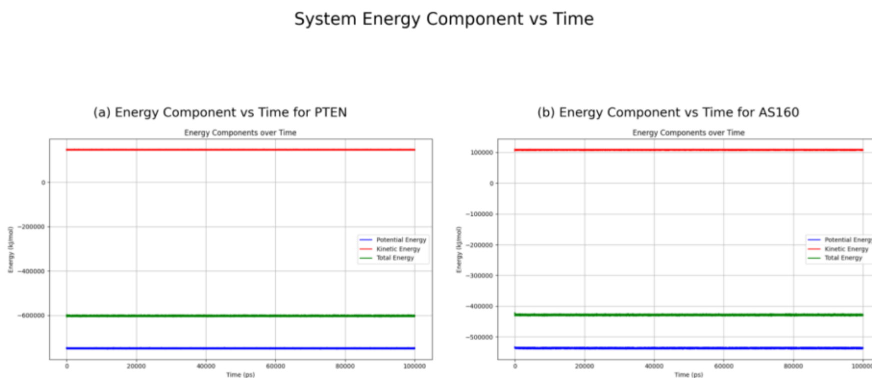

Figure 4.6: Different forms of energy, namely potential, kinetic, and total, from the MD simulations of apo PTEN and AS160.

(a) Energy component vs Time for PTEN, **(b)** Energy component vs Time for AS160. **X-axis:** The length of the simulation is 100 ns, shown in picoseconds ranging from 0 to 100000 ps. **Y-axis:** The energies in kJ/mol for the various energy components in the system according to the color legend.

4.2.1.2 Stability and Convergence of AS160 and PTEN: RMSD, RMSF, and Radius of Gyration (Rg) Analysis

MD can be a helpful tool in assessing the stability and conformational change of the simulated proteins, PTEN and AS160, either in solution or in complex with ligands. This section demonstrates the convergence of the MD simulations of the two apo proteins utilizing common stability measurements, including RMSD, RMSF, and Rg analyses.

After the initial equilibration period, the RMSD of PTEN oscillates around the values of 0.3 to 0.4 nm; this suggests the protein settles into a stable conformation. Because there are no significant changes in the overall conformation of the protein during this simulation, the steady value of the RMSD supports the fact that PTEN does not lose its structural integrity. This is an important requirement for understanding how PTEN interacts with ligands, as it allows pinpointing a defined binding site. AS160 also stabilizes after an initial brief fluctuation and has a lower RMSD of 0.2 nm, suggesting a more rigid or less flexible structure than PTEN in the simulated environment. This suggests that AS160 achieves conformational stability without undergoing significant structural changes.

The RMSF plot describes flexibility of individual amino acid residues in the proteins. The RMSF plot for PTEN has peaks reaching up to 0.5 nm, which correspond to highly flexible regions that are either loop regions or terminal ends. However, since most of the residues have lower RMSF values, it would be appropriate to say that PTEN has a stable core, with only a few segments showing notable variation. The RMSF plot for AS160 stands for a compact and stable protein structure, since only a few peaks are high. Thus, the protein is more rigid. The structural stability for both proteins is further supported by the analysis of the Radius of Gyration. During the course of the simulation, PTEN's Rg oscillates between 2.15 and 2.25 nm, which indicates stable compactness as the protein does not undergo significant expansion or contraction. In comparison with PTEN, Rg of AS160 stabilizes at a lower range of about 1.93 –1.97 nm, suggesting it to be more compact (or simply smaller). Such a stable and compact form of AS160 would meet its possible functional requirement toward keeping a rigid structure in cellular contexts.

In summary, these stability and convergence analyses suggest that, though some structural fluctuations exist in both PTEN and AS160 the integrity of both structures is maintained throughout the course of the simulations. PTEN is seen to be more flexible than AS160 in selected

regions, which may be related to functional or binding activities, whereas AS160 appears more compact and rigid. These structural differences are in tune with different biological functions: rigidity supports the structural role of AS160 in the signaling pathways, whereas the flexibility in PTEN might be supportive for the interaction with different ligands.

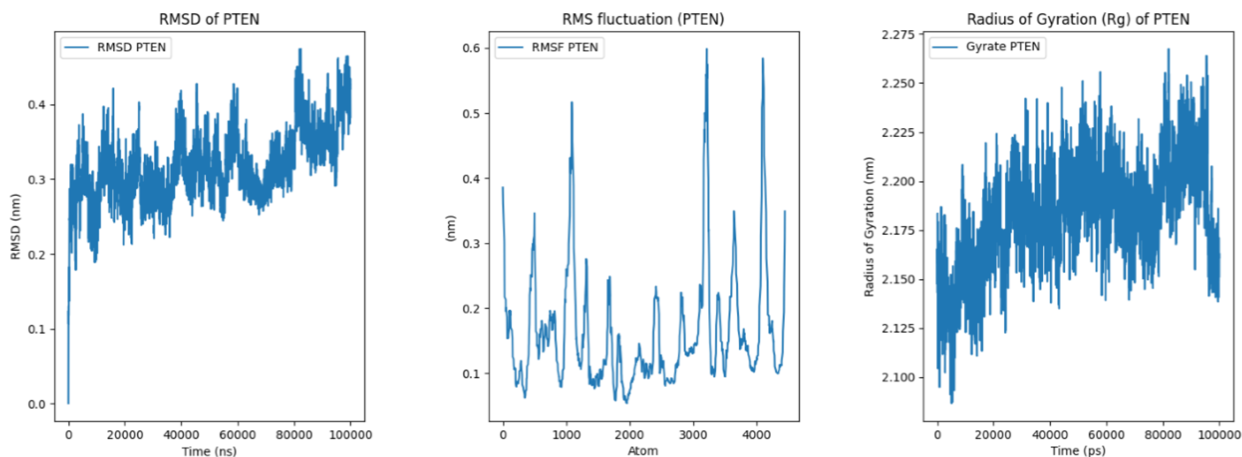

Figure 4.7: PTEN's Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), and Radius of Gyration (Rg).

The backbone atoms (N, Cα, and C) were selected for RMSD and RMSF calculations to monitor protein structural deviations and analyze residue-level flexibility.

