

Metabolic Analysis of Dihydroquercetin treated - Hepatocellular Carcinoma Cells



By:

Amir Anwar Tareen (00000323380)

Haniya Hammad (00000323732)

Subha Niazi (00000323953)

Atta-ur-Rahman School of Applied Biosciences (ASAB)

National University of Sciences and Technology (NUST)

2023

Metabolic Analysis of Dihydroquercetin treated - Hepatocellular Carcinoma Cells



By:

Amir Anwar Tareen (00000323380)

Haniya Hammad (00000323732)

Subha Niazi (00000323953)

Supervised by: Dr. Maria Shabbir

Atta-ur-Rahman School of Applied Biosciences (ASAB)

National University of Sciences and Technology (NUST)

2023

Metabolic Analysis of Dihydroquercetin treated - Hepatocellular Carcinoma Cells



By:

Amir Anwar Tareen (00000323380)

Haniya Hammad (00000323732)

Subha Niazi (00000323953)

A thesis submitted to the National University of Science and Technology,
Islamabad, in partial fulfillment of the requirements of the degree of

Bachelor of Science in
Applied Biosciences

Thesis Supervisor: Dr. Maria Shabbir


Atta-ur-Rahman School of Applied Biosciences (ASAB)


National University of Sciences and Technology (NUST)


2023

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of BS FYP Thesis written by Mr./Ms. Amir Anwar Tareen, Haniya Hammad and Subha Niazi (Registration No. 00000323380, 00000323732, 00000323953, of (School/College/Institute) "Atta-ur- Rahman School of Applied Biosciences" has been vetted by undersigned, found complete in all respects as per NUST Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of BS degree. It is further certified that necessary amendments as pointed out, during final presentation of the scholar, have also been incorporated in the said thesis.

Signature: 
Name of Supervisor: Dr. Maria Shabbir
Date: 24.5.23

Signature (HOD): 
Date: 24.5.2023

Signature (Dean/Principal) 
Date: _____

AUTHOR'S DECLARATION

We, Amir Anwar Tareen, Haniya Hammad and Subha Niazi hereby state that our BS FYP thesis titled.

"Metabolic Analysis of Dihydroquercetin treated Hepatocellular Carcinoma Cell" is our own work and has not been submitted previously by us for taking any degree from this university 'National University of Sciences and Technology' or anywhere else in the country/world.

At any time if our statement is found to be incorrect even after our graduation, the university has the right to withdraw our BS degree.

Name of student:

Amir Anwar Tareen:



Haniya Hammad:



Subha Niazi:



Date: 24/5/23

Certificate for Plagiarism

It is certified that BS FYP Thesis Titled "Integrative Proteomics Profiling for Understanding the Role of Genetic Markers in Liver Cancer" by Amir Anwar Tareen, Haniya Hammad and Subha Niazi has been examined by me. I undertake the following: -

- a. Thesis has significant new work / knowledge as compared to already published or are under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph, or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.
- b. The work presented is original and own work of the author (i.e., there is no plagiarism). No ideas, processes, results, or words of others have been presented as Author's own work.
- c. There is no fabrication of data or results which have been compiled / analyzed.
- d. There is no falsification by manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.
- e. The thesis has been checked using TURNITIN (copy of originality report is attached) and found within limits as per HEC / NUST plagiarism Policy and instructions issued from time to time.

Date: 24/5/23

Dr. Maria Shabbir
Tenured Associate Professor
Department of Healthcare Biotechnology
Attache-Khanmaza School of Applied
Biosciences (MSAB), NUST Islamabad
Signature & Stamp of Supervisor

This thesis is dedicated to our beloved parents, friends, and family whose encouragement and guidance has been unmatched.

Acknowledgments

Above all, we would like to first of all thank our research supervisors, Dr Maria Shabbir and Dr Yasmin Badshah, whose guidance and dedication helped this project come to fruition. Their constant support has been much appreciated throughout the duration of our degree.

Secondly, with utmost gratitude, we thank Dr Hussnain A. Janjua and the entirety of the ASAB department for providing us with all of the required facilities that aided in the progress and eventual completion of our project.

We would also like to show particular gratitude to Ms. Khushbakht Khan and Ms Fizah who has aided us tremendously in the research work as well as our thesis write up and it's. Their guidance is sincerely valued.

Finally, and most importantly, the kindness and support of our parents and family, friends and fellow peers has been immensely appreciated. Without their encouragement, we would not be where we are today, and we thank them from the bottom of our hearts.

Table of Contents

Acknowledgments	8
List of Abbreviations.....	16
Abstract.....	19
CHAPTER 1.....	20
INTRODUCTION.....	20
1.1 Background.....	20
1.2 Development of cancer	20
1.3 Primary Liver Cancer.....	21
1.4 Epidemiology of Hepatocellular Carcinoma	21
1.5 Protein biomarkers.....	22
1.6 Top-down proteomics vs bottom-up proteomics	23
1.6.1 Top-down:.....	23
1.6.2 Applications	23
1.6.3 Bottom-up:.....	23
1.7 Role of GC/MS in cancer studies	25
1.8 Role of GCMS in Diseases	25
1.9 Aims and Objectives.....	26
CHAPTER 2.....	27
LITERATURE REVIEW.....	27
2.1 Liver Cancer	27
2.2 Distribution of Liver Cancer Cases and Deaths Globally.....	27
2.3 Distribution of Liver Cancer Cases and Deaths in Pakistan.....	30
2.4 Types of Liver Cancer	31

2.5 Diagnosis and Treatment of Liver Cancer	32
2.6 Cancer Biomarkers	33
2.7 Protein Biomarkers	33
2.8 Gas chromatography-mass spectrometer	34
CHAPTER 3.....	36
METHODOLOGY.....	36
3.1 In silico Part:	36
3.1.1 ADMET and Pharmacophore Analysis	36
3.1.2 Drug Energy Minimization	37
3.1.3 Retrieving Protein Structures and their Structural Accuracy Prediction	37
3.1.4. Drug and Protein Docking	38
3.1.5 Interaction Analysis	38
3.2 In-Vitro	39
3.2.1 Cell Lines pellet collection	39
3.3 Samples Preparation	40
3.3.1 Chemicals required.	40
3.4 Steps involved in the preparation of serum samples.....	40
3.4.1 Samples lysis.....	40
3.4.2 Centrifugation	41
3.4.3 Concentrating the metabolites.	41
3.4.4 Derivatization.....	41
3.4.5 Silylation.....	41
3.4.6 Incubation	42
3.5 GC/MS	42
3.6 GC/MS Analysis.....	43

3.7 GC-MS Data Analysis	43
3.8 Protein Pathway Analysis and Gene ID Extraction	43
3.9 Gene Ontology Analysis	44
3.10 Enrichment Analysis.....	44
CHAPTER 4.....	45
RESULTS.....	45
4.1 The Absorption, Distribution, Metabolism, Elimination, and Toxicity (ADMET) analysis of Dihydroquercetin	45
4.1.1 Physicochemical Properties	45
4.1.2 Pharmacokinetics	46
4.1.3 Toxicity	48
4.2 Pharmacophore Analysis	49
4.3 Energy Minimization	50
4.4 Docking.....	51
4.5 Interaction Analysis	52
4.6 Gene Ontology and Enrichment Analysis	60
4.7 Protein Pathway Analysis	64
CHAPTER 5.....	66
DISCUSSION.....	66
References.....	72

Table of Figures

Figure 1: Techniques used in proteomics. MALDI or electrospray ionization techniques are used to ionize the sample particles to make them easier to analyze.	22
Figure 2: Comparison between 2A and 2B top-down and bottom-up methods Top-down method uses large molecules that are ionized to lead them to analyzer. Bottom-up starts with small protein molecules, ionizing them and doing the analysis.	24
Figure 3: Number of new cases in 2020, both sexes, all ages. The pie chart shows cancer cases of different body parts. Breast cancer is the most prevalent. (Observatory).....	28
Figure 4: Number of deaths in 2020, both sexes, all ages. This chart shows the deaths being maximum from lung cancer, liver cancer one of the major death contributors. (Observatory)	28
Figure 5: Number of new cases in 2020, Pakistan, both sexes, all ages. In Pakistan, liver cancer accounts for 3% of cancer cases, breast cancer being the most prevalent. (Observatory)	30
Figure 6: Number of deaths in 2020 Pakistan, both sexes, all ages. deaths due to liver cancer account for around 4.4% of the total deaths in Pakistan. (Observatory).....	31
Figure 7 : By taking a sample of blood or another bodily fluid, liquid biopsy offers the chance to identify and evaluate cancers wherever they may exist in the body. The image includes key types of proteins and identifies some typical liquid biopsy targets(Landegren & Hammond, 2021)	34
Figure 8: GCMS setup explained; the figure gives a step-by-step illustration of the processes inside the GCMS setup for the protein analysis.	35
Figure 9 : Dihydroquercetin Pharmacophore Analysis. These graph shows (A) Blood-brain barrier permeability (LogBB), (B) Human intestinal absorption (HIA%), (C) hERGF affinity (pKi) and (D) hERG activity (pIC50).....	47
Figure 10: Pharmacophore Model. Figure 2 (A) and (B) is a Depiction of pharmacophore models produced through ZINCpharmer. Green Sphere show Hydrophobic Interaction, Purple Sphere indicates aromatic group, White Sphere indicate Hydrogen Donor, Yellow Sphere indicates Hydrogen Acceptor.....	49
Figure 11 : Energy Minimization. Dihydroquercetin after energy minimization by Chem 3D.....	51

Figure 12 : A schematic representation of the interactions between Dihydroquercetin and the AKT2 protein. Fig 12 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues with the ligand and (C) 2Dmodel of ligand obtained from LigPlot+. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding. 53

Figure 13: A schematic representation of the interactions between Dihydroquercetin and the Proto-oncogene cJun protein. Fig 13 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues with the ligand and (C) 2D model of ligand obtained from LigPlot+. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding..... 54

Figure 14: A schematic representation of the interactions between Dihydroquercetin and the Epidermal growth factor receptor. Fig 14 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues with the the ligand and (C) 2D model of ligand obtained from LigPlot+. Black colour atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding. 55

Figure 15: A schematic representation of the interactions between Dihydroquercetin and Eukaryotic translation initiation factor 2A. Fig 15 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues with the ligand and (C) 2D model of ligand obtained from LigPlot+. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding. 57

Figure 16: A schematic representation of the interactions between Dihydroquercetin and Serine/threonine-protein kinase mTOR. Fig 16 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues with the ligand and (C) 2D model of ligand obtained from LigPlot+. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding. 58

Figure 17: A schematic representation of the interactions between Dihydroquercetin and Phosphoinositide 3-kinase regulatory subunit 4. Fig 17 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues

residues with the ligand and (C) 2D model of ligand obtained from LigPlot+. Black colour atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding. 59

Figure 18: The enrichment overview of the metabolites. It refers to a systematic analysis of the biological pathways and functions associated with a set of metabolites. 63

Figure 19: Pathways generated through Cytoscape. 64

Figure 20 Mtor Signaling Pathway Involving Glycolysis 65

List of Tables

Table 1: Cancer incidence and mortality statistics worldwide and by region (Observatory)	29
Table 2 The six proteins and their respective Alpha Fold ID	38
Table 3: Physicochemical Properties of Dihydroquercetin	45
Table 4: Pharmacokinetic Analysis of Dihydroquercetin	46
Table 5: Toxicity Analysis of Dihydroquercetin	48
Table 6: Pharmacophore Class with positions, dimensions, and radius	50
Table 7: Protein and their Vina Score.....	52
Table 8: Showing the Molecular Function of metabolites.....	61
Table 9: Showing the Biological Process of metabolites.....	61
Table 10: Showing the cellular components of metabolites.	62
Table 11: Showing the Protein Class of metabolites.	62
Table 12: Pathways of metabolites.	63

List of Abbreviations

SNPs: Single-nucleotide polymorphism

CTCs: Circulating Tumor Cells

DTCs: Disseminated Tumor Cells

HCC: Hepatocellular Carcinoma

ICC: Intrahepatic Cholangiocarcinoma

HBV: Hepatitis B

HCV: Hepatitis C virus

WHO: World Health Organization

miRNA: MicroRNA

2-D: Two-Dimensional

PTM: Post-translational modifications

Lys C: Endo proteinase Lys-C

Arg C: Endopeptidase Arg-C

GC-MS: Gas chromatography–mass spectrometry

IEM: Inborn Errors of Metabolism

SLO: Smith-Lemli-Opitz syndrome

TB: Tuberculosis

CT: Computerized tomography

MRI: Magnetic resonance imaging

AFP: Alpha-fetoprotein

CA: Carcinoembryonic Antigen

ELISA: enzyme-linked immunosorbent assays

MS: Mass spectrometry (MS).

ADMET: Absorption, distribution, metabolism, excretion, and toxicity

SDF: Spatial Data File

PDB: Program Database format

Hep-G2: Hepatoblastoma cell line

PBS: Phosphate-buffered saline

TMCS: Trimethylchlorosilane

MSTFA: N-methyl-N-(trimethylsilyl) trifluoroacetamide

GC: Gas Chromatography

GLC: Gas-liquid chromatography

GSC: Solid phase Gas-solid chromatography

K_c: A distribution coefficient)

KEGG: Kyoto Encyclopedia of Genes and Genomes

PANTHER: Protein analysis through evolutionary relationships

hERG: Human ether-a-go-go related gene.

IC₅₀: Half-maximal inhibitory concentration

PI3k: Phosphoinositide 3-kinase

AKT: Protein kinase B

SAMTOR: S-adenosylmethionine sensor for the mTORC1 pathway

HIF1 α : Hypoxia-Inducible Factor 1 alpha.

ROS: Reactive oxidative species

DHQ: Dihydroquercetin

MTOR: Mammalian target of rapamycin

GLUT: Glucose Transporter

Abstract

In recent years, liver cancer has grown to become one of the most challenging and prevalent issues in the healthcare sector with millions of deaths each year. However, the issue with the disease lies in the late diagnosis and weak prognosis of the majority of patients. Moreover, conventional treatments are highly toxic and unaffordable for the majority of patients. Our study investigates the anticancer efficacy of Dihydroquercetin compound, on HepG2 cell lines of liver cancer shown promising antioxidant and anti-inflammatory properties. In silico and Invitro analysis of Dihydroquercetin ligand was performed using ADMET analysis, Pharmacophore analysis and GCMS analysis to understand the metabolomics of Liver cancer. Our results demonstrated strong binding affinity of DHQ with Serine/threonine-protein kinase mTOR which had a vina score of -8.3 which showed potential for being used as a target metabolite for the drug, and through this we can downregulate the glycolysis pathway in cancer cells. This research will help researchers to further investigate on creating more effective and anti-cancer drugs which will have a significant impact on survival rate.

CHAPTER 1

INTRODUCTION

1.1 Background

Despite being one of the most widely studied diseases of all time, cancer is an enigma. Hundreds, if not thousands, of drugs, have been tested on different cancers, and a few of them have been approved to be used clinically on humans as effective treatments over the years. Despite this, cancer still presents itself as a problem with its ever-evolving characteristics requiring constant modifications in treatment and prognosis mechanisms.

1.2 Development of cancer

Cancer develops due to multistep processes resulting in the accumulation of several genomic alterations, characterized by unrestricted proliferation of cells in the body, invasion of these cells, and metastasis of these cells to different parts of the organism.

There could be multiple reasons for such an occurrence which include but are not limited to chromosomal abnormalities, epigenetic changes, mutations and SNPs in mRNAs and miRNAs, cell cycle dysregulation such as bypassing of cells at certain checkpoints in the cell cycle hence allowing for uncontrolled division of cells, presence of hereditary oncogenes leading to cancers and many more.

The early spread of cancer cells is usually undetected by current imaging technologies. Patients with cancer in the initial stages sometimes exhibit no signs of metastases. Given this, specific sensitive methods have been developed to detect circulating tumor cells (CTCs) in the peripheral blood and disseminated tumor cells (DTCs) in the bone marrow. These technologies can be classified into cytometric and immunological and molecular approaches. It is essential to understand that cancer development not only changes cell morphology but also physiology. Many metabolic pathways are altered during the

development of cancers. This leads to important changes in the synthesis and modifications of metabolites of these pathways. Studying the changes in their expression, quantity, and post-transcriptional modifications could help us utilize them to manipulate these pathways for cancer prevention and treatment and develop potentially better diagnostic tools for the early detection of these cancers.

1.3 Primary Liver Cancer

Primary Liver Cancer is the 7th most common occurring cancer in the world and ranks 4th in mortality. There are two types of primary liver cancer hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (ICC). Persistent illness caused by hepatitis B (HBV) and hepatitis C virus (HCV) are the major causes of HCC; in the hepatocytes, there is inflammation due to oxidative stress, which causes chronic liver diseases. ICC happens in the epithelial cells of the bile duct, i.e. cholangiocytes. (Petrick & McGlynn, 2019).

