Investigating Pathogenic SNPs of PKCι in HCV Induced Hepatocellular Carcinoma

By

Naila Khan

00000320343

Supervised by

Dr. Maria Shabbir

Atta-Ur-Rahman School of Applied Biosciences

National University of Science and Technology

Islamabad Pakistan

2021

Investigating Pathogenic SNPs of PKCι in HCV Induced Hepatocellular Carcinoma

By

Naila Khan

00000320343

Supervised by
Dr. Maria Shabbir

Co-Supervised by
Mrs. Yasmin Badshah

Atta-Ur-Rahman School of Applied Biosciences

National University of Science and Technology

Islamabad Pakistan

2021

A thesis submitted to National University of Science and Technology, Islamabad, in partial fulfilment of the requirement for the degree of Master of Science in Healthcare Biotechnology

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS Thesis written by Ms. Naila Khan (00000320343), of Atta-Ur-Rahman School of Applied Biosciences (ASAB) 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 MS degree. It is further certified that necessary amendments as pointed by GEC members of the scholar have also been incorporated in the said thesis.

Signature: _____

Name of Supervisor: _____

Date: _____

Signature: _____

Name of Co-Supervisor: _____

Date: _____

Signature (HOD): _____

Date: _____

Signature (Dean/Principal): _____

Date: _____

Author's Declaration

I hereby declared that except where specific reference is made to the work of others, the content of this dissertation is original and have not been submitted in whole or in part for consideration for any other degree or qualification in this or any other University. This dissertation is the result of my own work and includes the outcome of the work done.

Naila Khan

Date: 26/08/2021

Certificate for Plagiarism

It is certified that MS Thesis Titled "Investigating Pathogenic SNPs of PKCι in HCV Induced Hepatocellular Carcinoma" by Naila Khan has been examined by me. I undertake the follows:

a. Thesis has significant new work/knowledge as compared 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 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 attached) and found within limits as per HEC plagiarism Policy and instructions issued from time to time.

Name & Signature of Supervisor

Supervisor name: _____

Signature: _____

Date: _____

Dedication

I dedicated this dissertation to all those who prayed for me and played their part in making me the person that I am today. May Allah bless you all!

ACKNOWLEDGMENTS

In the name of Almighty Allah, the most Beneficent, the most Merciful. All praises to ALLAH who is the sustainer of the universe and has control over the heavens and Prophet (PBUH), the ultimate source of guidance for humanity.

I truly want to pay sincere gratitude to my supervisor Dr. Maria Shabbir who stayed with us through every thick and thin during this whole project. Without her sheer hard work and dedication towards her students it would never have been possible. I would like to thank her for her continuous support, patience, motivation, and enthusiasm. Besides supervisor, I would like to extend my cordial gratitude to Dr. Yasmin Badshah for being the humblest co-supervisor ever. I would also like to thank my thesis committee members, Dr. Rumeza Hanif and Dr. Sobia Manzoor. It is a delight to mention about all my teachers who are responsible for my educational development, with deep emotions of generosity and thankfulness.

I would like to express my sincerest gratitude to my closest friends Uzma Hafeez, Natasha Naeem, Asma Talib Qureshi and Fakhru Nisa for their continuous support and affection whenever I was down. My thanks and appreciations also go to my seniors Khushbakht Khan and Hania Shah and lab mates Areeba Rehman, Talha Iqbal and Ayesha Raza for their cooperation, support, and encouragement through this whole tenure.

I am extremely thankful to Dr. Hussnain Janjua, Principal ASAB, and Dr. Touqeer Ahmad, Head of Department for his support in all matter concerning my studies and research

I am at dearth of words to express my gratitude to my parents and siblings who have been patient, supportive and caring throughout this complete journey. Thanks to my father Mukaram khan and mother Farhada for giving me strength to reach the stars and chase my dreams. Last but not the least, a wholehearted thanks to my sisters Rahim Bano, Bushra Khan, Maryam Khan, Hina Khan and Iqra Khan and brothers Aaqibullah and Abdullah who also contributed a lot towards this journey of success.

Naila Khan

Table of contents

List of Tables

List of Figures

List of Abbreviations

ADMET	Absorption, Distribution, Metabolism, Excretion, Toxicity
ALT	Alanine aminotransferase
AOM	Azoxymethane
ApcMin	Adenomatous polyposis coli Multiple intestinal neoplasia
ARMS-PCR	Amplification Refractory Mutation System Polymerase Chain Reaction
ATP	Adenosine Triphosphate
BAD	Bcl-2-associated death promoter
CAK	Cyclin-dependent kinase
CDC42	Cell division control protein 42 homologue
CDK7	Cyclin Dependent Kinase 7
CML	Chronic myeloid leukemia
CT	Cycle threshold
DAG	Diacylglycerol
DDG	Delta delta G
DNA	Deoxyribonucleic acid
ECT2	Epithelial Cell Transforming 2
EDTA	Ethylenediaminetetraacetic acid
ELK1	ETS Like-1 protein
ERK	Extracellular signal-regulated kinase
ESCC	Esophageal squamous-cell carcinomas
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C Virus
HOPE	Have Our Protein Explained
ICC	Intrahepatic cholangiocarcinoma
IKK	IκB kinase
IL-1	Interleukin-1
I-TASSER	Iterative Threading ASSEmbly Refinement

LIP	Lipid Interaction Protein
LSCC	Laryngeal squamous cell carcinoma
LTα	Lymphotoxin-alpha
LTβ	Lymphotoxin-beta
MAPK	Mitogen-activated protein kinase
MMP10	Matrix Metallopeptidase 10
mRNA	Messenger Ribonucleic acid
NADH	Nicotinamide adenine dinucleotide
NOTCH1	Notch homolog 1, translocation-associated
NRF2	Nuclear factor erythroid 2–related factor 2
NUMB	NUMB endocytic adaptor protein
p62	Ubiquitin-binding protein p62
Par-4	Prostate apoptosis response 4
PB1	Phox and Bem1
PDAC	Peptidoglycan-N-acetylmuramic acid deacetylase
PDK1	Pyruvate Dehydrogenase Kinase 1
PDZ	Post synaptic density protein
PhD-SNP	Predictor of human Deleterious Single Nucleotide Polymorphisms
PKC	Protein Kinase C
PKCι	Protein Kinase C Iota
PLS1	Plastin 1
PS	Pseudosubstrate
qRT-PCR	Quantitative Reverse Transcriptase-Polymerase Chain Reaction
Rac1	Ras-related C3 botulinum toxin substrate 1
RI	Reliability Index
RIP	Ribosome-inactivating proteins
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
RNA	Ribonucleic acid

SDF	System Data Format
SDS	Sodium Dodecyl Sulfate
SGPT	Serum glutamic pyruvic transaminase
SIFT	Sorting Intolerant from Tolerant
SKI	SKI Proto-Oncogene
SNAP	Scalable Nucleotide Alignment Program
SNP	Single Nucleotide Polymorphism
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TE	Tris EDTA
TNF-α	Tumor Necrosis Factor-alpha
TRAP6	Thrombin receptor-activating peptide-6
UV	Ultraviolet
ZIP	Zinc-regulated, Iron-regulated transporter-like Protein

Abstract

Hepatocellular carcinoma is a leading cause of cancer-related deaths due to its complexity in diagnosis, chemo-resistance, and aggressive nature. Identification of pathogenic SNP in PKCι can be a potential biomarker in early diagnosis and treatment of HCC. In this study high risk variants of PKCι associated with HCV induced HCC were identified and validated in the wet lab. The association of viral pathogenicity was checked with the identified potentially pathogenic variant of PKCι. The in silico structural and functional analysis of the selected SNP of PKCι was carried out using different online tools. Molecular dynamic simulations were carried out for further confirmation. Last of all, the analysis of molecular interaction of PKCι with potential inhibitor, Decahydrocyclopentachrysenediol, was performed to assess its therapeutic potential. The identified pathogenic SNP (rs ID1199530606) of PKCι showed significant association with hepatocellular carcinoma. So, it may be used for prognosis of hepatocellular carcinoma. Glycine changes into tryptophan at identified variant (rs ID1199530606) position and caused the overall stability of PKCι to decrease. Docking results of Decahydrocyclopentachrysenediol with PKCι revealed that Decahydrocyclopentachrysenediol can be a potential inhibitor of PKCι

Chapter 1: Introduction

1.1. Liver Cancer

Liver tumors may originate from the liver cells (Sia, Villanueva, Friedman, & Llovet), or it may spread from cancer in other parts of body (metastatic). Primary liver tumor is more prevalent in men than women (Bray et al., 2018). It is the sixth most frequently diagnosed cancer and the fourth leading cause of deaths worldwide, with an approximately 841,000 cases and 782,000 deaths in 2018 (Bray et al., 2018).

