A DISSERTATION ON

Modulatory impact of Millet based bioactive compounds in disrupting the Advance Glycolated End-products (AGE) formation

SUBMITTED TO THE
DEPARTMENT OF BIOENGINEERING
FACULTY OF ENGINEERING & INFORMATION TECHNOLOGY
INTEGRAL UNIVERSITY, LUCKNOW

IN PARTIAL FULFILMENT
FOR THE
DEGREE OF MASTERS OF TECHNOLOGY
IN BIOINFORMATICS

BY
Amna Tamimi
M. Tech Bioinformatics (IV Semester)
Roll No: 2101081001

UNDER THE SUPERVISION OF

Dr. Parthasarathi Ramakrishnan
Principal Scientist
CSIR-IITR

INTEGRAL UNIVERSITY, DASAULI, KURSI ROAD
LUCKNOW- 226026

DECLARATION FORM

I, **Amna Tamimi**, a student of **M. Tech Bioinformatics** (II Year/ IV semester), Integral University have completed my six months dissertation work entitled **"**Modulatory impact of millet based bioactive compounds in disrupting the Advanced Glycolated End-products (AGE) formation**"** successfully from the **CSIR- Indian Institute of Toxicology Research, Lucknow** under theable guidance of **Dr. Parthasarathi Ramakrishnan, Principal Scientist.**

I, hereby, affirm that the work has been done by me in all aspects. I have sincerely prepared this project report and the results reported in this study are genuine and authentic.

Amna Tamimi

Course Coordinator

सीएसआईआर-भारतीय विषविज्ञान अनुसंधान संस्थान
CSIR-INDIAN INSTITUTE OF TOXICOLOGY RESEARCH
वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद् | COUNCIL OF SCIENTIFIC & INDUSTRIAL RESEARCH
(विज्ञान एवं प्रौद्योगिकी मंत्रालय, भारत सरकार | MINISTRY OF SCIENCE & TECHNOLOGY, GOVT. OF INDIA)

6th Sept, 2023

To whomsoever it may concern

This is to certify that **Amna Tamimi**, a student of M. Tech. Bioinformatics (II Year/ IV Semester), Integral University, Lucknow has completed a six-month (2nd January 2023 - 1st July 2023) training program under my guidance at the **CSIR-Indian Institute of Toxicology Research.** She has concluded a study on the topic "**Modulatory impact of millet based bioactive compounds in disrupting the Advanced Glycolated End-products (AGE) formation".** The thesis embodies the results of original work and studies carried out by the student herself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or anybody else from this or any other University/Institution.

During the training, I found her hardworking, enthusiastic and self-motivated. I wish her all the best for future endeavors.

Best Regards

7/9/2023

डॉ. रामाकृष्णन पार्थसारथी / Dr Ramakrishnan Parthasarathi
प्रधान वैज्ञानिक / Principal Scientist
कम्प्यूटेशनल टॉक्सिकोलॉजी फैसिलिटी
Computational Toxicology Facility
सेन्टर फॉर इनोवेशन एण्ड ट्रान्सलेशनल रिसर्च
Centre for Innovation and Translational Research
सीएसआईआर–भारतीय विषविज्ञान अनुसंधान संस्थान
CSIR-Indian Institute of Toxicology Research
विषविज्ञान भवन,31,महात्मा गांधी मार्ग,लखनऊ–226001, उ.प्र.(भारत)
Vishvigyan Bhawan,31,M.G.Marg,Lucknow-226001,U.P.(INDIA)

INTEGRAL UNIVERSITY
LUCKNOW - INDIA
Phone No.: +91(0522) 2890812, 2890730, 3296117, 6451039, Fax No.: 0522-2890809
Kursi Road, Lucknow-226026 Uttar Pradesh (INDIA)

CERTIFICATE BY INTERNAL ADVISOR

This is to certify that **Amna Tamimi**, a student of **M. Tech Bioinformatics** (II Year/IV Semester), Integral University has completed her six months dissertation work entitled **"Modulatory impact of millet based bioactive compounds in disrupting the Advanced Glycolated End-products (AGE) formation"** successfully. She has completed this work from **Computational Toxicology Facility, Council of Scientific and Industrial Research - Indian Institute of Toxicology Research (CSIR-IITR), Lucknow** under the guidance of **Dr. Parthasarathi Ramakrishnan, Principal Scientist**. The dissertation was a compulsory part of her **M. Tech Bioinformatics.**

I wish her good luck and bright future.

Dr. Mohammad Kalim Ahmad Khan
Professor
Department of Bioengineering
Faculty of Engineering & Information Technology

INTEGRAL UNIVERSITY
LUCKNOW - INDIA
Phone No.: +91(0522) 2890812, 2890730, 3296117, 6451039, Fax No.: 0522-2890809
Kursi Road, Lucknow-226026 Uttar Pradesh (INDIA)

TO WHOM IT MAY CONCERN

This is to certify that **Amna Tamimi,** a student of **M. Tech Bioinformatics** (II Year/ IV Semester), Integral University has completed her six months dissertation work entitled **"Modulatory impact of millet based bioactive compounds in disrupting the Advanced Glycolated End-products (AGE) formation"** successfully. She has completed this work from **Computational Toxicology Facility, Council of Scientific and Industrial Research – Indian Institute of Toxicology Research (CSIR-IITR), Lucknow** under the guidance of **Dr. Parthasarathi Ramakrishnan, Principal Scientist**. The dissertation was a compulsory part of her **M. Tech Bioinformatics.**

I wish her good luck and bright future.

Dr. Alvina Farooqui
Professor and Head
Department of Bioengineering
Faculty of Engineering & Information Technology

ACKNOWLEDGEMENTS

I am greatly indebted to the **Almighty Allah**, for giving me the strength to accomplish this work successfully.

Words cannot express my sincere gratitude towards my guide **Dr. Parthasarathi Ramakrishnan,** Principal Scientist, Computational Toxicology Facility, CSIR-IITR, for his tremendous help during my dissertation. Hiseffortless guidance and encouragement have helped me learn valuable lessons and gain knowledge about the field.

It gives me great pleasure to express my deep sense of gratitude to **Prof. S.W Akhtar** (Founder & Chancellor Integral University, Lucknow), **Dr. Syed Nadeem Akhtar** (Pro-Chancellor Integral University, Lucknow), **Prof. Javed Musarrat** (Vice Chancellor Integral University, Lucknow) and **Prof. (Dr.) Mohd. Haris Siddiqui** (Registrar Integral University, Lucknow) for giving me the opportunity to join this esteemed university and providing me with all the necessary facilities.

I express my deepest sense of gratitude towards **Dr. Alvina Farooqui** (Professor and Head, Department of Bioengineering, Faculty of Engineering, Integral University), **Prof. (Dr.) Mohd. Kalim Ahmad Khan** and **Prof.** (**Dr.) Salman Akhtar** for their excellent guidance and valuable advice during my post-graduation. I am extremely grateful to my teachers and classmates at the Department of Bioengineering, Integral University, Lucknow for their direction, support, and encouragement during my course.

I would like to extend my thanks to **Dr. Shraddha Pandit** (Senior Project Associate, CSIR-IITR), for her kind support and suggestion provided to me during the course of this study. I would also like to extend my gratitude to **Ms. Tanya Jamal** and **Ms. Anamta Ali** for their unconditional help, love and support. I truly loved their pleasant company, which kept me motivated all throughout my study. Last but not the least, I am deeply obliged to my family for their continuous support, love, and encouragement. I could not have accomplished this work without their help and support.

AMNA TAMIMI

LIST OF CONTENTS

LIST OF ABBREVATIONS

AGE	Advanced Glycation End-products
NF-B	Nuclear transcription factor
ADMET	Absorption, Distribution, Metabolism, Excretion
STAT	Signal Transducer and Transcriptional Activator
TAGE	Toxic Advanced Glycation End-products.
MD	Molecular Dynamics
RAGE	Receptor of Advanced Glycation End-products.
ROS	Reactive Oxygen Species
NADPH	Nicotinamide Adenine Dinucleotide Phosphate.
AD	Alzheimer's Disease
TBI	Traumatic Brain Injury
ALS	Amyotrophic Lateral Sclerosis
GA	Glyceraldehyde
AG	Aminoguanidine
SPC	Simple Point Charge
RMSF	Root Mean Square Fluctuation
RMSD	Root Mean Square Deviation
Rg	Radius of Gyration
BBB	Blood Brain Barrier
HBA	Hydrogen Bond Acceptor
HBD	Hydrogen Bond Donor
B.E.	Binding Energy

LIST OF FIGURES

LIST OF TABLES

1. Introduction

The advanced glycation end products (AGEs) are organic compounds with a wide range of structural and functional characteristics that are produced in all living organisms [Perrone *et al.,* 2020] . Toxic substances that are formed when protein or fat interacts with sugar in the bloodstream. This process is known as glycation. In other words, AGEs are the irreversible products produced by nonenzymatic glycation, which contain potentially hazardous heterogeneous compounds.

Nonenzymatic glycation occurs when a reducing sugar's reactive carbonyl group comes into contact with proteins, lipids, or nucleic acids [Rungratanawanich W. *et al., 2021]*. AGE structures come in a wide variety, which points to a wide range of formation methods. In fact, AGEs can be produced from a variety of precursors through a number of different mechanisms, making them both endogenous and exogenous [A.Twarda-Clapa *et al.,* 2022]. The nuclear transcription factor (NF-B), mitogen-activated protein kinase (MAPK), and signal transducer and transcriptional activator (STAT) pathways are only a few of the many signaling pathways that AGEs control. Some metabolic disorders, such as diabetes, atherosclerosis, and Alzheimer's disease, become more severe by the pathogenic function that AGEs play in the development of these conditions [Q. Song *et al.,* 2021]. A subclass of AGEs known as toxic AGEs (TAGEs) interacts with the AGE receptors on specific cells and organs to cause harmful consequences [M. Takeuchi *et al.,* 2010]. Toxic AGEs (TAGE) are the name given to these AGEs because they are highly cytotoxic [M. Takeuchi *et al.,* 2020]. When intracellular proteins and glyceraldehyde react, toxic advanced glycation end-products (toxic AGEs, TAGEs) are produced, and their formation is a contributor to a number of cellular diseases.[M. Takeuchi *et al.,* 2022], such as pathogenesis of diabetes mellitus and its related vascular complications, also brain, liver, and heart cells are harmed by TAGE buildup [M. Takeuchi *et al.,* 2020][M. Takeuchi *et al.,* 2016] .

Benfotiamine, aminoguanidine, aspirin, tenilsetam, thiazolidinediones, antioxidant agents (carnosine, flavonoids, curcumin), and chelators with AGE inhibition properties (carnosine, pyridoxamine) are just a few of the many substances that have been reported to have a potential role against AGEs. The anti-AGEs efficacy of these substances was shown by preclinical evaluation study results; however, a precise mechanism of action has not yet been identified. Clinical trials are still necessary to confirm many of these drugs' effects [Chen *et al.,* 2018]. Thiazolidinediones have a number of unfavorable side effects, especially when used for an extended period of time. For example, edema, congestive hearts failure, weight gain, fractures, diabetic macular edema, and other conditions [JS Eggleton *et al.,* 202]. Anemia, liver abnormalities, gastrointestinal problems, influenza, lupus, and flu-like symptoms were among the adverse effects of aminoguanidine that were often reported. Considering the toxic and

potential side effects of these synthetic molecules, natural compounds are preferable because of their immense scaffold variety and structural complexity.

Natural remedies have long been available, with many effective medicines coming from plant sources. In the past, natural compounds and their structural analogues have significantly influenced pharmacology, particularly for the treatment of infectious disorders and many chronic diseases like cancer . Currently, several of these natural substances are being researched a spossible medication candidates. Natura lproducts remain to be a possible source of medicine despite all technological improvements made in the creation of synthetic pharmaceuticals because they possess a unique structural diversity that synthetic drugs do not, as compared to ordinary combinatorial chemistry [Dias *et al*.,2012]. They typically have more H-bond acceptors and donors, more sp^3 carbon and oxygen atoms than synthetic compounds, higher molecular masses, lower calculated octanol-water partition coefficients, and more rigid molecules which make them distinguishable.

Secondary plant metabolites known as phenolic compounds have been researched for their anti-oxidant and anti-glycation abilities [MA. Anis *et al.,* 2020]. The natural substance that we work with in the present study as an inhibitor of advanced glycation end products are millets. As millets are rich in nutrients, containing compounds like phenolics that possess anti-inflammatory, antioxidant, and anti-glycation activity, by neutralising carbonyl intermediates during the glycation reaction and chelating transition metals. These millet based bioactive compounds may provide protection by scavenging free radicals, hence preventing the production of AGEs [MA. Anis *et al.,* 2020]. The millet based bioactive compounds were collected from available databases and were subjected to high throughput screening to filter drug-like compounds which structurally mimics the standard drug . Further, the molecular docking was performed for control drugs and derived natural compounds in the binding site of target protein. The validation of the compounds efficacy was done using simultaneous docking with the target in the presence of TAGE. Later, the acquired results were validated by MD simulation study.

Objectives

1. The focal point of this study is to use Millet based bioactive compounds to inhibit the formation of Advanced Glycation End (AGE) products. With the help of literature evidence, target protein involved in AGE formation and millet based bioactive compounds were selected.
2. Predicting the natural compounds kinetics, drugability, toxicity and pharmacokinetic properties using prediction tools. Pharmacophore analysis, Comparative molecular interaction study and simultaneous docking of controls and natural compounds with the target in presence of TAGE was done.
3. Later, MD simulation assesses and analyzes the interaction of hits with the target receptor in comparison to control and TAGE compounds.

2. Review of literature

One can easily distinguish between two types of AGE physiological implications. (1) AGEs encourage the cross-linking of proteins and can serve as molecules that connect various peptide chains in proteins, which can lead to protein aggregation or increased tissue stiffness and a loss of function in both tissues and proteins. This crucial aspect of AGEs' physiological function is frequently explored in relation to ageing, skin conditions, neurological illnesses, diabetes-related cardiovascular disease, and chronic renal disease.(2) Protein-bound AGEs have the ability to directly enhance reactive oxygen species (ROS) formation and have an impact on cell physiology by activating cell membrane receptors such RAGE, OST-48 (OST-48, also known as AGER1), 80 KH protein (AGER2), and galectin-3 (AGER3) [A.J. Lin *et al*., 2018].

Proteins or lipids that have been nonenzymatically glycated by glucose, or other reducing sugars and their derivatives, such as glyceraldehyde, glycolaldehyde, methylglyoxal, and acetaldehyde, are known as advanced glycation end products (AGEs).A number of diseases, including diabetes and its complications, retinopathy, neuropathy, neurological disorders (such as Parkinson's disease and Alzheimer's disease), atherosclerosis, hypertension, cancers and various other chronic diseases that are influenced by the pathological mechanisms that AGEs perform.Maillard reactions,lipid peroxidation and the polyol pathway are the three main methods of AGE generation [Kuzan *et al.*, 2021].

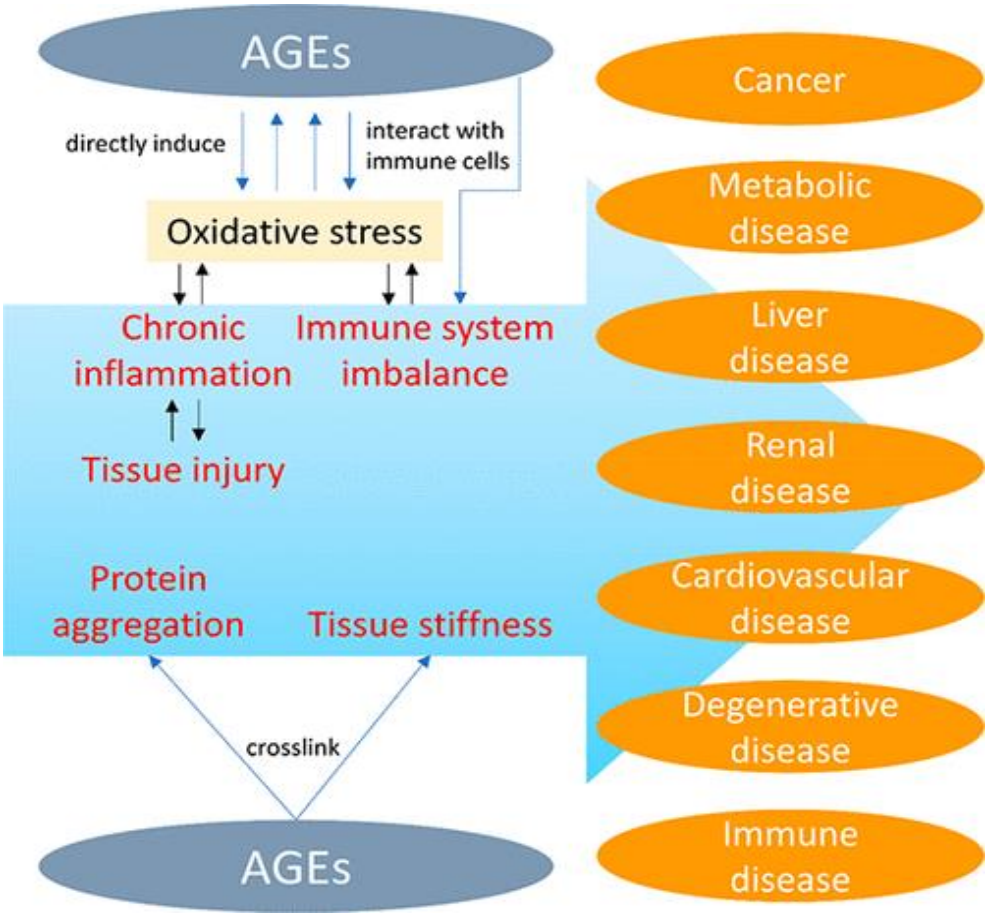


Figure 1. *AGE induced biological effects and co-related diseases. (Lin et al., 2018)*

The formation and accumulation of advanced glycation end products (AGEs) in the kidney is worsened by protracted hyperglycemia, dyslipidemia, and oxidative stress. The nonenzymatic Maillard reaction, sometimes known as the "browning reaction," results in AGEs when amino groups from proteins, lipids, or nucleic acids interact with aldehyde groups from sugars and carbonyls. The Amadori product is the first stable byproduct of this reaction. AGEs are created as a result of its subsequent oxidation, fragmentation, rearrangement, and dehydration.The fact that practically all proteins have some AGE changes as a means of indicating molecular age, however, is due to the fact that glucose levels must always be present to support brain function. [MC *et al.*, 2011]. The primary source of vascular oxidative stress in Type 2 diabetes, nicotinamide adenine dinucleotide phosphate (NADPH) oxidase, is expressed and activated at higher levels in human endothelial cells as a result of AGEs.Increased NADPH oxidase activity causes the cellular antioxidants catalase, glutathione, glutathione peroxidase, superoxide dismutase, and glutathione to become depleted, resulting in the production of free radicals.Numerous reactive alpha and beta unsaturated aldehydes are produced by free radicals as a result of the lipid peroxidation of polyunsaturated fatty acids such as glucose are capable of producing advanced lipid peroxidation end products (ALE), which have the potential to serve as glycemic status markers [Dias *et al.*, 2014].

In a hyperglycemic state, the polyol route for glucose metabolism is active, which promotes the generation of AGEs. Aldose reductase, the first and rate-limiting enzyme in this pathway, uses NADPH as the electron donor to catalyze the conversion of glucose into sorbitol. Then sorbitol dehydrogenase converts NAD to NADH and sorbitol into fructose at that point.In the non-enzymatic glycosylation, fructose produced as a byproduct can be phosphorylated into fructose-3-phosphate and subsequently degraded into 3-deoxyglucosone, both of which can be employed as precursors to make AGEs [Kang *et al.*, 2022].