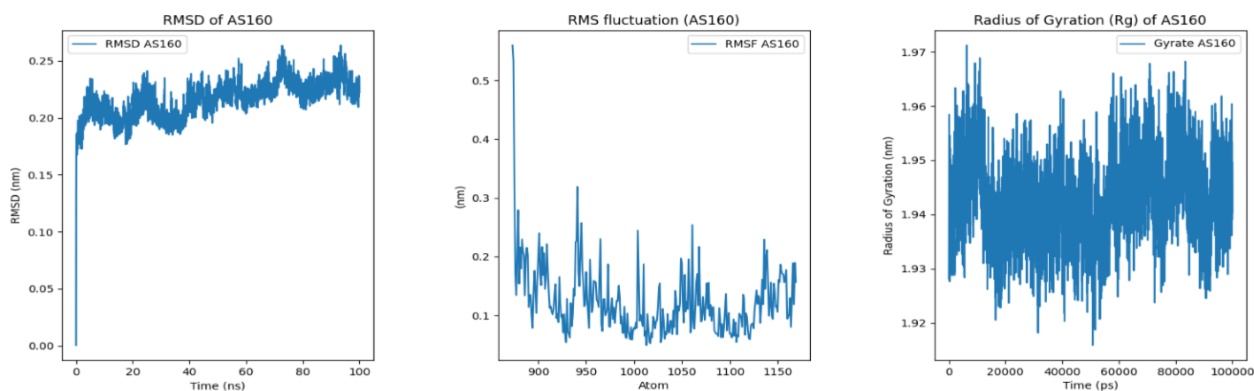

Figure 4.8: AS160 Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), and Radius of Gyration (Rg).

The backbone atoms (N, Cα, and C) were selected for RMSD and RMSF calculations to monitor protein structural deviations and analyze residue-level flexibility.

4.2.2 MD of PTEN with Phytochemicals

4.2.2.1 Setup of PTEN Simulation with Phytochemicals

Atomic molecular dynamics simulation of PTEN in complex with six different phytochemicals was performed using the GROMACS package and the AMBER14SB force field. The interactions of PTEN with each ligand were studied under physiological conditions. Each ligand was specifically parameterized to reflect its individual specific charge state at the physiological pH with the intent of accurately representing its interactions with PTEN. In view of the different functional groups present, it was important to probe for different charge states and different interaction profiles of each phytochemical screened in interacting with PTEN, to learn more about their binding mechanism and to inhibit enzymatic function by ligand binding. Gallic acid, 4-methoxysalicylic acid, 4-hydroxybenzoic acid, benzoic acid, protocatechuic acid, and p-anisic acid were the selections for this study. All compounds contain the carboxylic acid, hydroxyl, and methoxy functional groups that determine their physiological charge states. We found that the carboxylic acid groups in all six phytochemicals are deprotonated at physiological pH and carry a net negative charge.

4.2.2.2 Stability and Flexibility Analysis (RMSD, Rg and RMSF)

The RMSD analysis after the simulation of PTEN-ligand complexes gives vital information about the stability of these interactions in time as is shown in **Figure 4.9 t**he RMSD for ligands such as 3,4-Dihydroxybenzoic Acid (Protocatechuic Acid) varies highly, especially in later stages of the simulation. With such variation, the complex that would be formed between protocatechuic acid and PTEN would not be stable. This might be due to less favorable interactions with this ligand, which could diminish its ability for productive modulation of PTEN activity. By contrast, benzoic acid shows the least and most consistent RMSD values, which means that during the simulation, this molecule is in a tightly bound, stable conformation. That would suggest that benzoic acid might have a stronger effect on PTEN; such stability could be indicative of long-lasting interactions between the protein and ligand.

Moreover, low RMSD values were observed for gallic acid and 4-methyl salicylic acid; this means that these ligands make quite stable complexes with PTEN. These phytochemicals continued to

interact with PTEN's binding site consistently, these ligands are therefore good candidates for further research. Besides, though 4-methoxybenzoic Acid may bind decently with PTEN, its binding might not be as stable as the other ligands, resulting in sporadic conformational shifts.

In summary, ligands with similar constant and low RMSD values are represented by benzoic acid, gallic acid, and 4-methyl salicylic acid. These ligands can sustain a robust binding to PTEN, conferring a better influence on PTEN function. On the other hand, higher RMSD fluctuations of ligands such as 4-methoxybenzoic Acid and protocatechuic Acid suggest weaker or less stable binding, thus needing further optimization to better interact with PTEN. Thus, this present analysis will determine the ligand with the highest potential for therapeutic intervention against PTEN.

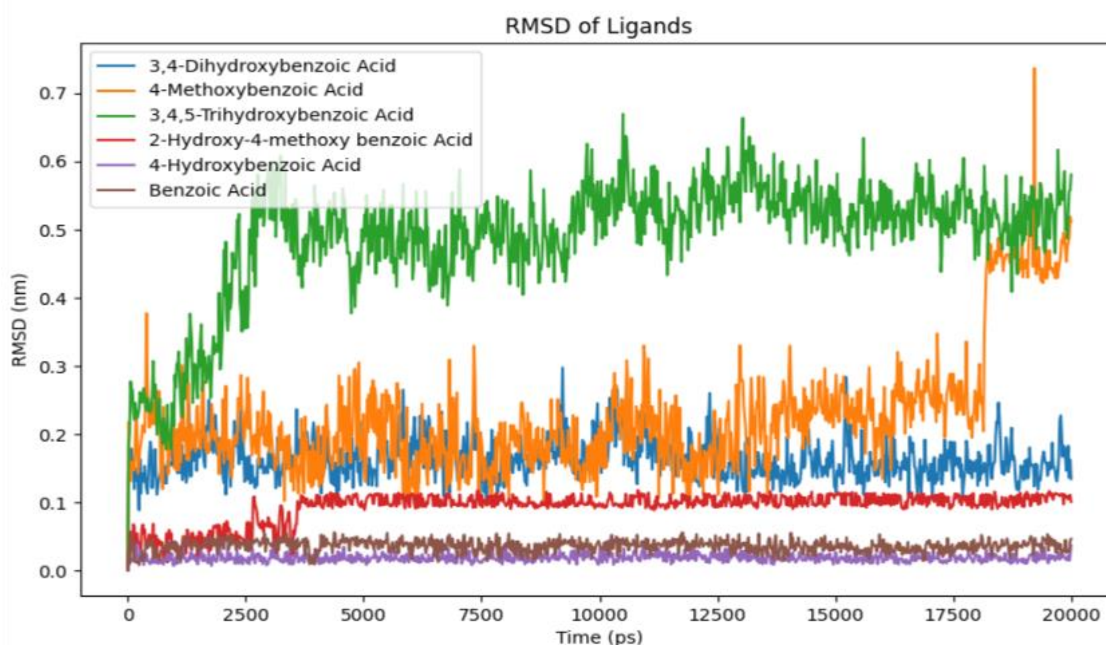

Figure 4.9: PTEN-Ligand Complexes RMSD Analysis.

The figure shows the Root Mean Square Deviation of six phytochemicals complexed with PTEN over molecular dynamics simulation. The x-axis is time in picoseconds while the y-axis is RMSD in nanometers, specifying structural fluctuations of the ligands with respect to their initial binding pose.