1.4 Epidemiology of Hepatocellular Carcinoma

Nearly 73.4% of hepatocellular carcinoma (HCC) cases are reported due to hepatitis C virus (HCV), and hepatitis B (HBV) causes two times as many cases as HCV. It was reported by the world health organization (WHO) 3.5% of the world's population was infected, and 88,700 mortalities were due to cirrhosis and HCC. The burden on the health care system due to HBV-associated HCC will decrease as vaccination is available worldwide. Hepatitis C virus (HCV) can be acquired during childhood, and the symptoms are asymptomatic and 17 times more likely to develop HCC than a non-infected control group. (Petrick & McGlynn, 2019).

In Pakistan, deaths and new cases of cancer are increasing, and the lack of quality and accessibility of the healthcare system also plays a significant role. The average age for hepatocellular carcinoma (HCC) in Pakistan is 7.6 per 100,000 persons annually for males and 2.8 for females. The statistics and knowledge on hepatocellular carcinoma (HCC) in Pakistan are very limited due to the lack of a national policy towards hepatocellular carcinoma (HCC) administration (Hafeez Bhatti et al., 2016).

1.5 Protein biomarkers

Proteomics analysis can be used to compare the protein sample of healthy and diseased individuals to study the possible biomarkers. Certain proteins are found to be present in the diseased person with a greater concentration in comparison to healthy individuals, and we assume that the presence of those proteins controls the tumor aggression and metastasis. among the potential biomarkers

Alpha-fetoprotein (AFP) is found to be prevalent in many HCC patients. Other miRNA panels, including miR-122, miR-192, miR-21, miR-223, miR-26a, miR-27a and miR-801, are very good diagnostics tools and biomarkers for liver cancer.

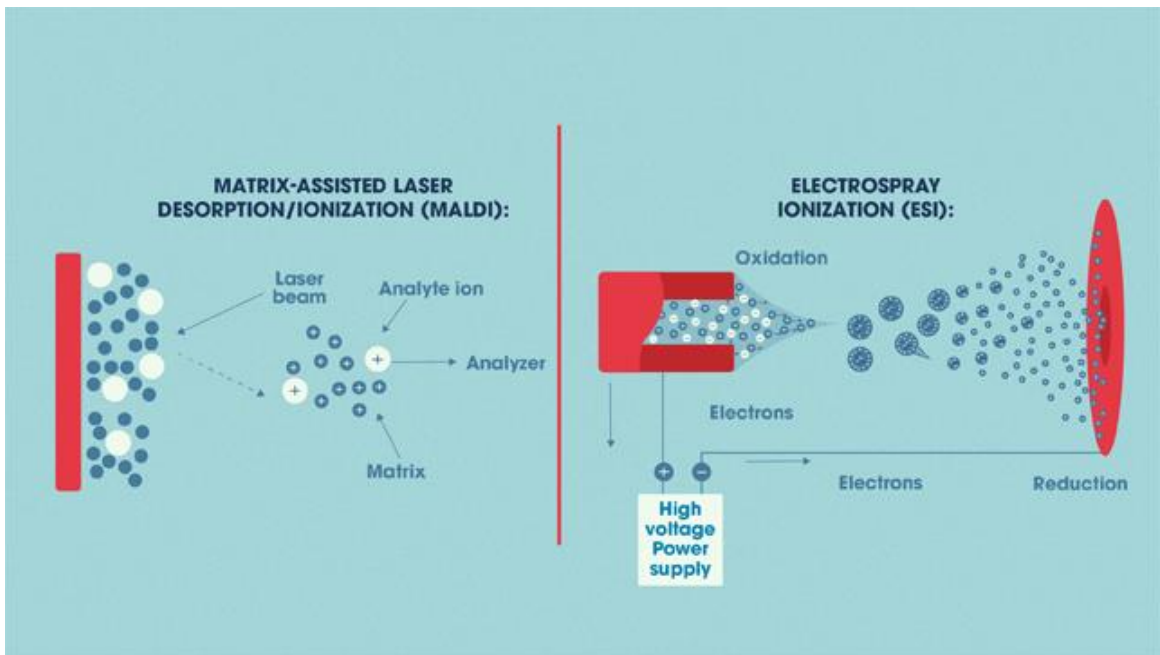


Figure 1: Techniques used in proteomics. MALDI or electrospray ionization techniques are used to ionize the sample particles to make them easier to analyze.

1.6 Top-down proteomics vs bottom-up proteomics

1.6.1 Top-down:

Proteins are separated based on charge and mass using 2-D gel electrophoresis or Mass Spectrometry. This technique is used for separating intact proteins from a sample. Electrospray ionization is used to ionize the sample or matrix-assisted laser desorption/ionization following fragmentation through different separation strategies, like electron-capture dissociation and electron transfer dissociation. As a promising elective procedure for protein identification, profiling, sequencing and PTM, top-down permits MS examination of proteins that have not been divided, meaning the labile primary protein qualities generally obliterated in bottom-up MS are protected.

1.6.2 Applications

Top-down proteomics, by its high-throughput analysis, can be utilized in various research fields like Personalized medicines, Biomarker discovery, Drug discovery and development, Systems biology, Agriculture, Food Science, Paleo proteomics, astrobiology clinical science, forensic science and many more.

1.6.3 Bottom-up:

These Proteomic approaches include proteolytic cleavage of proteins into short peptide sections by proteases. The most broadly utilized protein is a synthetically adjusted trypsin that divides peptide bonds C-terminal to lysine and arginine buildups. A benefit of trypsin is the formation of short doubly or triply charged peptides that can be dissolved in water and easily separated by cation exchange and reversed phase chromatography technique and ionized by electrospray ionization. To expand the quantity of identified peptides and protein sequences, protocols must involve proteases with various arrangement specificities, like Lys C, Arg C etc.

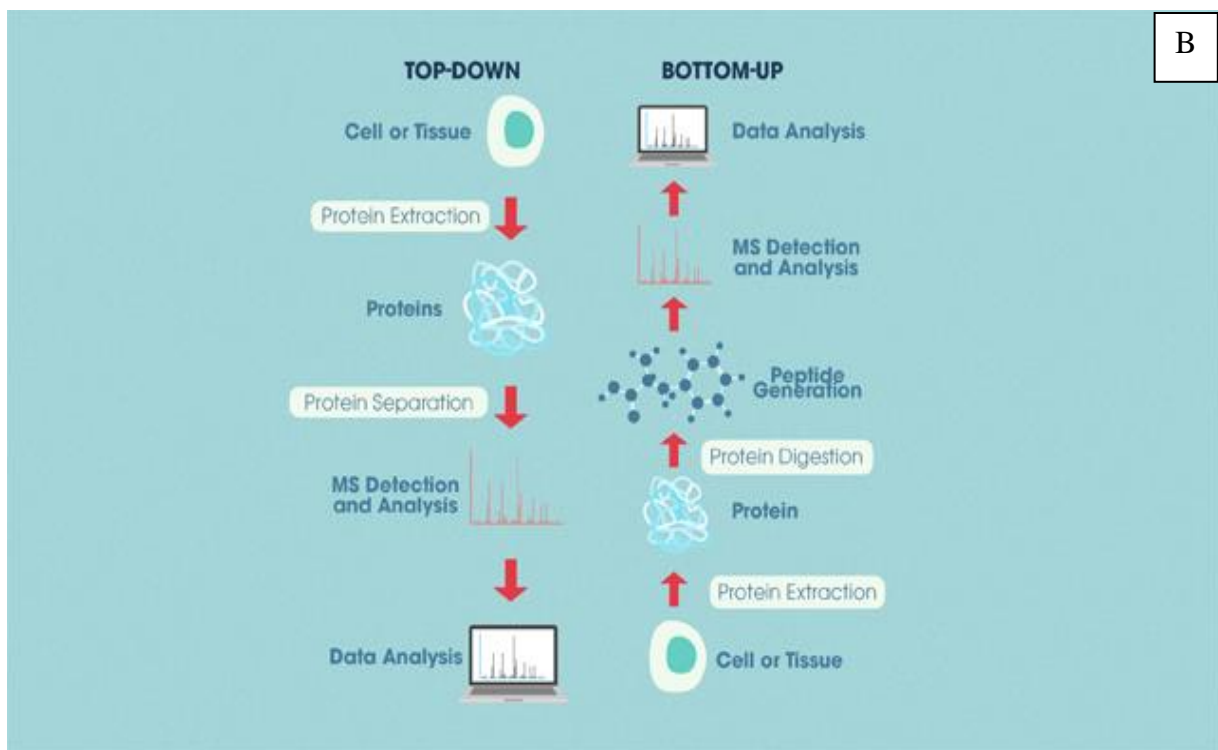
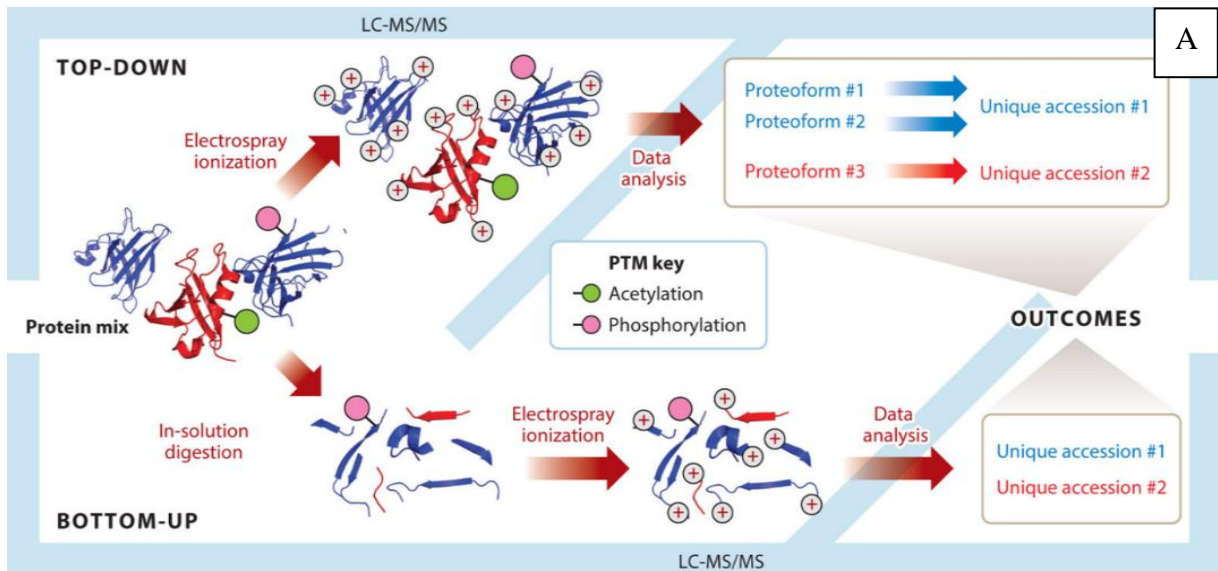


Figure 2: Comparison between 2A and 2B top-down and bottom-up methods. Top-down method uses large molecules that are ionized to lead them to analyzer. Bottom-up starts with small protein molecules, ionizing them and doing the analysis.

1.7 Role of GC/MS in cancer studies

Gas chromatography-mass spectrometry (GC-MS)-based metabolomics is ideal for identifying and quantitating small molecular metabolites (<650 Daltons), including small acids, alcohols, hydroxyl acids, amino acids, sugars, fatty acids, sterols, catecholamines, drugs, and toxins. GC-MS can identify and semi-quantify over 200 compounds per study in human body fluids (e.g., plasma, urine, or stool) samples.

In cancer research, metabolite changes in tumor cells are of immense importance, as discussed above. All these metabolites belong to the different classes of compounds mentioned above, which can be analyzed with GC/MS, proving to be a useful tool to aid such studies. Some general advantages of using GC/MS in such studies are mentioned in the table below. In a study conducted in the recent past, GC/MS-based metabolomic approaches were used to validate the role of urinary sarcosine and target biomarkers for human prostate cancer by microwave-assisted Derivatization. Urine metabolic fingerprinting was done using GC/MS revealing changes in metabolites in prostate cancer cells.

1.8 Role of GCMS in Diseases

Apart from gas chromatography-mass spectrometry (GC-MS) application in cancer, it is widely used for analyzing metabolites in different diseases. It can be used to analyze metabolic disorders; through urine analysis of infants, we can detect them early, such as Inborn Errors of Metabolism (IEM). More than 130 metabolic disorders can be detected through GC-MS (Kałużna-Czaplińska, 2011).

Based on aberrant concentrations of certain chemicals in biological fluids, GC-MS can be used to identify disease conditions like Smith-Lemli-Opitz syndrome (SLO), which can be caused by deficits in cholesterol synthesis and can be identified by GC-MS, which reveals elevated amounts of 7-dehydrocholesterol in the blood and tissues. Additionally, the method has been applied to find pulmonary TB indicators in human breath. GC-MS results can be used to differentiate between patients with and without an infection (Kałużna-Czaplińska, 2011).

1.9 Aims and Objectives

The aims and objective of this study were to identify Dihydroquercetin affects liver cancer cells. By using a gas chromatography-mass spectrometer to monitor the metabolite levels in plasma samples from HCC cases and individuals with liver cirrhosis, we assessed the impact (GCMS). Additionally, they will support the development of efficient cancer patient treatment regimens and treatments.

CHAPTER 2

LITERATURE REVIEW

2.1 Liver Cancer

Liver Cancer initiates in the liver and forms a tumor during a prolonged liver disease or cirrhosis. Liver cancer is the most difficult cancer to medicate, and hepatocellular carcinoma (HCC) is the most occurring primary cancer. Hepatitis B and Hepatitis C are the leading risk factors for HCC. Hepatitis B causes 60%, and Hepatitis C accounts for 33% of cancer (C.-Y. Liu, Chen, & Chen, 2015).

Worldwide cancer ranks as the primary cause of mortality and causes a decrease in life expectancy. In 2020 an estimated 19.3 million new cancer cases and 10.0 million deaths were recorded due to cancer. An increase of 47% in the global cancer burden is expected in 2020; therefore, preventive measures and providing cancer care are important for global cancer control (Sung et al., 2021).

2.2 Distribution of Liver Cancer Cases and Deaths Globally

According to 2020 assessments, liver cancer is the sixth most common type. In men, it is the fifth; in women, it is the ninth most common type of cancer. In 2020 there were 905,677 new cases of liver cancer (Sung et al., 2021). The below images summarize the number of deaths and new cases of liver cancer globally in 2020.