Several diseases are prone to developing or worsening as a result of glycoxidation processes that occur under oxidative stress including diabetes, atherosclerosis, and Alzheimer's disease (AD) as well as in the secondary stages of traumatic brain injury (TBI).The deactivation of numerous enzymes by AGEs, which take the form of intra- and interprotein cross links, accelerates the development of medical condition.Additionally, the downstream inflammatory cascade events are produced by AGEs interactions with their receptors.Particularly in cases of Alzheimer's disease and other neurodegenerative disorders, such as traumatic brain injury (TBI) and amyotrophic lateral sclerosis (ALS), the overexpression of AGE receptors interacts [Reddy *et al.*, 2022].

Further interest in the potential health effects of AGEs was sparked by a number of papers showing an increase in the levels of circulating AGEs in persons with diabetes and chronic renal

disease. [Prasad *et al.*, 2019]. This emerged after a large number of empirical studies showing a connection between AGEs and a number of conditions, including memory loss as we age, the pathophysiology of eye diseases, polycystic ovary syndrome, wound healing, cardiovascular issues, bone health, periodontitis, erectile dysfunction, anaemia in older community-dwelling women, slow walking speed in older adults, peripheral neuropathy, peripheral artery disease, obstructive sleep apnea , islet β-cell dysfunction, cancer, elevated cellular oxidative and inflammatory state, schizophrenia, Alzheimer's disease, and risk for metabolic syndrome in adults and children.

The excessive generation of glyceraldehyde (GA), an intermediary of sugar metabolismin neuronal cells, hepatocytes, and cardiomyocytes is caused by the frequent consumption of high amounts of sugar, which has been linked to the onset/progression of lifestyle-related disease.Toxic advanced glycation end-products (toxic AGEs, TAGE) a subset of AGE, are created when GA reacts with intracellular proteins, accumulating various diseases.Through the receptor for AGEs, the cellular leakage of TAGE impacts the nearby cells,consequently encouraging the development of lifestyle-related diseases.TAGE seeped into the extracellular space, raising the amounts of extracellular TAGE in the circulating fluids, and the intracellular build-up of TAGE caused a variety of cellular diseases [Takeuchi*et al.,* 2021]. Only a few numbers of synthetic AGE inhibitors work by interfering with the first interaction between reducing sugars and amino groups to prevent the development of Schiff bases and AGEs in the glycation process. For instance, acetylsalicylic acid, an ingredient in aspirin, has been shown to block the glycation process by acetylating free amino groups in proteins by preventing the reducing sugars from attaching. Whereas most synthetic inhibitors play a major role in the latter stage of glycation by blocking the formation of Amadori products. A lot of research has been done on aminoguanidine (AG) and pyridoxamine, which are thought to be powerful carbonyl scavengers, as conventional AGE inhibitors. Even though these synthetic compounds show strong and powerful inhibitory effects, they can also cause severe side effects like gastrointestinal disturbance, anaemia, and flu-like symptoms which caused safety concerns that led to the termination of ACTION II clinical trial [Peng *et al.*, 2011].

As natural compounds are comparatively safer than the synthetic compounds for human consumption, it attracted more interests from researchers. These compounds also possess good inhibitory activity against AGE formation and appear to have lesser toxic effects. Recently, there has been an increase in interest in searching natural substances for fresh glycation inhibitors. [*Song Q, Liu J, Dong L, Wang X, Zhang X*, 2021], [*Peng, X., Ma, J., Chen, F., & Wang, M,* 2011]. Polyphenols, polysaccharides, terpenoids, vitamins, alkaloids, and peptides are the six types of

naturally occurring substances that may potentially prevent the production of AGEs based on their structural characteristics. Due to the diversity in the structural and functional unit of natural compounds, they suppress AGE development through a variety of different ways. The seven groups of mechanisms that prevent AGE formation include blocking aldose reductase, protecting protein glycation sites, scavenging oxidative free radicals, controlling AGE receptors, trapping active dicarbonyl compounds, chelating metal ions, and lowering blood sugar levels [*Song Q, Liu J, Dong L, Wang X, Zhang X*, 2021].

Millets are reported to contain phenols with potent anti-crosslinking and anti-glycation effects on collagen along with several biological processes, including those that are anti-oxidative, anti-hypertensive, anti-tumor, and that prevent the production of AGEs.Additionally, small millets phenolic compounds showed significant suppression of two important postprandial hyperglycaemia-related enzymes (amylase and glucosidase) as well as antioxidant capabilities in several routes [Anis *et al.*, 2020]. Millets may include protein glycation inhibitors that are much better at preventing the production of AGE and protein glycoxidation while also offering advantages over synthetic medicines without having negative side effects.

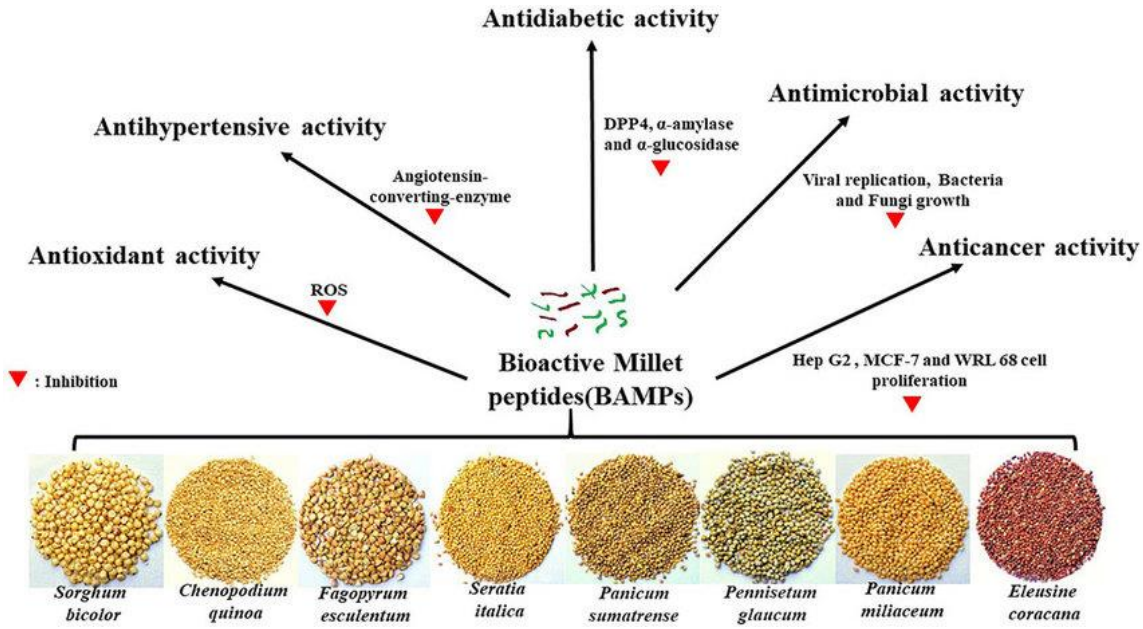


Figure 2. *Health promoting biological activities of bioactive millet . (A.Majid and P. Priyadarshani., 2020)*

3. Materials and Methods

3.1 Tools used

RCSB PDB –

RCSB Protein Data Bank (RCSB PDB) **(** http://www.rcsb.org/ **)** enables advances in science and education by giving access as well as tools for exploration, visualization, and analysis of: experimentally determined 3D structures from the Protein Data Bank (PDB) archive Computed structure models (CSM) from AlphaFold DB and ModelArchive. Examining this data in light of external annotations provides a structural view on biology. The Protein Data Bank (PDB) is a crucial tool for study and instruction in basic biology, health, energy, and biotechnology. The RCSB PDB is the first open access digital data source in biology and medicine and the US data center for the PDB archive. Through an online information portal and a downloadable data archive, PDB provides access to details about the 3D structures of the molecules that make up life. Since the PDB contains a large amount of 3D structure data and has made significant advancements in protein architecture, protein structure prediction is becoming more and more crucial. The procedure has been accelerated using AI and deep learning methods. As a result, it is now able to analyze huge datasets and more quickly find new therapeutic targets. These techniques have been shown to drastically reduce the time and resources needed for protein structure prediction. On top of that, novel protein folds and functions that were previously unknown have been found thanks to the application of AI and deep learning [W. Tian *et al.,* 2018].

PubChem –

A component of the Molecular Libraries Roadmap Initiatives of the US National Institutes of Health (NIH), PubChem **(**https://pubchem.ncbi.nlm.nih.gov**)** was established in 2004 as a public archive for data on chemical compounds and their biological activity. The scientific research community can now access a substantial system called PubChem that provides chemical information. Substance, Compound, and BioAssay are three databases that are connected together in PubChem. Both the Compound database and the Substance database store distinct chemical structures that have been taken from the former [S. Kim *et al.,* 2016]. The Substance database houses chemical information that has been deposited by individual data contributors to PubChem. The BioAssay database holds the biological activity information of chemical compounds that have been examined in assay experiments.

SWISS ADME –

Analysing a vast number of molecular structures in order to identify those that have the best chance of treating patients is a step in the process of developing new medications. Pharmacokinetics, which includes the ADME characteristics (Absorption, Distribution, Metabolism, and Excretion) can each be evaluated individually using specialized methodologies, determines how a therapeutic chemical will behave in the body. By estimating early ADME in the discovery phase, the percentage of pharmacokinetics-related failure in the clinical phases is reduced. Computer models have been advocated as an effective substitute for experimental techniques for the prediction of ADME. Chemicals that are probably to be unstable, reactive, poisonous, or interact with biological experiments are removed from chemical libraries using substructure searches and physicochemical parameters. The molecular fingerprint (FP) and the FP2 technique, two molecular descriptors created by cheminformaticians and utilized in classification models for ADME behaviors, are two examples of the molecular descriptors. For the use of machine learning technology, CADD is a pioneering field. An easy way to submit and analyse CADD results is by using the free web tool SwissADME**(** http://www.swissadme.ch.). It has cutting-edge methods and tools, as well as a number of input options, multi-molecule computing, and the capability to present the findings [H.M. Berman *et al.,* 2000].

Discovery Studio -

The Discovery Studio Visualizer Client is a graphical user interface for accessing Discovery Studio Science. **(**https://mybiosoftware.com/**).** This free, user-friendly visualisation tool is the ideal choice for managers and academics who need to engage with modellers but do not need access to the expert-level analytical tools in Discovery Studio [D. Studio., 2008]. It is possible to automate and customise common modelling processes using the Perl-based scripting API, which provides a large selection of common jobs and commands. Various formats that adhere to industry standards are available to view and distribute data on proteins and small compounds. It offers unequalled abilities for collaborating on workflows, sharing data, and utilising computational resources. Additionally, it offers scriptable functions like surface diffraction, chirality and valency checks, secondary structure prediction, access to electrostatics tools, and management of constraints and restrictions. Protein modelling, pharmacophore analysis, and structure-based design may all be easily integrated with one another using the open operating system SciTegic Pipeline PilotScitegic Enterprise Server platformTM, which is the foundation of Discovery Studio.

Autodock Vina –

The inventor and developer of Autodock Vina is Dr. Oleg Trott, of the Molecular Graphics Lab at The Scripps Research Institute (now CCSB). The current version of Autodock Vina is 1.2.0. **(**https://github.com/ccsb-scripps/AutoDock-Vina**).** The input and output file formats used by Vina and AutoDock are both PDBQT molecular structure files. PDBQT files can be produced using MGLTools (manually or interactively), and they can also be viewed. Additional files like the AutoDock and AutoGrid parameter files (GPF, DPF), grid map files, and so forth are not required.Vina's design philosophy avoids asking the user to comprehend its implementation specifics, adjust exotic search parameters, cluster results, or have knowledge of complex mathematics (quaternions). All that is required is a description of the search space, including the binding site, and the structures of the molecules being docked. Grid map computation and atom charge assignment are not required. As it can utilise many CPUs or cores to speed up operation, Vina is quicker and more effective than AutoDock 4. Now that numerous ligands may be docked simultaneously in Vina, fragment-based drug design may benefit from this feature. High-throughput virtual screenings can be accelerated with AutoDock Vina 1.2.0's batch processing and Python bindings. The command line interface or Python can be used to access its simultaneous multiple-ligand docking functionality against a single target structure [J. Eberhardt *et al.,*2021].

With the aid of unbound and MD simulations, homology modelling, and other methods, molecular docking is used to predict the noncovalent binding of macromolecules and ligands. To uncover possible leads for drug development, docking is used to screen digital libraries of compounds that resemble drugs. It is crucial to improve precision while decreasing the amount of time it takes because it uses a lot of computer resources. The need for manually choosing the atom types for grid maps, making grid map files with AutoGrid, choosing the "search parameters," and clustering the results after docking is also eliminated by Vina because it calculates its own grid maps quickly and automatically and does not store them on the disc. It also clusters and ranks the results without providing any intermediate information to the user. [Trott and Olson, 2010].

R Studio –

Visualising and analysing multivariate data are the objectives of principal component analysis. PCA only works with quantitative variables, in contrast to the other approaches. In addition to economics, image processing, healthcare, and security, there are many other applications for principal component analysis in daily life. PCA can help identify important variables that contribute to the variation in data, making it a useful tool for decision-making in a variety of

industries. Through R Studio **(**http://www. rstudio.com**)**, we can create various PCA plots, including Biplot, Screeplot, and Boxplot. Large datasets' dimensionality can also be decreased using PCA, which makes it simpler to analyse and understand complex data [T. RStudio, 2020].

Cygwin –

In order to compile and run Unix or Linux programmes on a Microsoft Windows operating system (OS), Cygwin is a set of free and open-source tools**(**https://www.cygwin.com/**)**. Users can enjoy a Linux-like experience in a Windows environment thanks to Cygwin [D. Lazenby, 2000]. This functionality makes it simpler for developers to support their applications that operate on the Windows platform and aids in the migration of applications from Unix or Linux-based systems to Windows-based systems.

The dynamic link library (DLL) cygwin1.dll serves as the brains behind the Cygwin suite of tools. The Portable OS Interface, or POSIX, system call capability is provided by the DLL, which acts as an emulation layer. Along with an X server and a comprehensive range of X applications, the Cygwin distribution also includes a sizable number of free utilities, such as the majority of GNU and numerous Berkeley Software Distribution products. Incorporating the words Cygnus and Windows, the term Cygwin was born [C. Xiao *et al., 2016]*.

3.2 METHODOLOGY

3.2.1 Collection of Millet based bioactive compounds

With the help of literature survey structural information about millet based bioactive compounds were obtained. The 92 structure of these compounds were retrieved from PubChem. A dataset of compounds along with their Compound name, PubChem ID, IUPAC name and canonical smiles was prepared. Later these structures were used in the molecular docking studies.

3.2.2 Protein and Ligand preparation

The crystal structure of target protein Aldose reductase (**PDB ID: 4QXA**) was obtained from the RCSB Protein Data Bank (https://www.rcsb.org/structure/4qxa). In order to prepare the target protein and to avoid steric clashes, the structure was pre-processed before docking by removing the heteroatoms, which include water ions, small molecules, and other residues. All the pre-processing is achieved by using Discovery Studio. The chemical structures were obtained from the PubChem **(**https://pubchem.ncbi.nlm.nih.gov/**)** database. In addition, 3D conformations of all 94 chemical compounds, including control and phytochemicals were converted from .sdf to .pdb, for further studies.

3.2.3 Virtual Screening

The success of a drug's development process depends on a drug's potency or effectiveness as well as acceptable qualities for absorption, distribution, metabolism, and excretion (ADMET) and toxicity (T). As a result, it aids in the prediction of the compound's drug likeness. In silico ADMET studies were performed on the 92 millet based bioactive compounds with regards to their ADMET, mutagenicity, and carcinogenicity profiles were computed using two tools i.e., Swiss ADME**(Daina&A.,2017)**and the ADMET module and TOPKAT (TOxicity Prediction by Kobmputer Assisted Technology) module of Biovia Discovery Studio software 4.0.**(B. L., &Rohane, 2021)**

3.2.4 Pharmacophore Analysis

Pharmacophore have been proven to be successful for the selection of targeted sets of compounds. There are two categories of pharmacophores: structure-based pharmacophores and ligand-based pharmacophore. As we know pharmacophore models are widely used as virtual screening filters, ligand based pharmacophore was done to separate out compounds that have druglike properties. In this study, a pharmacophore model of controls was generated in which only 29 compounds consist of pharmacophoric features such as hydrogen bond acceptors

(HBAs), hydrogen bond donors (HBDs), hydrophobic areas (Hs), aromatic rings (ARs), positively/negatively ionizable groups (PIs/NIs), and exclusion volumes (XVOLs).

3.2.5 Molecular Interaction Analysis

Intermediary procedures were completed using the graphical user interface programme AutoDockTools (**Huey & R., 2012)** including the creation of grid boxes and pdbqt files for protein and ligand preparation (ADT). The prepared file was saved by AutoDock in PDBQT format. Auto Grid was used for the preparation of the grid map using a grid box. The grid size was set to 68 × 76 × 70 xyz points with grid spacing of 0.375 Å and grid center was designated at dimensions (x, y, and z): -12.237, 25.864, -3.284.

The ligand structure is used to generate a scoring grid that will speed up processing. AutoDockVina was used for docking, utilising grid box parameters from the configuration file, protein and ligand information, and so on. The iterated local search global optimizer is used by AutoDockVina. For further examination, the configuration with the lowest binding energy or affinity was extracted and aligned with the receptor structure.

3.2.6 Multiple Ligand Simultaneous Docking

In the present study, an attempt was made on multiple ligand simultaneous docking (MLSD) to mimic the real molecular recognition processes better. Docking multiple substrates simultaneously at the enzymatic sites of PPO would reflect the real biochemical process, gaining mechanistic insights of the concerted action of multiple substrates/cofactors when they approach the binding pockets. Binding dynamics, including possible intermediates and transition states, could even be revealed energetically **(Li and Li, 2015).**We then validated the binding affinities obtained by docking using, Molecular Dynamics (MD) simulation to cross-check the stability of the complex formed.

3.2.7 Molecular Dynamic Simulations

Molecular dynamic simulation can be performed to check the stability of the complex. It can be done using Gromacs(https://www.gromacs.org/) and will be visualizedusing UCSF Chimera and Xmgrace to analyze and check the physical movement of the atoms and understand themoleculeis stable ornotinreal system. The classical approach of molecular dynamics (MD) method **[**Alder and Wainwright, 1957**]** one simply solves numerically Newton's equations of motion for the interacting many-particle system [Rapaport*et al.*, 2004].

In order to fully comprehend the conformational modifications and dynamic features of the interactions, MD simulations of the strongest protein-ligand complex in comparison to the TAGE

compounds were performed. A total of 4 compounds were selected and simulated using GROMACS 2018 [Loschwitz&Jäckering,2021] (Groningen Machine for Chemical Simulations) with GROMOS96 43a1 force field. The production of ligand parameters related to the selected force field was done using the ATB server. The protein-ligand complexes in the dodecahedron boxes were dissolved using simple point charge (SPC) water molecules, and the system was subsequently neutralised by the addition of counterions. In addition, energy minimization was applied by running the steepest descent minimization integrator at 50000 steps with a force convergence of about 10 kJ/mol. Then, utilising NVT (canonical) and NPT (isothermal-isobaric) ensembles at a temperature of 300 K, all complexes were brought into equilibrium. Thereafter, the systems were simulated for the 10 ns and were analyzed by using different parameters such as Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius Gyration and a number of hydrogen bonds interactions of the protein-ligand complexes etc.