1.1.1 Types of Liver Cancer

Liver cancer has several types based on the type of cells that becomes cancerous.

i. Hepatocellular carcinoma (HCC)

Hepatocellular carcinoma (HCC), also known as hepatoma, begins in hepatocellular cells. HCC encompasses around 75% of all liver cancer cases. The risk factors associated with it include [i]environmental exposures, cirrhosis of the liver triggered by alcoholism and hepatitis B and C viruses (Petrick & McGlynn, 2019).

ii. Fibrolamellar Carcinoma

Fibrolamellar HCC is an uncommon type of HCC and is usually more receptive to treatment than any other types of liver cancer (Sia et al., 2017).

iii. Cholangiocarcinoma

Cholangiocarcinoma also known as bile duct cancer takes place in the tiny, tube-like bile ducts within the liver that transport bile to the gallbladder. Intrahepatic bile duct cancer starts in ducts situated in the liver. Extrahepatic bile duct cancer grows in ducts outside the liver. Intrahepatic cholangiocarcinoma (ICC) accounts for around 12-15% of liver cancer cases (Petrick & McGlynn, 2019). ICC is less studied type of liver cancer as compared to HCC. The risk factors reported commonly for cholangiocarcinoma in Asia include viral hepatitis, cirrhosis, liver fluke infections and primary sclerosing cholangitis. (Clements, Eliahoo, Kim, Taylor-Robinson, & Khan, 2020; Khan, Tavolari, & Brandi, 2019)

iv. Angiosarcoma

Angiosarcoma, also known as hemangiosarcoma, is responsible for almost 1 percent of all liver cancers. Angiosarcomas starts in the blood vessels of the liver and expand rapidly. They typically are detected at a later stage.

v. Hepatoblastoma

Hepatoblastoma is a rare type of tumor that starts in liver. The symptoms associated with it are swellings, lumps, or pain in the abdomen. Blood test, biopsy or imaging is used for its diagnosis (Sia et al., 2017).

vi. Biliary cystadenocarcinoma

Biliary cystadenocarcinoma, an uncommon cystic tumor that occurs in the liver or, less often, in the extrahepatic biliary system. Congenital liver cysts, biliary cystadenoma,

hepatoduodenal ligament, and bile ducts are some of the common points of origin of Biliary cystadenocarcinoma (Sia et al., 2017).

1.1.2 Liver metastasis

Liver metastasis, also known as secondary liver cancer, takes place when primary cancer from another part of the body extend to the liver. Colon or colorectal cancer is the common cause of most of the liver metastases.

Liver cancer occurrence rates changes from 5.1 per 100,000 person-years in Europe to 17.7 per 100,000 person-years in eastern Asia (Ferlay et al., 2018). This change in figures shows regional differences in the occurrence of risk factors, particularly for HCC (Kulik & El-Serag, 2019). HCC is the most frequently occurring type of liver cancer in developing countries like Pakistan due to poor health facilities (Srivatanakul, Sriplung, & Deerasamee, 2004).

1.2. Hepatitis C virus induced HCC

According to a survey by World Health Organization, HCC caused second most casualties across the world. In addition to other risk factors, one of the prominent risk factors associated with HCC progression is Hepatitis C Virus (HCV). Chronic HCV infection is deemed as a main cause of HCC in developed countries. Chronic liver inflammation is induced by HCV, which start multiple changes comprising production of oxidative stress, steatosis, progressive fibrosis, cirrhosis and eventually HCC. The variations are either due to direct virus infection or immune system mediated response (during inflammation chronic consequences). Furthermore, HCV promotes generation of certain cytokines like Lymphotoxin alpha (LTα) and beta (LTβ) that are directly

associated with HCC. Metabolism of lipids are also altered by HCV that leads to more deposition of fats that is then linked with HCC incidence (Vescovo et al., 2016).

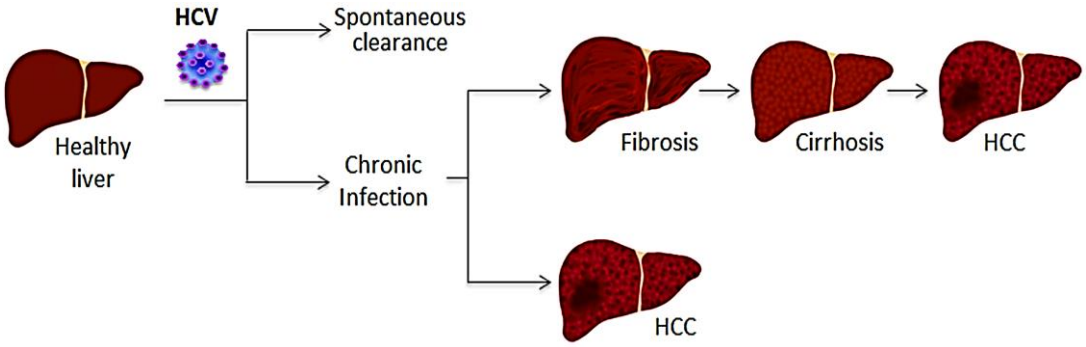

Figure 1. Schematic diagram of HCC in HCV infected patients

1.3. PKC Iota

The PRKCI gene encodes an important protein, PKCι, which is a lipid-dependent, serine/threonine kinase. PKCι participates in several signaling pathways that regulate cell survival (Sanz, Sanchez, Lallena, Diaz-Meco, & Moscat, 1999; Wooten, Seibenhener, Neidigh, & Vandenplas, 2000; Xie, Guo, Zhu, Wooten, & Mattson, 2000), differentiation (Wooten et al., 2000), polarity (Joberty, Petersen, Gao, & Macara, 2000), and microtubule dynamics in the early secretory pathway (Tisdale, 2002). PRKCI is a 5′ to 3′ kinase that belongs to the evolutionarily conserved PKC family. The PRKCI gene is located on chromosome 3 (3q26.2) between the polyhomeotic homolog 3 gene in centromeric position and the SKI-like oncogene in telomeric position. There are 18 exons and 83618 bases on the plus strand. The PRKCI transcript holds 4887 bases, and the open reading frame extends from 239 to 2029. PKCι protein have 596 amino acids and its molecular mass is 68262 Da. Lipid second messengers (ceramide, phosphatidylinositol 3, 4, 5-P3, and phosphatidic acid), phosphoinositide-dependent

kinase (PDK1), tyrosine phosphorylation and specific protein-protein interactions can regulate PKCι activity.

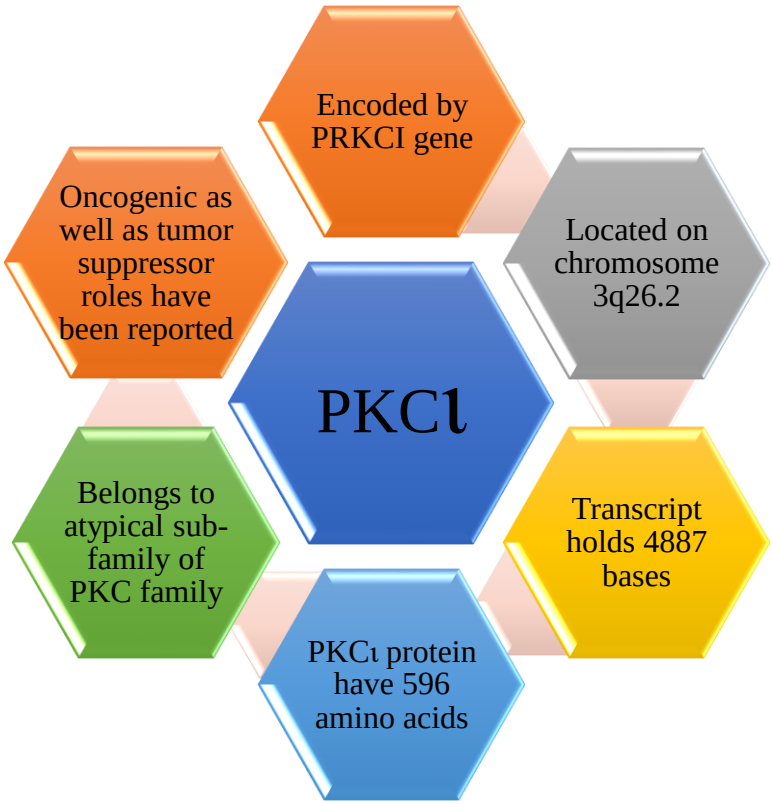

Figure 2. Protein Kinase C Iota (PKCι)