4.2.2.3 Binding Behavior

The hydrogen bonding interaction between PTEN and six phytochemicals has been analyzed in this study, based on the presence or absence of the carboxyl group in the ligands. In analyzing the results, two figures were used: **Figure 4.10** represents hydrogen bonding interactions in PTEN without ligand carboxyl groups, and **Figure 4.11** represents hydrogen bonding interactions in PTEN with ligand carboxyl groups. These differences show the importance of the carboxyl group in the long-term viability of the ligand-protein interaction.

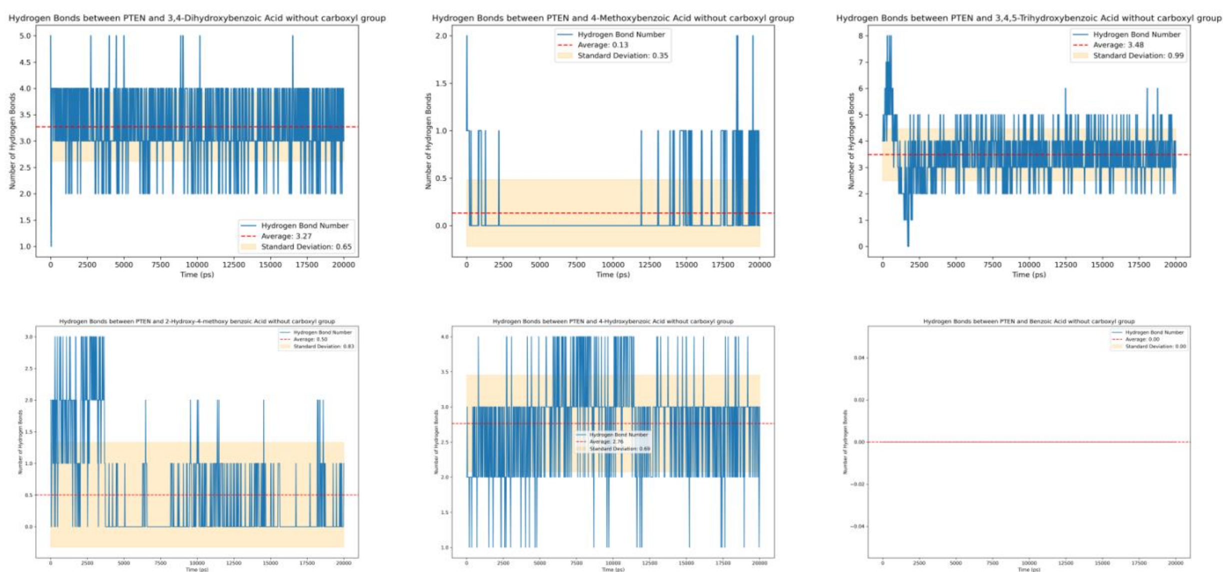

Figure 4.10: Hydrogen bond analysis in the absence of ligands carboxyl group (deprotonated state).

X-axis: the simulation time in picoseconds. Y-Axis: The number of hydrogen bonds formed in the absence of a carboxyl group of the ligands. Blue Line: The number of hydrogen bonds that each ligand formed during the simulation. The red dashed line represents the average number of hydrogen bonds during the simulation run. The orange-shaded area describes fluctuation within occurrences of hydrogen bonds from the simulation run that reflects the steadiness of this interaction.

Binding behaviors can be viewed in Figure 10, which illustrates the hydrogen bonding of PTEN with ligands devoid of a carboxyl group. In these simulations, the ligand 3,4-dihydroxybenzoic acid forms an average of 3.27 hydrogen bonds throughout, with small fluctuations in the number,

as reflected by the Standard Deviation (SD = 0.65). This suggests that hydroxyl groups play a key role in preserving hydrogen bonds with PTEN, since it presents a stable and reliable interaction. 3,4,5-trihydroxybenzoic acid has a higher variability, SD = 0.99, but an average of 3.48, indicating that even more hydrogen bonds are formed. Thus, the extra hydroxyl group facilitates stronger binding even though the bond dynamics is greater.

By contrast, the 4-methoxybenzoic acid complex exhibits extremely poor binding, with an average of only 0.13 hydrogen bonds. This suggests that the methoxy group alone does not favor hydrogen bonding with PTEN. The 2-hydroxy-4 methoxy benzoic acid complex exhibited weak but marginally more stable bonding than 4 methoxy benzoic acid, with an average of 0.50 hydrogen bonds and greater variability (SD = 0.83), suggesting intermittent hydrogen bonding over the course of the simulation. Finally, the carboxyl group-depleted benzoic acid does not interact through hydrogen bonds with PTEN: the average is zero.

In summary, **Figure 4.10** shows how the deprotonation of the carboxyl group alters hydrogen bonding. As shown above, the average number of hydrogen bonds is not severely deviated throughout but still maintains reasonable stability due to significant fluctuations - orange shaded region. Such variability in interaction does signify those moderate levels of instability of the network, propagating through and affecting ligand-interaction overall uniformity.

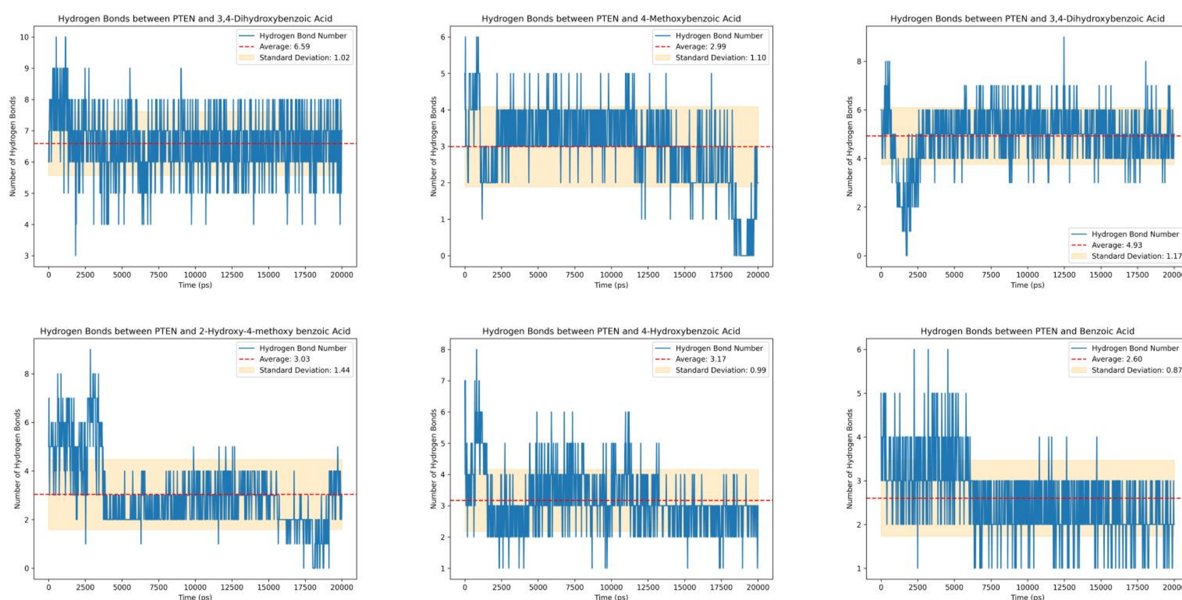

Figure 4.11: Hydrogen bond analysis in the presence of a carboxyl group (Protonated).