4.Results & Discussions

4.1 Pharmacokinetics and toxicity prediction of ligand molecules

Using the ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) descriptors of Discovery Studio software 4.0 and Swiss ADME, the pharmacokinetic profile of the millet-based bioactive compound was calculated.Mostly in discipline of computer-aided drug design, the discovery of "drug-like" molecules using molecular databases is commonly favoured. The Lipinski's rule of 5 is the most often used drug-likeliness filter. If the violation is 1 or 0, it means that the ligand adheres to the receptor effectively. If the number of violations was greater than two, the compound was eliminated from further consideration. This study revealed that out of 92 compounds, only 48 compounds fulfilled the Lipinski rule of five.

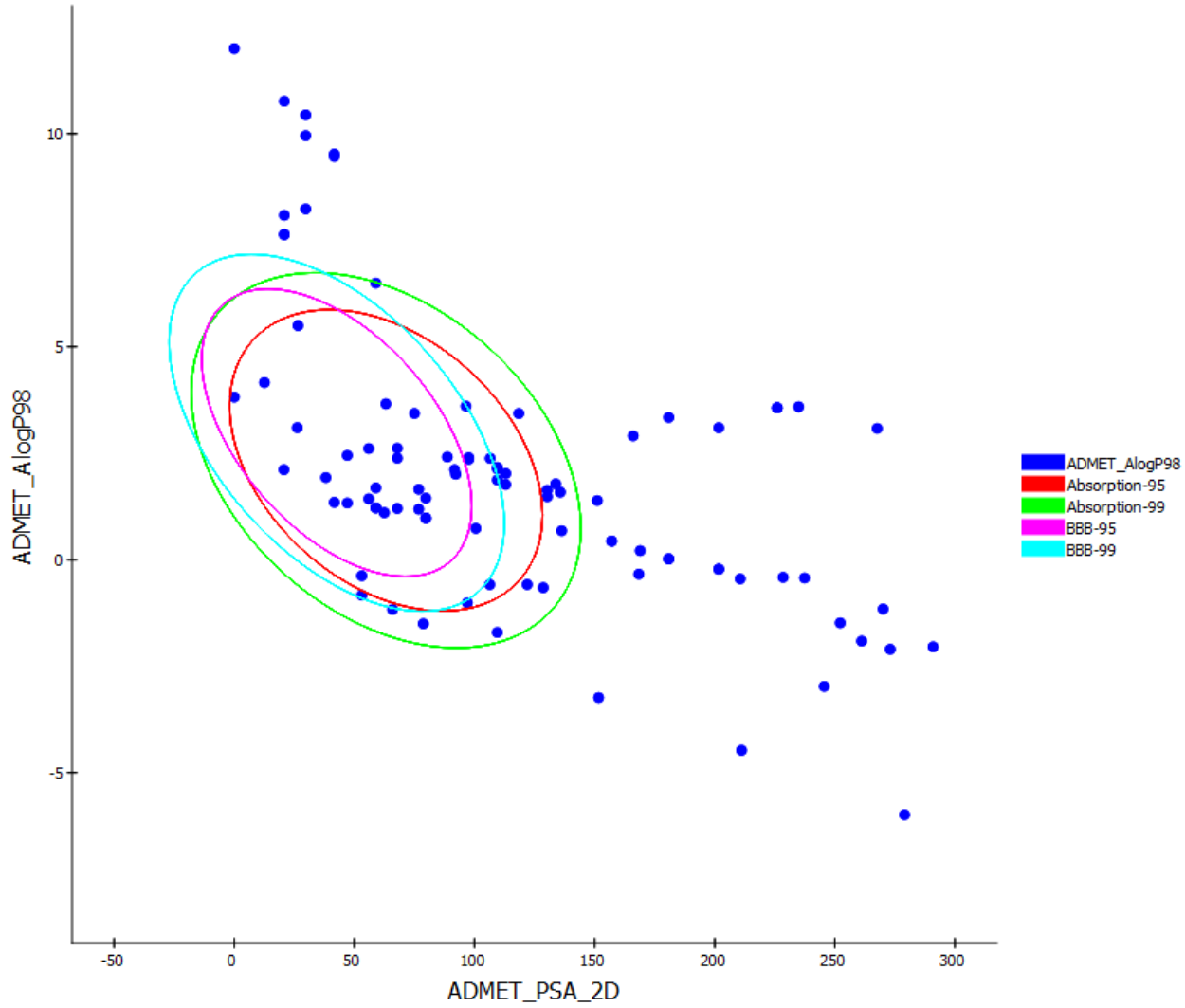


Figure 3. *ADMET plot of 2D polar surface area against calculated Alogp98.*

The BOILED-Egg method of calculating the penetration through the brain or intestines, as an accurate predictive model, was also used to analyse the link between the lipophilicity and the polarity of the bioactive chemicals. These characteristics are essential when looking for a new

drug, hence almost half of the ligands make strong candidates for being potential medications.The plot is showing green and blue colouredeclipses of 99% confidence limits for intestinal absorption and blood-brain barrier (BBB), as well as red and pink coloured eclipses of 95% confidence limits for intestinal absorption and BBB

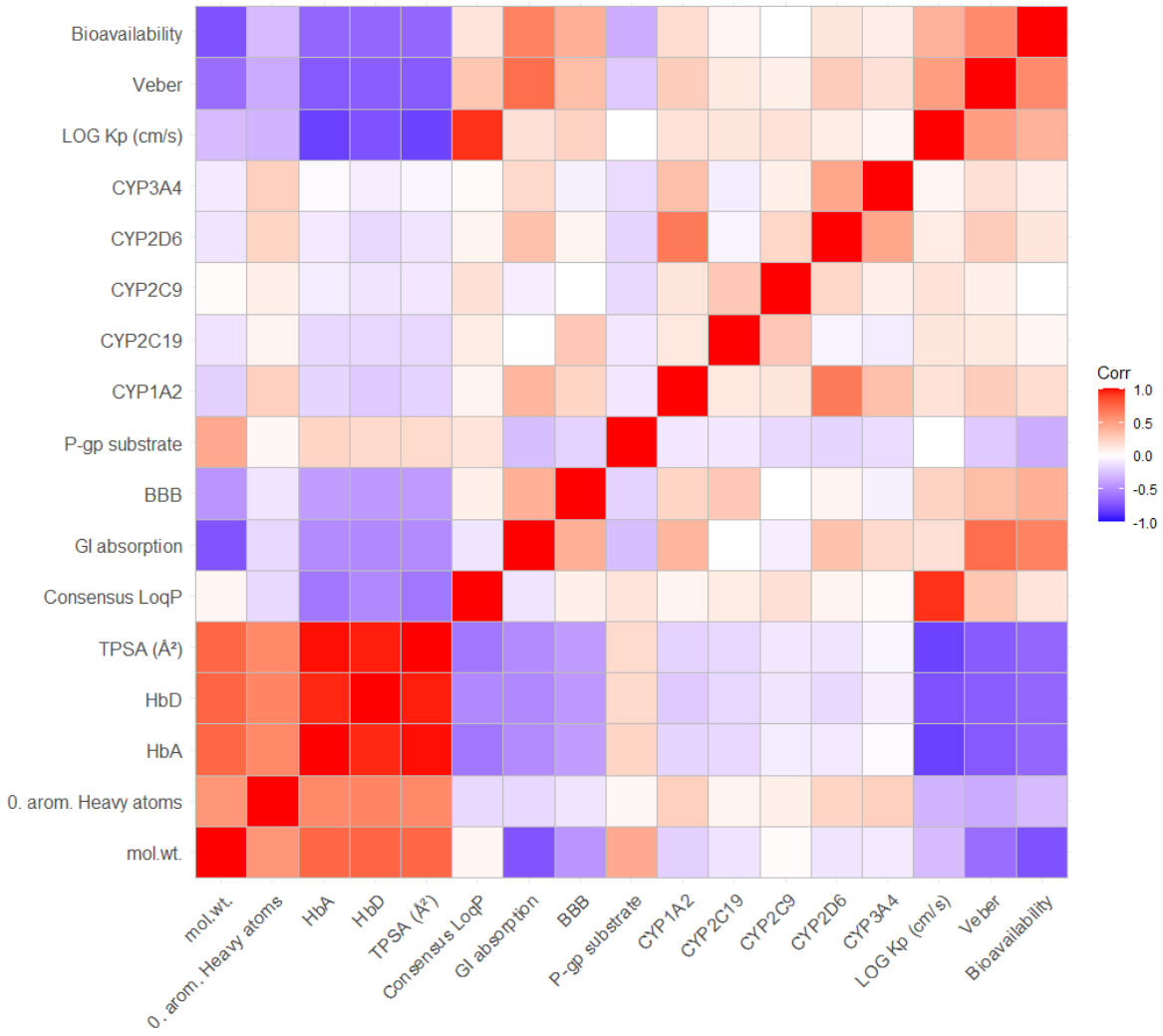


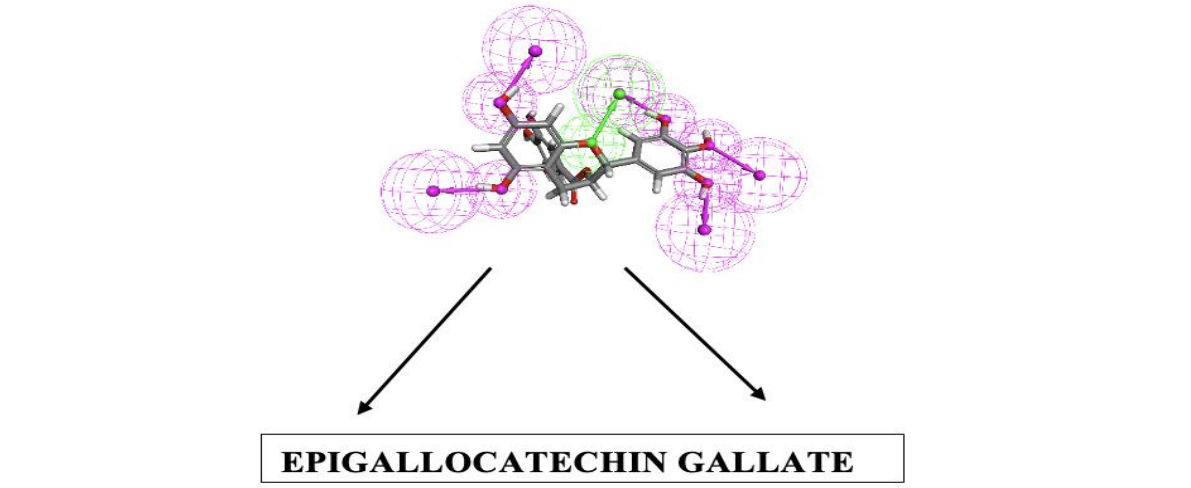
Figure 4. *Heatmap summary of in-silico ADMET properties of millet based bioactive compounds using R studio.*

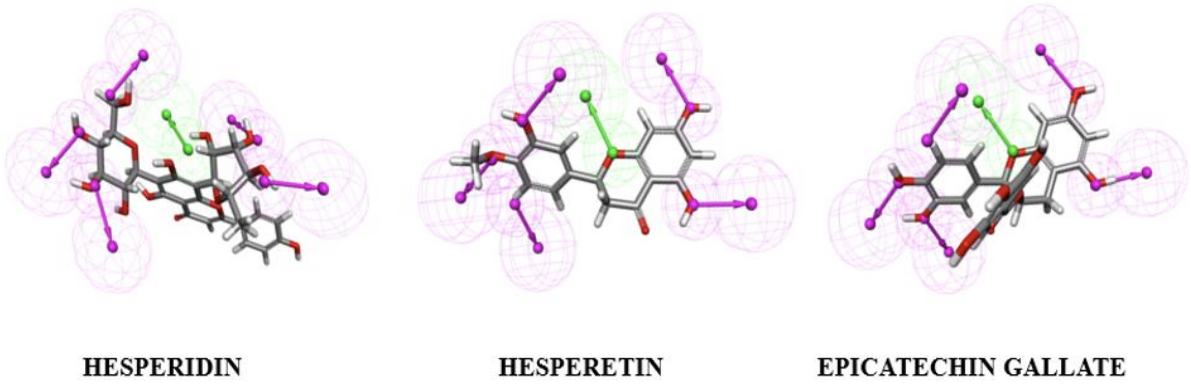
Figure. 4., shows the summarization of ADMET properties of 92 compounds. The compounds that passed all of the ADMET parameters are shown by the colour red, whereas the compounds that failed all of the parameters are represented by the colour purple.

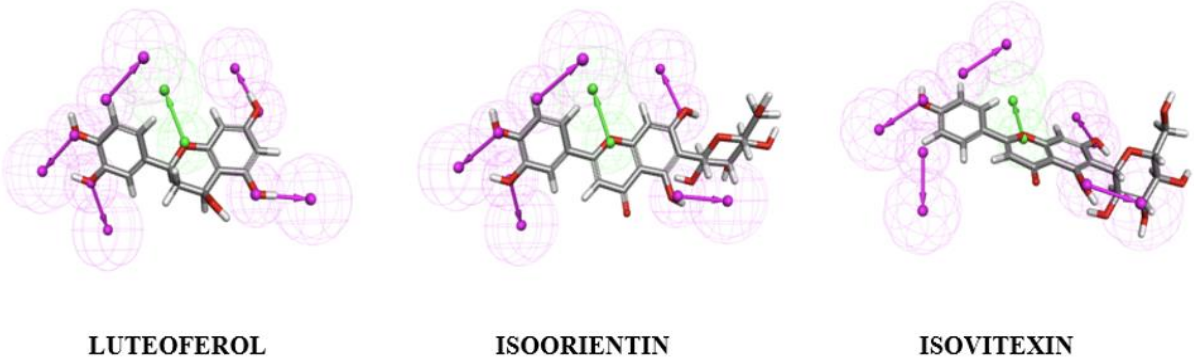
4.2 Generation and Validation of Ligand-Based Pharmacophore

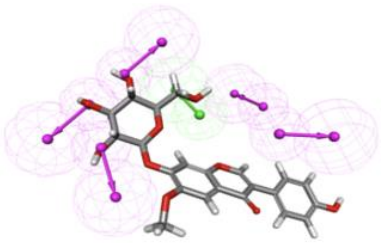
The controls, ascorbic acid and epigallocatechin gallate were used to construct the pharmacophore. 29 out of 48 millet based bioactive compounds showed reliable pharmacophoric features. These features are identified using colour coded :- Hydrogen bond acceptor (HBA),

green; hydrogen bond donor (HBD), violet; hydrophobic (H), light blue; positive ionizable (pos), red; negative ionizable (neg), blue. The bioactive conformations of the ligands with similar binding modes may be identified using a trustworthy pharmacophore model. The conformation chosen for each compound, which is taken to be the bioactive conformation, corresponds to the conformation that fit the pharmacophore the best. This clearly shows that the filtered 29 compounds have mapped well on the reported pharmacophore model. Thus, verifying its potential for virtual screening.