1.3.1. Structure of PRKCI

PKCι is a single polypeptide protein having regulatory N-terminal and catalytic C-terminal. The regulatory domain in the N-terminal have PB1 domain which facilitates protein-protein interactions between PKCι and other proteins having PB1 domain such as Par-6 (partitioning-defective 6) (Joberty, Petersen, Gao, & Macara, 2000; Lin et al., 2000; Noda et al., 2001; Qiu, Abo, & Steven Martin, 2000), ZIP/p62 (Hirano et al., 2004; Puls, Schmidt, Grawe, & Stabel, 1997), and MEK5 (MAPK (mitogen-activated protein kinase)/ERK (extracellular-signal-regulated kinase) kinase 5) (Diaz-Meco &

Moscat, 2001; Hirano et al., 2004). The PKCι PS (auto-inhibitory pseudo substrate sequence) in the inactive state is in the substrate binding cavity in the kinase domain and is exiled upon PKCι activation. The C1 domain is bind by Phosphatidylserine to anchor PKCι to the membrane. The PKCι lack C2 domain. This might be the reason of aPKCs being insensitive to Ca2+, diacylglycerol and phorbol esters, which are powerful activators of the other isoforms. The catalytic domain of PKCι is further divided into the C3 and C4 domains that facilitate ATP-binding and substrate binding.

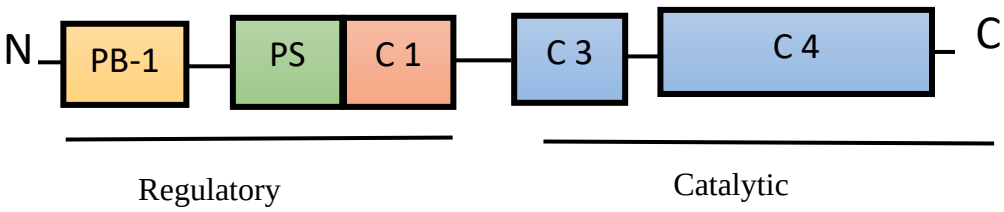

Figure 3. PKCι schematic diagram. Key: Phox Bem1 (PB1), Pseudo substrate sequence (PS)

Compared to other isozymes of PKC, PRKCI is more conserved from evolutionary point of view. The amino acid sequence homology of PKCι and PKCζ is 72% while the kinase domain is 86% identical. With other isoforms of PKC, PKCι shows less homology. Even in the highly conserved catalytic domain it is only 53% identical (Selbie et al., 1993).

Till date no germline mutation has been identified in PRKCI gene yet. However, there exist somatic mutations and PKCι gene is amplified in esophageal (Y. L. Yang et al., 2008), lung (R. P. Regala et al., 2005) and ovarian (A. M. Eder et al., 2005; L. Zhang et al., 2006) cancers. In a metastatic melanoma sample, a P118L mutation was found (Greenman et al., 2007). Although different PKC isozymes are linked to different

aspects of transformation, which include uncontrolled proliferation, invasion, migration, metastasis, angiogenesis, and resistance to apoptosis, but strong evidence are available only about atypical PKC iota as a bona fide oncogene. In liver cancer, overexpression of PKCι has been observed (Du et al., 2009). Cytoplasmic PKCι in hepatocellular carcinoma (HCC) has been linked with decreased cell-cell contact, loss of both adherens and tight junction formation, decreased E-cadherin expression, and rise in cytoplasmic beta-catenin (Du et al., 2009).

A tumor suppressor role for PKCι in hepatocellular carcinoma has also been reported recently. Cellular metabolic reprogramming is caused by loss of PKCι in obesity-driven liver cancer patients. This cellular metabolic reprogramming then induces two reciprocally sustainable processes, oxidative phosphorylation, and autophagy, which stimulate liver-tumor aggressiveness in an NRF2-dependent manner (Ramadori, Li, & Heikenwalder, 2020).

The protein PKC Iota has been positively associated with HCC in several studies. However, those studies usually indicate its involvement in the carcinogenic process at functional level. Any genetic association, such as that of the presence of SNPs in the Protein Kinase C Iota Coding (PRKCI) gene, between PKC Iota and HCC susceptibility and progression is yet to be found.

1.4. Single nucleotide polymorphisms (SNPs)

Single nucleotide polymorphism, or SNPs, are one of the commonly occurring genetic variation among individuals. It is basically a change at a single position in DNA sequence occurring in more than 1% of population. These may change the functional consequences of proteins. SNPs are present in different parts of gene and can act as

susceptibility or causative factor in cancer. Of all the SNPs, non-synonymous SNP (nsSNPs)/missense SNPs residing in the coding region are very crucial and accounts for residual change which may have neutral or deleterious effect on protein (Capriotti & Altman, 2011; Collins, Guyer, & Charkravarti, 1997). Non-synonymous SNPs alter the function of proteins by altering protein activity, solubility, and protein structure. These can also introduce premature termination in the protein sequence (Deng, Zhou, Fan, & Yuan, 2017). Polymorphism in certain genes like those controlling DNA mismatch repair, cell cycle regulation, immunity and metabolism etc. can lead to cancer development (Deng et al., 2017).

SNPs have arisen as genetic markers for diseases, and there are numerous SNP markers available in public databases. Previous studies have made known the importance of defining mutations as deleterious or non-deleterious and their association with certain diseases, thus pinpointing pathogenic SNPs that are functionally weaken or damaged due to structure-damaging properties (Kamaraj & Bogaerts, 2015; Kamaraj & Purohit, 2013; Karchin, 2009).

1.5. Decahydrocyclopentachrysenol

Decahydrocyclopentachrysenol is a triterpenoid aldehyde extracted from Nepeta hindostana. Scientific studies on the crude extract reported antiatherosclerotic, anticancer, antifungal, cardioprotective, antibacterial and antioxidant activities. In this study we will particularly evaluate the anti-cancer role of Decahydrocyclopentachrysenediol.

HCC is one of the most prevalent cancers in the world and PKCι have been associated with many cancers. Any genetic association specifically that of PKC Iota SNPs with

the increased susceptibility of developing hepatocellular carcinoma was yet to be found. Therefore, the focus of this study was to identify the potentially damaging SNPs present in PKC Iota and investigate and find their association with HCC first by computational approaches and then validate the results in wet lab. In the current study various tools were utilized for SNPs analysis and domain studies of Protein Kinase C Iota (PRKCI). In addition to that, we studied the interaction of Decahydrocyclopentachrysenediol with PKCι to study its possible therapeutic role against HCC. We performed *insilico* research to investigate the anti-cancer role of Decahydrocyclopentachrysenediol. For protein-drug interaction, LigPlot+ and PyMOL were used.

1.6. Problem Statement

Hepatocellular carcinoma is a leading cause of cancer-related deaths due to its complexity in diagnosis, chemo-resistance, and aggressive nature, therefore aim of the study was to identify PKCι as a potential biomarker in early diagnosis and druggable target in the treatment of HCC.

1.7. Objectives

- Identification of pathogenic non-synonymous SNPs in PKCι
- Prediction of PKCι tertiary structure, uniqueness, and localization
- Analysis of protein's overall structure, & dynamic behavior
- Analysis of PKCι variant in HCV induced HCC
- Relation of PKCι genotypes with viral load
- Analysis of molecular interaction of PKCι with inhibitor compound (Decahydrocyclopentachrysenediol)

Chapter 2: Literature Review

2.1. Hepatocellular carcinoma and its prevalence

Hepatocellular carcinoma (HCC) is the most prevalent type of liver cancer. It begins in hepatocellular cells and accounts for 75-85% of all liver cancer cases. Liver cancer occurrence rates changes from 5.1 per 100,000 person-years in Europe to 17.7 per 100,000 person-years in eastern Asia (Ferlay et al., 2018). This change in figures shows regional differences in the occurrence of risk factors (Kulik & El-Serag, 2019). According to world health organization (WHO), in the year 2020, the estimated number of new cases of liver cancer in world is 710,359 and in Pakistan it is 4354. While the estimated number of deaths are 638,911 in world and 4365 in Pakistan in 2020. HCC is the most frequently occurring type of liver cancer in developing countries like Pakistan (Srivatanakul, Sriplung, & Deerasamee, 2004).