X-axis: Time, ps - simulation time in picoseconds. The Y-axis describes the number of hydrogen bonds between PTEN and the ligand's carboxyl group at each time step. Blue Line: The number of hydrogen bonds that take place for every ligand throughout the simulation. The red dotted line shows the average number of hydrogen bonds formed during the simulation. The dispersion of standard deviation is shown with an orange shading to indicate variability from the mean number of hydrogen bonds.

Most of the ligands showed an observable enhancement in hydrogen bonding in **Figure 4.11** for the ligands with a carboxyl group. The average number of hydrogen bonds formed by 3,4-dihydroxybenzoic acid is 3.32, which is more than in its non-carboxylate form, but it has slightly more fluctuations, SD = 0.75. Thus, the slight increase implies that interaction stability is marginally improved by the carboxyl group. The carboxyl group acts significantly in hydrogen bond formation with PTEN, since the most notable improvement happened in 4-methoxybenzoic acid, which formed an average of 2.86 hydrogen bonds now, with an SD = 1.19. This reflects how important the carboxyl group is in allowing the ligands to interact with PTEN more robustly after they had previously formed weak bonds.

Interestingly, this carboxyl addition radically decreases the hydrogen bonding present within 3,4,5-trihydroxybenzoic acid to only 1.44 on average. It is presumed that the addition of the carboxyl group creates unfavorable steric hindrances for the ligand binding to PTEN, hindering its ability to act well through hydrogen bonding interactions. And the ligand deviated from its initial docking pose as the RMSD plot shown. With 2-hydroxy-4-methoxybenzoic acid, the carboxyl enhances the bonds, increasing its hydrogen bonding average to 2.53 (SD= 0.98). This would suggest that the carboxyl group confers better alignment or exposure for hydrogen bonding. The benzoic acid is similarly greatly enhanced in forming an average of 2.6 hydrogen bonds, SD = 0.87, again pointing to the key role of the carboxyl group in allowing hydrogen bonding to PTEN.

In general, a carboxyl group would enhance hydrogen bonding interactions between PTEN and its ligands; nonetheless, the overall effect really depends upon the general structure and functional group placement on the ligand, as seen for 3,4,5-trihydroxybenzoic acid.

The protonation of the carboxyl group notably affects the interaction between the hydrogen bonding of PTEN protein and their ligands. Upon protonation, polarity along with their capabilities to donate the hydrogen bonds transforms from carboxylic acids into carboxylates ($-COO^-$).

Stability would be improved while strengthening the interaction between ligands and the proteins through more prominent hydrogen bonds along with polar adjacent side chains within the protein in the case of protonated carboxyl's. Conversely, carboxylates could interact differently by hydrogen bonding or be weaker in their deprotonated state, thereby decreasing the overall binding affinity. The functional groups play a key role in the biological activity of the ligands because the protonation state determines the strength and effectiveness of ligand-protein interactions.

Table 4.2: PTEN Binding Site Residues with six Phytochemicals after MD run.

This table summarizes binding site residues of PTEN related to various phytochemicals and ligands.

Phytochemicals/Ligand	Binding site residues for PTEN
3,4 Dihydroxybenzoic Acid	92 ASP, 93 HIS, 128 LYS, 129 GLY, 130 ARG, 166 VAL, 167 THR, 168 ILE, 171 GLN
4-Methoxybenzoic Acid	126 ALA, 129 GLY, 130 ARG, 166 VAL, 167 THR, 168 ILE, 171 GLN
3,4,5-Trihydroxybenzoic Acid	92 ASP, 93 HIS, 124 CYS, 128 LYS, 129 GLY, 130 ARG
2-Hydroxy-4-methoxy benzoic Acid	92 ASP, 124 CYS, 126 ALA, 128 LYS, 129 GLY, 130 ARG 167 THR, 168 ILE, 171 GLN
4-Hydroxybenzoic Acid	92 ASP, 93 HIS, 126 ALA, 128 LYS, 129 GLY, 130 ARG, 169 THR, 168 ILE, 171 GLN
Benzoic Acid	124 CYS, 128 LYS, 126 ALA, 130 ARG

The binding site residues presented in Table 3 differ from the list of residues documented in the docking analysis since they were inferred from MD simulations and specific selection criteria. Residues were selected after analysis of the MD trajectories at a 4 Å cut-off distance to define the ligand-protein interaction. The cut-off occupancy value used was 0.5 Å. This approach captures

dynamic behavior and stability of interactions by identifying residues based on their persistent proximity to the ligand throughout the simulation.

Table 4.2 gives the binding of six phytochemicals to the PTEN protein, as obtained from molecular dynamics simulations. The specific set of residues that each phytochemical binds in the active site of PTEN determines how the ligand may affect PTEN's biological function. A 0.5 Å proximity criterion has been used for identifying the important interactions between ligands and protein residues, as it stands for the distance threshold within which the atoms from the ligands and protein residues are assumed to form stable and meaningful interactions.

3,4 Dihydroxybenzoic Acid targets a wide variety of PTEN residues, for example, ASP92, HIS93, LYS128, GLY129, ARG130, VAL166, THR167, ILE168, and GLN171. These residues more than likely contribute significantly to the ligand's stability inside the binding pocket of PTEN. In this state, hydrophobic contacts through VAL166, ILE168, and THR167 add more stability to the ligand due to hydrophobic interactions, while charged groups like ASP92 and ARG130 are very much possible to take part in ionic or hydrogen bonds. This maintains the balancing interaction that assists the complexation of the ligand.

In contrast, the binding of 4-Methylbenzoic acid is given by residues such as ALA126, GLY129, ARG130, VAL166, THR167, ILE168, and GLN171. As in 3,4 Dihydroxybenzoic Acid, hydrophilic as well as hydrophobic residues are also found here, entailing this ligand–protein complex, too. Salt bridges or hydrogen bonds that help further stabilize the complex may necessitate ARG130 and GLN171, whereas the presence of ALA126 and GLY129 renders flexibility, aiding the direction of placement in the pocket for the ligand.