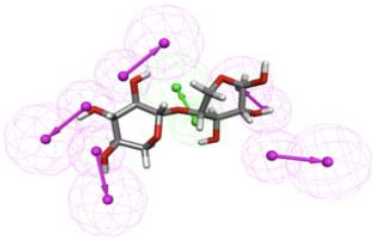




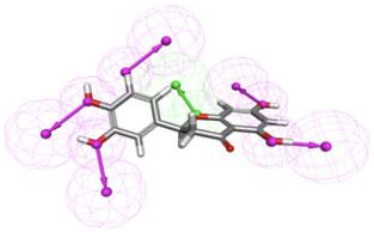




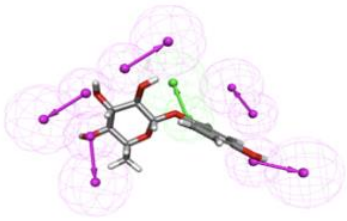
GLYCITIN



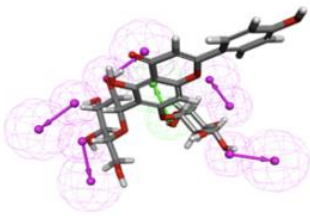
XYLOBIOSE



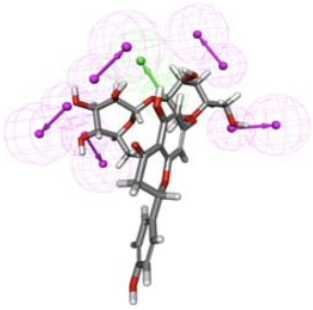
ERIODICTYOL



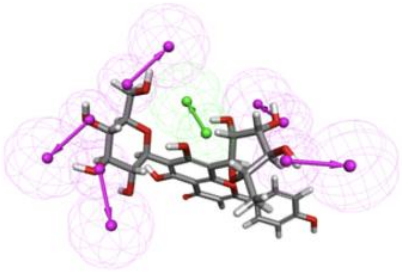
ARBUTIN



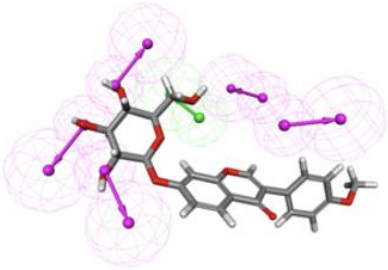
SAPONARIN



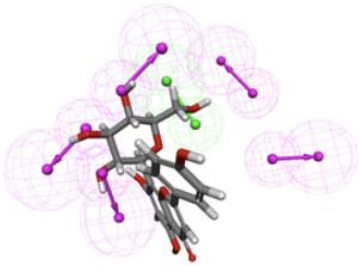
NARINGIN



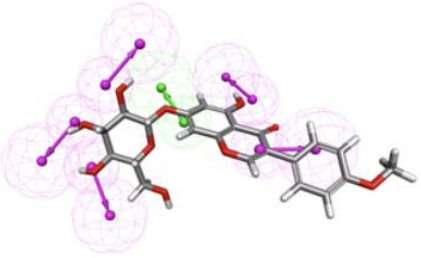
VIOLANTHIN



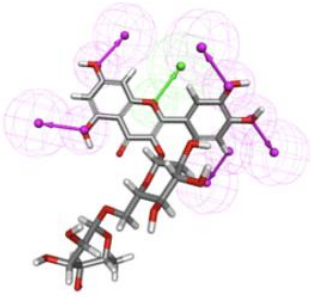
ONONIN



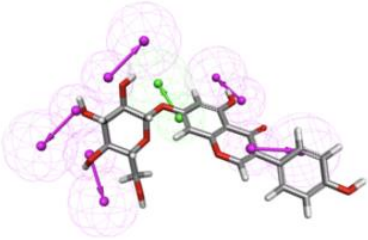
VITEXIN



SISSOTRIN



RUTIN



GENISTIN

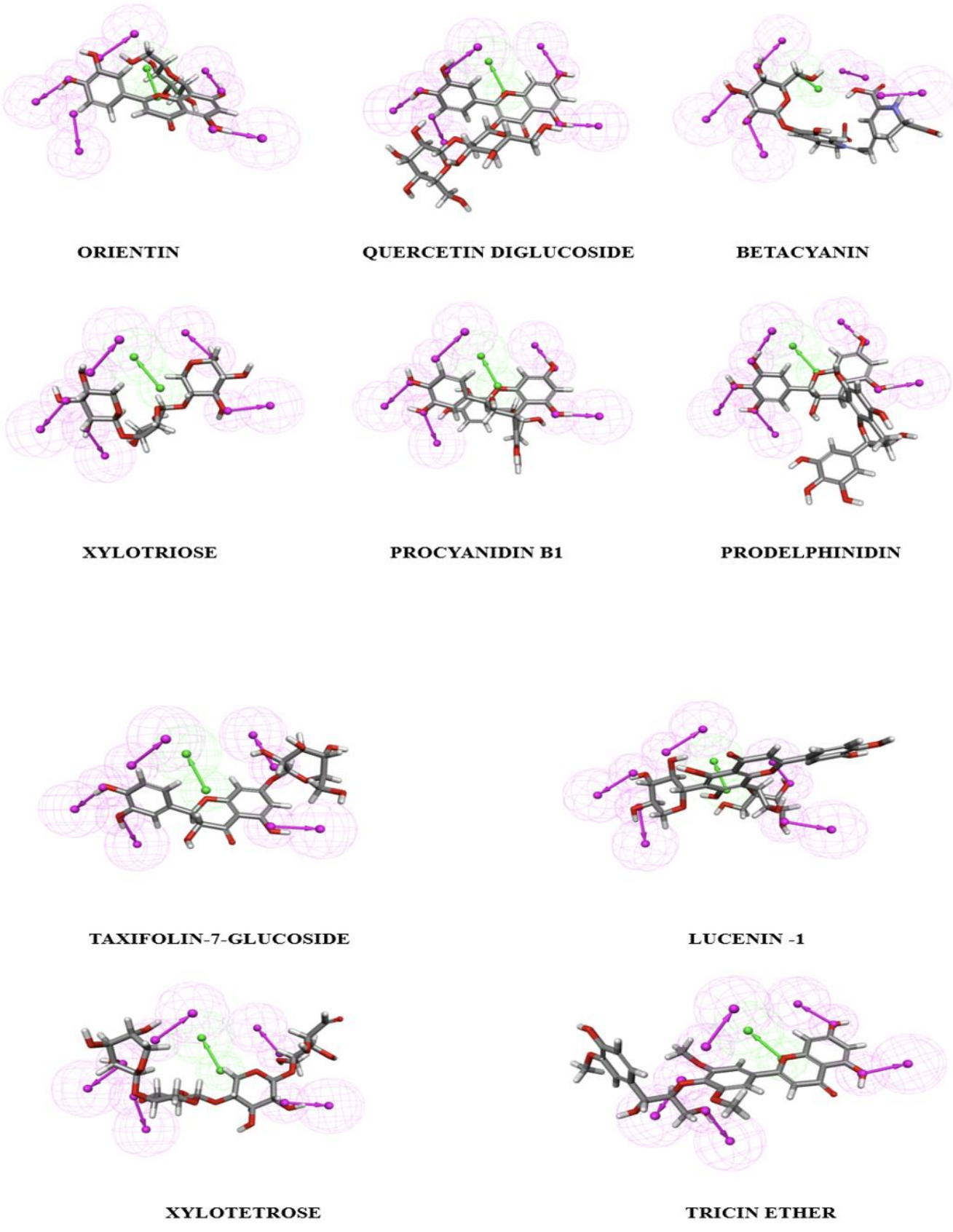
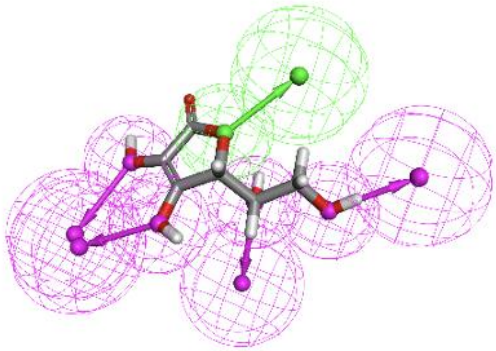
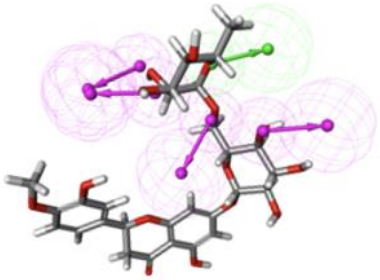


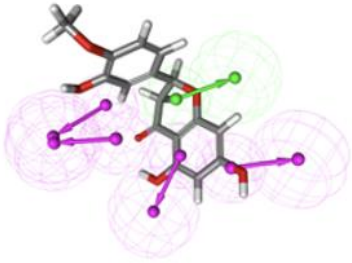
Figure 5. *Generated ligand-based pharmacophore model of control Epigallocatechin gallate (65064) with bioactive compounds of millets.*



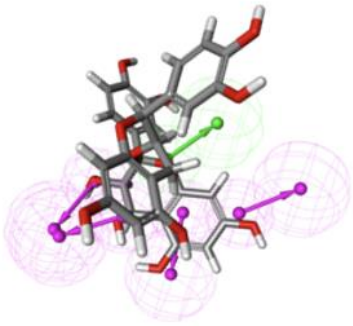
ASCORBIC ACID



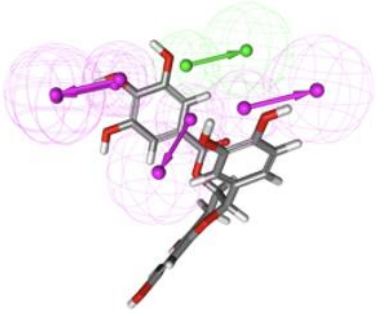
HESPERIDIN



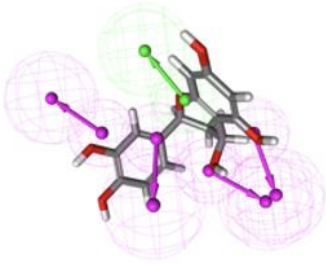
HESPERETIN



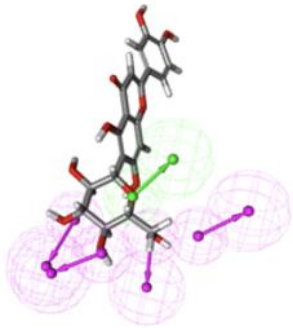
PROANTHOCYANIDIN



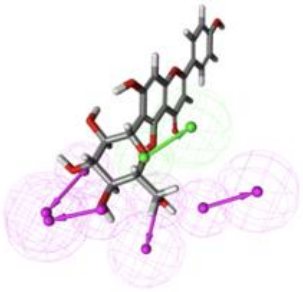
EPICATECHIN GALLATE



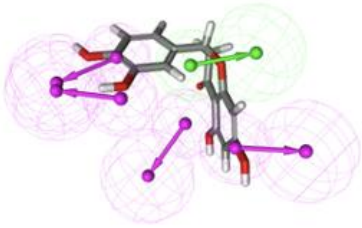
LUTEOFEROL



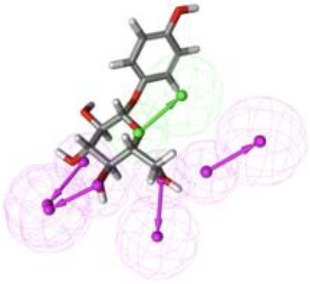
ISOORIENTIN



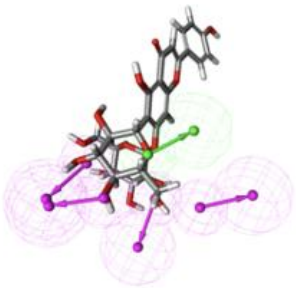
ISOVITEXIN



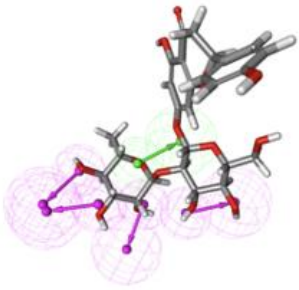
ERIODICTYOL



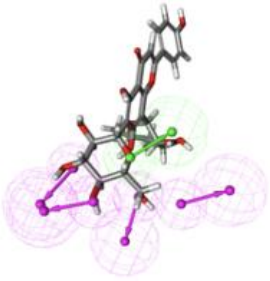
ARBUTIN



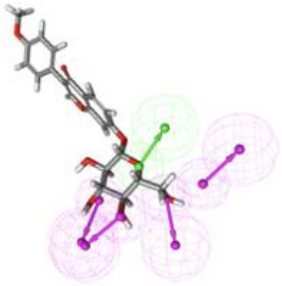
SAPONARIN



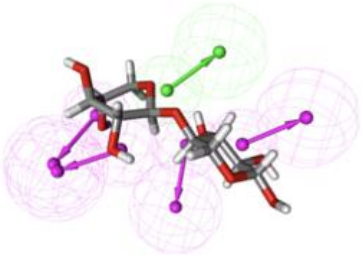
NARINGIN



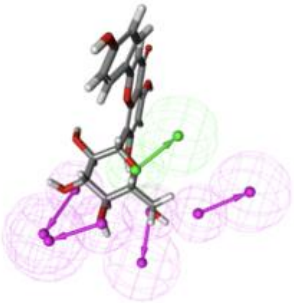
VIOLANTHIN



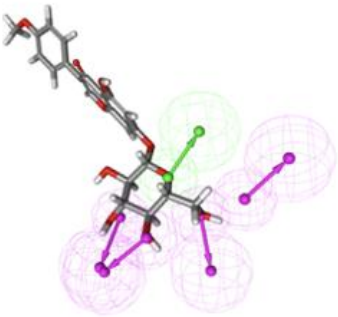
ONONIN



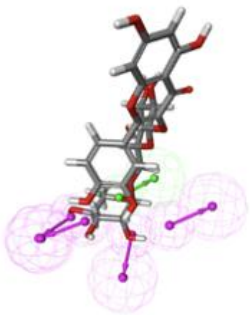
XYLOBIOSE



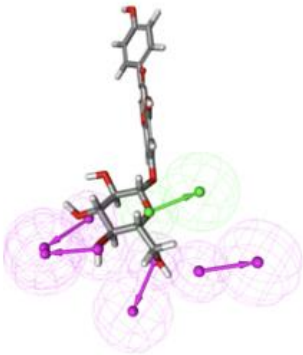
VITEXIN



SISSOTRIN



RUTIN



GENISTIN

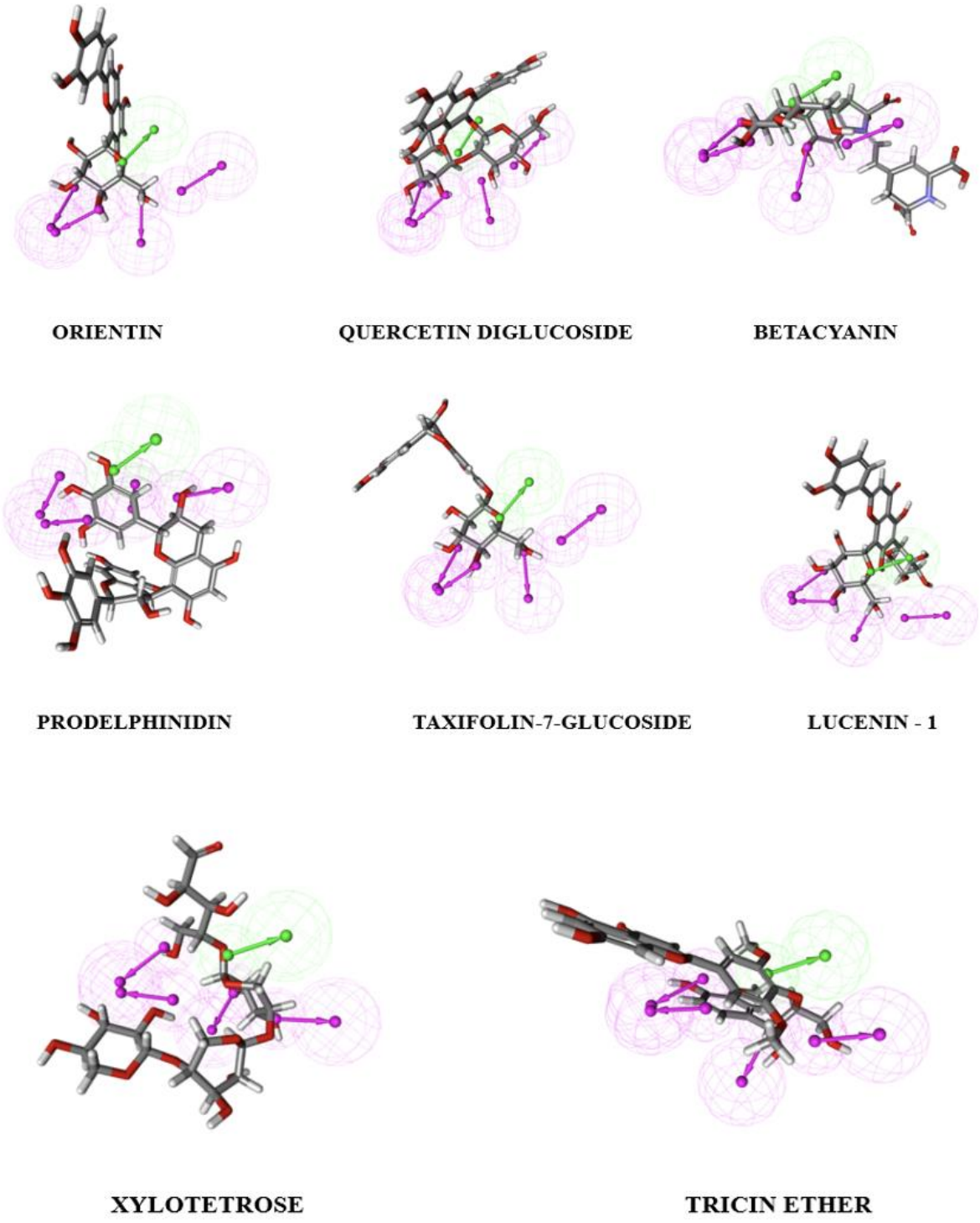


Figure 6. *Generated ligand-based pharmacophore model of control Ascorbic acid (54670067) with bioactive compounds of millets.*

4.3 Molecular Docking Analysis

A docking study was utilized for better understanding of the structural basis of protein-ligand interactions.The structure of Aldose reductase (PDB ID-4QXA) established by means of X-ray crystallography was used for the docking analysis. The filtered 29 bioactive compounds were docked against the target Aldose reductase (PDB ID-4QXA). The binding energy of these compounds are discussed in the table below.

Table 1. Binding energy of docked ligands and its active site pocket residues

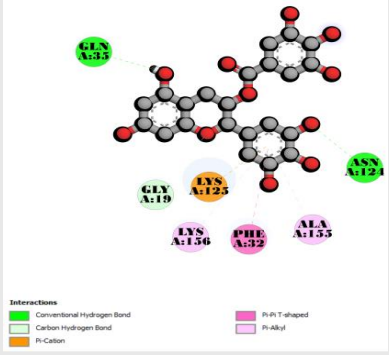
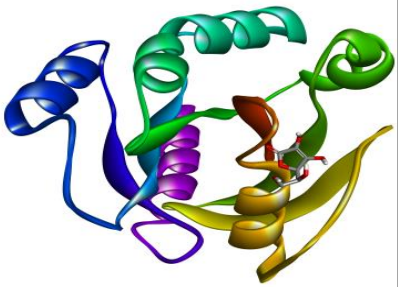
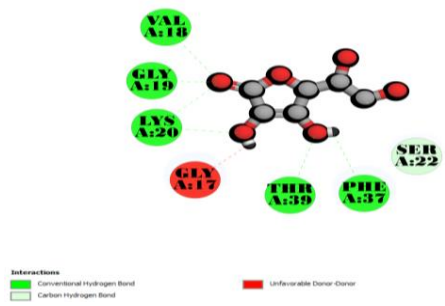
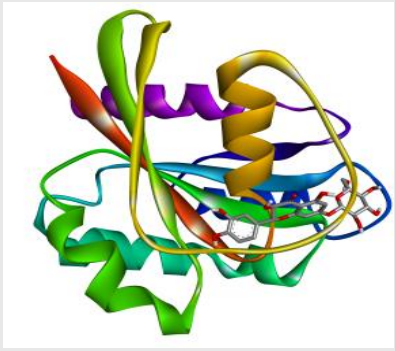
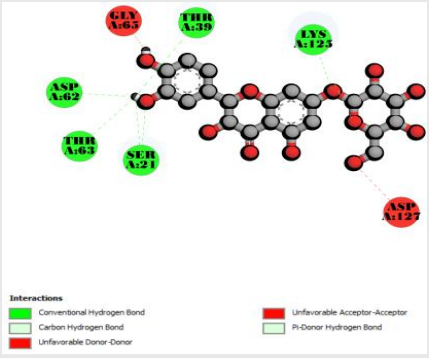
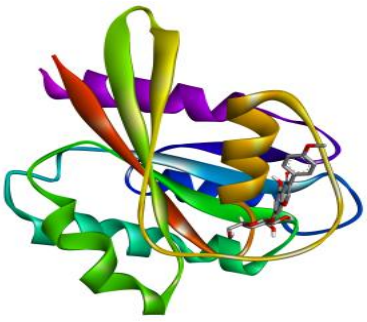
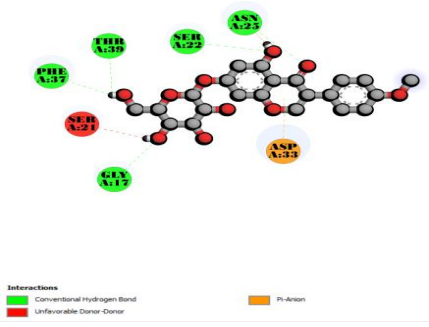
S.No	Compound	CID	Target	Binding Energy (kcal/mol)	Residue
1	Epigallocatechin gallate (Control)	65064	4QXA	-8.4 kcal/mol	GLY 19, PHE 32, GLN 35, LYS 123, ASN 124, ALA 133, LYS 136
2	Ascorbic acid (Control)	54670067	4QXA	-6.3 kcal/mol	VAL 18, GLY 19, GLY 17,LYS 20, SER 22, PHE 37,THR 39
3	Taxifolin 7-glucoside	14282775	4QXA	-9.7 kcal/mol	SER 21, THR 39, ASP 62, THR 63, LYS 123
4	Sissotrin	5280781	4QXA	-9.1 kcal/mol	GLY 17, SER 22, ASN 23, ASP 33, PHE 37, THR 39
5	Rutin	5280805	4QXA	-9.1 kcal/mol	PHE 32, PHE 37, LYS 123,ASP 127, ALA 133,LYS 136
6	Genistin	5281377	4QXA	-8.9kcal/mol	GLY 17, LYS 31, ASP 33,PHE 37
7	Ononin	442813	4QXA	-8.8 kcal/mol	LYS20,LYS31,ASP33
8	Luteoforol	114505	4QXA	-8.8 kcal/mol	GLY 19 , PHE 32, SER34,GLN35,LYS123,ALA133,LYS136