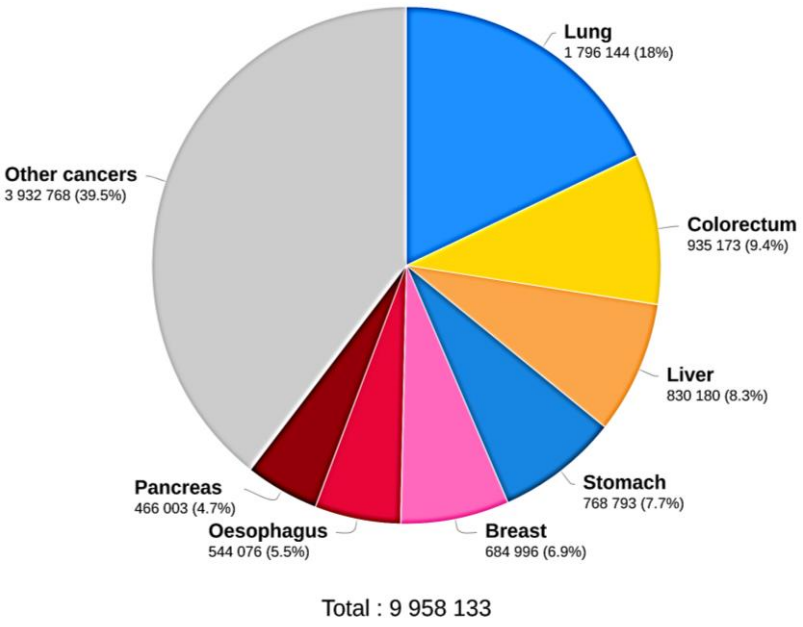

Figure 4. Estimated number of deaths in 2020, worldwide, both sexes, all age (world health organization, 2020)

2.2. Protein Kinase C

PKCs are an exclusive family of serine threonine kinases. It functions in phosphorylation of hydroxyl group of serine and threonine residues of proteins. Increased concentration of DAG (Diacylglycerol) or calcium ion (Ca+2) can activate them. PKCs function in tumor formation and progression. A lot of research from the past two decades indicate that various isoforms of PKC are somehow involve in causing cancer (Martiny-Baron & Fabbro, 2007).

There are different isozymes of PKCs family which are categorized into three subfamilies i.e., conventional, or classical PKCs, Novel PKCs and atypical PKCs. This grouping is based on second messenger requirements. The conventional or classical PKCs include α, βI, βII, and γ isoforms and their activation require calcium ion (Ca+2), DAG and phospholipids like phosphatidylserine. Novel PKCs on the other hand require DAG but do not require calcium ion (Ca+2) for activation. These include δ, ε, η, and θ isoforms. The third sub-family i.e., atypical sub-family include ζ and ι/λ isoforms. They require neither calcium ion (Ca+2) nor phorbol esters or DAG but instead require phosphatidyl serine for their activation. The prime focus of this study is the ι / λ isoforms of the atypical sub-family of PKCs (Parker, Justilien, Riou, Linch, & Fields, 2014).

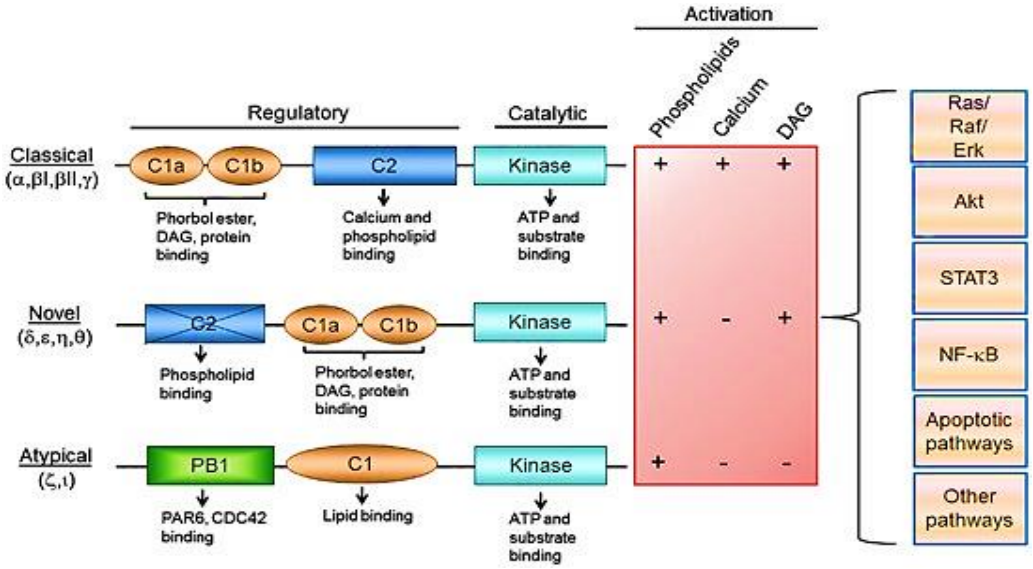

Figure 5. Structure of PKC isozymes (Nicole R Murray, Kalari, & Fields, 2011)

2.3. PRKCI: A unique member of PKC family

The PKCs are important elements of various signaling pathways that control different cellular functions like cell migration, cell cycle, survival, proliferation, cell migration, differentiation, and polarity. Studies have shown that when PKC isozymes are manipulated genetically, they carry out different and surplus cellular functions. In cancer, PKC functions, localization, phosphorylation, and expression changes and different PKC isozymes are linked to different aspects of transformation, which include uncontrolled proliferation, invasion, migration, metastasis, angiogenesis, and resistance to apoptosis. Although all isozymes of PKC are somehow linked to cancer, but strong evidence are available only about atypical PKC iota as a bona fide oncogene. The presence of PB1 domain in PKCι is exclusive to other atypical PKCs. This makes it a striking therapeutic target for treatment of cancer (Parker et al., 2014)

The protein kinase C (PKC) family of serine/threonine kinases were initially discovered around 40 years ago. And since their discovery they have been an attractive area of research in the field of cancer as cellular receptors for tumor facilitating phorbol esters

(Fields & Regala, 2007). Then around Twenty years later atypical (aPKCs) subfamily was discovered and became an area of great interest for researchers. Several studies have reported that the members of this subfamily play very crucial roles in signaling pathways governing cell growth, differentiation, and survival (Moscat & Diaz-Meco, 2000). In 1993 PKCι was identified in insulin-secreting cell line RINm5F (Selbie, Schmitz-Peiffer, Sheng, & Biden, 1993). Interesting studies in 2011 by Murray, Fields and colleagues recognized PKCι as an oncogene, vital for transformed growth of different human cancer cell types (Fields & Regala, 2007; Nicole R Murray, Kalari, & Fields, 2011)

2.3.1. Cellular Localization of PKCι

PKCι is primarily expressed in the cytoplasm. Once expressed, it then translocates to the cell membrane because of exposure to second messengers and colocalizes with p62/ZIP in lysosome-targeted endosomes (Sanchez, De Carcer, Sandoval, Moscat, & Diaz-Meco, 1998). And finally, translocation of PKCι into the nucleus takes place after Src phosphorylation (White, Seibenhener, & Wooten, 2002) where Cdk7 forms a complex with it (Win & Acevedo-Duncan, 2008)

2.3.2. PKCι: Role in cell cycle

PKCι and PKCζ normally are reported to have role in cell cycle progression. Consistent with this, in a study conducted by Paramio et al. it was reported that inhibition of PKCζ leads to blockade of cell cycle progression (Paramio, Segrelles, Ruiz, & Jorcano, 2001). In a Ras-mediated upregulation of cyclin D1, PKCι is present upstream of PKCζ (Kampfer et al., 2001). CAK is activated and phosphorylated by PKCι in response to PI-3K signaling in glioma and neuroblastoma cells (Acevedo-Duncan, Patel, Whelan,

& Bicaku, 2002; Desai et al., 2011; Pillai et al., 2011). In ovarian cancer, PKCι may target cyclin E (A. M. Eder et al., 2005).