The interactions between PTEN and 3,4,5-Trihydroxybenzoic Acid, 2-Hydroxy-4-methoxybenzoic Acid, 4-Hydroxybenzoic Acid, and benzoic Acid are characterized by the presence of both common and different binding residues. The highly conserved residues in binding the multiple ligands include ASP92, HIS93, LYS128, GLY129, and ARG130. This, therefore, indicates that these residues must be important for stabilizing the ligand-protein interactions. Other residues implicated in binding are THR167, ALA126, and CYS124, all of which would appear to play a less consistent role, depending on the ligand in question.

Overall, these results thus suggest that these phytochemicals share a similar mode of binding to PTEN with a major contribution of electrostatic interactions, hydrogen bonds, and hydrophobic contacts despite their structural diversity. Some key residues recurring in the binding of different ligands thus point to important functional regions of PTEN for ligand binding that may also be relevant to modulate PTEN biological activity. These molecular insights into PTEN targeting are pointing toward various potential therapeutic strategies by which PTEN can be manipulated in cancer, metabolic disorders, and abnormal cell growth. From the point of view of therapies, these findings also pointed toward the particularly essential information on how small molecules would be generated which will interact selectively with PTEN and alter its activity.

4.2.3 MD of AS160 with Phytochemicals

4.2.3.1 Stability and Flexibility Analysis (RMSD)

Analysis of the ligand RMSD for the different complexes shows that none of the AS160-ligand complexes was stable during the MD simulations. All these complexes have high RMSD values, indicating that the ligands left the binding poses as predicted. In stable complexes, a plateau at much smaller RMSD values is reached after an initial phase of equilibration. (**see Figure 4.12**).

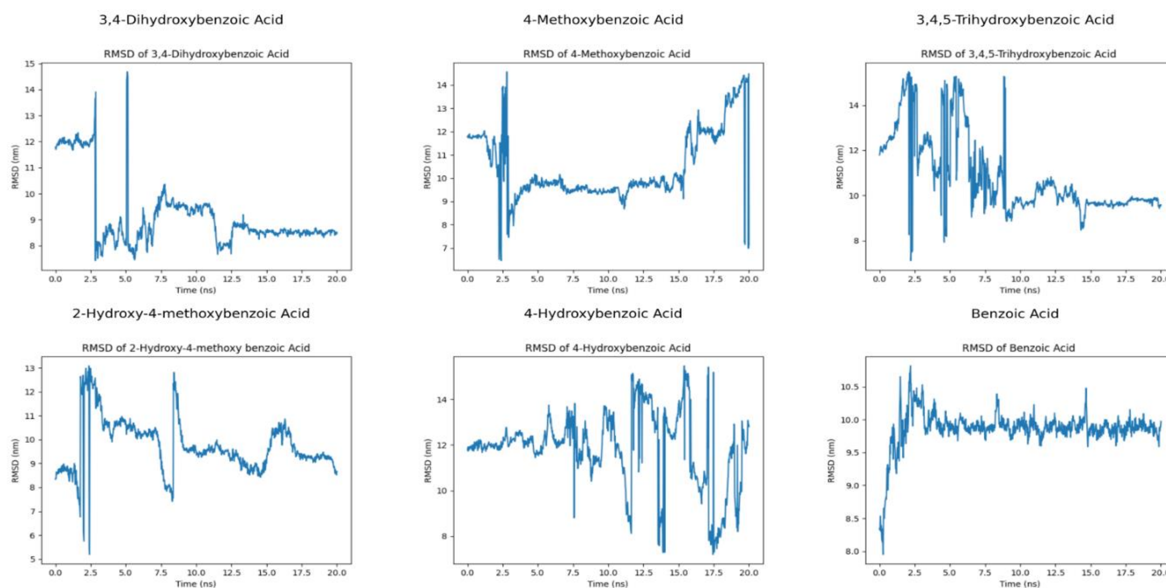

Figure 4.12: RMSD Analysis of AS160-Ligand Complexes.

RMSD: Measures average atomic deviations of the ligand with respect to its initial position obtained from docking, monitoring complex stability: lower values suggest stability, and higher values indicate instability vs Time (ns): X-axis represents simulation duration.

A closer look reveals that most of the complexes have RMSD peaks higher than 10 Å, indicating large structural deviations of the ligands from their initial conformation, especially those involving ligands such as 3,4-Dihydroxybenzoic Acid and 2-Hydroxy-4-methoxybenzoic Acid. Such deviations show that the ligands are not forming stable interactions with the target protein, hence the instability of these complexes.

These results suggest that further molecular dynamics analysis is unlikely to yield more useful information. Any conclusions based on such data would be flawed if the ligands were changing conformation significantly or were not binding appropriately. Additional analysis is not needed at this stage; instead, the MD simulations suggest that the phytochemical under study does not bind to AS160.

While the current simulations revealed unstable complexes, several things might be done in future, in particular revisiting some docking parameters that were applied. Besides considering different ligand conformations or various binding positions might also allow alternative interpretation of modes by which the ligand interacts with the protein. An alternative would be to look for better binding ligands, including the design of new ligands.

The RMSD analysis, in summary, has illuminated the instability of the ligand-protein complexes for AS160 and hence suggests that the phytochemicals under study are not able to modify AS160 function.

Chapter Five: Discussion

5.1 Molecular Docking

5.1.1 Results Overview

Protein-ligand docking provides so-called docking scores, which can be interpreted as binding free energy values that help to describe the binding strength and stability of the protein-ligand complex; with lower binding free energy values indicating stronger binding affinity and stable interaction. The docking results for PTEN and AS160 when interacting with the six phytochemicals show a significant difference in binding free energies. Results indicate that PTEN has a stronger binding free energy to the six phytochemicals under study than AS160, which suggests that the six phytochemicals can interact more effectively with PTEN.

3,4-Dihydroxybenzoic acid (-7.4 kcal/mol) and 2-Hydroxy-4-Methoxybenzoic acid (-7.4 kcal/mol) are the most promising phytochemicals due to their highest binding free energies when interacting with PTEN and suggest that these phytochemicals could be a good inhibitor for PTEN, which should be further studied.

However, AS160 docking results show lower binding free energies with the six phytochemicals. 3,4-Dihydroxybenzoic acid has the strongest binding free energy with -4.3 kcal/mol free binding energy, which indicates weaker binding when compared to PTEN. Benzoic acid had the weakest binding free energy with both AS160 and PTEN with binding free energy of -3.6 kcal/mol and – 6.5 kcal/mol, respectively. Benzoic acid has a simple structure with no extra functional groups other than carboxylic acid group, which suggests that the functional groups of the other phytochemicals under study have key roles in the current protein-ligand interactions.