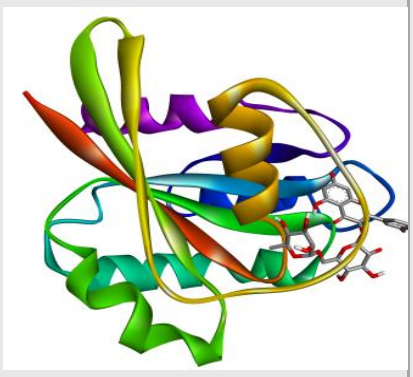
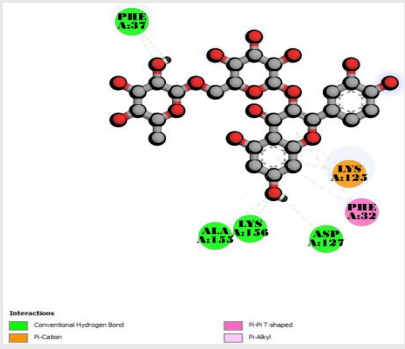
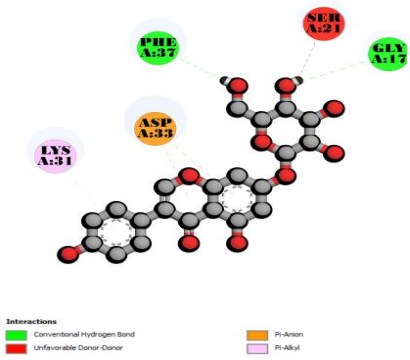
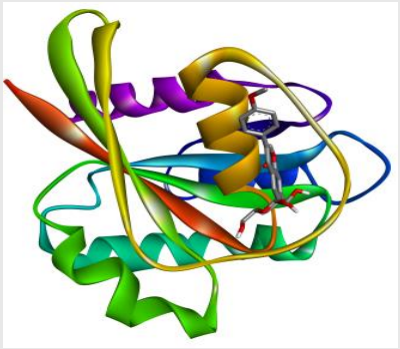
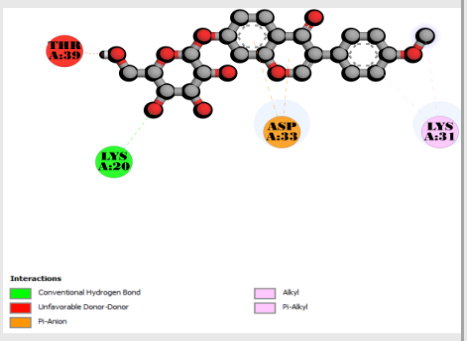
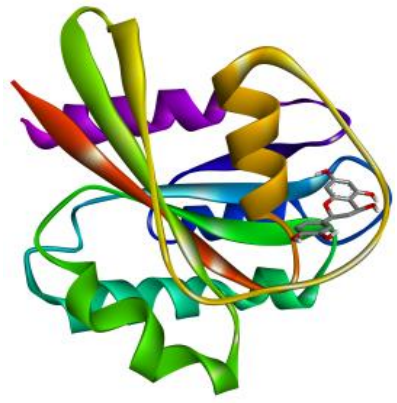
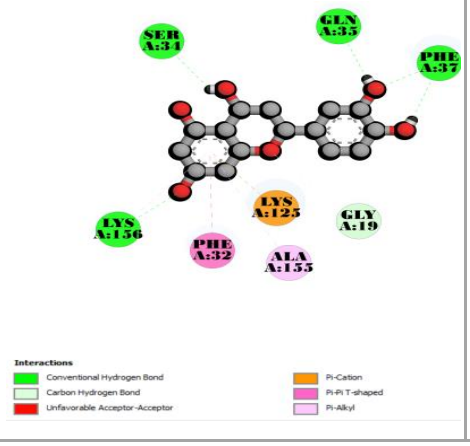
9	Hesperidin	10621	4QXA	-8.8 kcal/mol	SER 22, ASN 25, LYS 31,ASP 33, GLN 35, PHE 37
10	Prodelphinidin	13831068	4QXA	-8.7 kcal/mol	SER 21, PHE 32, SER 34,PHE 37, ASP 90, LYS 125,ALA 155, LYS 156
11	Eriodictyol	440735	4QXA	-8.7 kcal/mol	GLY 19, PHE 32, PHE 37,LYS 123, ALA 155,LYS 156
12	Glycitin	187808	4QXA	-8.7 kcal/mol	ASP 33, PHE 37, THR 39
13	Isoorientin	114776	4QXA	-8.7 kcal/mol	SER21,THR39,LYS123,ALA13
14	Procyanidin B1	11250133	4QXA	-8.6 kcal/mol	LYS 20, SER22,PHE 37,THR 39
15	Xylotriose	10201852	4QXA	-8.6 kcal/mol	GLY17,PHE32,GLY35, LYS 125, ALA155
16	Hespretin	72281	4QXA	-8.5 kcal/mol	ASP 81, ALA 109, VAL 111,LYS 112, GLU 113,SER 116, THR 173, ASP 177
17	Proanthocyanidins	107876	4QXA	-8.4 kcal/mol	LYS 20, LYS 31, ASP 33,LYS 123, ALA 155, LYS 156
18	Tricin ether	101633864	4QXA	-8.3 kcal/mol	GLY17,PHE32,GLN35, LYS 125, ALA 155
19	Betacyanins	6324775	4QXA	-8.3 kcal/mol	VAL 18, GLY 19,GLY 17,LYS20,SER22,PHE37,THR 39
20	Orientin	5281675	4QXA	-8.2 kcal/mol	PHE32,SER34,PHE37,GLU113THR 173,
21	Naringin	442428	4QXA	-8.2 kcal/mol	LYS31,ASP33,LYS123,ASN 124, ALA 133, LYS 136

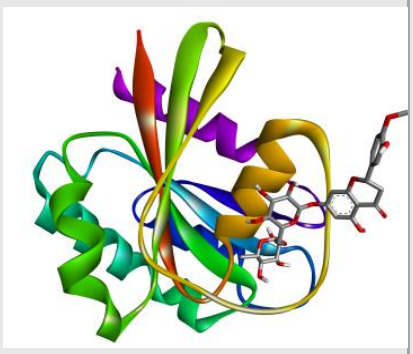
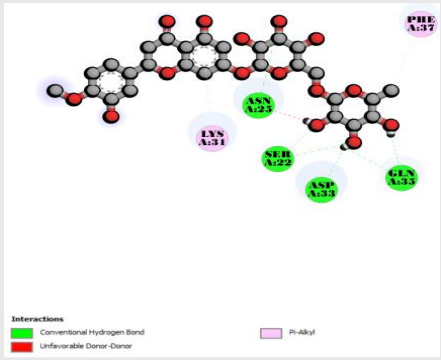
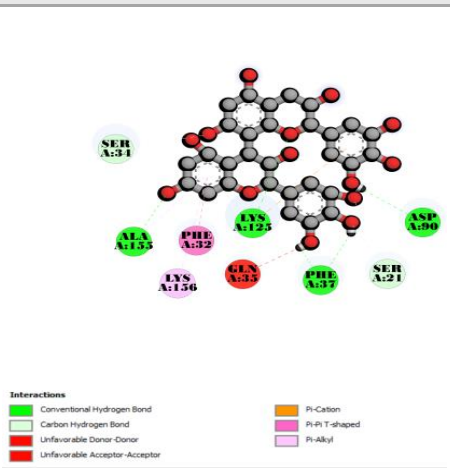
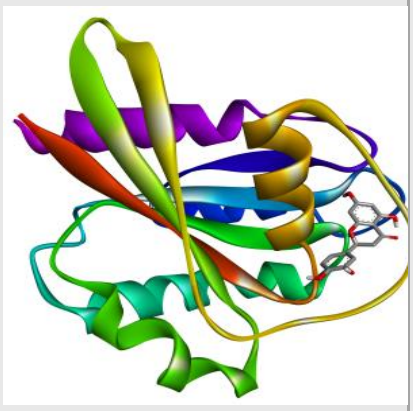
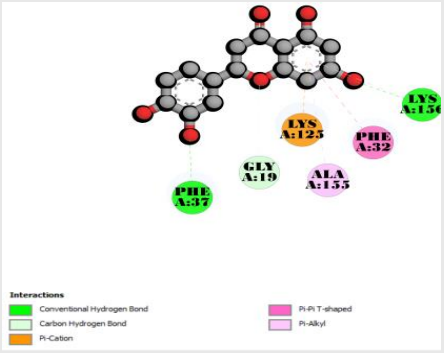
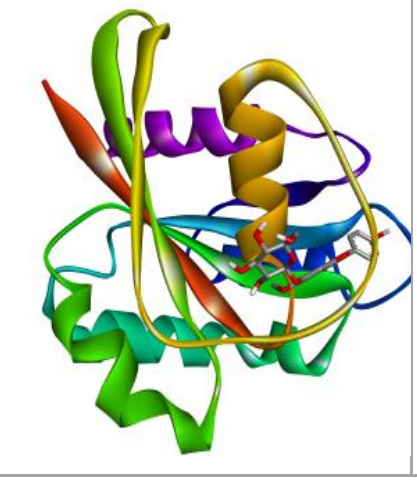
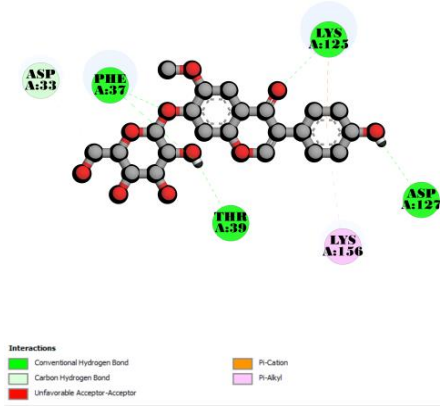
22	Isovitexin	162350	4QXA	-8.2 kcal/mol	THR 63, LYS 123 SER 116, THR 173, ASP 177
23	Vitexin	5280441	4QXA	-7.9 kcal/mol	LYS 20, SER 22, VAL 18, Gly137,Gly138
24	Xylotetrose	101601989	4QXA	-7.7 kcal/mol	LYS 31, ASP 33, PHE 37, LYS 125, ASP 127, LYS 136
25	Saponarin	441381	4QXA	-7.6 kcal/mol	GLY 19, SER 22, PHE 37, ASN 124, LYS 125, ASP 127, LYS 136
26	Epicatechin gallate	107905	4QXA	-7.5 kcal/mol	GLY 17, GLY 19,ASN 23, GLN 35 ,PHE 37
27	Violanthin	442665	4QXA	-7.4 kcal/mol	LYS 31, ASP 33
28	Arbutin	440936	4QXA	-7.2 kcal/mol	GLY 17 , PHE32,GLN 35, LYS 125,ALA 155
29	Xylobiose	4439538	4QXA	-6.8 kcal/mol	GLY17,LYS20,SER22 ,ASN 23,PHE37, THR 39
30	Lucenin-1	101316994	4QXA	-6.7 kcal/mol	GLU 132, GLN 139, ASP 147, PRO 149
31	Quercetin diglucoside	5282166	4QXA	-6.6 kcal/mol	ASP 81, ALA 109, VAL 111, LYS 112, GLU 113, SER 116, THR 173, ASP 177

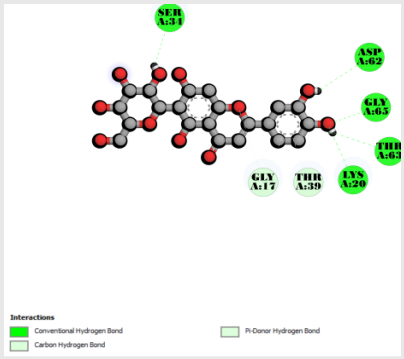
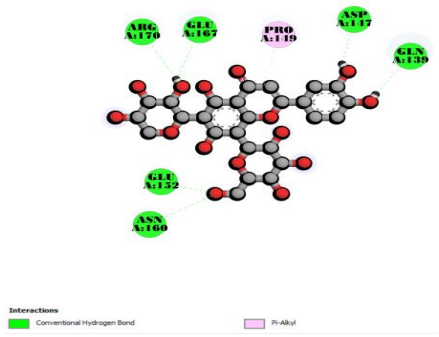
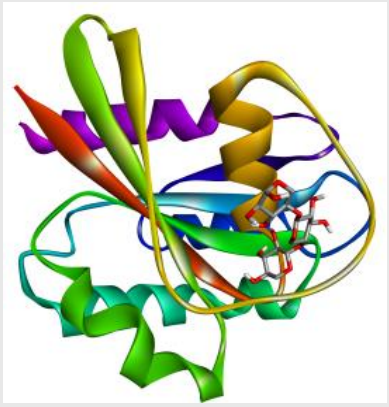
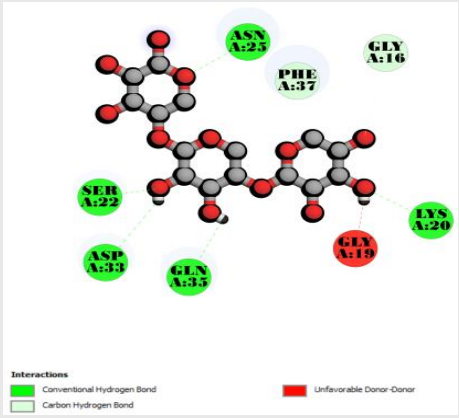
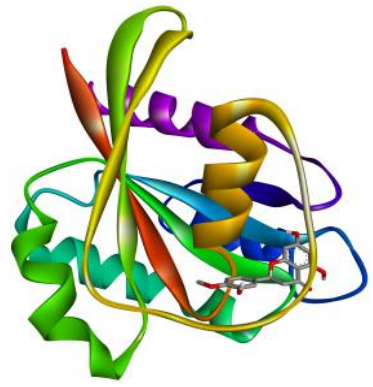
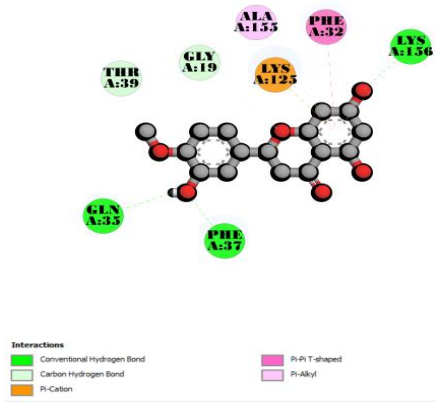
The compounds Taxifolin-7-Glucoside (14282775) with binding energy (B.E)-9.7 kcal/mol, Sissotrin (5280781) with -9.1 kcal/mol, Rutin (5280805) with -9.1 kcal/mol and Genistin (5281377) with -8.9 kcal/mol showed most promising binding affinity compared to the control (Epigallocatechin gallate (65064) with B.E -8.4 kcal/mol and ascorbic acid (54670067) with B.E -6.3 kcal/mol).The interactions with significant amino acids and the hydrophobic interactions between these molecules and amino acids are potential explanations for the higher binding scores of these compounds. For the purposes of sorting leads, this was taken intoconsideration.

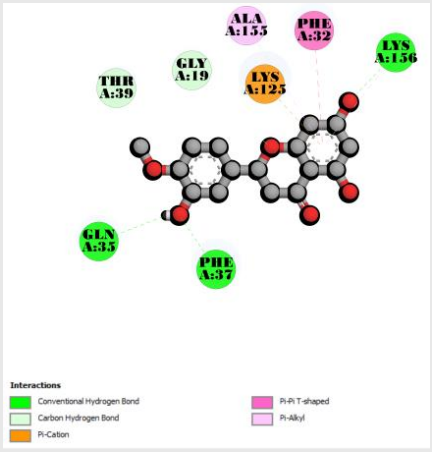
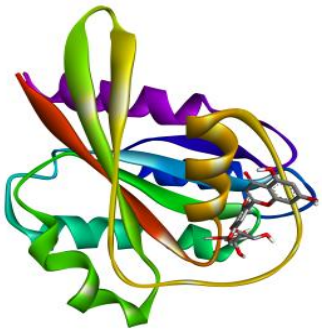
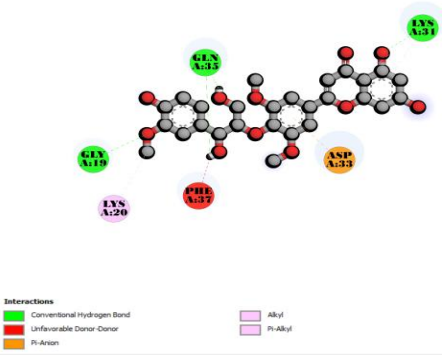
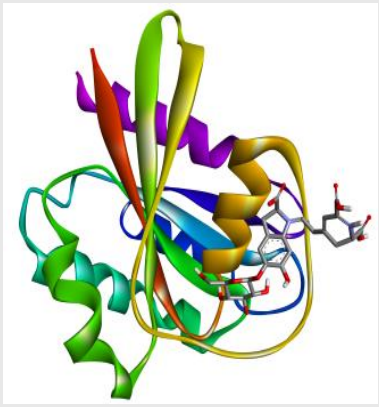
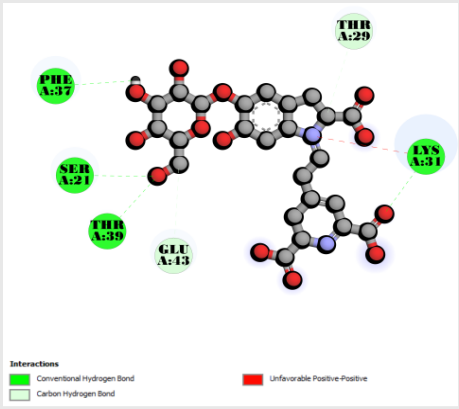
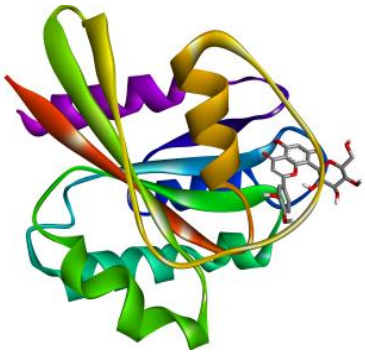
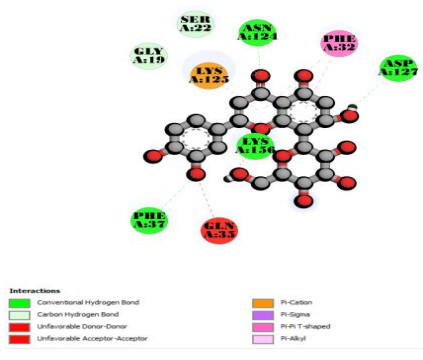
Table 2. Representation of complex formed along with the active site residues.