2.3.3. Effect of over expression of PKC iota

The PKCι gene (PRKCI) is commonly overexpressed in human cancers. The increase in PRKCI copy number are detected in ~80% of human primary lung squamous cell carcinomas (LSCC) (Roderick P Regala et al., 2005), ~70% of serous epithelial ovarian cancer (Astrid M Eder et al., 2005) and ~53% of ESCC tumors (Yi-Ling Yang et al., 2008). These findings were further proved by the analysis of the data from Cancer genome atlas and other large-scale sequencing projects. The copy number gains are prominent in some cancer types like head, cervical, neck, lung squamous and ovarian serous cancer than the others like bladder, breast, kidney, lung adenocarcinoma, stomach, and uterine cancers. Analysis of gene expression data for the same tumor types also reveal a strong relation between the PRKCI copy number gains and increased PKCι mRNA expression for these tumor types. Thus, cancer specific gene copy number gains in PRKCI are the important genetic process carrying out PKCι expression in human cancers.

2.3.4. Effect of loss of PKC iota

Loss of PRKCI genetically leads to the inhibition Kras-driven hyperplasia and lung tumor formation in vivo (Jin et al., 2005). In another study conducted in 2000 it was reported that loss of function of Drosophila aPKC leads to the death of premature embryos even before they are cellularized (Wodarz, Ramrath, Grimm, & Knust, 2000). Another study reported in 2015 state that genetic loss of PRKCI causes increase in stem/progenitor cells. Consistent with these findings we can say that inhibition of PRKCI maybe valuable for the increased production of pluripotent and multipotent

stem cells or can even be possible for increased primordial germ cells production. This process may need inactivation and localization of NUMB driving the activation of NOTCH1 and downstream effectors (Mah, Soloff, Hedrick, & Mariani, 2015)

2.3.5. Molecular targets and interactions of PKCι

The aPKC family of kinases is made of two members ζPKC and λ/ιPKC. These show sequence similarity in their catalytic domain to other PKC isoforms, but the regulatory domain is very different. Many studies report that these kinases play major role in controlling cell growth and survival (Berra et al., 1993; Bjorkoy, Perander, Overvatn, & Johansen, 1997; Díaz-Meco et al., 1996; N. R. Murray & Fields, 1997; Wooten, 1999), probably by their regulation of important signaling pathways i.e. those that activate the AP-1 and NF-κB transcription factors (Akimoto et al., 1996; Berra et al., 1995; Diaz-Meco et al., 1993; Liao, Monia, Dean, & Berk, 1997; Sontag, Sontag, & Garcia, 1997; Takeda et al., 1999; Wooten, 1999). Initially it was thought that aPKC do it through Ras signaling, but with the advancement of knowledge in this field it was revealed that there are other complex interactions with many different proteins that do not take part in the Ras signaling. Scientists have discovered several proteins that interact with aPKC whose understanding is very helpful in learning the function and mechanism of action of these kinases. These include p62, Par-6, MEK5, Par-4, Bcl-XL and ERK (Moscat & Diaz-Meco, 2000).

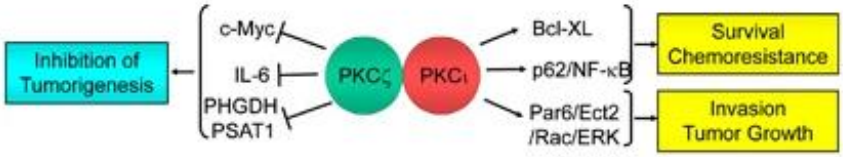

Figure 6. Atypical PKCs (PKCι and PKCζ) roles in cancer (Garg et al., 2014).

The AID (short stretch of acidic amino acids) site of Scaffold protein p62 binds to the V1 domain of aPKCs and link it to the tumor necrosis factor α (TNFα) and interleukin-1 (IL-1) receptor signaling through its interactions with RIP and TRAF6, respectively. Both RIP and TRAF6 have role in inflammatory response driven by the cytokines TNFα and IL-1 (Sanz, Sanchez, Lallena, Diaz-Meco, & Moscat, 1999). Another AID containing linking protein, α isoform of the kinase MEK5 (Diaz-Meco & Moscat, 2001) also interact with V1 domain of aPKC. These interactions are crucial for the activation of the MEK5/ERK5 pathway. These may be induced by mitogen. p62 and MEK5 both link aPKC to distinct signaling pathways through their AID sequence. Yet another AID site containing protein Par-6 (partitioning-defective-6) is stimulated by CDC42 or Rac. The Par-6 molecule make a complex with V1 site of aPKC present near Par-3, which communicate with the PDZ domain of Par-6. This complex formation leads to the phosphorylation of Par-3 by the aPKCs, causing downstream effects on cytoskeletal arrangement. This communication shows the importance of aPKCs in both Ras- and CDC42-induced cell transformation (Bjorkoy et al., 1997; Qiu et al., 2000).

V1 domain of the aPKCs binds several adapter molecules. These interactions link aPKC to different signaling pathways directly or indirectly. This could be the reason of specificity of the actions of aPKC especially when PKC isoforms show similarity in catalytic domain.

Besides V1 domain binding sites adapter molecules, there are two zinc finger domain binding molecules as well. These are LIP (lambda-interacting protein) (Díaz-Meco et al., 1996) and Par-4 (prostate androgen response-4) (Díaz-Meco et al., 1996). LIP is reported to be an activator of aPKC enzymatic activity (Díaz-Meco et al., 1996) while Par-4 is an inhibitor (Díaz-Meco et al., 1996). The exact mechanism of action of LIP is

not explored well yet. However, Par-4 is a crucial molecule for cell function (Díaz-Meco et al., 1996). Par-4 upon expression, inhibits IKK and thus the activation of NF-κB by TNFα, leading to the TNFα-induced death of the cells that are normally resistant to apoptosis (Díaz-Meco et al., 1996).

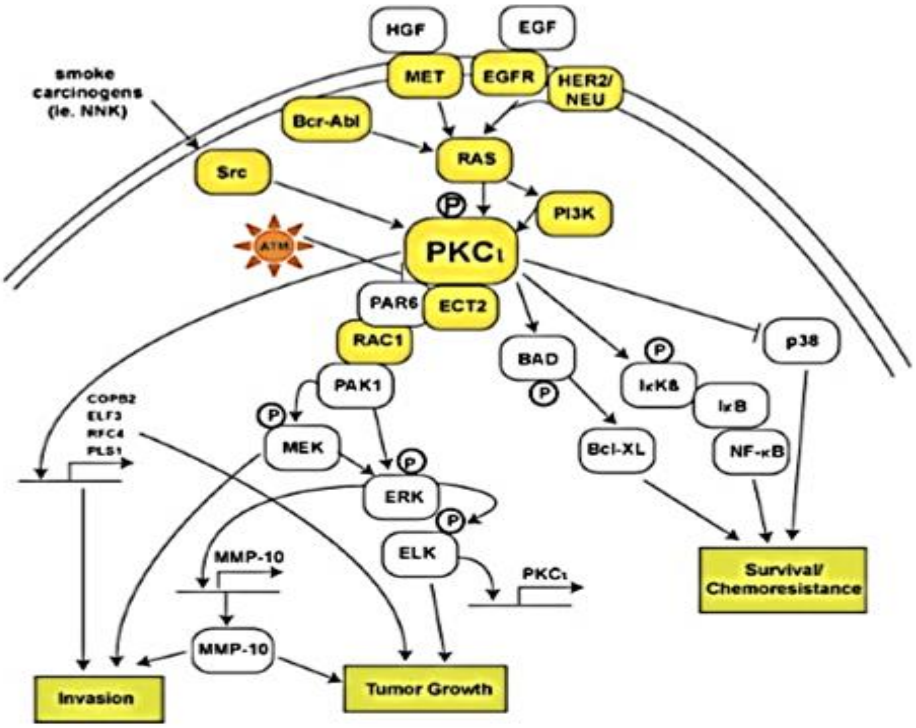

Figure 7. Schematic diagram of oncogenic signaling of PKCι (Murray, Kalari, & Fields, 2011)

2.3.6. PKCι: Dysregulation in cancer

The PKCι gene (PRKCI) is commonly overexpressed in human cancers. The increase in PRKCI copy number are detected in ~80% of human primary lung squamous cell carcinomas (LSCC) (Roderick P Regala et al., 2005), ~70% of serous epithelial ovarian cancer (Astrid M Eder et al., 2005) and ~53% of ESCC tumors (Yi-Ling Yang et al., 2008). These findings were further proved by the analysis of the data from Cancer genome atlas and other large-scale sequencing projects. The copy number gains are

prominent in some cancer types like head, cervical, neck, lung squamous and ovarian serous cancer than the others like bladder, breast, kidney, lung adenocarcinoma, stomach, and uterine cancers. Analysis of gene expression data for the same tumor types also reveal a strong relation between the PRKCI copy number gains and increased PKCι mRNA expression for these tumor types. Thus, cancer specific gene copy number gains in PRKCI are the important genetic process carrying out PKCι expression in human cancers. One important mutation reported in the substrate docking domain of PKCι is R471 (Linch et al., 2013).