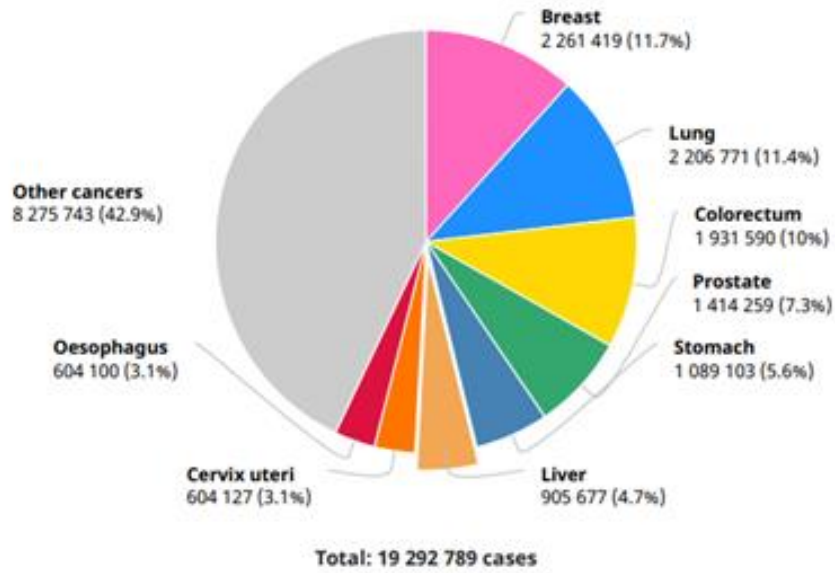


Figure 3: Number of new cases in 2020, both sexes, all ages. The pie chart shows cancer cases of different body parts. Breast cancer is the most prevalent. (Observatory)

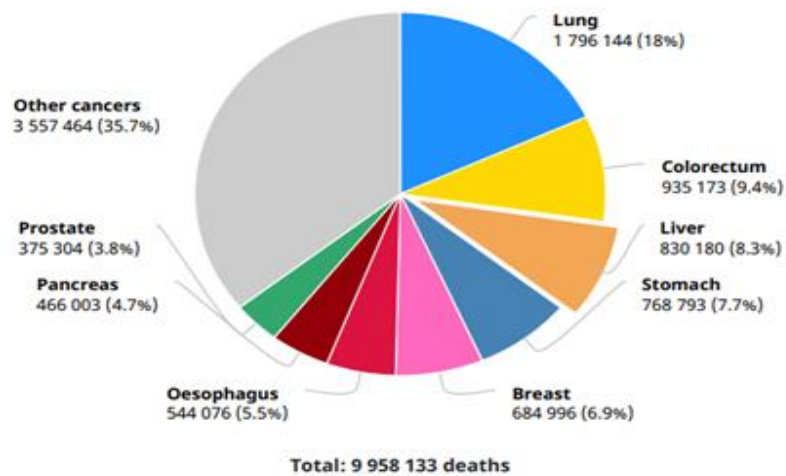


Figure 4: Number of deaths in 2020, both sexes, all ages. This chart shows the deaths being maximum from lung cancer, liver cancer one of the major death contributors. (Observatory)

Table 1: Cancer incidence and mortality statistics worldwide and by region (Observatory)

	Incidence						Mortality					
	Both sexes		Males		Females		Both sexes		Males		Females	
	New cases	Cum. risk 0-74 (%)	New cases	Cum. risk 0-74 (%)	New cases	Cum. risk 0-74 (%)	Deaths	Cum. risk 0-74 (%)	Deaths	Cum. risk 0-74 (%)	Deaths	Cum. risk 0-74 (%)
Eastern Africa	12 326	0.55	7 225	0.68	5 101	0.44	11 542	0.54	6 718	0.66	4 824	0.43
Middle Africa	6 072	0.64	4 204	0.90	1 868	0.40	5 716	0.62	3 956	0.88	1 760	0.39
Northern Africa	31 913	1.83	20 468	2.40	11 445	1.28	30 352	1.75	19 408	2.30	10 944	1.22
Southern Africa	2 601	0.50	1 626	0.72	975	0.32	2 447	0.48	1 574	0.72	873	0.29
Western Africa	17 630	0.94	11 619	1.29	6 011	0.62	16 887	0.92	11 084	1.26	5 803	0.62
Caribbean	3 383	0.61	1 928	0.78	1 455	0.47	3 182	0.57	1 808	0.72	1 374	0.44
Central America	11 819	0.72	5 895	0.78	5 924	0.66	11 231	0.69	5 584	0.74	5 647	0.64
South America	24 293	0.49	13 662	0.63	10 631	0.37	23 153	0.47	12 963	0.60	10 190	0.35
Northern America	46 599	0.86	32 505	1.30	14 094	0.45	34 818	0.58	23 021	0.85	11 797	0.33
Eastern Asia	491 687	2.03	357 292	3.04	134 395	1.02	449 534	1.83	326 903	2.76	122 631	0.91
South-Eastern Asia	99 265	1.55	71 496	2.39	27 769	0.79	95 668	1.50	69 084	2.32	26 584	0.75
South-Central Asia	54 698	0.35	36 213	0.48	18 485	0.23	52 769	0.35	34 973	0.47	17 796	0.22
Western Asia	11 342	0.55	6 998	0.74	4 344	0.37	10 927	0.53	6 737	0.71	4 190	0.36
Central and Eastern Europe	24 782	0.52	15 184	0.82	9 598	0.30	23 002	0.48	13 985	0.75	9 017	0.28
Western Europe	26 128	0.67	18 506	1.08	7 622	0.29	23 657	0.54	16 420	0.84	7 237	0.26
Southern Europe	24 796	0.80	16 938	1.26	7 858	0.37	21 243	0.59	14 446	0.94	6 797	0.27
Northern Europe	11 924	0.58	7 451	0.81	4 473	0.36	10 513	0.45	6 493	0.63	4 020	0.28
Australia and New Zealand	3 344	0.73	2 429	1.15	915	0.32	2 503	0.48	1 713	0.72	790	0.25
Melanesia	933	1.26	572	1.50	361	1.03	911	1.29	560	1.57	351	1.03
Polynesia	58	0.96	41	1.53	17	0.40	55	0.91	38	1.44	17	0.39
Micronesia	84	1.76	68	2.89	16	0.65	70	1.47	54	2.29	16	0.65
Low HDI	33 097	0.69	20 742	0.91	12 355	0.50	31 602	0.68	19 756	0.90	11 846	0.49
Medium HDI	99 994	0.55	69 180	0.78	30 814	0.33	95 859	0.54	66 358	0.75	29 501	0.32
High HDI	548 935	1.62	391 686	2.40	157 249	0.86	524 307	1.55	373 738	2.30	150 569	0.83
Very high HDI	223 321	0.84	150 489	1.29	72 832	0.43	178 107	0.59	117 470	0.91	60 637	0.31
World	905 677	1.11	632 320	1.65	273 357	0.60	830 180	1.01	577 522	1.49	252 658	0.55

2.3 Distribution of Liver Cancer Cases and Deaths in Pakistan

The below Images summarize the number of new cases of liver cancer in Pakistan in 2020, a total of 5331 cases were reported (2967 males, 2364 females), and 5109 (2829 males, 2280 females) deaths were reported due to liver cancer.

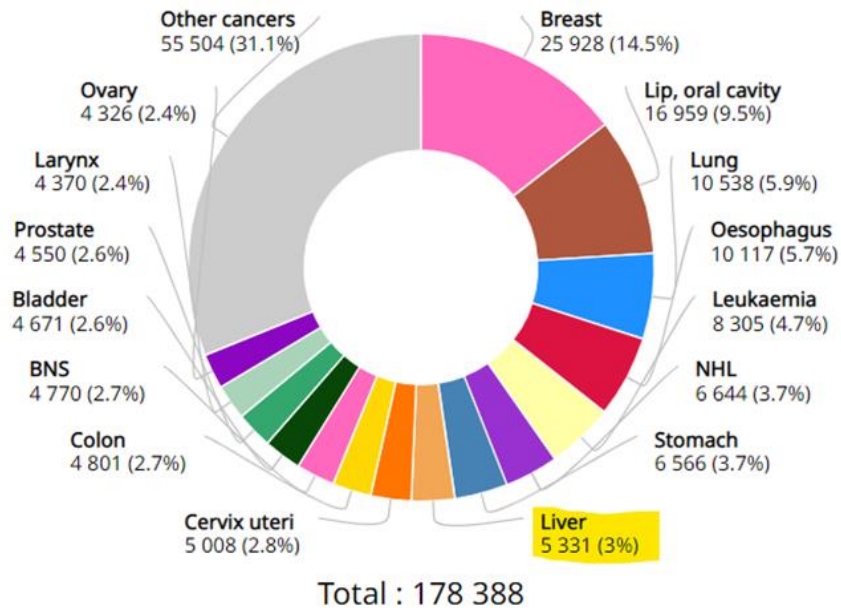


Figure 5: Number of new cases in 2020, Pakistan, both sexes, all ages. In Pakistan, liver cancer accounts for 3% of cancer cases, breast cancer being the most prevalent. (Observatory)

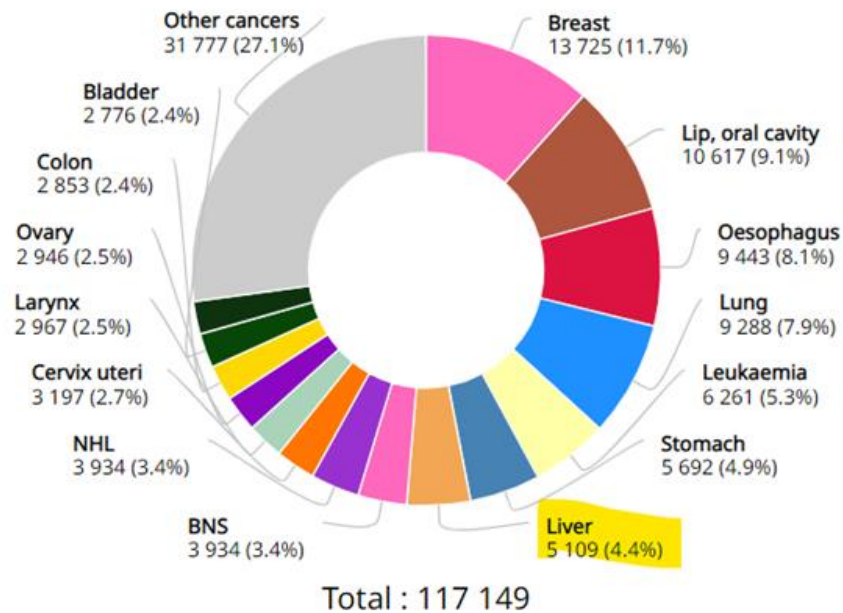
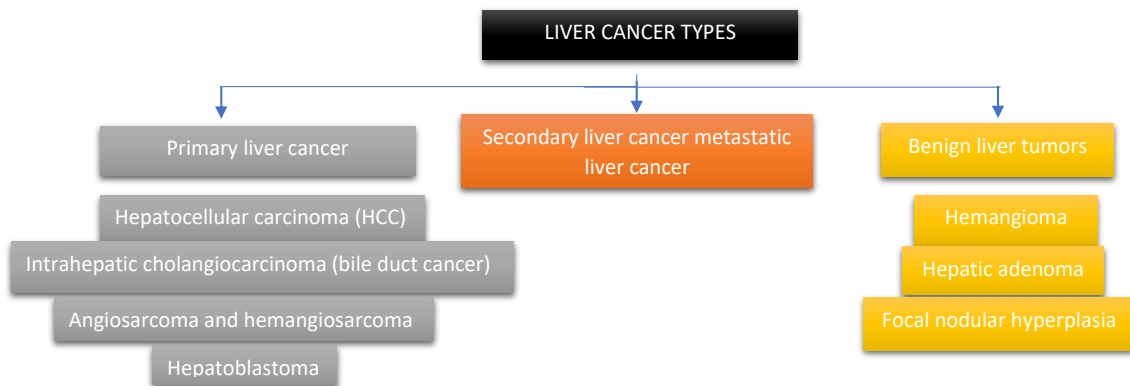


Figure 6: Number of deaths in 2020 Pakistan, both sexes, all ages. deaths due to liver cancer account for around 4.4% of the total deaths in Pakistan. (Observatory)

2.4 Types of Liver Cancer



Primary liver cancer is a term used to describe cancer that begins in the liver. There are many primary liver cancer types. Hepatocellular carcinoma (HCC) is the most widespread type of liver cancer in adults. It has different growth patterns. Some grow larger as a single tumor and later spread to other parts of the liver. The other type grows

as small cancer nodules in the liver (Balogh et al., 2016). Intrahepatic cholangiocarcinoma (bile duct cancer) this cancer starts in the tube, which carries bile to the gallbladder (Nakanuma et al., 2000). Angiosarcoma, hemangiosarcoma and hepatoblastoma are the rarest types of cancer. Angiosarcoma and hemangiosarcoma begin in the cell lining of liver blood vessels, and hepatoblastoma is found in children younger than four years old (Ananthakrishnan, Gogineni, & Saeian, 2006).

Secondary liver cancer is found in the liver but does not begin in the liver. It begins in other parts of the body and has metastasized from other parts such as the breast, lungs, stomach etc. Benign liver tumors grow large enough to cause problems, and they do not grow in adjacent tissues or spread to faraway body parts (Ananthakrishnan et al., 2006).

2.5 Diagnosis and Treatment of Liver Cancer

For the diagnosis of liver cancer, there are some tests and techniques. A Blood Test is done to identify liver function defects. The doctor may recommend imaging tests that are ultrasound, CT, or MRI. A biopsy is removing a piece of tissue from the liver to assess it in the laboratory. The treatment of Liver cancer depends on the age, the patient's stage, and personal preference. Treatment can be done through surgery. Different operations can be done, such as surgery to remove a tumor; for example, a doctor may remove a part of the liver that contains cancer and a portion of healthy liver tissue adjacent to the tumor. A liver transplant can also be done in which the diseased liver is replaced by a healthy liver; it is done for those patients with early-stage cancer (Abou-Alfa et al., 2020).

A liver cancer treatment delivered locally targets the cancer cells or their immediate surroundings. Localized liver cancer therapy options include Heating Cancer Cells (using electric current to heat cancer cells and kill them), Freezing Cancer Cells (using cold temperature to kill cancer cells), Injecting tumors with alcohol and chemotherapy drugs or placing beads with radiation in the liver. Radiation Therapy is performed with the help of X-rays and protons, which shrink tumors and kill them, when there are no other treatment options, and it is also done to reduce the symptoms (Abou-Alfa et al., 2020).

Targeted Drug Therapy mainly focuses on blocking specific defects present within cancer cells, which causes cancer cells to die. Immunotherapy uses the body's immune system to

destroy cancer cells. Cancer cell produces some proteins that protect cancer cells from the body's immune system, so immune therapy interferes with this method and causes the body's immune system to detect cancer cells as foreign and destroy them (Abou-Alfa et al., 2020).

2.6 Cancer Biomarkers

A biomarker is defined as a biological molecule present in the body (blood, body fluid, cells) which is an indication of a normal or aberrant process or a disease. Cancer biomarkers help us in risk evaluation, i.e., to differentiate between patients at a greater risk. It can help to detect cancer at an early stage and help in determining the stage of cancer. Most importantly, cancer biomarkers can be used to assess the efficacy of a certain therapy and to estimate the probability of recurrence in individuals who have recovered (Henry & Hayes, 2012).

2.7 Protein Biomarkers

Cancer diagnosis through protein biomarkers helps in early detection and observing the disease. Some protein biomarkers used to detect cancer are AFP (alpha-fetoprotein), CA-125, CEA (carcinoembryonic antigen), CA15-3, CA19-9 etc. These biomarkers are present in blood plasma and can be analyzed through liquid biopsy. There were many limitations to detecting cancer from protein biomarkers as they lacked specificity, and it was difficult to locate the origin of the disease. (Landegren & Hammond, 2021).

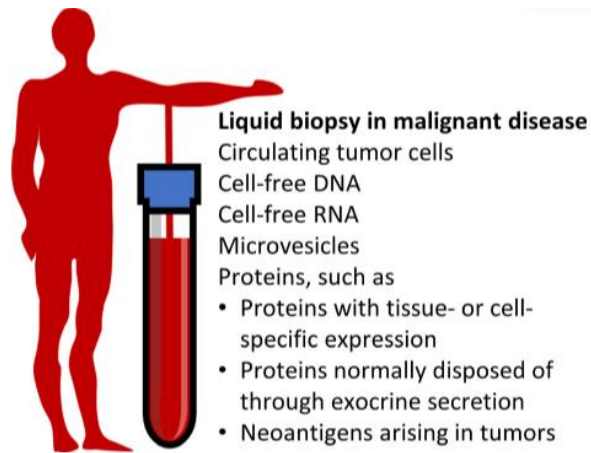


Figure 7 : By taking a sample of blood or another bodily fluid, liquid biopsy offers the chance to identify and evaluate cancers wherever they may exist in the body. The image includes key types of proteins and identifies some typical liquid biopsy targets(Landegren & Hammond, 2021)

However, these limitations were overcome with recent advancements in proteomics technology, such as high-density antibody microarrays, enzyme-linked immunosorbent assays (ELISA), and mass spectrometry (MS). Producing biomarker signatures, which are formed by combining many biomarkers, are crucial because they include more information than a single biomarker and will aid in more precise identification and cancer therapy (Borrebaeck, 2017).

2.8 Gas chromatography-mass spectrometer

The technique can be used to separate and analyze any state of the sample by first vaporizing them to the gas phase and then separating on the stationary phase on a capillary column; the sample is propelled using gases like helium and hydrogen, and then each compound comes out of different columns at different time based on their boiling points and polarity. The time is said to be the retention time of the sample. After leaving the column, the sample compounds are ionized and fragmented using the mass spectrometry technique. After being ionized, it is accelerated in the ion trap, where they are separated based on mass-to-charge ratios. The next step is data acquisition.

A complex sample produces different gas columns that give different peaks through the mass spectrometer, which helps us to identify and quantify a wide range of analytes.

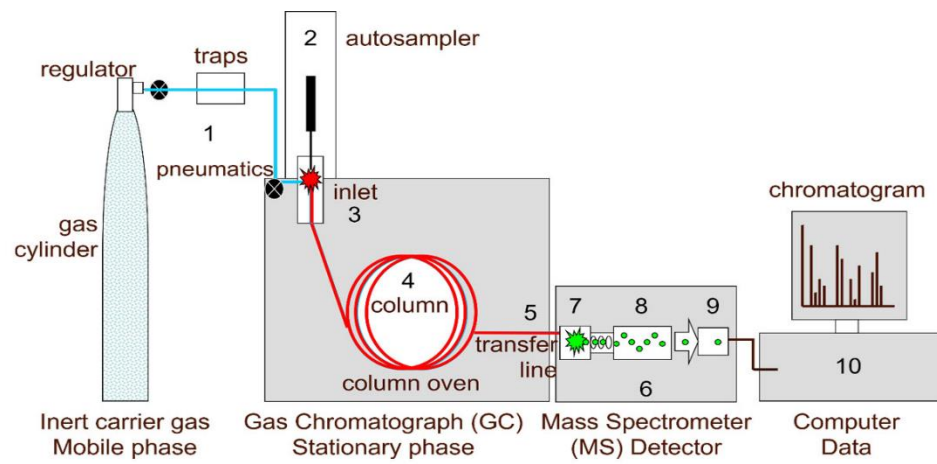


Figure 8: GCMS setup explained; the figure gives a step-by-step illustration of the processes inside the GCMS setup for the protein analysis.