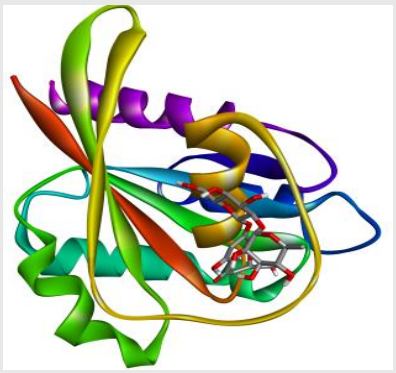
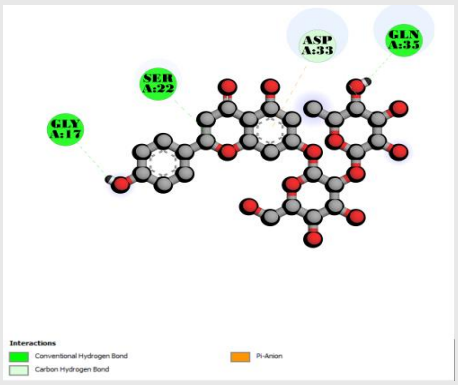
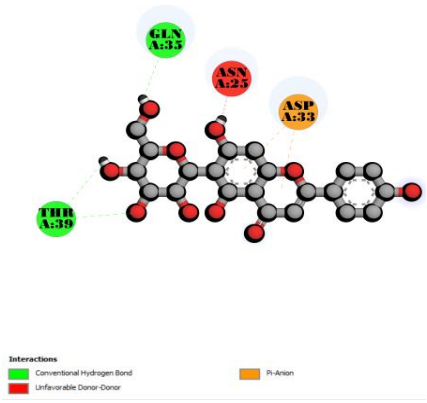
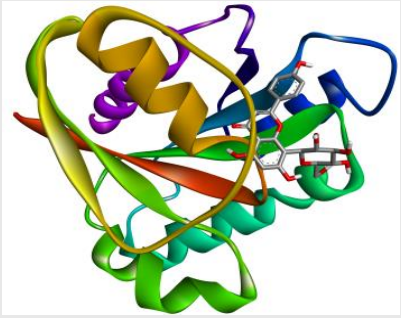
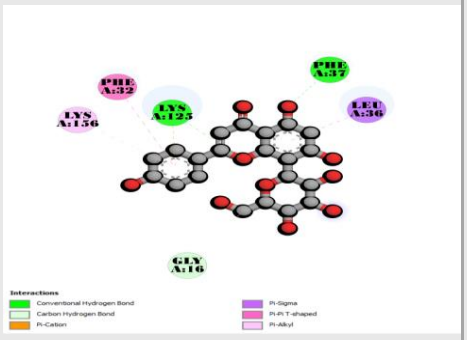
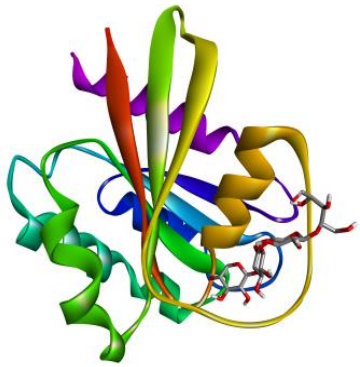
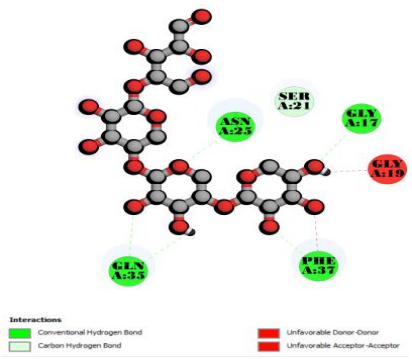
	Complex Formed	2D Ligand Interaction	Name
1			Epigallocatechin gallate(Control)-Target Aldose reductase(PDB ID-4QXA)
2			Ascorbic acid(Control)–Target Aldose reductase(PDB ID-4QXA)
3			Taxifolin-7-glucoside - Target Aldose reductase(PDB ID-4QXA)
4			Sissotrin – Target Aldose reductase(PDB ID-4QXA)

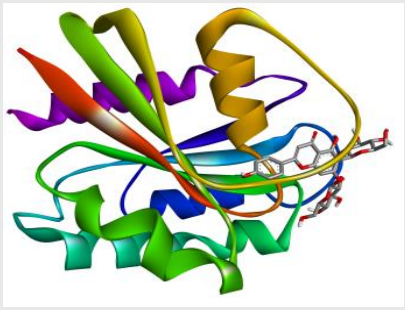
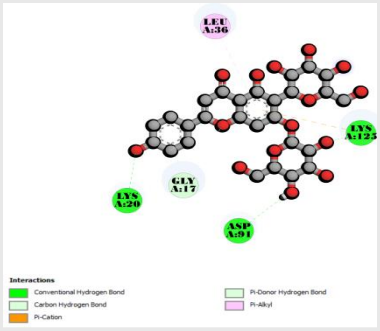
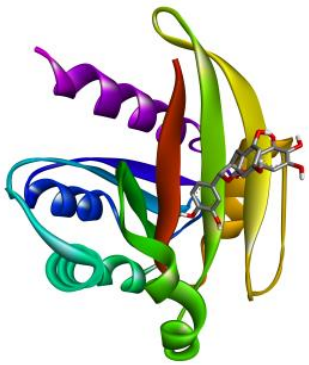
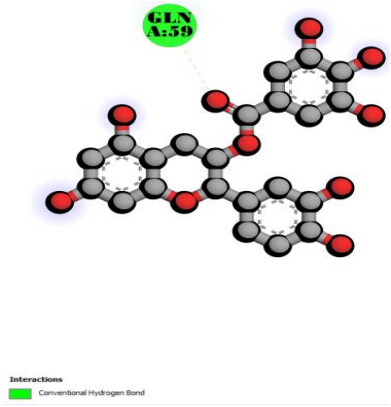
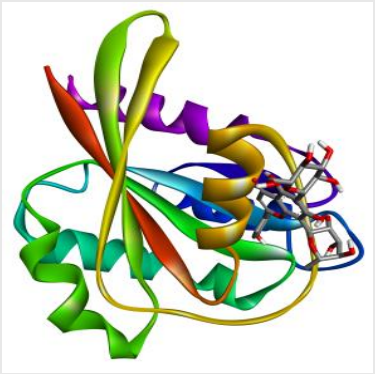
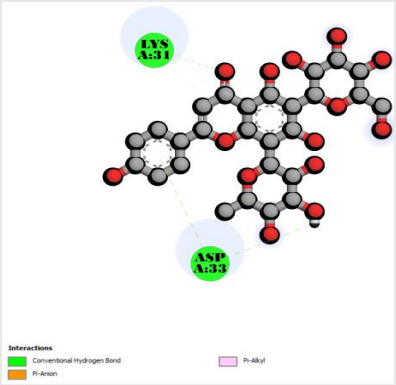
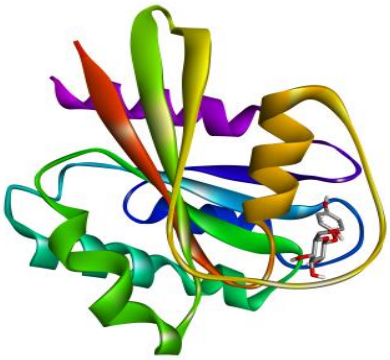
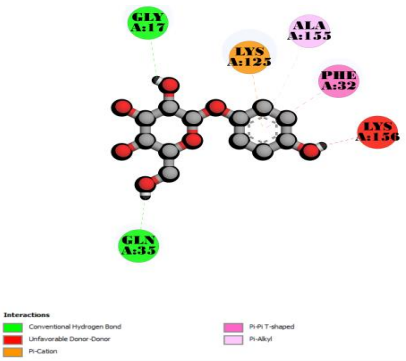
5			Rutin – Target Aldose reductase(PDB ID-4QXA)
6			Genistin – Target Aldose reductase(PDB ID-4QXA)
7			Ononin – Target Aldose reductase(PDB ID-4QXA)
8			Luteoforol – Target Aldose reductase(PDB ID-4QXA)

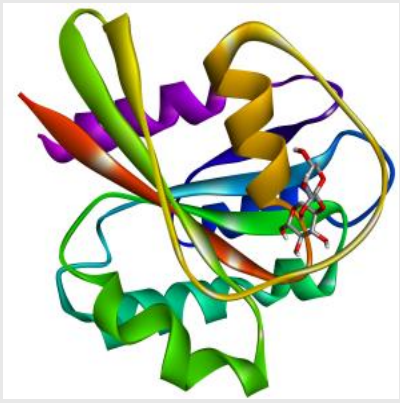
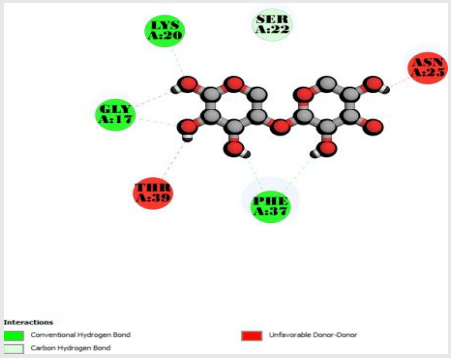
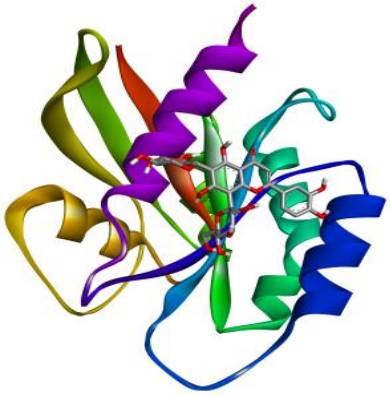
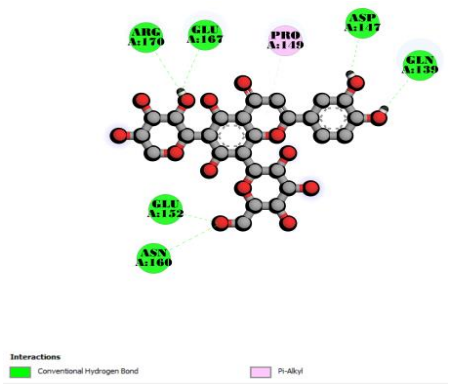
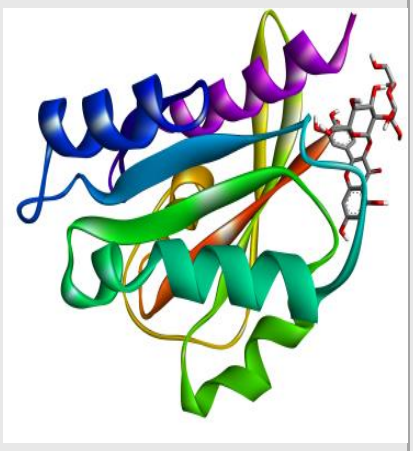
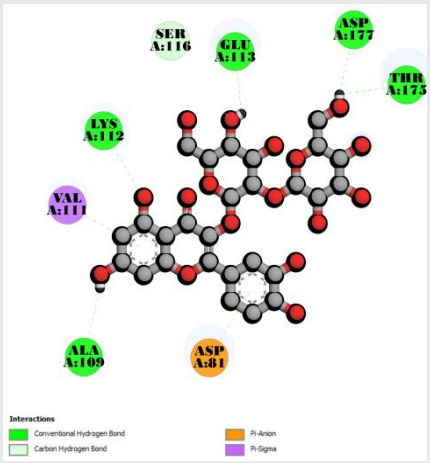
9			Hesperidin – Target Aldose reductase(PDB ID-4QXA)
10			Prodelphinidin – Target Aldose reductase(PDB ID-4QXA)
11			Eriodictyol – Target Aldose reductase(PDB ID-4QXA)
12			Glycitin – Target Aldose reductase(PDB ID-4QXA)

13			Isoorientin – Target Aldose reductase(PDB ID-4QXA)
14			Procyanidin B1 – Target Aldose reductase(PDB ID-4QXA)
15			Xylotriose – Target Aldose reductase(PDB ID-4QXA)
16			Hespretin - Target Aldose reductase(PDB ID-4QXA)

17			Proanthocyanidins – Target Aldose reductase(PDB ID-4QXA)
18			Tricin ether – Target Aldose reductase(PDB ID-4QXA)
19			Betacyanins – Target Aldose reductase(PDB ID-4QXA)
20			Orientin – Target Aldose reductase(PDB ID-4QXA)

21			Naringin – Target Aldose reductase(PDB ID-4QXA)
22			Isovitexin – Target Aldose reductase(PDB ID-4QXA)
23			Vitexin – Target Aldose reductase(PDB ID-4QXA)
24			Xylotetrose – Target Aldose reductase(PDB ID-4QXA)

25			Saponarin – Target Aldose reductase(PDB ID-4QXA)
26			Epicatechin gallate – Target Aldose reductase(PDB ID-4QXA)
27			Violanthin – Target Aldose reductase(PDB ID-4QXA)
28			Arbutin – Target Aldose reductase(PDB ID-4QXA)

29			Xylobiose – Target Aldose reductase(PDB ID-4QXA)
30			Lucenin -1 – Target Aldose reductase(PDB ID-4QXA)
31			Quercetin diglucoside – Target Aldose reductase(PDB ID-4QXA)

5.4 Multiple Ligand Simultaneous Docking (MLSD)

The concurrent interactions between the substrate and the macromolecule that are binding together in the presence of an inhibitor are studied using a computer method called multiple ligand simultaneous docking.In this study, MLSD of top 4 potential leads i.e., Taxifolin-7-Glucoside (14282775), Sissotrin (5280781), Rutin (5280805) and Genistin (5281377) against target Aldose reductase (PDB ID-4QXA) along with the TAGEs (Glyceraldehyde (751), Acetaldehyde (177) and Methylglyoxal (880)) was carried out to analyze how these potential leads will perform in the presence of toxic compounds.

Table.3. Binding energy and active site residues of simultaneously docked compounds in presence of TAGE Glyceraldehyde with Target Aldose reductase (PDB ID-4QXA).

S.No	Compound	CID	Complex Model	Binding Energy (kcal/mol)	Residue
1	Glyceraldehyde (TAGE 1)	751	1	-4.65 kcal/mol	GLY19, LYS20, SER 21, THR 39, ASP62, THR 63
2	Taxifolin 7-glucoside	14282775	3	-5.21 kcal/mol	GLY 17, GLY19, SER21, ASN25, ASP33, LYS 125, ASP 127
3	Sissotrin	5280781	4	-5.65 kcal/mol	GLY16,GLY17,VA18 ,LYS20,SER34,HIS3 8, LYS 125, LYS 156
4	Rutin	5280805	8	-1.89 kcal/mol	GLY17, GLN35, LEU36, PHE37, HIS38, SER88, ASP90, ASP91, LYS125
5	Genistin	5281377	6	-5.71kcal/mol	GLY17, LYS20, SER22, SER34, THR39, ASP62, THR63, ALA64, GLY65, LYS125

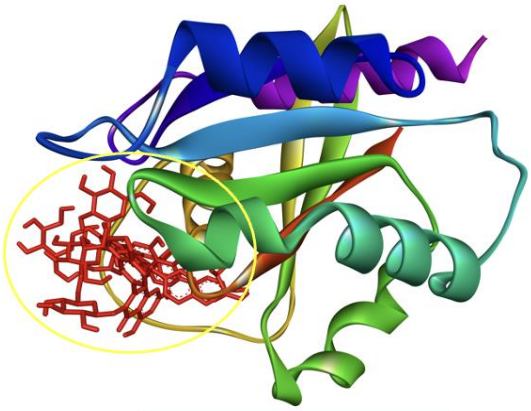
Table.4. Binding energy and active site residues of simultaneously docked compounds in presence of TAGE Acetaldehyde with Target Aldose reductase (PDB ID-4QXA).

S.No	Compound	CID	Complex Model	Binding Energy (kcal/mol)	Residue
1	Acetaldehyde (TAGE 2)	177	3	-3.32 kcal/mol	LYS20, GLY19, VAL18
2	Taxifolin 7-glucoside	14282775	4	-5.21 kcal/mol	GLY17, SER21, ASN25, ASP33, LYS125, ASP127
3	Sissotrin	5280781	2	-5.69kcal/mol	GLY16, GLY17, VAL18, LYS20, SER34, HIS38, LYS125, LYS156
4	Rutin	5280805	6	-1.89 kcal/mol	ASP15, ASP90, ASP91, GLN132, LYS125
5	Genistin	5281377	3	-5.70kcal/mol	GLY17,LYS20,SER22,SER34,THR39, ASP62,THR63,ALA64, GLY65, LYS125

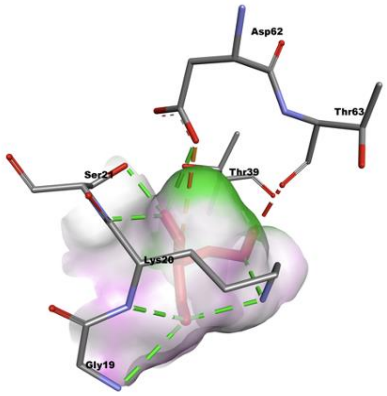
Table.5. Binding energy and active site residue of simultaneously docked compounds in presence of TAGE Methylglyoxal with Target Aldose reductase (PDB ID-4QXA).

S.No	Compound	CID	Complex Model	Binding Energy (kcal/mol)	Residue
1	Methylglyoxal (TAGE 3)	880	5	-4.53kcal/mol	GLY16,GLY17,VAL 18,GLY19,LYS20
2	Taxifolin 7-glucoside	14282775	7	-4.95 kcal/mol	LYS20,ASP33,SER3 4,GLN35,PHE37,HI S38,ASP90, LYS125
3	Sissotrin	5280781	2	-5.66kcal/mol	GLY16,GLY17,VAL 18, LYS20, SER 34, HIS38,LYS125,LYS1 56
4	Rutin	5280805	9	-1.87kcal/mol	GLY17,GLN35,LEU 36,PHE37,SER88,A SP90,ASP91,LYS125
5	Genistin	5281377	6	-5.70kcal/mol	GLY17,LYS20,SER2 2,SER34,THR39,AS P62,THR63,ALA64, GLY65

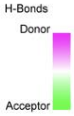
Compared to the earlier molecular docking results from this study, it can be deciphered that there have been declining changes in the stability of the molecular interaction and binding energy of the potential leads. The compound Rutin (5280805) was eliminated as it was weakly bound and produced insignificant results. Further validation of the binding affinities obtained by docking will be done by using Molecular Dynamics (MD) simulation to cross-check the stability of the complex formed.