i. Non-Small Cell Lung Cancer (NSCLC)

In NSCLC patients, PRKCI gene is amplified and act as an oncogene. Amplification leads to over-expression of PRKCI gene in NSCLC tumors (L. A. Frederick et al., 2008; R. P. Regala et al., 2005). Any Changes or loss of PRKCI gene inhibits oncogenic growth and transformation of lung stem cells induced by Kras. In addition to that PKCι phosphorylate pro-apoptotic protein BAD and cause an increased resistance of NSCLC to NNK-driven apoptosis (Jin, Xin, & Deng, 2005). The presence of PB1 domain in PKCι is exclusive to other atypical PKCs. It is required for oncogenic PKCι-Par6-Rac1-MMP10 signaling axis that facilitate anchorage-independent growth and invasion of human NSCLC cells in vitro and tumorigenicity in vivo. PKCι is also linked genetically to ECT2 (Justilien, Jameison, Der, Rossman, & Fields, 2011). Upon phosphorylation of ECT2 by PKCι, an oncogenic PKCι-Par6-Ect2 complex forms. This complex activates Rac 1 which then transform NSCLC cell (Justilien et al., 2011).

The expression of COPB2, ELF3, RFC4, and PLS1 in primary lung adenocarcinoma is also regulated by PKCι (Erdogan, Klee, Thompson, & Fields, 2009).

ii. Colon cancer

Over expression of PKCι takes place in human colon tumors. Murray et al., reported AOM-driven colon tumors in mice (N. R. Murray et al., 2004). The number of AOM-induced colon tumors increases due to the increased expression of PKCι and hence the benign adenoma grows into a malignant intramucosal carcinoma (N. R. Murray et al., 2004). PKCι is also involved in the oncogenic Ras-driven transformation of the intestinal epithelium APCM in/+ mice (N. R. Murray, Weems, Braun, Leitges, & Fields, 2009).

iii. Ovarian cancer

In patients with ovarian cancer, PKCι expression is elevated (A. M. Eder et al., 2005; Weichert, Gekeler, Denkert, Dietel, & Hauptmann, 2003; L. Zhang et al., 2006). The level of expression gives information about the stage of the cancer. This shows that role of PKCι in tumor progression and aggressiveness (A. M. Eder et al., 2005; Weichert et al., 2003; L. Zhang et al., 2006). As the level of expression of PKCι decreases, it reduces anchorage-independent growth of ovarian cancer cells. On the other hand, increased expression of PKCι facilitate transformation of murine ovarian surface epithelium (L. Zhang et al., 2006).

iv. Pancreatic cancer

Significant increase in expression of PKCι takes place in patients with human pancreatic cancer. Disruption in the expression of PKCι inhibit the transformation and invasion of human Pancreatic Ductal Adenocarcinoma (PDAC) cells (Scotti, Bamlet, Smyrk, Fields, & Murray, 2010). Tumorigenicity of PDAC cell tumors injected orthotopically into the pancreas is also inhibited by the disrupted PKCι expression (Scotti et al., 2010). These findings show that PKCι is the culprit causing tumors and metastases of kidney, liver, diaphragm and mesentery (Scotti et al., 2010). The pathway

responsible for transformed growth and invasion in PDAC cells is Rac1-MEK/ERK1/2 (Scotti et al., 2010).

v. Gliomas

PKCι is responsible for the facilitation of survival in glioblastoma cells by p38 mitogen-driven protein kinase signaling which provide protection to these cells from cytotoxicity of chemotherapeutic agents (Baldwin, Parolin, & Lorimer, 2008). Overexpression of PKCι takes place in glioblastoma multiforme. As a result of genetic knock out of PKCι, sensitization of glioblastoma cells to cisplatin takes place (Baldwin et al., 2008). RNAi driven knocking out of PKCι also inhibits division and invasiveness of glioma cell lines in vitro (Baldwin et al., 2008; Patel et al., 2008).

vi. Chronic myelogenous leukemia

PKCι act as a survival gene in chronic myelogenous leukemia (CML). Over-expression of PKCι takes place in human K562 leukemia cells. Ras/Mek/Erk pathway is activated by chimeric tyrosine kinase oncogene Bcr-Abl. This activation then stimulates PKCι expression by an Elk1 transcription factor site in the promoter of PKCι (Gustafson et al., 2004). The activation of PKCι by Bcr-Abl is crucial for apoptotic resistance to chemotherapy in K562 CML cells (N. R. Murray & Fields, 1997).

vii. Esophageal cancer

In esophageal squamous cell carcinomas (ESCC) patients PRKCI amplification and hence PKCι expression takes place (Y. L. Yang et al., 2008). A study conducted by Yang et al., reported that tumor progression and metastasis in ESCC is the result of PRKCI overexpression (Y. L. Yang et al., 2008).

2.3.7. PKCι inhibitors

The presence of PB1 domain in PKCι is exclusive to other atypical PKCs. It is required for oncogenic PKCι-Par6-Rac1-MMP10 signaling axis that facilitate anchorage-independent growth and invasion of human NSCLC cells in vitro and tumorigenicity in vivo (L. Frederick et al., 2008). So PB1-PB1 domain interaction between PKCι and Par6 is a promising target for therapeutic drugs. Currently there are a few FDA approved gold containing compounds such as aurothioglucose (ATG), aurothiomalate (ATM) and auranofin (ANF) that can dislocate the PB1-PB1 domain contact between PKCι and Par6 and thus inhibit them. The inhibition of PKCι-Par6 binding by these compounds is dose dependent, show good antitumor activity. These also inhibit PKCι-facilitated Rac1 activation and anchorage independent growth of NSCLC cells in vitro and tumorigenicity in vivo (Stallings-Mann et al., 2006). The inhibitory effectiveness of ATM and ANF have been tested and further trial are ongoing for PKCι-Par6 inhibition.

Another inhibitor reported for PKCι is methyl dihydrogen phosphate [4-(5-amino-4-carbamoylimidazol-1-yl)-2,3-dihydroxycyclopentyl] (Pillai et al., 2011). It is more specific and compete with myelin basic protein in vitro. It also blocks the division of neuroblastoma cells through CDK7 (Pillai et al., 2011).

Many small molecule inhibitors for PKCι PIF pocket have been recognized (Lopez-Garcia et al., 2011), Still more research need to be done for further developments and better understanding of the mechanism of such approaches to PKCι.

Triterpenoid ursolic acid an inhibitor reported for aPKCζ-sequestosome 1 interaction might as well be an inhibitor of aPKCι because the antibodies that detect the priming site phosphorylation cannot differentiate aPKCι and aPKCζ (Huang, Huang, Lin-Shiau, & Lin, 2009). It has not proven yet whether ursolic is the basis of an aPKC intervention.

A study conducted in 2013 reported an aPKC specific ATP-competitive inhibitor compound CRT0066854, a thieno[2`,3-d] pyrimidine (Kjær et al., 2013). It inhibits substrate phosphorylation in cells and prevents division in 2D and 3D cultures. It is also involved in decreased polarized morphogenesis and directional migration (Linch et al., 2014). More research still needs to be done on how these compounds work in clinical setting.

Chapter 3: Materials and Methods

The use of computational methodological approaches for getting biological understanding is well-known. Various studies support the concept that the use of different algorithms increases prediction accuracy (Khan & Ansari, 2017). In the current study, *in-silico* tools and methods were first employed to predict the structure of normal PKCι and to identify the potential damaging missense SNPs in the protein which are further investigated and validated as biomarkers for hepatocellular carcinoma through wet lab procedures.