CHAPTER 3

METHODOLOGY

3.1 In silico Part:

3.1.1 ADMET and Pharmacophore Analysis

Zinc 15 is an online database containing commercially available compounds that can be virtually seen through this database. Using Zinc15, we extracted the Dihydroquercetin smiles that are required in SwissADME.

Absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties of industrial chemicals, insecticides, and drug candidates are heavily reliant during Drug development and environmental danger evaluation. To perform the ADMET analysis of Dihydroquercetin, we used SwissADME and admetSAR (Cheng et al., 2012), and also used an online service that predicts ADMET properties which has been developed by the Laboratory of Medicinal Chemistry in the Department of Chemistry at Lomonosov Moscow State University (Radchenko, Karpov, & Sosnin, 2016).

Through SwissADME and admetSAR, we got the physicochemical properties, lipophilicity, water solubility, pharmacokinetics, drug-likeness, and medicinal chemistry. Gastrointestinal absorption is high for dihydroquercetin, which means the drug will adequately absorb from the GI tract, reaching its target more quickly and effectively, thereby improving treatment effects. It doesn't violate any of the five rules of Lipinski. Since it can more promptly pass biological membranes, it is more likely to have excellent oral bioavailability.

Pharmacophore modelling is used in drug development to identify the essential chemical properties, or pharmacophores, needed for a molecule to bind to a particular target protein or receptor. These pharmacophores contain crucial structural or functional elements required for the molecule's biological activity. For the Pharmacophore analysis,

we used ZincPharmer (Koes & Camacho, 2012). Using ZincPharmer's "search Molport" option, we uploaded the Dihydroquercetin structure. With this, we learned how many hydrogens acceptor/donor, Hydrophobic and aromatic groups are present.

3.1.2 Drug Energy Minimization

After successful results of the ADMET and Pharmacophore analysis of the Drug, retrieved the Drug/Inhibitor structure in SDF (Spatial Data File) format from PubChem (PubChem CID: 439533). After using Chem 3D, we minimized the energy. PerkinElmer created the software programme Chem3D (Cousins, 2005), which is used for molecular modelling, visualization, and analysis. It is a component of the ChemOffice Suite, which also consists of other computer programs for managing data and sketching and analyzing chemicals. Using this tool, we minimized the drug energy, which is the process of finding a molecule's lowest energy state by adjusting the atoms' positions. Later, it was used for docking.

3.1.3 Retrieving Protein Structures and their Structural Accuracy Prediction

We obtained the six protein structures in PDB (Program Database format) using Protein Data Bank (Rose et al., 2016). For structural accuracy analysis Alpha fold (Jumper et al., 2021) was used, through this we could comprehend a protein's function, connections with other molecules, and part in numerous biological processes, it is essential to be able to anticipate a protein's 3D structure with accuracy. After the accurate analysis using Pymol, water and hetro atoms from proteins were removed.

Table 2 The six proteins and their respective Alpha Fold ID

Proteins	Alpha Fold ID
Epidermal growth factor receptor	AF-P00533-F1
AKT2 protein	AF- Q6P4H3-F1
Phosphoinositide-3 kinase regulatory subunit 4	AF- D6RBB7-F1
Serine/threonine-protein kinase mTOR	AF- P42345-F1
Proto-oncogene cJun	AF-A0A510GAI3-F1
Eukaryotic translation initiation factor 2A	AF- F8WF18-F1

3.1.4. Drug and Protein Docking

Ligand and Drug docking was performed through CB-Dock (Y. Liu et al., 2020) .It is a computational technique for predicting small compounds' binding mechanism and affinity to protein targets and molecular docking's-Dock employs a hybrid docking technique that integrates shape complementarity, electrostatics, and knowledge-based scoring systems to forecast a ligand's binding posture and affinity to a protein target. Dihydroquercetin and different proteins involved in liver cancer were docked, multiple vina scores were obtained, and the data were filtered by selecting the most negative vina score because the more negative the vina score is, the better the binding.

3.1.5 Interaction Analysis

LigPlot+ (Laskowski & Swindells, 2011) is a Java-based tool with many applications in biological research. It is used to visualize and examine protein-ligand interactions. Researchers in structural biology and drug development frequently utilize this upgraded and improved version of the original LigPlot programme. After the docking, using ligplot+ we performed the interaction analysis. LigPlot+ generated a schematic diagram on the protein and drug interactions. We analyzed a wide range of protein-ligand

interactions through the schematic diagram, including hydrogen bonds, hydrophobic interactions, and electrostatic interactions.

3.2 In-Vitro

3.2.1 Cell Lines pellet collection

Cell line revival and maintenance

To revive and maintain a cell line, We Obtained a stock culture of the cell line and thaw the frozen cells. Plate the thawed cells in an appropriate growth medium and incubate at the recommended temperature and atmospheric conditions. Subculture the cells when they reach 80-90% confluence. This involves removing the medium, washing the cells with PBS, and adding fresh growth medium. Passage the cells every 2-3 days or as necessary to maintain cell health and prevent senescence. a viable cell line was maintained by regularly checking for contamination (e.g., bacteria, fungi, mycoplasma), performing a cell viability assay, and storing stocks of the cells. We maintained low passage numbers and avoided prolonged exposure to high levels of stress or toxic agents to prevent genetic instability. the cell line was cryopreserved by gradually freezing the cells in a growth medium containing a cryoprotectant.

Invitro drug treatment

After the revival and maintenance of the Hep-G2 cell line, we performed in vitro drug delivery. We Prepare the dihydroquercetin drug solution in an appropriate vehicle, such as sterile saline or a cell culture medium. Aspirate the culture medium from the cells and wash the cells with a sterile solution, such as phosphate-buffered saline (PBS). Added the drug solution to the cells, ensuring the cells are evenly coated. the cells were incubated with the drug solution for the specified time at the appropriate temperature. After that we observed the cells for any changes in phenotype or behavior, such as changes in cell viability, cell proliferation, or apoptosis.

Centrifugation and pellet collection

After drug treatment and incubation, we Washed the cells with a sterile solution, such as phosphate-buffered saline (PBS), to remove any residual drug. Centrifuged the cell suspension at a low speed (e.g., 400-800 g) for 5-10 minutes to form a cell pellet. Carefully aspirate the supernatant, being careful not to disturb the cell pellet. Resuspend the cell pellet in a solution as desired, such as a cell lysis buffer or another buffer for downstream experiments.

3.3 Samples Preparation

3.3.1 Chemicals required.

The chemicals required for sample preparation were ordered from Thermofisher Scientific. This included Extraction solvent, which is methanol: water in a 1:1 v/v ratio, Methoxamine HCL dissolved in pyridine, 1% trimethylchlorosilane (TMCS) and MSTFA (N-methyl-N-(trimethylsilyl) trifluoroacetamide)

3.4 Steps involved in the preparation of serum samples.

3.4.1 Samples lysis

The separation of polar and non-polar metabolites from other biological components requires sample lysis. To avoid the influence of varied cell numbers, each sample included the same number of cells. A total of 300 microliters of an extraction solvent (methanol: water, 1:1 v/v) and 100 microliters of samples were combined together. The polar metabolites are separated by the extraction solvent. A non-polar solvent is required to separate the non-polar metabolites. To achieve quantitative extraction of metabolites, the samples were vortexed for 2 minutes. For one hour, the samples were maintained on ice. During this time, the samples were vortexed twice every 15 minutes for 2 minutes.

3.4.2 Centrifugation

To separate the insoluble cell matrices, the materials were centrifuged at 13000 rpm for 10 minutes at -4°C. The supernatants were collected in separate GC vials after centrifugation, and the pellets were discarded.

3.4.3 Concentrating the metabolites.

After collecting the supernatants in separate GC vials, the samples were dried for 1 hour in vacuum concentrators at 37°C. Heat, centrifugal force, and vacuum are used in these concentrators to remove moisture from samples. This is widely used in sample preparation. When a vacuum is introduced, the centrifugal force created in 30 prevents the liquids from escaping. The heat is applied indirectly through the walls of the vacuum chamber.

3.4.4 Derivatization

Derivatization is the process by which non-volatile or polar metabolites are converted into volatile or non-polar metabolites. It also increases the thermal stability of the products, the detector response by introducing functional groups, and the GC separation performance. To derivatize the dry materials, 25 microliters of 20mg/ml methoxamine hydrochloride dissolved in pyrimidine were utilized. After adding this derivatization reagent, the samples were vortexed for 2 minutes and then kept at 25° C for around 6 hours.

3.4.5 Silylation

Prior to GC, another process of Derivatization is Silylation. This is also done to decrease the polarity of the metabolites, which improves their stability and behavior in GC. For Silylation, 25 microliters of MSTFA + 1% TMCS were dissolved in pyridine and utilized as a sialylation mixture. After adding this sialylation reagent, the samples were vortex mixed for 2 minutes.

3.4.6 Incubation

After that, the samples were incubated for 30 minutes at a temperature of 50°C. They were then transferred into 200 microliter micro inserts. The samples were then prepared for GC analysis.

3.5 GC/MS

In the GCMS process, the sample first passes through a Gas Chromatography unit. This is where the separation of compounds in a mixture takes place. The molecules are then moved through the Mass Spectrometry unit, where the process of ionization of molecules through electron ionization takes place. The detector identifies the molecular fragments based on their m/z ratio. In mobile phase gas chromatography, there exists a gas supply (N₂, H₂ or He are used as carrier gases). It also contains a flow control to control a carrier gas, a heat-controlled oven, a heat-controlled injector, a detection system, and a data recording device. The stationary phase may have either a polymeric liquid phase Gas-liquid chromatography (GLC) or a solid phase Gas-solid chromatography (GSC). The stationary phase on the inner wall of the capillary column. In GC, the inert carrier gas carries the analytes of interest present in the mixture of compounds. Dependent on their vapor pressures, these analytes are vaporized or pumped through the stationary phase. Because of their physical and chemical natures, they interact with the stationary phase differently and hence elute at different retention times from the column. The distribution of the compound between the stationary phase and the mobile phase directs the interaction between an analyte and the stationary phase. This is indicated by the K_c (a distribution coefficient). $K_c = C_s/C_m$ C_s indicates the concentration of analyte present in the stationary phase, while C_m indicates the concentration of analyte present in the mobile phase. A large K_c relates to a longer retention time of a compound in a stationary phase or inside of the column. The temperature of the column and the chemical properties of the stationary phase control the distribution coefficient.

3.6 GC/MS Analysis

To analyze metabolomics, a GC-MS-QP 2010 ultra-system equipped with lab solutions GCMS software was used, and an SH- Rxi-5Sil MS column was employed to separate the metabolites. The GC column utilized 99.9% helium at a flow rate of 1.0 ml/min as a carrier gas. Initially, the oven temperature was set at 60°C for 2 minutes and then raised to 310°C by 50°C per minute during the analysis. Both the interface and ionization temperatures were kept at 250°C. The analysis was conducted in full scan mode, within the range of 50 to 650 amu. A 10-microliter sample was injected in split less mode using an AOC-20i injector. In order to have a good resolution of the metabolites the ramp rate was set at 5 °C/min. The eV scan range was set at 50- 650 m/z GC total ion chromatograms and fragmentation patterns of the compound were analyzed using the NIST/NIH/EPA Mass spectral library. Each sample took 43.67 minutes to run.

3.7 GC-MS Data Analysis

The results of GC-MS were received in the form of a spreadsheet. Containing data on different metabolites found in treated and control samples and their area was also mentioned. The KEGG IDs were retrieved for the metabolites through the Kyoto Encyclopedia of Genes and Genomes (KEGG). Then the common metabolites in control and treated were selected, which were 30 metabolites. Of these 30 metabolites, the average area was calculated for both the control and treated sample. Fold change was calculated using this formula: average concentration of proteins in disease samples ÷ average concentration of proteins in the samples. The p-value was calculated for these 30 metabolites; metabolites with a p-value less than 0.05 were selected for further filtration. This yielded a total of 18 metabolites.

3.8 Protein Pathway Analysis and Gene ID Extraction

The open-source software project Cytoscape (Saito et al., 2012) aims to combine high-throughput expression data with biomolecular interaction networks and other pertinent molecular states. As a result, a coherent conceptual framework that makes it easier to analyze and understand complicated biological systems is produced. Following filtering,

Cytoscape was used to examine the proteins for visualization and pathway analysis. To discover relationships between them and investigate their function in the regulation of disease, the pathways of 18 proteins were combined.

3.9 Gene Ontology Analysis

PANTHER (protein analysis through evolutionary relationships) (Mi, Muruganujan, Ebert, Huang, & Thomas, 2019) is a bioinformatics tool that uses a database system to classify gene and protein families. Its primary function is to aid researchers in studying proteins' evolutionary relationships and functional characteristics. Using Panther, we uploaded the gene ids we got from cytoscape and we the got the localization and exact location of the 18 metabolites in the mammalian cells, which is helpful in pathway generation.

3.10 Enrichment Analysis

An online bioinformatics application called MetaboAnalyst (Pang et al., 2021) offers a full array of analysis and visualization tools for high-throughput metabolomics data. Using it, we did an enrichment analysis which investigated the substantial enrichment of a collection of metabolites that are functionally linked. The enrichment Analysis result contained Bar Chart, Network View, Dot Plot and Summary Table.

CHAPTER 4

RESULTS

4.1 The Absorption, Distribution, Metabolism, Elimination, and Toxicity (ADMET) analysis of Dihydroquercetin

For ADMET properties, we retrieved the smiles ID using Zinc 15. Smiles ID of Dihydroquercetin: O=C1c2c(O)cc(O)cc2O[C@@H](c2ccc(O)c(O)c2)[C@H]1O. Tools used to predict the ADMET properties, two tools were used 1) SwissAdme and 2) admetSAR. These two tools gave us physicochemical properties, lipophilicity, and water solubility. The results from SwissAdme and admetSAR are shown in the tables below.

4.1.1 Physicochemical Properties

Below table 1,2 and 3 summarizes the physiochemical properties of Dihydroquercetin. It shows the drug molecular weight, the number of heavy atoms, aromatic heavy atoms and different groups that are present.

Table 3: Physicochemical Properties of Dihydroquercetin

Physicochemical Properties	
Formula	C15H12O7
Molecular weight	304.25 g/mol
Num. heavy atoms	22
Num. Aromatic heavy atoms	12
Num. rotatable bonds	1
Num. H-bond acceptors	7
Num. H-bond donors	5
Regression analysis	Value
Water solubility	-2.999
Plasma protein binding	1.057
Acute Oral Toxicity	2.146

4.1.2 Pharmacokinetics

Gastrointestinal absorption is high for dihydroquercetin, which means the drug will adequately absorb from the GI tract, reaching its target more quickly and effectively, thereby improving treatment effects. It doesn't violate any of the five rules of Lipinski. Since it can more promptly pass biological membranes, it is more likely to have excellent oral bioavailability.

Table 4: Pharmacokinetic Analysis of Dihydroquercetin

Pharmacokinetics	
Subcellular localization	Mitochondria
GI absorption	High
BBB permeant	No
1	No
CYP1A2 inhibitor	No
CYP2C19 inhibitor	No
CYP2C9 inhibitor	No
CYP2D6 inhibitor	No
CYP3A4 inhibitor	No
Log Kp (skin permeation)	-7.48 cm/s
OATP2B1 inhibitor	No
OATP1B1 inhibitor	Yes
OATP1B3 inhibitor	Yes
MATE1 inhibitor	No
OCT2 inhibitor	No
BSEP inhibitor	No
P-glycoprotein inhibitor	No
P-glycoprotein substrate	No
CYP inhibitory promiscuity	Yes
Drug likeness	
Lipinski	Yes; 0 violation
Ghose	Yes
Veber	Yes
Egan	Yes
Muegge	Yes
Bioavailability Score	0.55

These Graphs below were made using an online service for ADMET properties, it contains graphical information on the blood-brain barrier, human intestinal absorption, hERG activity and affinity. The Log of the Blood-Brain Barrier is -0.23 which indicates

that the barrier function is high and Dihydroquercetin has low permeability across the blood-brain barrier. Human intestinal absorption is 100% which makes the drug ideal as it will enter the bloodstream without any loss or degradation. hERG channels play an important role in heart rhythm regulation and the results related to hERG are concerning because a pKi of 4.12 suggests that the compound has a moderate to high affinity for hERG channels, a pIC50 of 3.97 indicates that the compound has an inhibitory potency towards the channel. Further research should be done for example electrophysiology assays, are important to investigate drugs effect on hERG channels and to assess its safety profile.