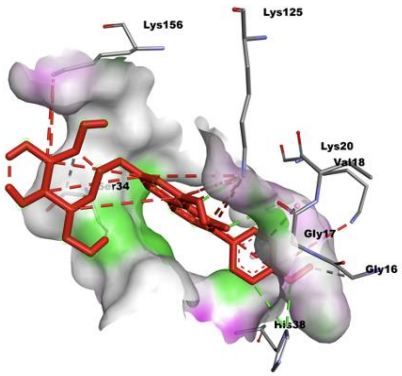


Aldose reductase
(PDB ID – 4QXA)

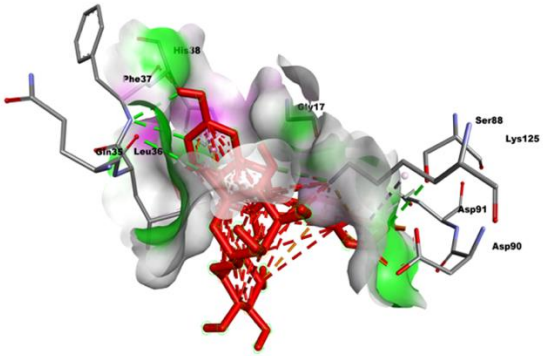


Glyceraldehyde
(TAGE -1)

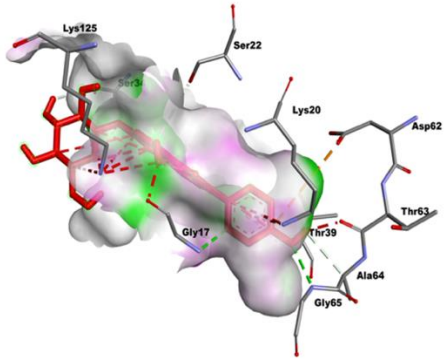




Sissotrin



Rutin



Genestin

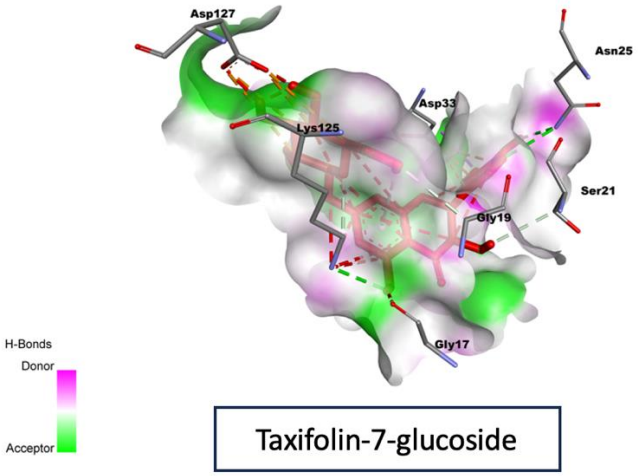
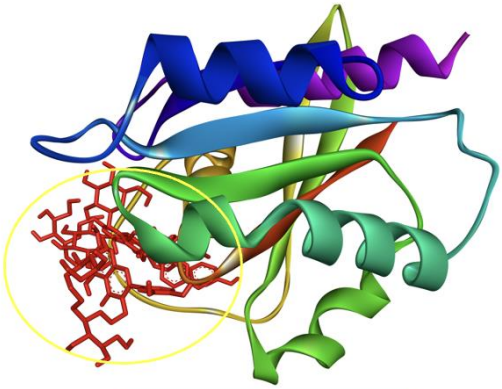
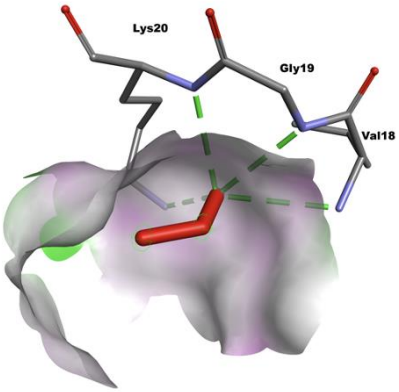
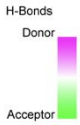


Figure 7. *MLSD results of potential leads Taxifolin-7-Glucoside, Sissotrin, Rutin and Genistin with Target Aldose reductase (PDB ID-4QXA) in presence of Glyceraldehyde. (TAGE)*

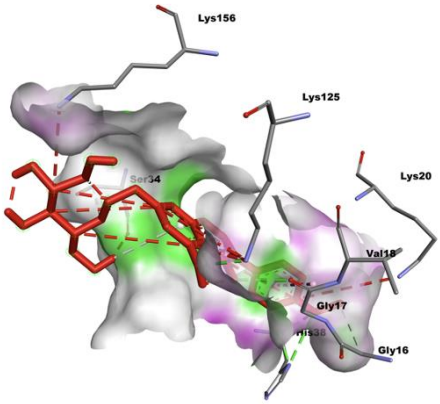


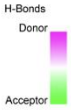
Aldose reductase
(PDB ID – 4QXA)



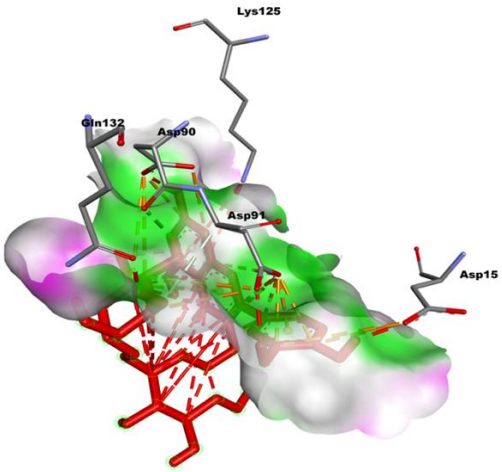


Acetaldehyde
(TAGE-2)

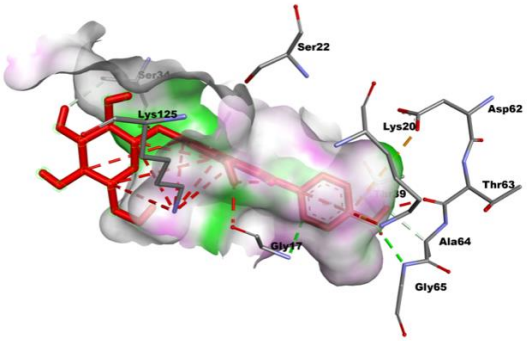




Sissotrin



Rutin



Genestin

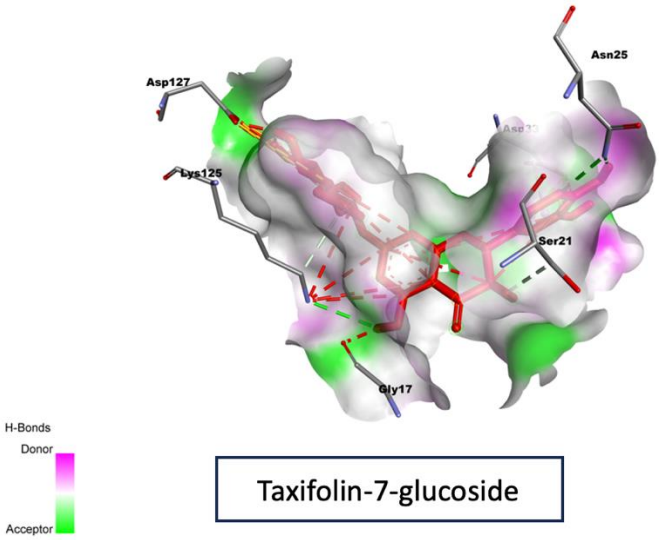
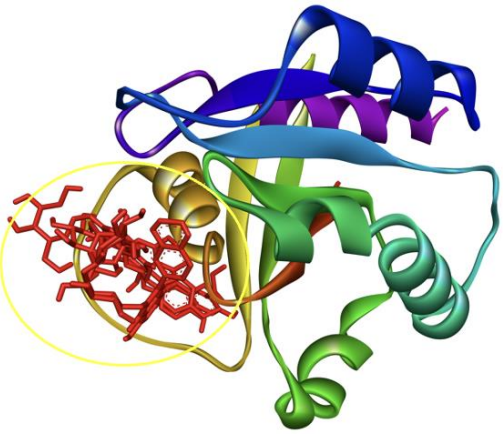
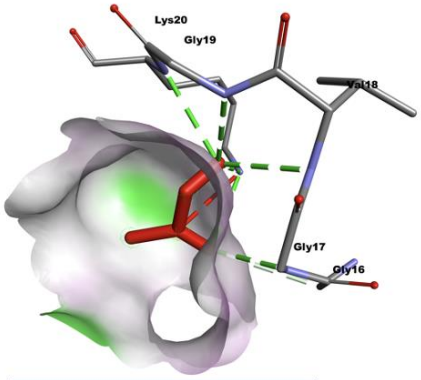


Figure. 8. *MLSD results of potential leads Taxifolin-7-Glucoside, Sissotrin, Rutin and Genistin with target Aldose reductase (PDB ID-4QXA) in presence of Acetaldehyde. (TAGE)*

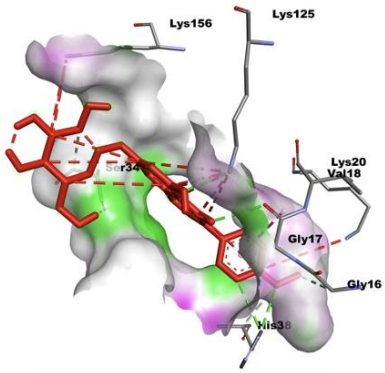


Aldose reductase
(PDB ID – 4QXA)



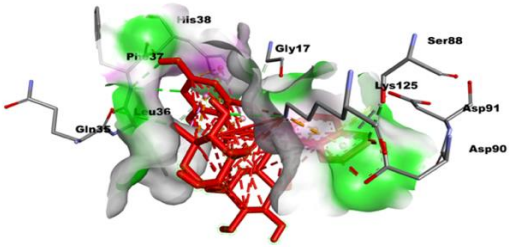
H-Bonds
Donor
Acceptor

Methylglyoxal
(TAGE-3)



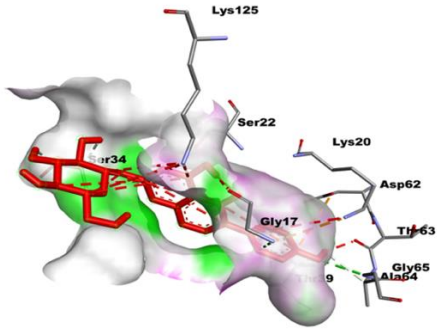
H-Bonds
Donor
Acceptor

Sissotrin



H-Bonds
Donor
Acceptor

Rutin



H-Bonds
Donor
Acceptor

Genestin

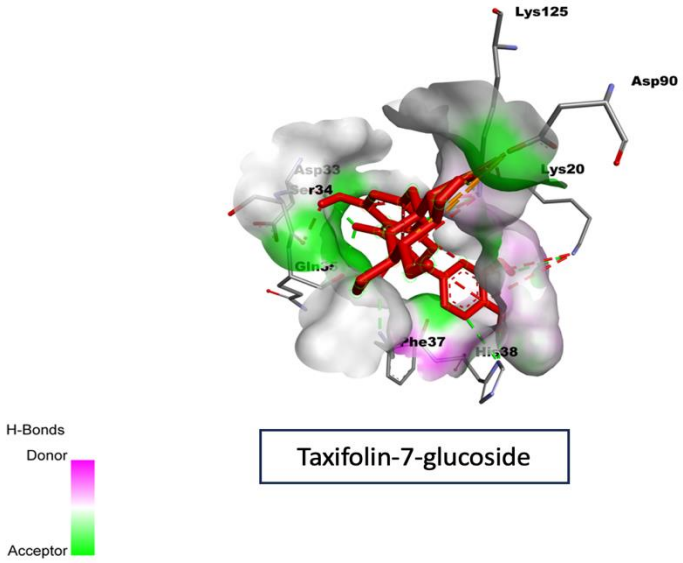


Figure. 9. *MLSD results of potential leads Taxifolin-7-Glucoside, Sissotrin, Rutin and Genistin with target Aldose reductase (PDB ID-4QXA) in presence of Methylglyoxal. (TAGE)*

4.5. Molecular Dynamic Simulation Analysis

The three bound complexes, (Genistin-Target Aldose reductase (PDB ID-4QXA), Sissotrin-Target Aldose reductase (PDB ID-4QXA) and Taxifolin-7-glucoside-Target Aldose reductase (PDB ID-4QXA)) along with two control complexes, (Ascorbic acid-Target Aldose reductase (PDB ID-4QXA) and Epigallocatechin gallate-Target Aldose reductase (PDB ID-4QXA)) and a TAGE complex (Glyceraldehyde-Target Aldose reductase (PDB ID-4QXA)) were subjected to 10 ns all-atom MD simulations. Any compound can cause significant conformational changes when it binds to its Target Aldose reductase (PDB ID-4QXA). The structural dynamics of proteins are estimated using root-mean-square deviation (RMSD).

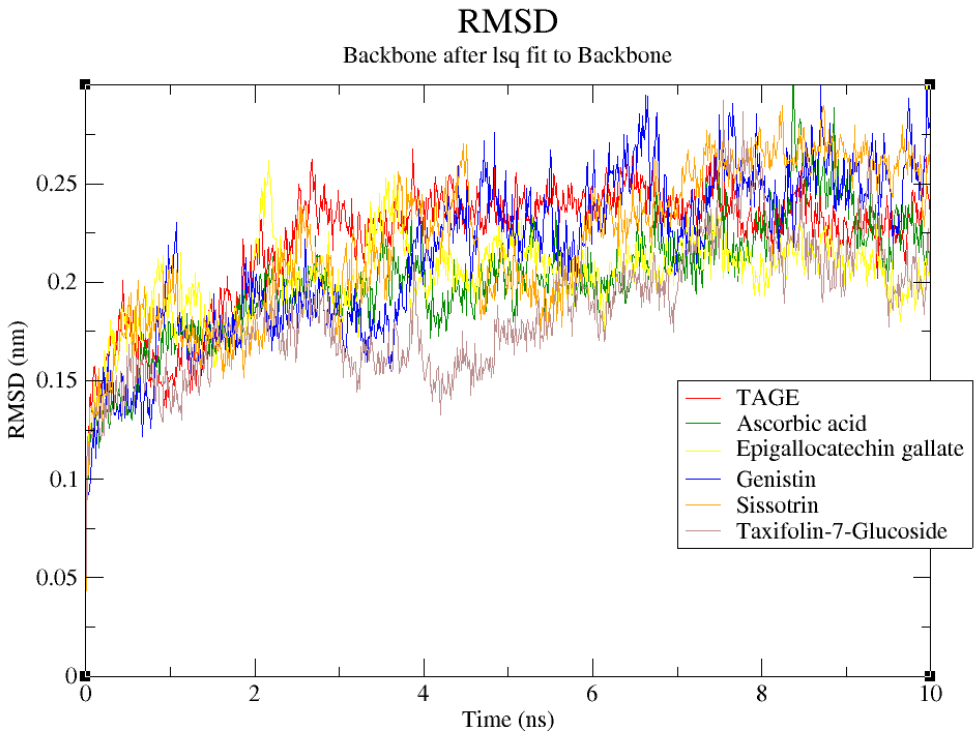


Figure.10. *RMSD plot of complex Glyceraldehyde(TAGE)-Target Aldose reductase (PDB ID-4QXA), Ascorbic acid-Target Aldose reductase (PDB ID-4QXA), Epigallocatechin gallate-Target Aldose reductase(PDB ID-4QXA), Genistin-Target Aldose Reductase (PDB ID-4QXA), Sissotrin-Target Aldose reductase (PDB ID-4QXA) and Taxifolin-7-glucoside-Target Aldose reductase (PDB ID-4QXA).*

We computed the RMSD of all the systems to analyze the structural dynamics of target with two control, TAGE and three potential leads. **(Figure.10.)** the six complex had average RMSD value that lies between 0.25 to 0.255 nm, respectively. Similarly, the average RMSF of two control was 0.133 nm and 0.152 nm and TAGE compound was 0.148, while of genistin, sissotrin and taxifolin-7-glucoside was calculated to be 0.142, 0.154 and 0.137 nm, respectively.

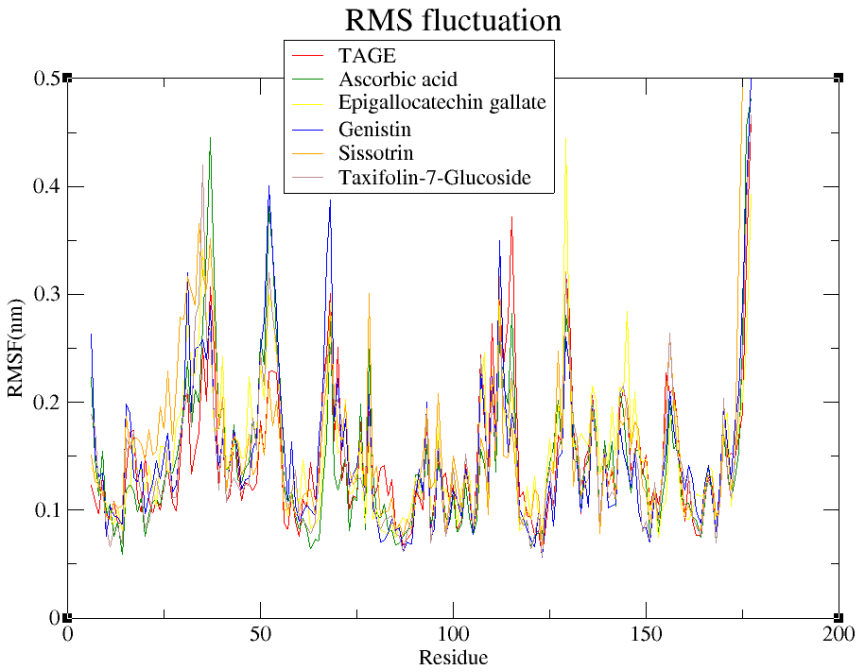


Figure.11. *RMSF plot of complex Glyceraldehyde (TAGE)-Aldose reductase, Ascorbic acid-Aldose reductase, Epigallocatechin gallate-Aldose reductase, Genistin-Aldose reductase, Sissotrin-Aldose reductase and Taxifolin-7-glucoside-Aldose reductase.*

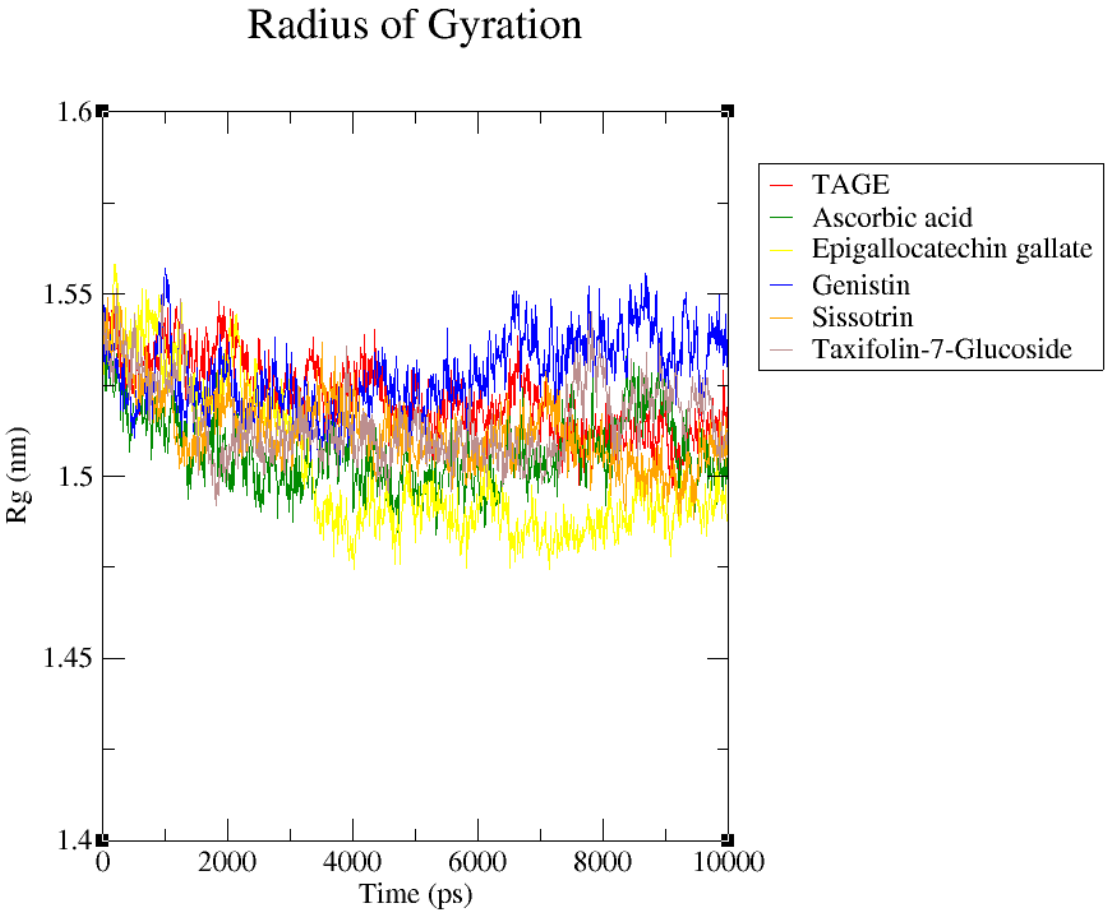


Figure.12. *Time evolution of the Radius of gyration.*

The radius of gyration (Rg) is a structural variable related to protein's three-dimensional (3D) structure and overall conformation has been used to investigate its compactness and folding tendency (**Figure.12.**). The mean Rg values of the TAGE-Target Aldose reductase (PDB ID-4QXA) were estimated about 1.548 nm, while controls Ascorbic acid-Target Aldose reductase (PDB ID-4QXA) and Epigallocatechin gallate-Target Aldose reductase (PDB ID-4QXA) were estimated about 1.52 nm and 1.492 nm, and potential leads Genistin-Target Aldose reductase (PDB ID-4QXA), Sissotrin-Target Aldose reductase (PDB ID-4QXA) and Taxifolin-7-glucoside-Target Aldose reductase (PDB ID-4QXA) 1.535 nm, respectively. The variation in Rg plot of genistin shows a little variation as compared to the other compounds complex with genistin showing a higher deviation. This change in Rg values might be attributed to target aldose reductase (PDB ID-4QXA) packing when binding to potential leads (genistin, sissotrin and taxifolin-7-glucoside) and controls and TAGE represents the loose or tight packaging of protein during the simulation run. run. In this case, no significant structural change was detected in target following binding to compound, where Rg achieves a stable equilibrium state, implying that the complexes are stable along the path. .