The PTEN docking result indicates that PTEN binding pockets can interact with the phytochemicals strongly, potentially because of the type and arrangement of the residues present in its structure. The binding site in AS160 compared to PTEN seems to be less optimal in this respect, which accounts for its weaker binding interactions. The binding free energy shows that the six phytochemicals interact with both AS160 and PTEN, but only some of the compounds are expected to achieve biological effects because of their strong and stable binding properties.

5.1.2 Interaction Patterns with AS160

In the AS160 binding site, the six phytochemicals adopt different poses but interact with common binding residues within the binding site. Even though the phytochemicals have common binding residues, they have different binding free energies, which likely result from differences in binding depth and orientation of the ligands.

Leu-968, Gly-972, Arg-973, and Gln-984 are the common residues in the AS160 binding site that phytochemicals tend to interact with. These residues are involved in hydrogen bonding and hydrophobic interaction, therefore helping in stabilizing the protein-ligand complex. Hydrogen bond analysis demonstrated that 3,4-dihydroxybenzoic acid and 3,4,5-trihydroxy benzoic acid have multiple hydrogen bonds formed with the AS160 binding site, suggesting that these two ligands form the most strong and stable complex with AS160.

5.1.3 Interaction Patterns with PTEN

In the PTEN binding site, all six phytochemicals show a different orientation within the binding site like with AS160. There is a common binding residue within the PTEN binding site, such as Asp-92, Lys-128, Arg-130, and Gln-171 that are responsible for supporting protein-ligand complex stability and strength through forming hydrogen bonds and hydrophobic interactions with ligands.

Like AS160, hydroxylated or/and methoxylated ligands tend to make multiple hydrogen bonds with the PTEN binding site, which supports the suggestion that the functional groups in phytochemical enhance protein-phytochemical interaction through hydrogen bonding or hydrophobic interaction with a key residue in the binding site.

5.2 Molecular Dynamic Simulation

5.2.1 Apo Proteins simulations

To ensure the intrinsic stability of the proteins, MD simulations were performed for AS160 and PTEN in apo forms (without ligands). RMSD analysis of these simulations shows that PTEN is more flexible than AS160, with RMSD values of 0.2 nm for AS160 and 0.3 to 0.4 nm for PTEN, in other words, AS160 has a more rigid structure than PTEN. RMSF analysis presents that both proteins have a stable structure, but PTEN has higher fluctuations, and these fluctuations are found in loop regions within the protein. Furthermore, Rg analysis also supports RMSD and RMSF

findings and presents that PTEN has a less compact structure and allows conformational flexibility more than AS160.

5.2.2 Insight into PTEN-phytochemical interactions

Studying PTEN-phytochemicals interaction after MD simulation will be helpful in identifying ligand-induced stability changes. After the MD run, RMSD analysis of the ligands demonstrated that 3,4,5-Trihydroxybenzoic acid and 2-Hydroxy-4-Methoxybenzoic acid form the most stable complexes with PTEN, with low RMSD values. Other phytochemicals show high RMSD values suggesting unstable complexes. The RMSF analysis suggested that PTEN has differential flexibility of conformations with various ligand bindings. Some of the ligands, such as 3,4,5-trihydroxybenzoic acid and 2-hydroxy-4-methoxybenzoic acid, show moderate to high PTEN residue fluctuations, indicating conformational flexibility due to localized ligand binding; in contrast, benzoic acid binding results show lower RMSF values, suggesting a more rigid interaction. This implies that functionalized ligands enable PTEN to take on more flexible conformations potentially enhancing its binding adaptability whereas more simple ligands like benzoic acid stabilize the protein in a rigid state.

5.2.3 Insight into AS160-phytochemical interactions

MD simulations of the AS160-ligand complexes demonstrated an unstable interaction, conversely to PTEN-ligands interactions, which can be seen from high RMSD values for the ligands. 3,4-Dihydroxybenzoic acid and 2-Hydroxy-4-Methoxybenzoic acid resulted in RMSD values higher than 10 Å, which present unstable and weak interactions.

These results suggest that the AS160 protein does not have the right binding pocket for these ligands and there is a need to look for either other binding sites with AS160 or other phytochemicals. After the instability seen in the RMSD values, no further analysis of the MD simulations of the AS160-ligand complexes was performed as it would have been useless; instead, a novel study to refine the ligands or looking for alternative binding sites within AS160 is suggested.

5.2.4 Hydrogen bond analysis

The MD simulation analysis revealed that ligands with hydroxyl or methoxy groups such as 3,4-Dihydroxylbenzoic acid and 2-Hydroxy-4-Methoxybenzoic tend to form multiple hydrogen bonds in comparison with ligands without such functional groups such as benzoic acid, which suggests that these functional groups play a critical role in hydrogen bonding and therefore support complexes stability and strength. To further study and confirm this suggestion, ligands with hydroxyl functional groups were stripped off these groups, i.e., dehydroxylated, to compare the number of hydrogen bonds formed between protein and ligands with and without hydroxyl group. The results support the assumption because the number of hydrogen bonds is critically reduced after dehydroxylation. So, the presence of this functional group has a key role in hydrogen bond formation and complex stabilization.

5.3 Future Directions

As the results show that 3,4-Dihydroxylbenzoic acid, 3,4,5-Trihydroxybenzoic acid, and 2-Hydroxy-4-Methoxybenzoic acid are promising phytochemicals for binding to PTEN, further studies on them as potential therapeutic drugs that target PTEN in the insulin signaling pathway for diabetes treatment is encouraged. Considering that PTEN is involved in many signaling pathways, such as in cell proliferation and other metabolic pathways, these phytochemicals could also be useful to develop other drugs to treat cancer and metabolic disorders.

For AS160, future studies should focus on refining the ligands or trying other phytochemicals with AS160. Furthermore, the possibility of an alternative binding site in AS160 should be tested using structure-based drug design or enhanced MD simulation methods.

5.4 Conclusion

Overall, molecular docking and MD simulation studies helped in understanding the binding interactions between AS160 and PTEN with the selected phytochemicals. The results revealed a strong and stable interaction between several of these ligands and PTEN. Conversely, the AS160-ligands complexes showed an unstable and weak interaction. These results indicated that further studies in PTEN as a therapeutic target and the promising phytochemicals as inhibitors to treat diabetes are promising and worthwhile. However, studies on AS160 and the current ligands revealed that shift in focus towards ligand improvement or alternative AS160 binding sites is required.