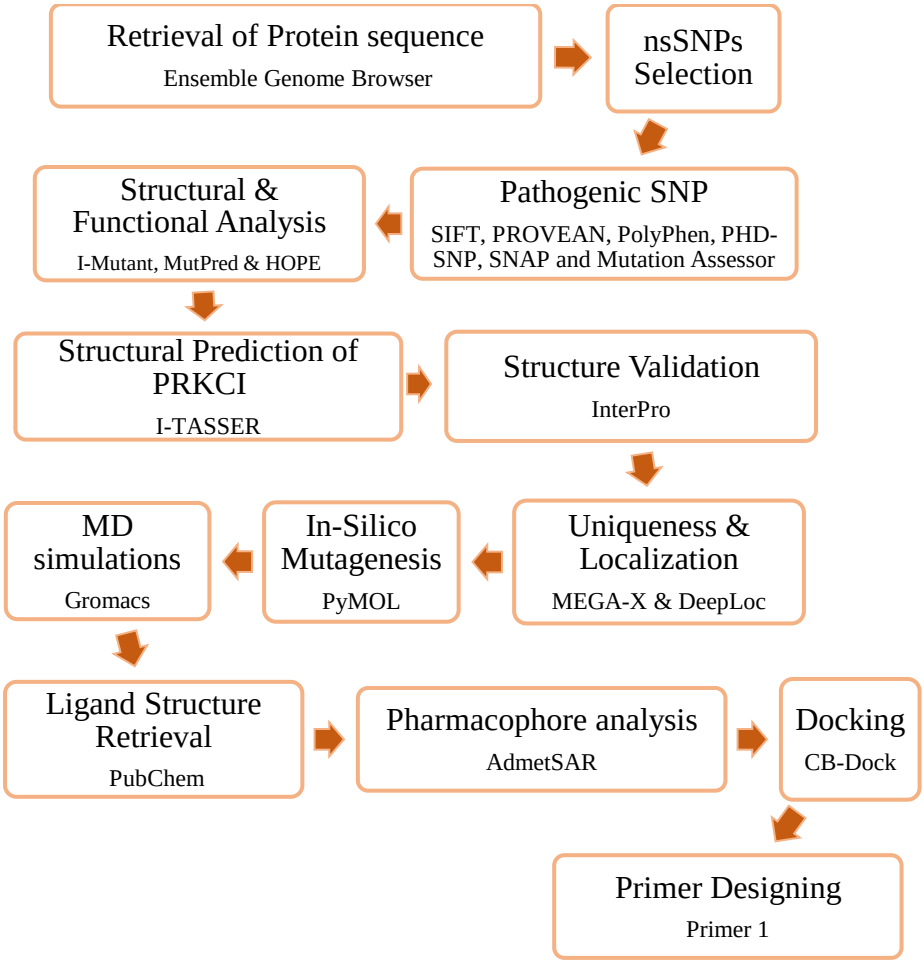

Figure 8. Schematic diagram of different *insilico* methods employed in this study

3.1. Protein Structure Prediction and Visualization

The protein sequence of PRKCI gene with transcript ID: PRKCI-201 ENST00000295797.5 was obtained from Ensembl database in FASTA format. This sequence was then submitted to I-TASSER (Iterative Threading ASSEmbly Refinement) (Roy, Kucukural, & Zhang, 2010) which is an online tool for prediction of protein structures based on the threading approach of protein modelling and generates each predicted protein model with a confidence score ranging from -5 to 2 (Yang & Zhang, 2015; Zhang, Freddolino, & Zhang, 2017). The predicted models were then visualized with the help of PyMOL molecular visualization system. In addition, the predicted models by I-TASSER were cross-checked using InterPro database (Blum et al., 2021) and other literature sources available, regarding the structural features of already studied and determined similar proteins.

3.2. Uniqueness and Localization

To predict the localization of the PKC iota, the amino acid sequence of the protein was submitted to a web-based tool known as DeepLoc to analysis where exactly the protein is compartmentalized in the cell, while for uniqueness the amino acid sequences of all the members of PKC family were retrieved and then submitted to a software MEGA-X to predict the phylogenetic tree of the protein and its uniqueness.

3.3. SNPs Identification

For the identification of potential damaging SNPs, data about all the missense SNPs of PKCι was taken from the ENST00000295797.5 transcript ID of PKCι found in the ensemble database. The data was then sorted and subjected to online tools like SIFT (Sim et al., 2012), PROVEAN (Choi & Chan, 2015), PolyPhen (Adzhubei et al., 2010), PHD-SNP, SNAP and Mutation Assessor (Reva, Antipin, & Sander, 2011). Each of these tools employs an algorithm which scores and categorizes the SNPs into different levels of harmfulness. SIFT scores from 0 – 1

with zero scoring SNP being the deleterious and one being the tolerated. PolyPhen score also range from 0 – 1 but classifies SNPs having score 0.9 and above as probably damaging, 0.4 – 0.8 as possibly damaging and below 0.4 as benign SNPs. While Mutation Assessor places them in four groups of high (0.9), medium (0.5 – 0.9), low (0.2 – 0.4), and neutral (0.0 – 0.1) SNPs. The data from these tools was analyzed and compared to see which potentially harmful SNPs were present among all of them.

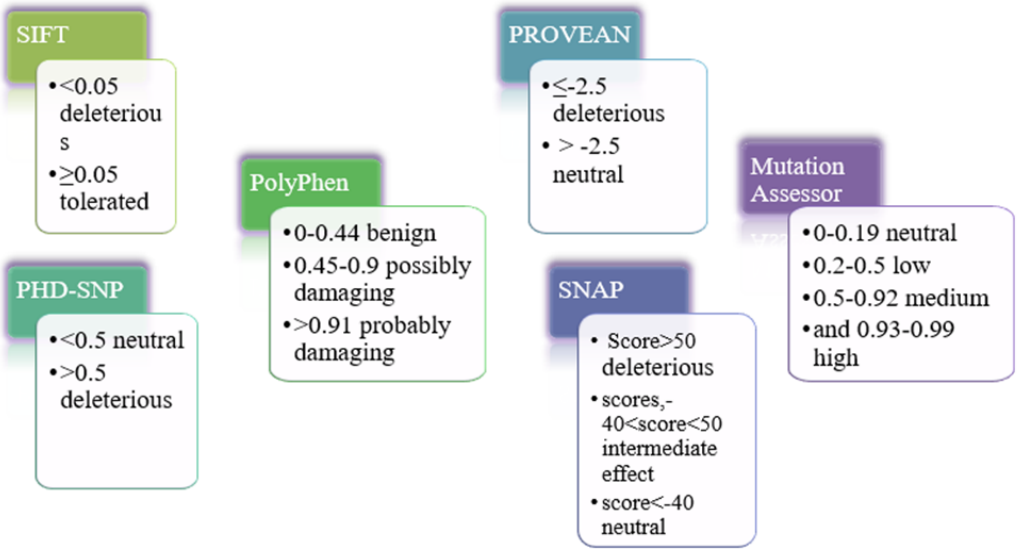

Figure 9. Tools and their selection criteria used for selection of deleterious missense SNPs

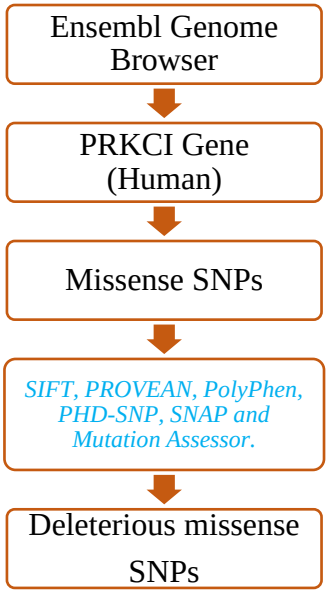

Figure 10. Flow sheet of the in-silico prediction of damaging missense SNPs in PRKCI gene.

3.4. SNPs Analysis

Online tools for analyzing the selected potentially harmful SNPs were also applied. These tools include I-Mutant, MutPred and HOPE. After entering the protein sequence and specifying the SNPs, I-Mutant gives its results in terms of DDG Value (DG (New Protein) – DG (Wild Type) in Kcal/mole) and RI (Reliability Index). The DDG value smaller than -0.5 indicates large decrease of stability while the DDG greater than 0.5 shows large increase of stability in the protein structure. If the value is greater than or equal to -0.5 or smaller than or equal to 0.5, it means neutral stability of the protein structure.

MutPred on the other hand generates the output as a general probability score which if greater than 0.5, indicates pathogenicity of the protein, in addition to the probability scores and P-values for the gain or loss of certain structural and functional properties

of the proteins. The P-value of 0.05 or smaller points to the significance of predicted change, rejecting the null hypothesis.

HOPE analyses and evaluates the SNPs with reference to the information available about the protein on other databases such as UniProt and generates a detailed report enlisting all the possible effects the respective SNP might have on the proteins structure and conserved properties.

3.5. *In silico* Mutagenesis

The mutated model for the PKCι protein (predicted by I-TASSER) parallel to the relevant amino acid substitution was made by Swiss-Pdb Viewer (v4.10). The amino acid in the wild sequence i.e., Glycine is replaced with the variant i.e., Tryptophan and the .pdb files were saved to be used in further *in silico* processes.