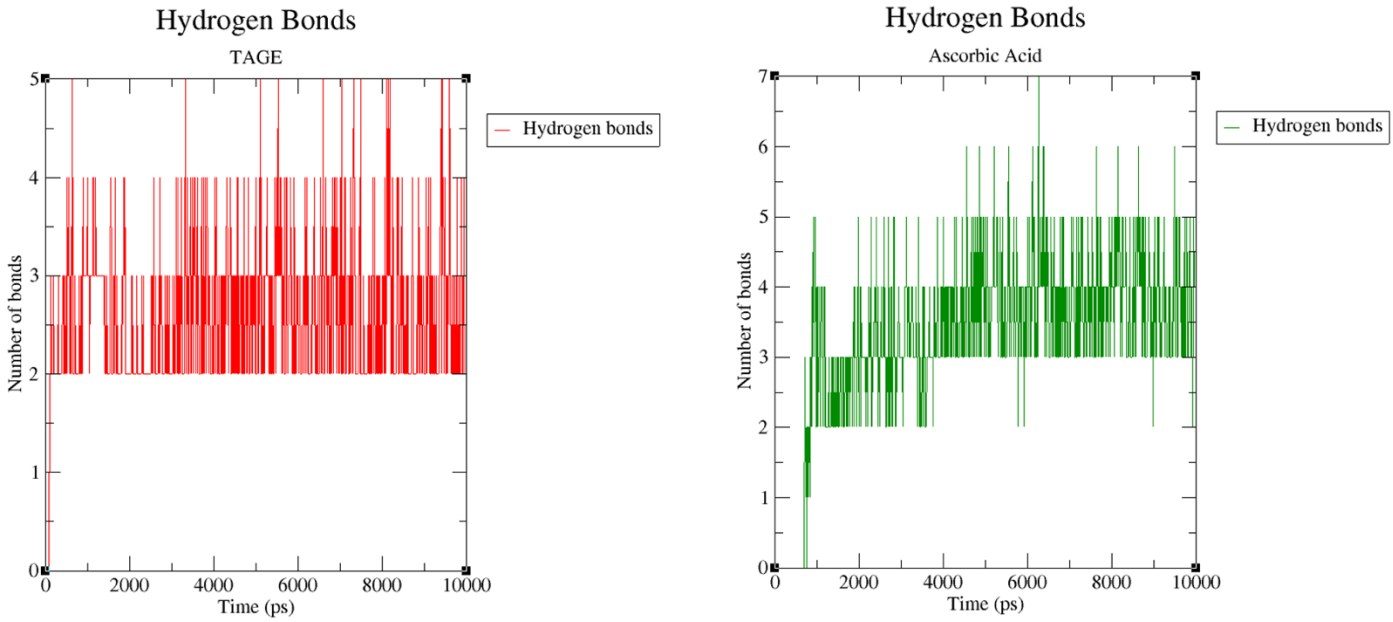


Figure.13. *Intramolecular Hydrogen Bond between Glyceraldehyde (TAGE)- Target Aldose reductase (PDB ID-4QXA) and Ascorbic acid-Target Aldose reductase(PDB ID-4QXA).*

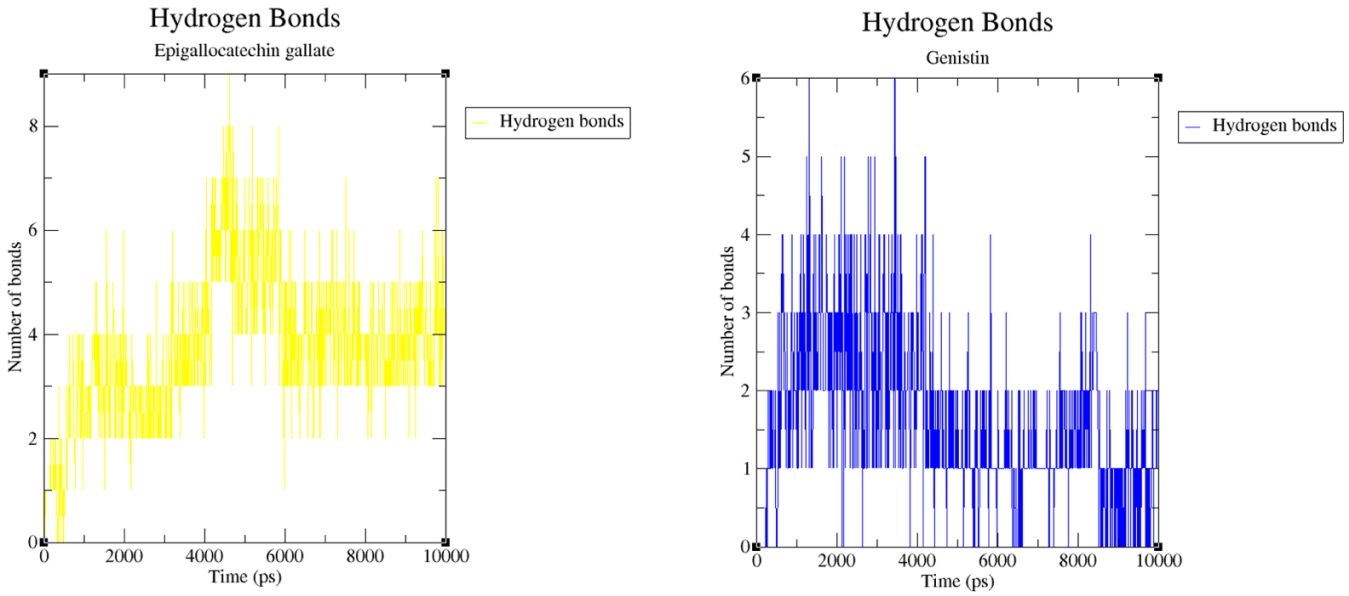


Figure.14. *Intramolecular Hydrogen Bond between Epigallocatechin gallate-Target Aldose reductase(PDB ID-4QXA) and Genistin-Target Aldose reductase(PDB ID-4QXA).*

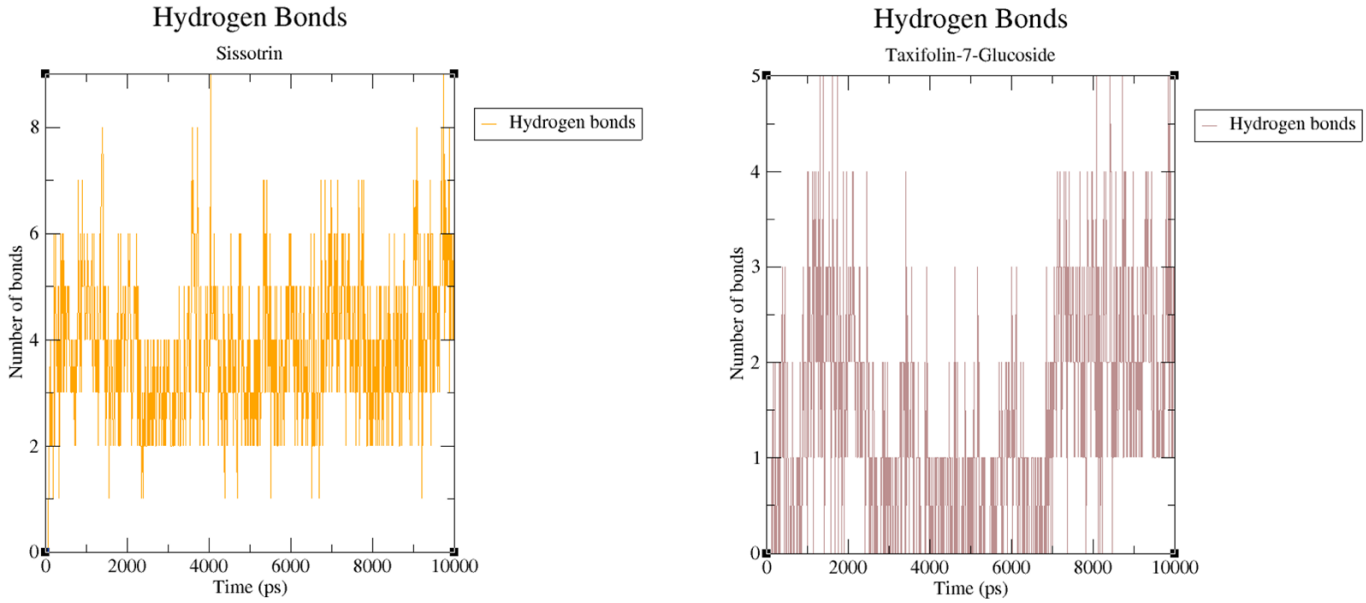


Figure.15. *Intramolecular Hydrogen Bond between Sissotrin-Target Aldose reductase(PDB ID-4QXA) and Taxifolin-7-glucoside-Target Aldose reductase(PDB ID-4QXA).*

Within a protein, intramolecular hydrogen bonds (H-bonds) are critical for maintain its 3D structure, overall shape, and to also validate the stability of the complex. Moreover, the number of intermolecular H-bonds of the three potential leads Genistin-Target Aldose reductase (PDB ID-4QXA), Sissotrin-Target Aldose reductase (PDB ID-4QXA) and Taxifolin-7-glucoside-Target Aldose reductase (PDB ID-4QXA) was evaluated to be 2, 4 and 2, respectively (Figure. 14 and Figure. 15). Hence, the higher number of bonds involvement represents the complex stability.

5. Conclusion

Unmasking the previously unidentified bioactive compounds from diverse natural sources has been a focus of research due to the paradigm change toward the development of medications from natural sources. Numerous reports have shown that millet-based bioactive compounds are good natural candidates for blocking various health problems; one of them is Advanced glycation End products (AGE) formation. Advanced glycation End-products (AGEs) are a group of irreversible heterogeneous compounds produced by the breakdown reactions involving reducing sugars and amino acids.At various phases of the glycation process, millet-based bioactive compounds can prevent the production of AGE and protein glycation.

Natural substances have recently gained popularity as superior possible medicines due to their usefulness in boosting health while having fewer negative effects. Given the costly and time-consuming nature of new medication development, pharmaceutical companies face a pressing need to explore new avenues for drug research and development.Pharmaceutical companies will adopt a global trend of using natural medicines as a source of innovative ingredients for pharmaceuticals.

In this study, the focus is on the modulatory impact of those millet-based bioactive compounds that disrupt the AGE formation. The compounds were screened virtually based on ADMET properties and their pharmacophoric models were generated. Further, molecular interaction studies and molecular dynamic simulation have deciphered that genistin,sissotrin, and taxifolin-7-glucoside possess some therapeutic properties that can prevent the formation of AGE products. This has been established depending upon the binding of the bioactive compounds with the protein target as well as the stability of formed complexes.These compounds inhibit thiol group oxidation, conformational changes, and structural modification in target aldose reductase (PDB ID-4QXA) by scavenging free radical, metal chelation activities, capturing active carbonyl compounds, covering the glycation sites of proteins, and lowering blood glucose levels.

In conclusion, research indicates that millet-derived bioactive compounds offer an appropriate framework and a broader field of investigation for the discovery of potential AGE inhibitors.

6. References

Anis, M. A., &Sreerama, Y. N. (2020). Inhibition of protein glycoxidation and advanced glycation end- product formation by barnyard millet (Echinochloafrumentacea) phenolics. *Food Chemistry*, *315*, 126265

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... &Bourne, P. E. (2000). The protein data bank. *Nucleic acids research*, *28*(1), 235-242.

Chen, J. H., Lin, X., Bu, C., & Zhang, X. (2018). Role of advanced glycation end products in mobilityand considerations in possible dietary and nutritional intervention strategies. *Nutrition & metabolism*, *15*(1), 72.

Daina, A., Michielin, O., &Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific reports*, *7*(1), 42717

Dias, D. A., Urban, S., &Roessner, U. (2012). A historical overview of natural products in drug discovery. Metabolites, 2(2), 303–336.

Dias, I. H., & Griffiths, H. R. (2014). Oxidative stress in diabetes–circulating advanced glycation endproducts, lipid oxidation and vascular disease. *Annals of clinical biochemistry*, *51*(2), 125-127.

Eberhardt, J., Santos-Martins, D., Tillack, A. F., & Forli, S. (2021). AutoDock Vina 1.2. 0: New docking methods, expanded force field, and python bindings. *Journal of chemical information and modeling*, *61*(8), 3891-3898.

Eggleton JS, Jialal I. Thiazolidinediones. [Updated 2023 Feb 20]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan.

Huey, R., Morris, G. M., & Forli, S. (2012). Using AutoDock 4 and AutoDock vina with AutoDockTools: a tutorial. *The Scripps Research Institute Molecular Graphics Laboratory*, *10550*(92037), 1000.

Jejurikar, B. L., &Rohane, S. H. (2021). Drug designing in discovery studio.

Kang, Q., Dai, H., Jiang, S., & Yu, L. (2022). Advanced glycation end products in diabetic retinopathyand phytochemical therapy. *Frontiers in Nutrition*, *9*, 1037186.

Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., ... & Bryant, S. H. (2016).PubChem substance and compound databases. *Nucleic acids research*, *44*(D1), D1202-D1213.

Kuzan, A. (2021). Toxicity of advanced glycation end products. *Biomedical Reports*, *14*(5), 1-8.

Lazenby, D. (2000). Cygwin: For Windows NT. *Linux Journal*, *2000*(75es), 14-es.

Li, D., Liu, M. S., & Ji, B. (2015). Mapping the dynamics landscape of conformational transitions inenzyme: the adenylate kinase case. *Biophysical journal*, *109*(3), 647-66

Lin, J. A., Wu, C. H., & Yen, G. C. (2018). Perspective of advanced glycation end products on humanhealth. *Journal of agricultural and food chemistry*, *66*(9), 2065-2070.

Loschwitz, J., Jäckering, A., Keutmann, M., Olagunju, M., Olubiyi, O. O., &Strodel, B. (2021). Datasetof AMBER force field parameters of drugs, natural products and steroids for simulations using GROMACS. *Data in Brief*, *35*, 106948.

Majid, A., & Priyadarshini CG, P. (2020). Millet derived bioactive peptides: A review on their functional properties and health benefits. *Critical reviews in food science and nutrition*, *60*(19), 3342-3351

Peng, X., Ma, J., Chen, F., & Wang, M. (2011). Naturally occurring inhibitors against the formation ofadvanced glycation end-products. *Food & function*, *2*(6), 289-301.

Perrone, A., Giovino, A., Benny, J., & Martinelli, F. (2020). Advanced glycation end products (AGEs):biochemistry, signaling, analytical methods, and epigenetic effects. *Oxidative medicine and cellular longevity*, *2020*.

Prasad, C., Davis, K. E., Imrhan, V., Juma, S., &Vijayagopal, P. (2019). Advanced glycation end products and risks for chronic diseases: intervening through lifestyle modification. *American journal of lifestyle medicine*, *13*(4), 384-404.

Rapaport, D. C., & Rapaport, D. C. R. (2004). The art of molecular dynamics simulation. Cambridgeuniversity press.

Reddy, V. P., Aryal, P., &Darkwah, E. K. (2022). Advanced glycation end products in health and disease. *Microorganisms*, *10*(9), 1848.

RStudio, T. (2020). RStudio: integrated development for R. *Rstudio Team, PBC, Boston, MA URL http://www. rstudio. com*.

Rungratanawanich, W., Qu, Y., Wang, X., Essa, M. M., & Song, B. J. (2021). Advanced glycation end products (AGEs) and other adducts in aging-related diseases and alcohol-mediated tissueinjury. *Experimental & molecular medicine*, *53*(2), 168-188.

Song Q, Liu J, Dong L, Wang X, Zhang X. Novel advances in inhibiting advanced glycation end product formation using natural compounds. Biomed Pharmacother. 2021

Song, Q., Liu, J., Dong, L., Wang, X., & Zhang, X. (2021). Novel advances in inhibiting advanced glycation end product formation using natural compounds. *Biomedicine & Pharmacotherapy*, *140*, 111750.

Studio, D. (2008). Discovery studio. *Accelrys [2.1]*.

Takeuchi, M. (2016). Serum levels of toxic AGEs (TAGE) may be a promising novel biomarker for the onset/progression of lifestyle-related diseases. *Diagnostics*, *6*(2), 23.

Takeuchi, M. Toxic AGEs (TAGE) theory: a new concept for preventing the development of diseasesrelated to lifestyle. *DiabetolMetabSyndr* **12**, 105 (2020).

Takeuchi, M., Sakasai-Sakai, A., Takata, T., Takino, J. I., & Koriyama, Y. (2022). Effects of toxic AGEs(TAGE) on human health. *Cells*, *11*(14), 2178.

Takeuchi, M., Sakasai-Sakai, A., Takata, T., Takino, J. I., Koriyama, Y., Kikuchi, C., ... & Matsunaga,

T. (2021) Intracellular toxic AGEs (TAGE) triggers numerous types of cell damage. *Biomolecules*, *11*(3), 387.

Takeuchi, M.; Takino, J.; Yamagishi, S. Involvement of the Toxic AGEs (TAGE)-RAGE System in thePathogenesis of Diabetic Vascular Complications: **2010**, *11*, 1468–1482.

Twarda-Clapa, A., Olczak, A., Białkowska, A. M., &Koziołkiewicz, M. (2022). Advanced glycation end-products (AGEs): Formation, chemistry, classification, receptors, and diseases related to AGEs. *Cells*, *11*(8).

Xiao, C., Zhang, W., &Lv, Z. (2016, May). Multibeam Bathymetric Data Processing and ApplicationsBased on Cygwin and MB-System. In *2016 2nd Workshop on Advanced Research and Technology inIndustry Applications (WARTIA-16)* (pp. 190-194). Atlantis Press.

AMNATAMIMI20032023

By: AMNA TAMIMI

As of: Aug 24, 2023 01:43:03 PM

Similarity Index
8%

sources:

1 14 words / 1% - Internet from 13-Jan-2023 12:00AM pubmed.ncbi.nlm.nih.gov

2 18 words / 1% - Internet from 24-Jun-2019 12:00AM open.library.ubc.ca

3 59 words / 1% - from 13-Jun-2023 12:00AM iopscience.iop.org

4 37 words / 1% - Internet from 11-Feb-2019 12:00AM dspace.mit.edu

5 30 words / 1% - Internet from 24-Sep-2022 12:00A www.jbcrs.org

6 19 words / < 1% match - Internet from 04-Aug-2015 12:00AM www.frontiersin.org

7 76 words / < 2% match - Internet from 31-Mar-2020 12:00AM www.annualreviews.org

8 15 words / < 1% match - Internet from 17-Oct-2021 12:00AM nch.adam.com

9 88 words / < 2% match – Internet from 21-Dec-2014 12:00AM www.ncbi.nlm.nih.gov