References

Ahmed, M., Maldonado, A. M., & Durrant, J. D. (2023). From byte to bench to bedside: Molecular dynamics simulations and drug discovery. *BMC Biology, 21*(1), 299. https://doi.org/10.1186/s12915-023-01791-z

Akhtar, A., & Sah, S. P. (2020). Insulin signaling pathway and related molecules: Role in neurodegeneration and Alzheimer's disease. *Neurochemistry International, 135*, 104707. https://doi.org/10.1016/j.neuint.2020.104707

American Diabetes Association. (2023). Standards of medical care in diabetes—2023. *American Diabetes Association*. https://doi.org/10.2337/dc23-S001

Aneed, I. A., Ali, L. S., & Daragh, M. A. (2024). Cytological method in diagnosing diabetes: A qualitative analysis of oral cavity epithelium. *Journal of Medical and Clinical Research, 12*(2), 45–56. https://doi.org/10.15226/2374-6890/6/4/001139

Aneed, N., Ali, A., & Daragh, M. (2024). New cytological method to diagnose diabetes mellitus from oral mucosa. *Journal of Bioscience and Applied Research, 10*(2), 228–234. https://doi.org/10.21608/jbaar.2024.357660

de Boer, I. H., Bangalore, S., Benetos, A., Davis, A. M., Michos, E. D., Muntner, P., Rossing, P., Zoungas, S., & Bakris, G. (2017). Diabetes and hypertension: A position statement by the American Diabetes Association. *Diabetes Care, 40*(9), 1273–1284. https://doi.org/10.2337/dci17-0026

Hanhineva, K., Törrönen, R., Bondia-Pons, I., Pekkinen, J., Kolehmainen, M., Mykkänen, H., & Poutanen, K. (2010). Impact of dietary polyphenols on carbohydrate metabolism. International Journal of Molecular Sciences, 11(4), 1365–1402. https://doi.org/10.3390/ijms11041365

Hilal, Z., Antonescu, C. N., Randhawa, V. K., & Klip, A. (2008). Insulin action on glucose transporters through molecular switches, tracks, and tethers. *Biochemical Journal, 413*(2), 201–215. https://doi.org/10.1042/BJ20080723

International Diabetes Federation. (2021). *IDF diabetes atlas* (10th ed.). Retrieved February 15, 2025, from https://diabetesatlas.org/

Knowler, W. C., Barrett-Connor, E., Fowler, S. E., Hamman, R. F., Lachin, J. M., Walker, E. A., & Nathan, D. M. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine, 346*(6), 393–403. https://doi.org/10.1056/NEJMoa012512

Latva-Rasku, A., Hirvonen, M.-J., Stancáková, A., Koistinen, H., Kuusisto, J., Guan, L., et al. (2017). A partial loss-of-function variant in AKT2 is associated with reduced insulin-mediated glucose uptake in multiple insulin-sensitive tissues: A genotype-based callback positron emission tomography study. *Diabetes, 67*(2), 334–342. https://doi.org/10.2337/db17-1142

Lee, J. O., Yang, H., Georgescu, M. M., Di Cristofano, A., Maehama, T., Shi, Y., Pandolfi, P., & Pavletich, N. P. (1999). Crystal structure of the PTEN tumor suppressor: Implications for its phosphoinositide phosphatase activity and membrane association. Cell, 99(3), 323–334. https://doi.org/10.1016/S0092-8674(00)81663-3

Machado, M. R., & Pantano, S. (2020). Split the charge difference in two! A rule of thumb for adding ions in MD simulations. Journal of Chemical Theory and Computation, 16(3), 1367–1372. https://doi.org/10.1021/acs.jctc.9b00953

Mitophagy: A potential therapeutic target for insulin resistance. (2022). *Frontiers in Physiology, 13*, Article 957968. https://doi.org/10.3389/fphys.2022.957968

Patel, D. K., Prasad, S. K., Kumar, R., & Hemalatha, S. (2012). An overview on antidiabetic medicinal plants having insulin mimetic property. Asian Pacific Journal of Tropical Biomedicine, 2(4), 320–330. https://doi.org/10.1016/S2221-1691(12)60032-X

Pawar, V., Sontakke, A., & Pawar, S. (2023). Interconnection between oxidative stress and type 2 diabetes mellitus. *Journal of Bioscience and Applied Research, 9*(4), 192–198. https://doi.org/10.21608/jbaar.2023.325677

Petersen, M. C., & Shulman, G. I. (2018). Mechanisms of insulin action and insulin resistance. *Physiological Reviews, 98*(4), 2133–2223. https://doi.org/10.1152/physrev.00063.2017

Ramasubbu, K., & Devi Rajeswari, V. (2023). Impairment of insulin signaling pathway PI3K/Akt/mTOR and insulin resistance induced AGEs on diabetes mellitus and neurodegenerative diseases: A perspective review. *Molecular and Cellular Biochemistry, 478*, 1307–1324. https://doi.org/10.1007/s11010-022-04587-x

Schlein, L. (2024, November 13). Study: More than 800 million have diabetes worldwide, many untreated. *VOA Learning English*. https://learningenglish.voanews.com/a/study-more-than-800-million-have-diabetes-worldwide-many-untreated/7869300.html

Richter, E. A., & Hargreaves, M. (2013). Exercise, GLUT4, and skeletal muscle glucose uptake. *Physiological Reviews, 93*(3), 993–1017. https://doi.org/10.1152/physrev.00038.2012

Sakamoto, K., & Holman, G. D. (2008). An emerging role for AS160/TBC1D4 and TBC1D1 in the regulation of GLUT4 traffic. *American Journal of Physiology-Endocrinology and Metabolism, 295*(1), E29–E37. https://doi.org/10.1152/ajpendo.90331.2008

Sapra, A., & Bhandari, P. (2023). Diabetes. In *StatPearls (Internet)*. StatPearls Publishing. Retrieved January 16, 2025, from https://www.ncbi.nlm.nih.gov/books/NBK551501/

Sapra, A., Vaqar, S., & Bhandari, P. (2019). Diabetes mellitus. In *StatPearls Knowledge Base*. StatPearls Publishing. https://www.researchgate.net/publication/339297854

Sano, H., Kane, S., Sano, E., Miinea, C. P., Asara, J. M., Lane, W. S., & Lienhard, G. E. (2003). Insulin-stimulated phosphorylation of a Rab GTPase-activating protein regulates GLUT4 translocation. Journal of Biological Chemistry, 278(17), 14599–14602. https://doi.org/10.1074/jbc.C300063200

Scripps Research. (2019, April 25). AutoDockFR (ADFR): Flexible receptor docking software. Retrieved from https://ccsb.scripps.edu/adfr/

Seong, M. (2024). Insulin resistance mechanisms, implications, and therapeutic targets. *Journal of Diabetic Complications & Medicine, 9*(5). https://doi.org/10.37421/2475-3211.2024.9.286

Selph, S., Dana, T., Bougatsos, C., Blazina, I., Patel, H., & Chou, R. (2015). Screening for abnormal glucose and type 2 diabetes mellitus: A systematic review to update the 2008 U.S. Preventive Services Task Force recommendation (Report No. 13-05190-EF-1). *Agency for Healthcare Research and Quality*. https://pubmed.ncbi.nlm.nih.gov/25973510/