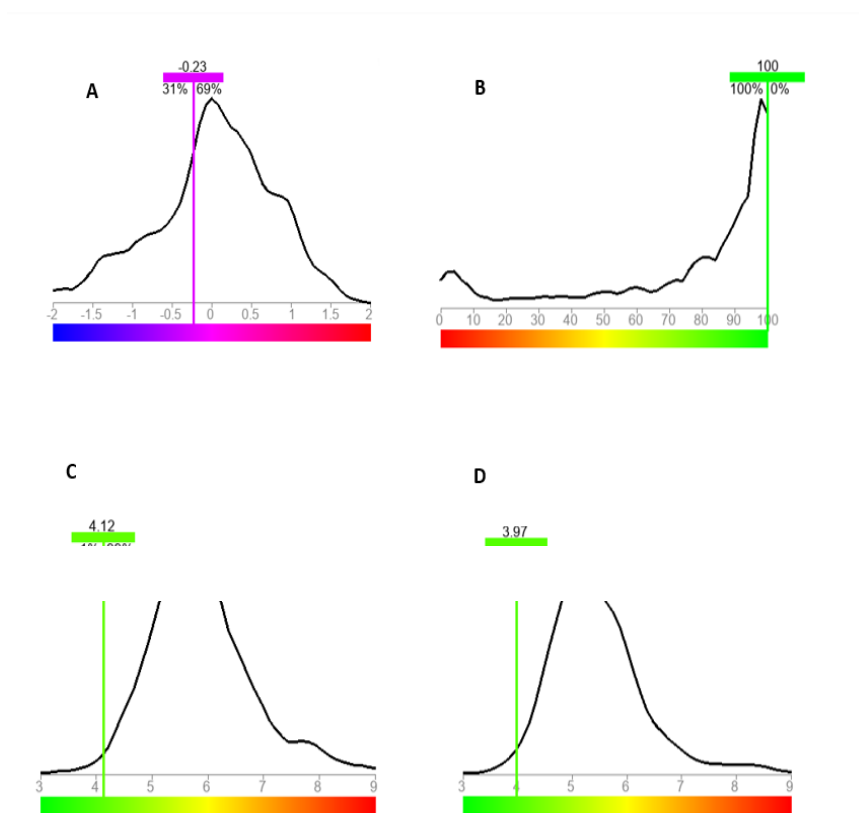


Figure 9 : Dihydroquercetin Pharmacophore Analysis. These graph shows (A) Blood-brain barrier permeability (LogBB), (B) Human intestinal absorption (HIA%), (C) hERGF affinity (pKi) and (D) hERG activity (pIC50).

4.1.3 Toxicity

The toxicity result is shown below. Dihydroquercetin shows no eye corrosion, there is no hepatotoxicity or any skin sensation. However, it causes respiratory reproductive and mitochondrial toxicity. However, it can bind the estrogen receptor, androgen receptor, thyroid receptor, and glucocorticoid receptor.

Table 5: Toxicity Analysis of Dihydroquercetin

Medicinal Chemistry	
PAINS	1 alert: catechol_A
Brenk	1 alert: catechol
Lead likeness	Yes
Synthetic accessibility	3.51
Carcinogenicity (binary)	-
Carcinogenicity (trinary)	Non-required
Eye corrosion	-
Eye irritation	+
Ames mutagenesis	+
Human Ether-a-go-go-Related Gene inhibition	-
Micronuclear	+
Hepatotoxicity	-
skin sensitization	-
Respiratory toxicity	+
Reproductive toxicity	+
Mitochondrial toxicity	+
Nephrotoxicity	+
Acute Oral Toxicity (c)	II
Estrogen receptor binding	+
Androgen receptor binding	+
Thyroid receptor binding	+
Glucocorticoid receptor binding	+
Aromatase binding	+
PPAR gamma	+

4.2 Pharmacophore Analysis

The pharmacophore model by Zincpharmer is a three-dimensional depiction of the chemical and structural characteristics required for a molecule to interact with a biological target and have a pharmacological effect. This model of Dihydroquercetin can be used to direct the design of novel molecules with improved features and we get to know the pharmacophore classes that are attached with their dimension and radius.

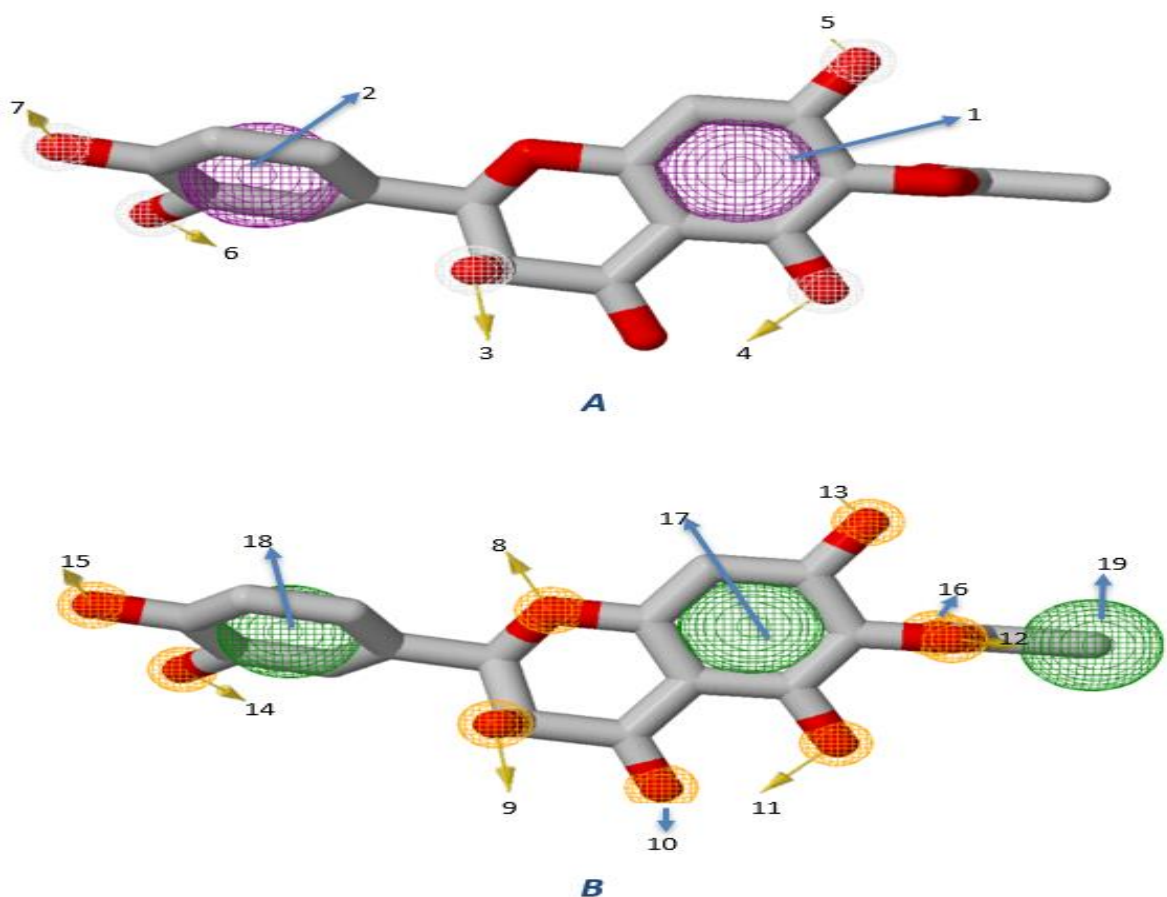


Figure 10: Pharmacophore Model. Figure 2 (A) and (B) is a Depiction of pharmacophore models produced through ZINCpharmer. Green Sphere show Hydrophobic Interaction, Purple Sphere indicates aromatic group, White Sphere indicate Hydrogen Donor, Yellow Sphere indicates Hydrogen Acceptor.

Table 6: Pharmacophore Class with positions, dimensions, and radius

Position	Pharmacophore Class	x	y	z	Radius
1	Aromatic	2.1	0.48	0.37	1.1
2	Aromatic	-4.24	0.38	-0.08	1.1
3	Hydrogen Donor	-1.37	-1.57	1.83	0.5
4	Hydrogen Donor	3.28	-2.01	0.23	0.5
5	Hydrogen Donor	3.7	2.73	0.5	0.5
6	Hydrogen Donor	-5.91	-0.45	-2.11	0.5
7	Hydrogen Donor	-6.93	0.96	0.06	0.5
8	Hydrogen Acceptor	-0.65	0.84	0.4	0.5
9	Hydrogen Acceptor	-1.37	-1.57	1.83	0.5
10	Hydrogen Acceptor	0.87	-2.97	0.28	0.5
11	Hydrogen Acceptor	3.28	-2.01	0.23	0.5
12	Hydrogen Acceptor	4.84	0.22	0.33	0.5
13	Hydrogen Acceptor	3.7	2.73	0.5	0.5
14	Hydrogen Acceptor	-5.91	-0.45	-2.11	0.5
15	Hydrogen Acceptor	-6.93	0.96	0.06	0.5
16	Hydrogen Acceptor	4.78	0.36	-2	0.5
17	Hydrophobic	2.1	0.48	0.37	1
18	Hydrophobic	-4.24	0.38	-0.08	1
19	Hydrophobic	6.89	0.09	-0.87	1

4.3 Energy Minimization

Energy minimization is an essential computational method in chemistry for optimizing a chemical molecule's molecular structure. Using Chem 3D reduced the potential energy of the system. Through this, we got the most stable and energetically advantageous conformation or arrangement of atoms in dihydroquercetin.

Total Energy: -3.0493 kcal/mol

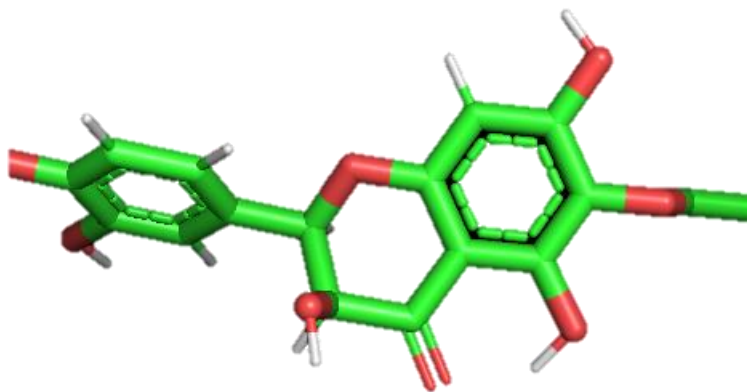


Figure 11 : Energy Minimization. Dihydroquercetin after energy minimization by Chem 3D

4.4 Docking

After the pharmacophore analysis, using PyMol, the attached water molecules and ligands were removed with the proteins. Water or any other heteroatom is removed as it would take part in the docking, water molecules are present in the active site of a protein and can form hydrogen bonds with both the protein and ligand. So, with the removal docking is done efficiently.

Dihydroquercetin and different proteins (Epidermal growth factor receptor, AKT2 protein, Phosphoinositide 3-kinase regulatory subunit 4, Serine/threonine-protein kinase mTOR, Proto-oncogene cJun, Eukaryotic translation initiation factor 2A) involved in liver cancer were docked using CB-Dock 2, multiple vina scores were obtained, the data were filtered by selecting the most negative vina score because the more negative the vina score is, the better the binding. The results obtained are shown below.

Table 7: Protein and their Vina Score

Protein	Vina Score
Epidermal growth factor receptor	-8.1
AKT2 protein	-7
Phosphoinositide 3-kinase regulatory subunit 4	-6.8
Serine/threonine-protein kinase mTOR	-8.3
Proto-oncogene cJun	-4.5
Eukaryotic translation initiation factor 2A	-5.7

4.5 Interaction Analysis

The Ligand and Protein interaction was analyzed through ligplot and 2D (ligplot+) and 3D results (pymol) diagrams of interaction were obtained. Through these diagrams we analyzed the specific interaction between Dihydroquercetin and the six proteins (AKT2 protein, Proto-oncogene cJun, Epidermal growth factor receptor, Eukaryotic translation initiation factor 2A, Serine/threonine-protein kinase mTOR and Phosphoinositide 3-kinase regulatory subunit 4). We could easily identify the amino acid residues that played an important role in the interaction with the ligand. In addition, the bond length and type of interaction was also assessed. The results are shown below in Fig 4, Fig 5, Fig 6, fig 7, Fig 8, Fig 9.

1. Protein: **AKT2 protein**

Ligand ID: **AF-Q6P4H3-F1**

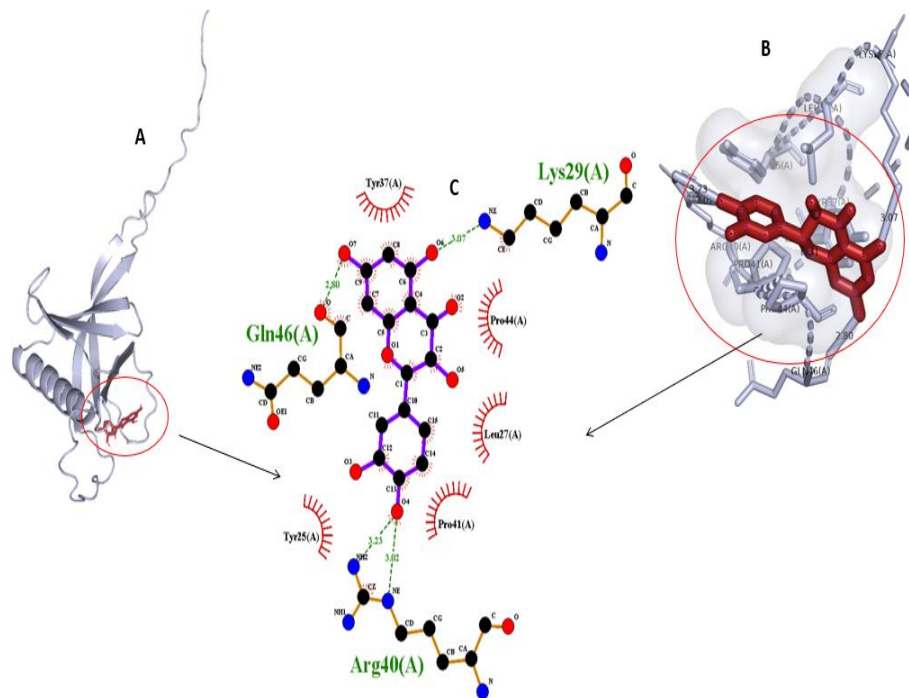


Figure 12 : A schematic representation of the interactions between Dihydroquercetin and the AKT2 protein. Fig 12 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues with the ligand and (C) 2Dmodel of ligand obtained from LigPlot+. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding.

Explanation of Fig 12 (C): The diagrams of Dihydroquercetin and AKT2 protein generated by Ligplot + and Pymol are shown above. Ligand (Dihydroquercetin) is colored purple. There is a total of 4 hydrogen bonds between Dihydroquercetin and AKT2. 1) Hydrogen Bond between Oxygen of Glutamine 46 (A) and Oxygen of Dihydroquercetin with a bond length of 2.80 Angstrom, 2) Hydrogen Bond between Nitrogen of Lysine 29 (A) and Oxygen of Dihydroquercetin with a bond length of 3.07 Angstrom, 3) Hydrogen Bond between Nitrogen of Arginine 40 (A) and Oxygen of Dihydroquercetin with a bond length of 3.23 Angstrom and 4) Hydrogen Bond between Nitrogen of Arginine 40 (A) and Oxygen of Dihydroquercetin with a bond length of 3.02 Angstrom. 5) Hydrophobic bonds are present between amino acid residues (Tyrosine 25 (A), Proline 41 (A), Leucine 27 (A) and Proline 44 (A) and ligand. Black color atoms are

carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding.

2. Protein: **Proto-oncogene cJun**

Ligand ID: **AF- A0A510GAI3-F1**

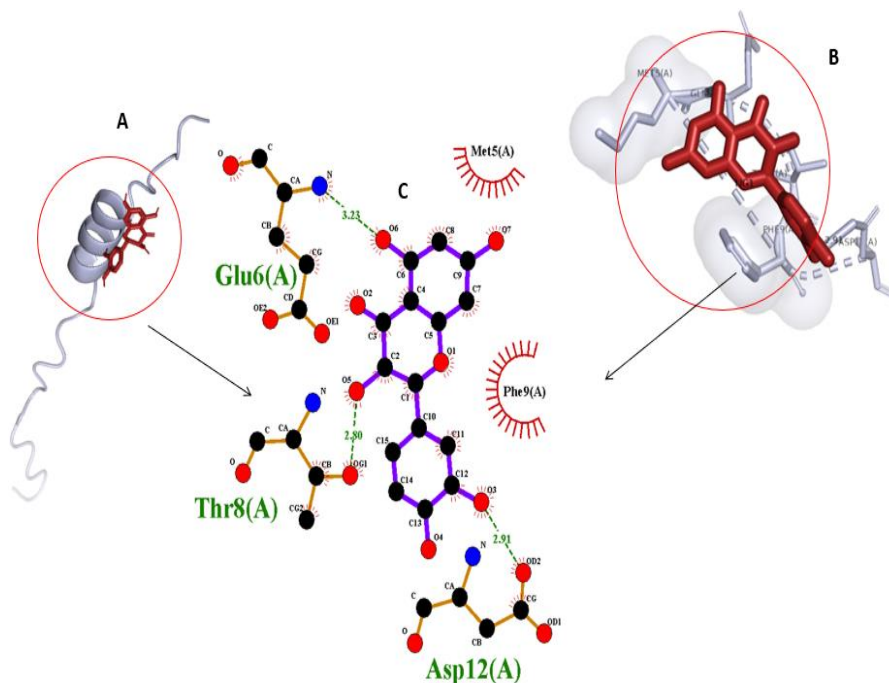


Figure 13: A schematic representation of the interactions between Dihydroquercetin and the Proto-oncogene cJun protein. Fig 13 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues with the ligand and (C) 2D model of ligand obtained from LigPlot+. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding.