Shanak, S., Bassalat, N., Barghash, A., Kadan, S., Ardah, M., & Zaid, H. (2022). Drug discovery of plausible lead natural compounds that target the insulin signaling pathway: Bioinformatics approaches. *Bioinformatics, 2022*, Article 2832889. https://doi.org/10.1155/2022/2832889

Shanak, S., Bassalat, N., Barghash, A., Kadan, S., & Zaid, H. (2022). Drug discovery of plausible lead natural compounds that target the insulin signaling pathway: Bioinformatics approaches. *BioMed Research International*, Article ID 123456. https://www.researchgate.net/publication/359380668

Thayer, S. M., Lo, J. O., & Caughey, A. B. (2020). Gestational diabetes. In *The National Center for Biotechnology Information (NCBI) Bookshelf*. https://www.ncbi.nlm.nih.gov/books/NBK551501

van Gerwen, J., Shun-Shion, A. S., & Fazakerley, D. J. (2023). Insulin signaling and GLUT4 trafficking in insulin resistance. *Biochemical Society Transactions, 51*(3), 1057–1069. https://doi.org/10.1042/BST20221066

World Health Organization. (2016). *Global report on diabetes*. World Health Organization. https://iris.who.int/handle/10665/204871

Yameny, A. A. (2024). Diabetes mellitus overview. *Journal of Basic and Applied Research, 12*(1), 1–10. https://jbaar.journals.ekb.eg/article_382794.html

Zaid, H., Tamrakar, A. K., Razzaque, M. S., & Efferth, T. (2018). Diabetes and metabolism disorders medicinal plants: A glance at the past and a look to the future. *Evidence-Based Complementary and Alternative Medicine, 2018*, Article 5843298. https://doi.org/10.1155/2018/5843298

الآليات الجزيئية التي تتحكم في إشارات الأنسولين: مناهج حاسوبية ومعلوماتية حيوية.

سوار ساطي أبو سيفين

أ.د. هلال زيد

أ.د. برجيت سترودل

د. صبا شنك

ملخص

الخلفية: يُعد مرض السكري ومقاومة الأنسولين من القضايا الصحية العامة العالمية الرئيسية، حيث يؤثران بشدة على الرفاهية والصحة العامة. لا يمكن تحديد أهداف علاجية جديدة لهذه الأمراض إلا من خلال فهم أعمق للآليات الجزيئية الخاصة بإشارات الأنسولين. لقد وُجد أن المواد الكيميائية النباتية النشطة بيولوجيًا (Phytochemicals) تؤثر على مسارات إشارات الأنسولين. ويُعد احتمال تداخل هذه المركبات مع بروتينات حيوية رئيسية في مسار الإشارة مثل AS160 وPTEN نافذة واعدة للتدخل العلاجي. ويمكن أن توفر الطرق الحاسوبية والمعلوماتية الحيوية تقارير مفيدة جدًا حول هذا النوع من التفاعلات الجزيئية.

الهدف: تهدف الدراسة إلى فهم التفاعلات الجزيئية بين ست مركبات نباتية مختارة وبروتينات رئيسية في مسار إشارات الأنسولين (AS160 وPTEN). ويكمن الهدف في دراسة إمكانيات هذه المركبات كأهداف علاجية محتملة لمقاومة الأنسولين والسكري من خلال المحاكاة الحاسوبية والدراسات الجزيئية الالتحاميه .

المنهجية: أُستخدم نهج حاسوبي يشمل المحاكاة الجزيئية للالتحام الجزيئي (Molecular Docking) ومحاكاة الديناميكيات الجزيئية (MD) للتنبؤ بألفة الارتباط (طاقة الارتباط الحرة) وأنماط التفاعل بين ست مركبات نباتية مختارة (حمض 3,4-ديهيدروكسي بنزويك، حمض 4-ميثوكسي بنزويك، حمض 3,4,5-تراي هيدروكسي بنزويك، حمض 2-هيدروكسي-4-ميثوكسي بنزويك، حمض 4-هيدروكسي بنزويك، وحمض البنزويك) مع بروتينات AS160 و PTEN. تم الحصول على الهياكل ثلاثية الأبعاد للمركبات من قاعدة بيانات Pub Chem ، بينما تم تحميل الهياكل البلورية للبروتينين من بنك بيانات البروتين (PDB) بالمعرفات 1D5R لـ PTEN و3QYB لـ AS160. أُستخدم برنامج Auto Dock FR لإجراء محاكاة الالتحام من أجل تحديد أنماط وطاقة ارتباط المجمعات بين البروتينات والمركبات. كما أُستخدم برنامج GROMACS لمحاكاة الديناميكيات الجزيئية لتقييم استقرار ومرونة هذه المجمعات على مدى فترة زمنية معينة.

النتائج: كشفت تجارب الالتحام أن المركبات النباتية المختارة أظهرت ارتباطًا أقوى مع PTEN مقارنة بـ AS160. من بين المركبات المختبرة، أظهر حمض 3,4-ديهيدروكسي بنزويك وحمض 2-هيدروكسي-4-

ميثوكسي بنزويك أعلى طاقة ارتباط بلغت -7.4 كيلو كالوري/ مول. أظهرت محاكاة MD أن مجمعات PTEN مع المركبات كانت مستقرة، خصوصًا مع المركبين المذكورين، حيث كانت قيم RMSD منخفضة. أما مجمعات AS160 مع المركبات فلم تُظهر تفاعلات مستقرة وكانت قيم RMSD أعلى. تشير تحليلات الروابط الهيدروجينية إلى أن الاستقرار في المجمع يرجع إلى وجود مجموعات الميثوكسي والهيدروكسيل، حيث ينخفض الاستقرار بشكل كبير عند غياب الروابط الهيدروجينية.

الاستنتاج: ساعدت دراسات الالتحام الجزيئي ومحاكاة الديناميكيات الجزيئية في فهم تفاعلات الارتباط بين AS160 و PTEN مع المركبات النباتية المختارة. أظهرت النتائج وجود تفاعل قوي ومستقر بين عدة مركبات و PTEN. في المقابل، كانت مجمعات AS160 مع المركبات غير مستقرة وضعيفة. تشير هذه النتائج إلى أن إجراء المزيد من الدراسات على PTEN كهدف علاجي، والمركبات النباتية الواعدة كمثبطات لعلاج السكري، يُعد واعدًا وجديرًا بالاهتمام. ومع ذلك، فإن الدراسات المتعلقة بـ AS160 والمركبات الحالية تشير إلى الحاجة إلى تحسين المركبات أو البحث عن مواقع ارتباط بديلة في AS160.

الكلمات المفتاحية: السكري، Auto Dock FR،GROMACS ، الالتحام الجزيئي، الديناميكيات الجزيئية.