Explanation of Fig 13 (C): The diagrams of Dihydroquercetin and Proto-oncogene cJun protein generated by Ligplot + and Pymol are shown above. Ligand (Dihydroquercetin) is colored purple. There is a total of 3 hydrogen bonds between Dihydroquercetin and cJUN. 1) Hydrogen Bond between Nitrogen of Glutamic Acid 6 (A) and Oxygen of Dihydroquercetin with a bond length of 3.23 Angstrom, 2) Hydrogen Bond between the

oxygen of Threonine 8 (A) and Oxygen of Dihydroquercetin with a bond length of 2.80 Angstrom, 3) Hydrogen Bond between Oxygen of Aspartate 12 (A) and Oxygen of Dihydroquercetin with a bond length of 2.91 Angstrom. 2 Hydrophobic bonds are present between amino acid residues (Methionine 5 (A), Phenylalanine 9 (A)) and ligand. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding.

3. Protein: Epidermal growth factor receptor

Protein ID: AF-P00533-F1

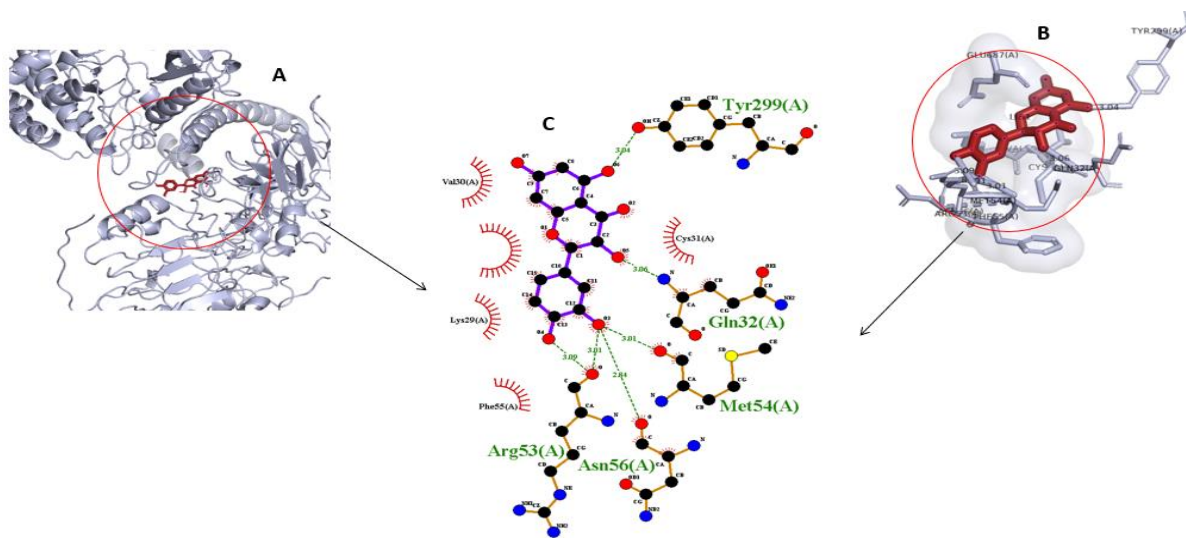


Figure 14: A schematic representation of the interactions between Dihydroquercetin and the Epidermal growth factor receptor. Fig 14 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues with the the ligand and (C) 2D model of ligand obtained from LigPlot+. Black colour atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding.

Explanation of Fig 14 (C): The diagrams of Dihydroquercetin and Epidermal growth factor receptor protein generated by Ligplot + and Pymol is shown above. Ligand (Dihydroquercetin) is colored purple. There is a total of 6 hydrogen bonds between Dihydroquercetin and EGFR. 1) Hydrogen Bond between Oxygen of Tyrosine 299 (A)

and Oxygen of Dihydroquercetin with a bond length of 3.04 Angstrom. 2,3) Oxygen of Arginine 53 (A) is making two hydrogen bonds with two Oxygens of Dihydroquercetin with bond lengths of 3.09 and 3.01 Angstrom, 4,5) Hydrogen Bond between Oxygen of Aspartate 12 (A) and Oxygen of Dihydroquercetin is making two hydrogen bonds with side chains of Asparagine 56 (A) and Methionine 54 (A) with a bond length of 2.84 and 3.01 Angstrom respectively. 6) Hydrogen Bond between Nitrogen of Glutamine 32 (A) and Oxygen of Dihydroquercetin with a bond length of 3.06 Angstrom. 5 Hydrophobic bonds are present between amino acid residues (Phenylalanine 55 (A), Lysine 29 (A), Glutamine 687 (A), Valine 30 (A), Cysteine 31 (A)) and ligand. Black color atoms are carbon atoms, yellow Sulphur, purple represents ligand binding, and brown represents non-ligand bonding.

4. Protein: Eukaryotic translation initiation factor 2A

Ligand ID: AF- F8WF18-F1

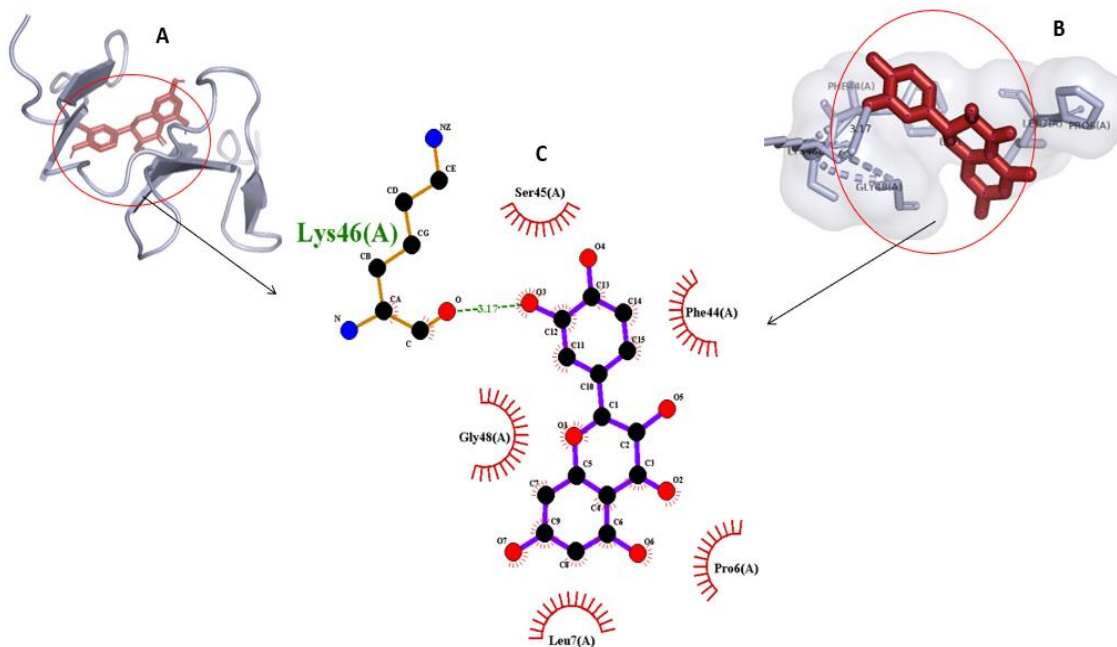


Figure 15: A schematic representation of the interactions between Dihydroquercetin and Eukaryotic translation initiation factor 2A. Fig 15 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues with the ligand and (C) 2D model of ligand obtained from LigPlot+. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding.

Explanation of Fig 15 (C): The diagrams of Dihydroquercetin and Epidermal growth factor receptor protein generated by Ligplot + and Pymol are shown above. Ligand (Dihydroquercetin) is colored purple. There is a total of 1 hydrogen bond between Dihydroquercetin and eIF2. 1) Hydrogen Bond between Oxygen of Lysine 46 (A) and Oxygen of Dihydroquercetin with a bond length of 3.17 Angstrom. 5 Hydrophobic bonds are present between amino acid residues (Leucine 7 (A), Proline 6 (A), Phenylalanine 44 (A), Serine 45 (A), Glycine 48 (A)) and ligand. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding.

5. Protein: Serine/threonine-protein kinase mTOR

Ligand ID: AF- P42345-F1

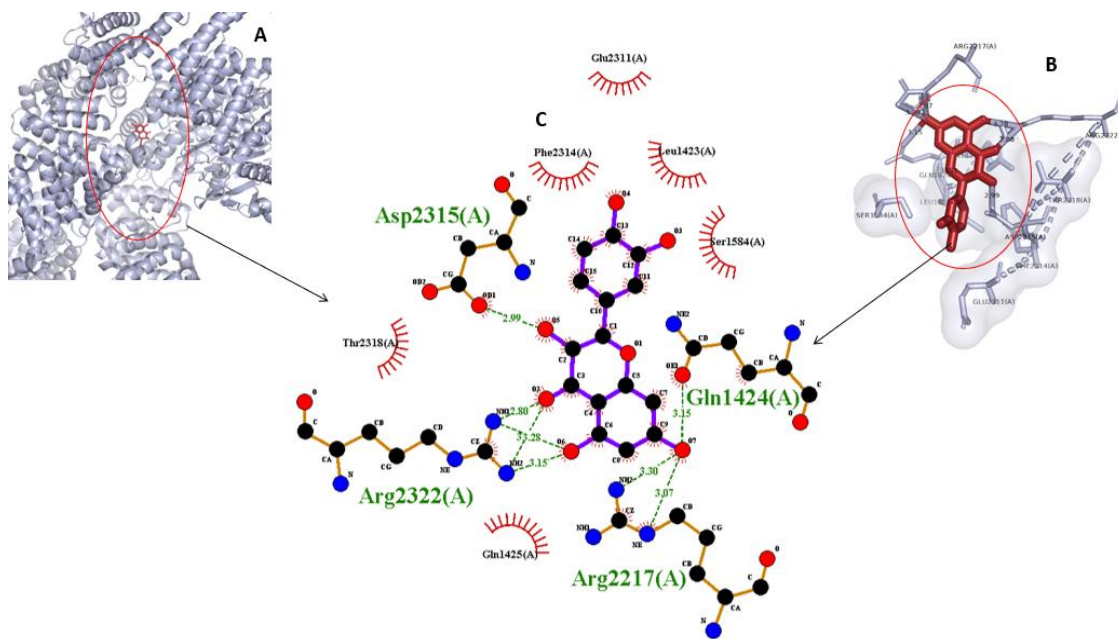


Figure 16: A schematic representation of the interactions between Dihydroquercetin and Serine/threonine-protein kinase mTOR. Fig 16 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues with the ligand and (C) 2D model of ligand obtained from LigPlot+.

Explanation of Fig 16 (C): The diagrams Dihydroquercetin and Serine/threonine-protein kinase protein generated by Ligplot + and Pymol is shown above. Ligand (Dihydroquercetin) is colored purple. There is a total of 8 hydrogen bonds between Dihydroquercetin and mTOR. 1) Hydrogen Bond between Oxygen of Aspartate 2315 (A) and Oxygen of Dihydroquercetin with a bond length of 2.99 Angstrom. 2) Hydrogen Bond between Oxygen of Glutamine 1424 (A) and Oxygen of Dihydroquercetin with a bond length of 3.15 Angstrom. 3,4) Oxygen of Dihydroquercetin made two hydrogen bonds with two nitrogens of the side chain of Arginine 2217 (A). 5,6,7,8) Two Oxygen of Dihydroquercetin made four hydrogen bonds with the nitrogen of arginine 2322 (A) side chain with bond lengths of 2.80, 3.328 and 3.15 Angstrom. 6 Hydrophobic bonds are

present between amino acid residues (Glutamine 1425 (A), Threonine 2318 (A), Phenylalanine 2314 (A), Glutamic Acid 2311 (A), Leucine 1423 (A), Serine 1584 (A)) and ligand. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding.

6. Protein: **Phosphoinositide 3-kinase regulatory subunit 4**

Ligand ID: **AF- D6RBB7-F1**

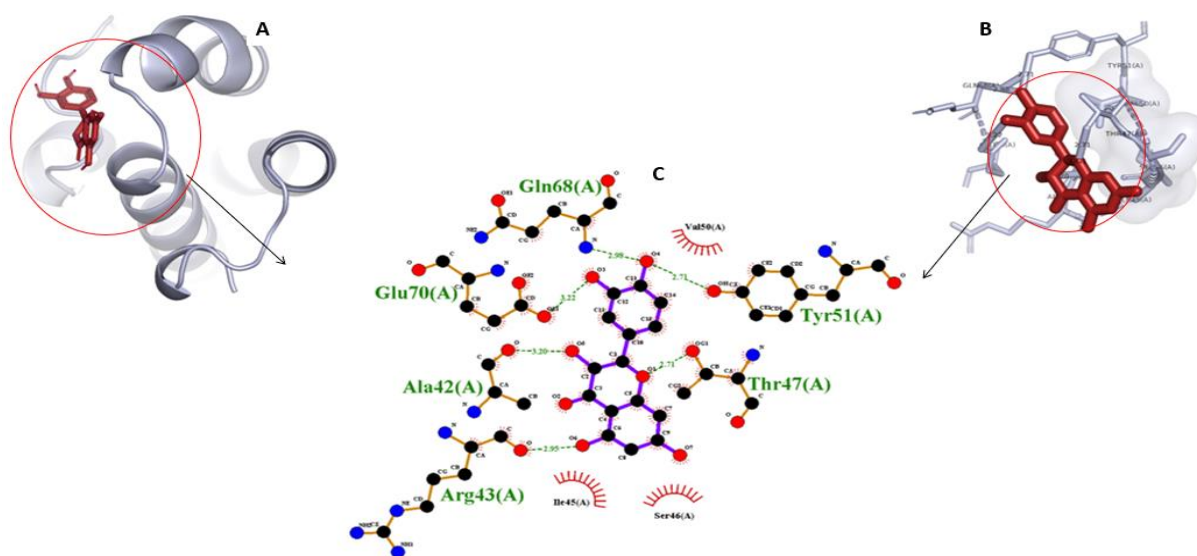


Figure 17: A schematic representation of the interactions between Dihydroquercetin and Phosphoinositide 3-kinase regulatory subunit 4. Fig 17 (A) shows the protein and ligand interaction obtained from CB Dock, (B) shows the interaction of the amino acid residues residues with the ligand and (C) 2D model of ligand obtained from LigPlot+. Black colour atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding.

Explanation of Fig 17 (C): The diagrams of Dihydroquercetin and Phosphoinositide 3-kinase regulatory subunit protein generated by Ligplot + and Pymol are shown above. Ligand (Dihydroquercetin) is colored purple. There is a total of 6 hydrogen bonds between Dihydroquercetin and PI3K. 1) Hydrogen Bond between Oxygen of Tyrosine 51 (A) and Oxygen of Dihydroquercetin with a bond length of 2.71 Angstrom. 2) Hydrogen Bond between Oxygen of Threonine 47 (A) and Oxygen of Dihydroquercetin with a bond length of 2.71 Angstrom. 3) Hydrogen Bond between Oxygen of Arginine 43 (A) and Oxygen of Dihydroquercetin with a bond length of 2.95 Angstrom (A). 4) Hydrogen Bond between Oxygen of Alanine 42 (A) and Oxygen of Dihydroquercetin with a bond length of 3.20 Angstrom. 5) Hydrogen Bond between Oxygen of Glutamic Acid 70 (A) and Oxygen of Dihydroquercetin with a bond length of 3.32 and 6) Hydrogen Bond between Oxygen of Glutamine 68 (A) and Oxygen of Dihydroquercetin with a bond length of 2.98 Angstrom. 3 Hydrophobic bonds are present between amino acid residues (Isoleucine 45 (A), Valine 50(A), Serine 46 (A)) and ligand. Black color atoms are carbon atoms, purple represents ligand binding, and brown represents non-ligand bonding.

4.6 Gene Ontology and Enrichment Analysis

Through gene ontology analysis we got the localization and exact location of the 18 metabolites in the mammalian cells, which was later used in pathway generation and the results from the enrichment analysis showed that the 18 functionally linked metabolites are significantly enriched. Results from both gene ontology and enrichment analysis are shown below.

Table 8: Showing the Molecular Function of metabolites.

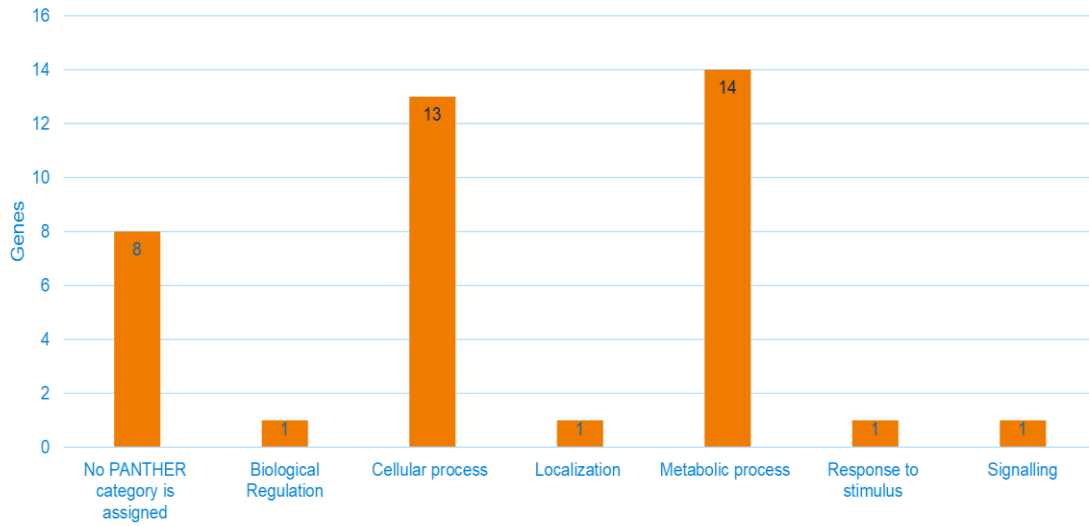


Table 9: Showing the Biological Process of metabolites.

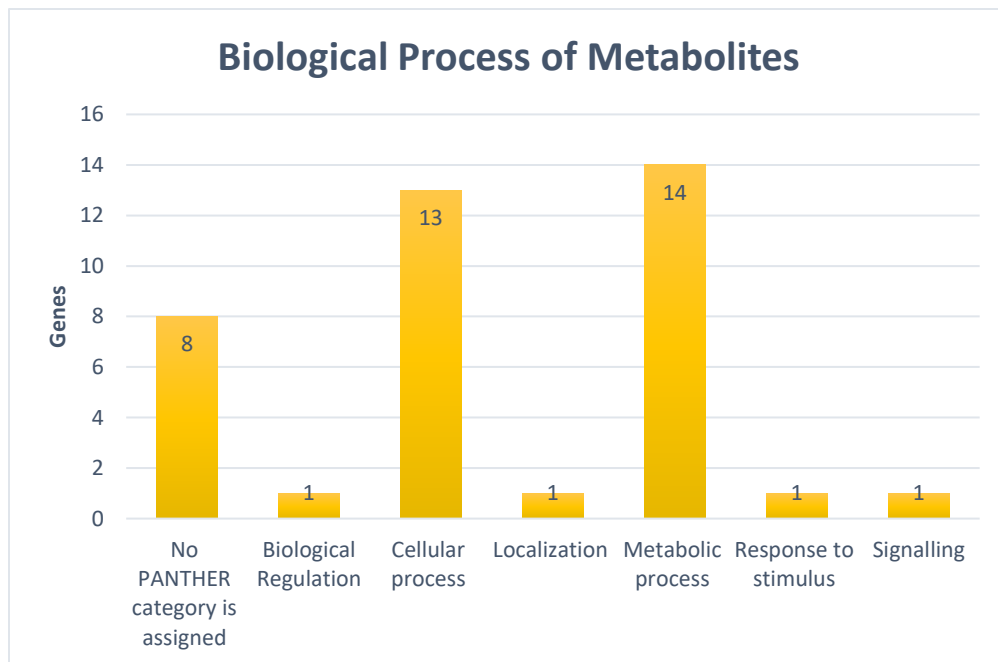


Table 10: Showing the cellular components of metabolites.

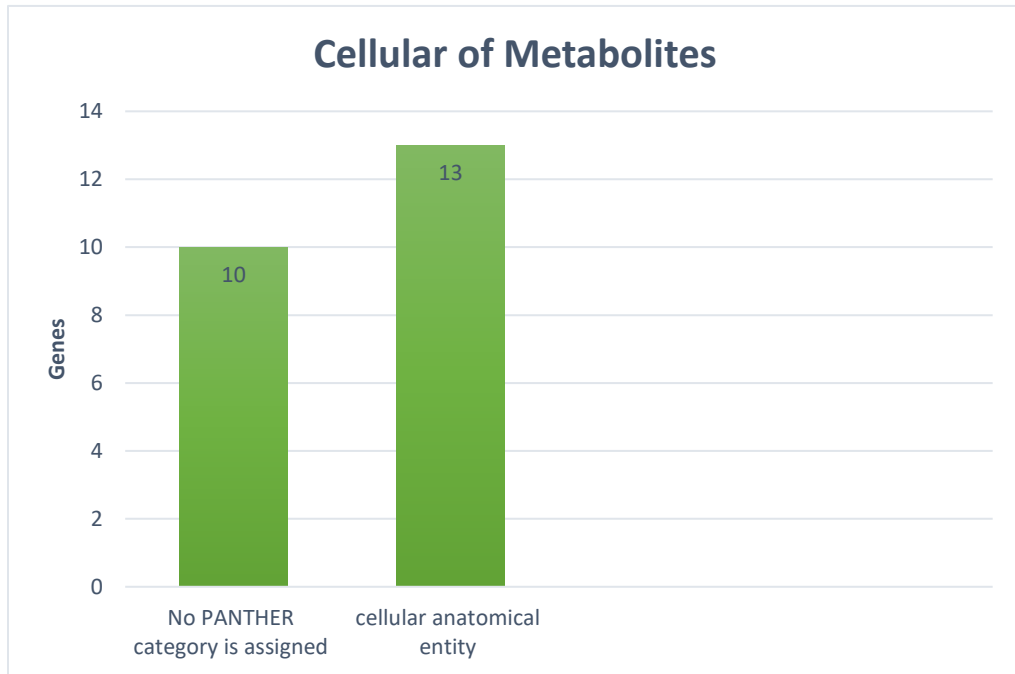


Table 11: Showing the Protein Class of metabolites.

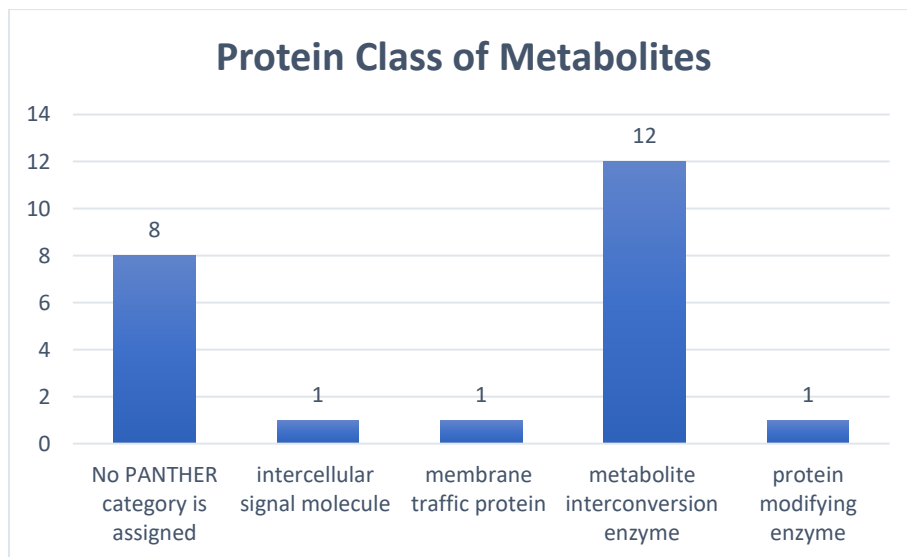


Table 12: Pathways of metabolites

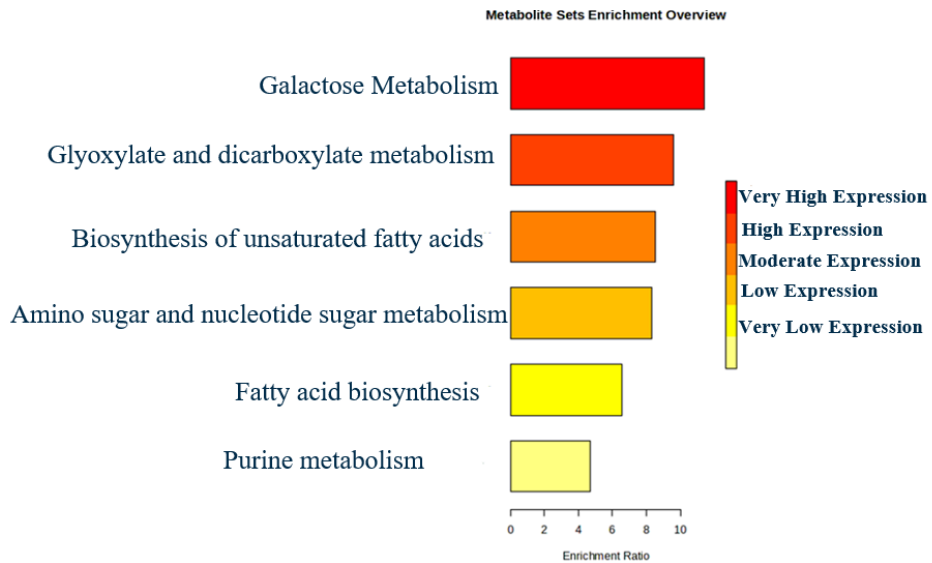
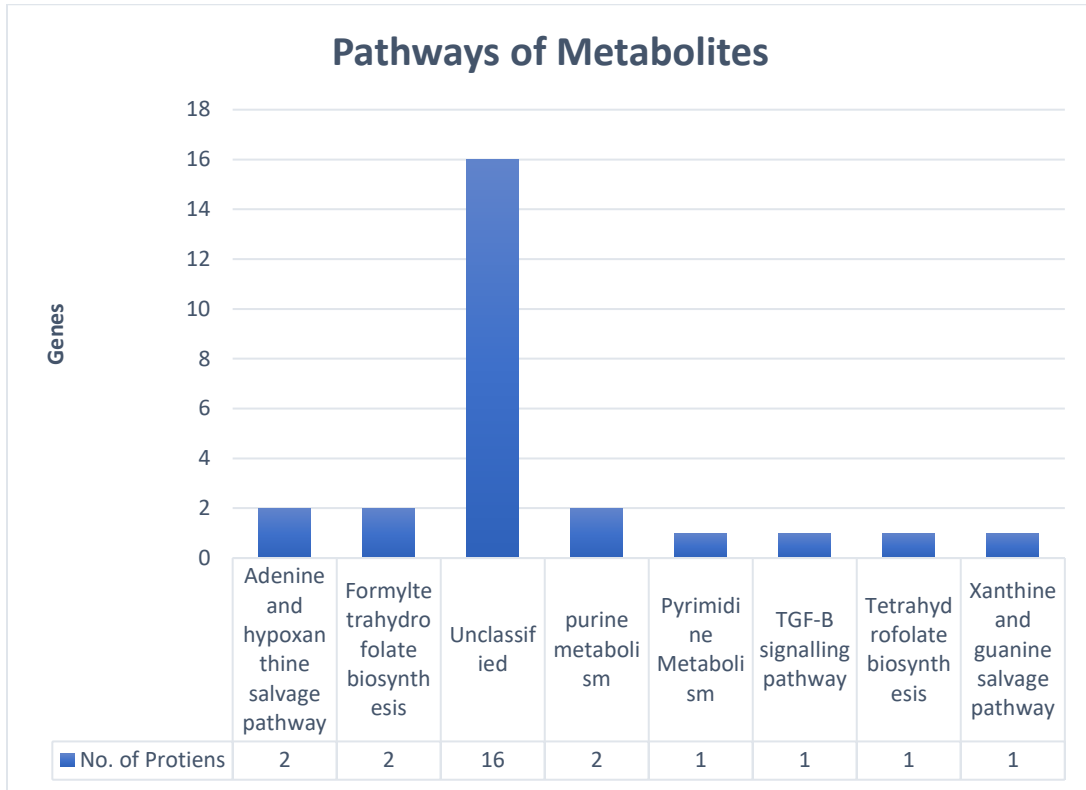


Figure 18: The enrichment overview of the metabolites. It refers to a systematic analysis of the biological pathways and functions associated with a set of metabolites.

MTOR signaling pathway has been identified to play a vital role in folate metabolism involving the glycolysis steps. Various metabolic MTOR is activated by a series of proteins like PI3k, and AKT, SAMTOR, which further activates the metabolic proteins like Fox 13 Myc, and HIF1a.

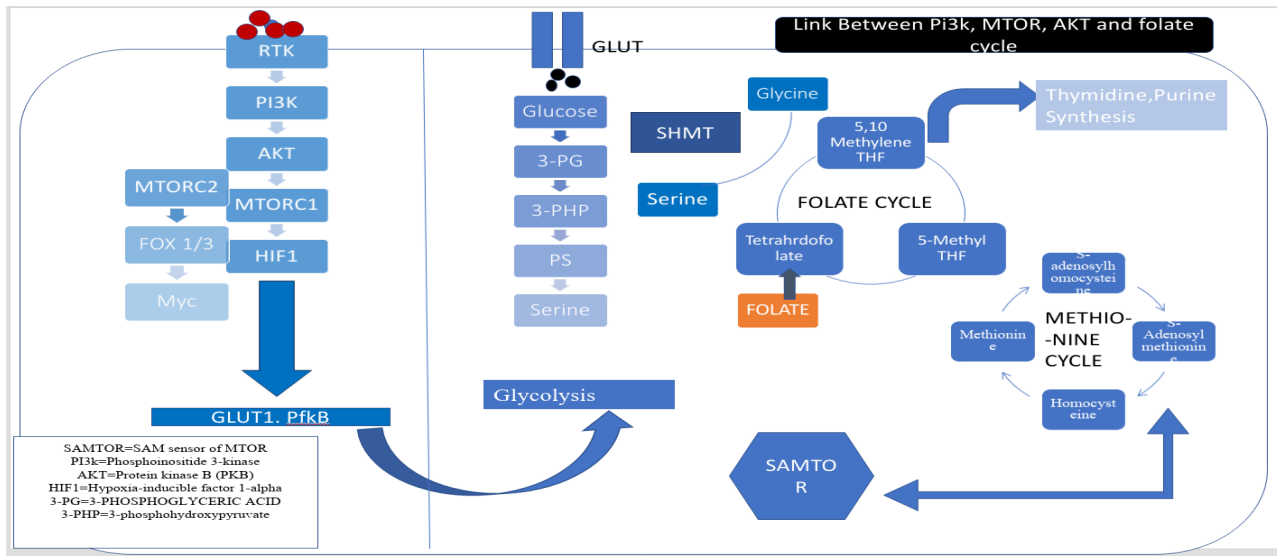


Figure 20 Mtor Signaling Pathway Involving Glycolysis

CHAPTER 5

DISCUSSION

Despite being one of the most widely studied diseases of all time, cancer is an enigma. Hundreds, if not thousands, of drugs, have been tested on different cancers, and a few of them have been approved to be used clinically on humans as effective treatments over the years. Despite this, cancer still presents itself as a problem with its ever-evolving characteristics requiring constant modifications in treatment and prognosis mechanisms. Nearly 73.4% of hepatocellular carcinoma (HCC) cases are reported due to hepatitis C virus (HCV), and hepatitis B (HBV) causes two times as many cases as HCV. It was reported by the world health organization (WHO) 3.5% of the world's population was infected, and 88,700 mortalities were due to cirrhosis and HCC. The burden on the health care system due to HBV-associated HCC will decrease as vaccination is available worldwide. Hepatitis C virus (HCV) can be acquired during childhood, and the symptoms are asymptomatic and 17 times more likely to develop HCC than a non-infected control group. In Pakistan, deaths and new cases of cancer are increasing, and the lack of quality and accessibility of the healthcare system also plays a significant role. The average age for hepatocellular carcinoma (HCC) in Pakistan is 7.6 per 100,000 persons annually for males and 2.8 for females.

The aims and objective of this study were to identify Dihydroquercetin affects liver cancer cells. By using a gas chromatography-mass spectrometer to monitor the metabolite levels in plasma samples from HCC cases and individuals with liver cirrhosis, we assessed the impact (GCMS). Additionally, they will support the development of efficient cancer patient treatment regimens and treatments.

The multiple health advantages of eating fruits and vegetables, researchers have been looking into the existence of various bioactive substances and how they affect human

health. Bioactive substances derived from plants have gained recognition in recent years as novel agents that are essential in preventing and/or treating a number of human illnesses, including as cancer, inflammation, cardiovascular disease, and neurological disorders. Flavonoids are the most beneficial bioactive compounds, A benzo-pyrone ring is included in the structure of flavonoids, which are simply hydroxylated phenolic compounds. Studies have shown that flavonoids intake reduces mortality risk. The therapy of carcinoma is significantly influenced by flavonoids. Because they contain hydroxyl groups, which have a strong antioxidant effect and can counteract the effects of free radicals, they can also chelate metal ions. Anticancer activities such as the prevention of angiogenesis, cytochrome P450 enzymes, P-glycoprotein, reactive oxidative species (ROS), and cell cycle regulators, as well as triggering of apoptosis (Das, Baidya, Chakraborty, Samanta, & Roy, 2021),(Madunić, Madunić, Gajski, Popić, & Garaj-Vrhovac, 2018; Raffa, Maggio, Raimondi, Plescia, & Daidone, 2017).

The pharmacophore and pharmacokinetic results of the present study indicated that Dihydroquercetin has the ability to cure liver cancer and is safe to use. It binds different receptors such as estrogen receptor, androgen receptor, thyroid receptor, glucocorticoid receptor, and aromatase binding (Inoue et al., 2019). Highlighting that the drug has a broad mechanism of action and has multiple physiological effects. As it does not violate any rule of Lipinski this indicates that it is used as a medication to treat liver cancer. The results show that the blood brain barrier function is high, and the absorption is 100%. These results are supported with previous studies, as in previous studies regarding Dihydroquercetin it has grabbed a lot of attention due to its antioxidant effect and most importantly it has shown great anti-cancer activity compared to other activities as investigated in vitro and in vivo with little or no side effects on human body (Das et al., 2021).

By the molecular docking of Dihydroquercetin with the (AKT2 protein, Proto-oncogene cJun, Epidermal growth factor receptor, Eukaryotic translation initiation factor 2A, Serine/threonine-protein kinase mTOR and Phosphoinositide 3-kinase regulatory subunit

4). We analyzed the interaction of amino acid residues binding affinity and binding modes. We were able to identify the type of interactions, type of bonds present with their bond length for example There were a total of 4 hydrogen bonds between Dihydroquercetin and AKT2. 1) Hydrogen Bond between Oxygen of Glutamine 46 (A) and Oxygen of Dihydroquercetin with a bond length of 2.80 Angstrom. 5 Hydrophobic bonds are present between amino acid residues (Tyrosine 25 (A), Proline 41 (A), Leucine 27 (A) and Proline 44 (A) and ligand. Therefore, we could identify the molecules that bind the drug with high affinity, and help us in selection of the lead compounds, we can optimize the drug structure to enhance the binding and selectivity. For past 30 years docking has been used to discover novel ligands, with time several docking methods were introduced (Shoichet, McGovern, Wei, & Irwin, 2002). By vina scoring system we compared the binding affinity of the ligand and protein. Vina Score is calculated by the sum of empirical free energy of binding and the entropic cost of binding. An improved binding mode is indicated by a lower Vina score, which is used to rank the binding modes of ligands, Affinity that falls between -8 to -11 indicate the best binding affinity(Guedes, de Magalhães, & Dardenne, 2014). Vina Scoring system is so far the most accurate method for molecular docking (Quiroga & Villarreal, 2016).

Metabolomic analysis of the treated cancer cells has proven to be a useful technique in identifying changes in the metabolic pathways of cancer cells compared to normal cells, providing comprehensions into the mechanisms of cancer development and metastasis. By utilizing different methods of metabolomic analysis identifying specific metabolites that are altered in treated vs non-treated cancer cells, the metabolic reprogramming that occurs and change of pathways occurring is better understood. Current metabolomics research aims to take this considerably further by looking at groups of metabolites or indeed the metabolome as a whole. These collections of data will contain patterns that then represent the metabolic signature of the sample, which can be compared to the patterns of other samples without the need to identify any of the individual molecules. This has the advantage of incorporating known and unknown metabolites of all the upstream events: gene expression and activated cellular pathways from the tumor; reactive and immunological responses from the host; as well as integrated signaling

pathway cross talk and environmental influences, by far a more comprehensive picture, albeit embedded in a vast sea of other metabolite data (Hart) Furthermore, metabolomics approach can be combined with various other approaches such as genomics and proteomics to understand how disease progression affects the genetic makeup of abnormal cells.

Dihydroquercetin was docked with six proteins that are involved in liver cancer which are as follows: Epidermal growth factor receptor, AKT2 protein, Phosphoinositide 3-kinase regulatory subunit 4, Serine/threonine-protein kinase mTOR, Proto-oncogene cJun, Eukaryotic translation initiation factor 2A. Initially the pharmacokinetics and toxicity analysis of the drug was studied that clarified that despite of not being able to cross the blood brain barrier, it had high absorption in the gastrointestinal tract so could be monitored as an effective oral drug. The toxicity analysis revealed that it did not show eye corrosion, hepatotoxicity, or any skin condition. However, it causes respiratory reproductive and mitochondrial toxicity.

The activity of these proteins when docked with Dihydroquercetin was studied thoroughly using CB-Dock2 and it was deduced that the most suitable protein ligand for the molecule was the Serine/threonine-protein kinase mTOR as it had the most negative vina score which means that it can be used as potential target for treatment with the drug. Furthermore, upon protein pathway analysis, it was discovered that there were four pathways that were involved. 1)Biotpterin metabolism, 2) Purine metabolism, 3) Urea cycle and metabolism of arginine, proline, glutamate, aspartate, and asparagine, and 4) Vitamin B9 (folate) metabolism. This means that these pathways can be manipulated and the metabolites that are part of these pathways can be potential targets of our drug. Similar models have been used before wherein GC/MS analysis is done on cancer cells that have been treated vs untreated cells with potential drugs. One such study involved developing an understanding of the effect of *Moringa oleifera* as an anti-cancer agent against breast and colorectal cancer cell lines. GC/MS analysis was done on the cancer

cell lines which revealed numerous known anti-cancer compounds in moringa, including eugenol, isopropyl isothiocyanate, D-allose, and hexadecenoic acid ethyl ester, all of which possess long chain hydrocarbons, sugar moiety and an aromatic ring.(Abdulrahman Khazim Al-Asmari, August 19, 2015). Another similar study reported the in vitro antioxidant and anticancer activity followed by in silico anticancer and estrogen-like activity of *Psidium guajava* L. essential oil against ER- α receptors which lead to potential inhibitory action against breast cancer pathways. Methods: The bioactive compounds in guava essential oil were screened using gas chromatography–mass spectrometry (GC-MS) (Mandal).

There are some limitations of the study conducted. The drug Dihydroquercetin is shown to be toxic to the respiratory reproductive and mitochondrial activity as mentioned in results. Further studies need to be conducted in order to develop dosage quantities that minimize the toxicity without compromising drug activity. Furthermore, ongoing research reveals new metabolites and genes which are added to the databases regularly so large amounts of data are undiscovered. In addition to that, Dihydroquercetin has limited clinical evidence in use in cancer therapies suggesting that there is room for further research and clinical trials to develop cancer therapy involving the drug. Alongside that, the interaction of DHQ with common chemotherapy agents is unknown therefore rendering risk of decreased efficacy of DHQ when used in conjunction with other chemotherapies.

MTOR signaling pathway has been identified to play a vital role in folate metabolism involving the glycolysis steps. Various metabolic MTOR is activated by a series of proteins like PI3k, and AKT, SAMTOR, which further activates the metabolic proteins like Fox 13 Myc, and HIF1a. The mTOR is activated by the protein SAMTOR produced as a result of the methionine cycle. These proteins combine to translate the GLUT 1, and PFKP genes in the nucleus which tend to activate the production of 3PG from glycolysis through a series of metabolic steps. The 3-PG and glucose are converted to make the serine which is incorporated into the Folate cycle to make 5:10 Methylene Tetra hydro folate. This reaction is catalyzed by Serine hydroxy methyl transferase. This is the first

step of the folate cycle. The cycle continues to initiate the methionine cycle again, which will produce more SAMTOR which further activates the mTOR signaling pathway.

In conclusion, treated and untreated (control) pellet samples of Hep-G2 cell line were collected. Through GC/MS and several in-silico tools, analysis was done in order to understand the effects of DHQ on cellular metabolites of liver cancer cells. It was discovered that the best protein for DHQ interaction was Serine/threonine-protein kinase mTOR which had a vina score of -8.3 which showed potential for being used as a target metabolite for the drug.

The current study should be replicated with multiple liver cancer cell lines and a greater sample size should be used to increase accuracy of the research. Since cancer is a globally distributed disease, similar studies should be conducted on different ethnic groups, age groups and genders. Proteins shown to have potent interaction with DHQ can be further mapped out in signaling pathways which in turn can be used to develop diagnostic biomarkers alongside treatment targets.

References

- Abdulrahman Khazim Al-Asmari, S. M. A., Md Tanwir Athar, Abdul Quaiyoom Khan, Hamoud Al-Shahrani, Mozaffarul Islam. (August 19, 2015). Moringa oleifera as an Anti-Cancer Agent against Breast and Colorectal Cancer Cell Lines.
- Abou-Alfa, G. K., Jarnagin, W., El Dika, I., D'Angelica, M., Lowery, M., Brown, K., . . . Crane, C. H. (2020). Liver and bile duct cancer. In *Abeloff's clinical oncology* (pp. 1314-1341. e1311): Elsevier.
- Ananthakrishnan, A., Gogineni, V., & Saeian, K. (2006). *Epidemiology of primary and secondary liver cancers*. Paper presented at the Seminars in interventional radiology.
- Balogh, J., Victor III, D., Asham, E. H., Burroughs, S. G., Boktour, M., Saharia, A., . . . Monsour Jr, H. P. (2016). Hepatocellular carcinoma: a review. *Journal of hepatocellular carcinoma, 3*, 41.
- Borrebaeck, C. A. K. (2017). Precision diagnostics: moving towards protein biomarker signatures of clinical utility in cancer. *Nature Reviews Cancer, 17*(3), 199-204. doi:10.1038/nrc.2016.153
- Cheng, F., Li, W., Zhou, Y., Shen, J., Wu, Z., Liu, G., . . . Tang, Y. (2012). admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. In: ACS Publications.
- Cousins, K. R. (2005). ChemDraw Ultra 9.0. CambridgeSoft, 100 CambridgePark Drive, Cambridge, MA 02140. www.cambridgesoft.com. See Web site for pricing options. In: ACS Publications.
- Das, A., Baidya, R., Chakraborty, T., Samanta, A. K., & Roy, S. (2021). Pharmacological basis and new insights of taxifolin: A comprehensive review. *Biomedicine & Pharmacotherapy, 142*, 112004.
- Guedes, I. A., de Magalhães, C. S., & Dardenne, L. E. (2014). Receptor–ligand molecular docking. *Biophysical reviews, 6*, 75-87.
- Hafeez Bhatti, A. B., Dar, F. S., Waheed, A., Shafique, K., Sultan, F., & Shah, N. H. (2016). Hepatocellular carcinoma in Pakistan: national trends and global perspective. *Gastroenterology Research and Practice, 2016*.
- Hart, C., Tenori, L., Luchinat, C., Di Leo, A. (2016). Hart, C., Tenori, Metabolomics in Breast Cancer: Current Status and Perspectives. In: Stearns, V. (eds) Novel Biomarkers in the Continuum of Breast Cancer. Advances in Experimental Medicine and Biology. In.
- Henry, N. L., & Hayes, D. F. (2012). Cancer biomarkers. *Mol Oncol, 6*(2), 140-146. doi:10.1016/j.molonc.2012.01.010
- Inoue, T., Saito, S., Tanaka, M., Yamakage, H., Kusakabe, T., Shimatsu, A., . . . Satoh-Asahara, N. (2019). Pleiotropic neuroprotective effects of taxifolin in cerebral amyloid angiopathy. *Proceedings of the National Academy of Sciences, 116*(20), 10031-10038.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., . . . Potapenko, A. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583-589.

- Kałużna-Czaplińska, J. (2011). Current medical research with the application of coupled techniques with mass spectrometry. *Med Sci Monit*, 17(5), Ra117-123. doi:10.12659/msm.881756
- Koes, D. R., & Camacho, C. J. (2012). ZINCPharmer: pharmacophore search of the ZINC database. *Nucleic acids research*, 40(W1), W409-W414.
- Landegren, U., & Hammond, M. (2021). Cancer diagnostics based on plasma protein biomarkers: hard times but great expectations. *Molecular Oncology*, 15(6), 1715-1726.
- Laskowski, R. A., & Swindells, M. B. (2011). LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. In: ACS Publications.
- Liu, C.-Y., Chen, K.-F., & Chen, P.-J. (2015). Treatment of liver cancer. *Cold Spring Harbor perspectives in medicine*, 5(9), a021535.
- Liu, Y., Grimm, M., Dai, W.-t., Hou, M.-c., Xiao, Z.-X., & Cao, Y. (2020). CB-Dock: A web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacologica Sinica*, 41(1), 138-144.
- Lliveras-Tenorio, A., Vinciguerra, R., Galano, E., Blaensdorf, C., Emmerling, E., Perla Colombini, M., . . . Bonaduce, I. (2017). GC/MS and proteomics to unravel the painting history of the lost Giant Buddhas of Bāmiyān (Afghanistan). *PloS one*, 12(4), e0172990.
- Madunić, J., Madunić, I. V., Gajski, G., Popić, J., & Garaj-Vrhovac, V. (2018). Apigenin: A dietary flavonoid with diverse anticancer properties. *Cancer Letters*, 413, 11-22.
- Mandal, A. K., Paudel, S., Pandey, A., Yadav, P., Pathak, P., Grishina, M., Jaremko, M., et al. Guava Leaf Essential Oil as a Potent Antioxidant and Anticancer Agent: Validated through Experimental and Computational Study.
- Mi, H., Muruganujan, A., Ebert, D., Huang, X., & Thomas, P. D. (2019). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic acids research*, 47(D1), D419-D426.
- Nakanuma, Y., Sripa, B., Vatanasapt, V., Leong, A., Ponchon, T., & Ishak, K. (2000). Intrahepatic cholangiocarcinoma. *World health organization classification of tumours Pathology and genetics of tumours of the digestive system*, 173-180.
- Pang, Z., Chong, J., Zhou, G., de Lima Morais, D. A., Chang, L., Barrette, M., . . . Xia, J. (2021). MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic acids research*, 49(W1), W388-W396.
- Parikh, N. D., Mehta, A. S., Singal, A. G., Block, T., Marrero, J. A., & Lok, A. S. (2020). Biomarkers for the early detection of hepatocellular carcinoma. *Cancer Epidemiology, Biomarkers & Prevention*, 29(12), 2495-2503.
- Petrick, J. L., & McGlynn, K. A. (2019). The changing epidemiology of primary liver cancer. *Current epidemiology reports*, 6(2), 104-111.
- Piñero, F., Dirchwolf, M., & Pessôa, M. G. (2020). Biomarkers in hepatocellular carcinoma: diagnosis, prognosis and treatment response assessment. *Cells*, 9(6), 1370.
- Quiroga, R., & Villarreal, M. A. (2016). Vinardo: A scoring function based on autodock vina improves scoring, docking, and virtual screening. *PloS one*, 11(5), e0155183.
- Radchenko, E., Karpov, P., & Sosnin, S. (2016). *System for prediction of pharmacokinetic properties and toxicity of drug compounds*. Paper presented at the XX Mendeleev Congress on general and applied chemistry.

- Raffa, D., Maggio, B., Raimondi, M. V., Plescia, F., & Daidone, G. (2017). Recent discoveries of anticancer flavonoids. *European journal of medicinal chemistry*, *142*, 213-228.
- Rose, P. W., Prlić, A., Altunkaya, A., Bi, C., Bradley, A. R., Christie, C. H., . . . Feng, Z. (2016). The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic acids research*, gkw1000.
- Saito, R., Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., Lotia, S., . . . Ideker, T. (2012). A travel guide to Cytoscape plugins. *Nature methods*, *9*(11), 1069-1076.
- Shoichet, B. K., McGovern, S. L., Wei, B., & Irwin, J. J. (2002). Lead discovery using molecular docking. *Current opinion in chemical biology*, *6*(4), 439-446.
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, *71*(3), 209-249.
- Tzartzeva, K., Obi, J., Rich, N. E., Parikh, N. D., Marrero, J. A., Yopp, A., . . . Singal, A. G. (2018). Surveillance imaging and alpha fetoprotein for early detection of hepatocellular carcinoma in patients with cirrhosis: a meta-analysis. *Gastroenterology*, *154*(6), 1706-1718. e1701.