Figure 11. Glycine mutates to Tryptophan

3.6. Protocol for MD simulation

Molecular Dynamic simulation abbreviated as MD simulation is a computer simulation method to get insight about the properties of materials at atomic or molecular level. It is considered as a bridge between theory and experiment and is used to predict or verify experiments. In the current study lysozyme in water tutorial of GROMACS has been used to carry out MD simulations of the wild PRKCI protein structure and that of

mutagenesis induced structure on supercomputer. The interface used for giving the commands and running simulations is PuTTY while WinSCP was used for the transfer of files between the supercomputer and personal computer.

3.7. Protein Docking and its Visualization

Molecular docking is a computational technique which is used for the prediction of the most promising ligand-target spatial configuration. There are many tools available for docking but in the current study CB Dock was used. CB-Dock is a protein-ligand docking method which automatically identifies the binding sites, calculates the center and size, customizes the docking box size according to the query ligands and then perform the molecular docking with Auto Dock Vina. The molecule selected for docking is Decahydrocyclopentachrysenol (fig 4.) which is a triterpenoid. Not much research has been done on this compound particularly but the crude extract of the plants from which this drug is extracted has potential anti-cancer properties. In addition to that, the properties predicted by AdmetSAR gave significant insight about the molecule's properties and its pharmacokinetics in human body and suggest that it have therapeutic potential.

The best Protein structure predicted by I-TASSER was used as protein input in pdb format and Decahydrocyclopentachrysenol (fig.4) obtained from PubChem with SID 103580173 in SDF format was used as a ligand input file for docking using CB dock. CB dock predicted five interactions between the ligand and the protein. We selected the most stable one with the least vina score i.e., -9.6. Docked structure obtained from CB Dock was first prepared on PyMOL for visualization and then to see detailed molecular interactions between the ligand and protein LigPlot+ was used. LigPlot+ generated the detailed 2D ligand-protein interaction diagram depicting the protein cavity to which the

ligands were bind and different chemical bonds and chemical interactions. 3D visualization of protein-ligand interaction was done by PyMOL.

3.8. Primer Designing

Primers against the identified potential common damaging SNPs were designed using primer 1 software for further wet lab procedures.

3.9. Sample Collection

Hundred samples from the patients of Hepatocellular carcinoma were collected for carrying out the wet lab validation procedures and confirming our hypothesis about the presence of the missense SNPs in PRKCI gene of those patients.

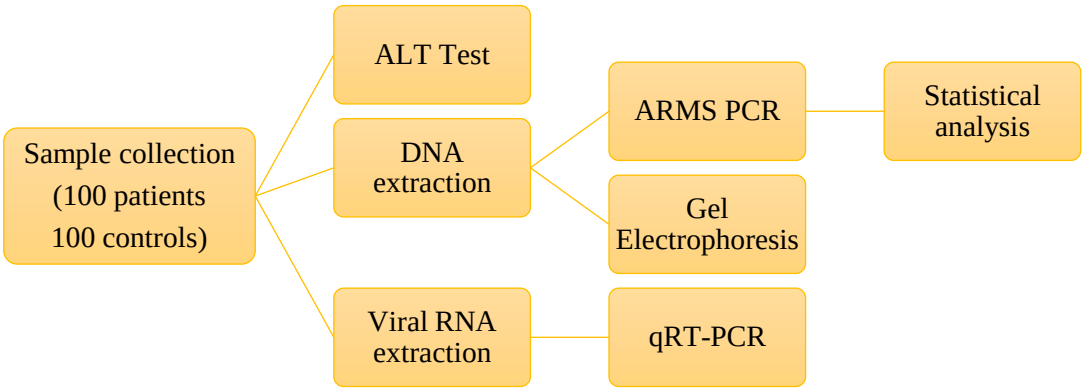

Figure 12. Schematic diagram of different invitro methods employed in this study

3.10. Genomic DNA Extraction from blood

DNA extraction is basically an analytical technique through which we isolate the DNA from the sample. DNA extraction from blood is important based on our needs and expected end results. The method used for extraction was phenol chloroform method. It is a two-day protocol.

Principal

DNA extraction involves many steps. The basic steps include the steps of disruption of cell membrane and nuclear membrane to release the DNA into solution followed by precipitation of DNA while ensuring removal of the contaminating biomolecules such as the proteins, polysaccharides, lipids, phenols, and other secondary metabolites. First, we need to separate the DNA from the cells. For separating the DNA, we need to lyse the cell membrane. For this purpose, we used extraction buffer i.e., SDS and EDTA. Both reagents have special roles such as EDTA is involved in removing Magnesium ions. These ions are present in the cell membrane and responsible for holding the structure of cell membrane. SDS is responsible for converting the DNA into linear form. It is also used for removing lipids present in cell membrane. Shaking by a vortex was carried out followed by adding the extraction buffer. Then it was subjected to centrifuge. All the cellular debris was separated as supernatant.

But many impurities are still present in the sample such as proteins. To remove them we used phenol and chloroform. These reagents can precipitate proteins. For DNA precipitation we used ethanol. After removing proteins, we were left with the RNA and DNA. RNA removal is achieved by using certain enzymes.

Apparatus

Following apparatus was used to extract DNA from blood

- Eppendorf tube
- Micropipettes
- Centrifuge
- Vortex machine

Chemicals

Following reagents were used in the extraction procedure:

- Patient's sample
- Ethanol (100%, 70%)
- Proteinase K
- Salt solution (3M sodium acetate)
- TE buffer 10X (PH 7.8)

The description of the solutions used in the extraction process along with their composition and functions are given below.

Table 1. Description of the solutions used in the extraction process along with their composition and functions

Solution	**Composition**	**Composition for 1 Liter**	**Function**
Solution-A	Sucrose (0.32M) 10mM Tris (7.5 pH) MgCl2 (5mM) Autoclave then add triton X-100(1% V/V)	109.44g 1.2114g 0.476g 10ml	It breaks the cell membrane allowing the release of DNA from the cell
Solution-B	10mM tris (pH7.5) NaCl (400mM) 2mM EDTA (pH 8.0)	1.2114g 23.37g 0.58g	It separates the proteins and precipitates DNA

Solution-C	Phenol	According to protocol	It separates the aqueous and organic phases in the solution clearly
Solution-D	Chloroform (24vol) Isoamyl-Alcohol (1vol)	48ml 2ml	It stabilizes the coagulation of proteins and reduces foaming so pure DNA can be isolated

Preparation of solutions

For the preparation of 500ml of solution A; 54.72g of sucrose (0.32mM), 0.238g of $MgCl_2$ (5mM) and 6.057g of Tris-base (10 mM) were added in nearly 450ml of distilled water. The pH of solution was then set to 7.5 – 8 using HCL (conc.) and NaOH (40% w/v solution). The volume was then brought up to a little less than 500ml using distilled water and the solution was autoclaved. 5ml of Triton X-100 (pure) was added after the solution was autoclaved because it is sensitive to heat.

Similarly, 500ml of solution B was also prepared. 11.685g of NaCl (400mM), 0.29g of EDTA (2mM) and 6.057g of Tris-base (10 mM) were added in nearly 450ml of distilled water. pH of the solution was again set to 7.5 – 8 in the same manner and the volume was brought up to 500ml before it was autoclaved. Solution C consisted of phenol only which is at risk of oxidation when exposed to light therefore, it was wrapped properly in the aluminum foil and carefully kept away from the light. 50ml of solution D was prepared with chloroform (48ml) and isoamyl alcohol (2ml).

Day 1 of DNA extraction Protocol

- Before starting the procedure of DNA extraction, the samples were taken out from 4 degree Celsius and incubated foe few minutes at room temperature.
- First 750µl of blood was added into 1.5ml centrifuge tube and 750µl of solution A was added in it.
- The tube is inverted 4-6 times to mix the blood sample with solution A and then incubated at room temperature for 5-10 min.
- The mixture was then centrifuge for 1 min at 13,000 rpm in a microcentrifuge.
- Supernatant is discarded and pallet was re-suspended in 750µl of solution A
- The mixture was centrifuge again at 13,000 rpm for 1 min. The supernatant was discarded, and pallet was re-suspended in 400µl of solution B (Dissolve the pallet completely by applying mechanical force through tapping repeatedly)
- Centrifuge the tube at 13000 rpm for 10 minutes.
- After centrifugation, discard supernatant and add 400 µl of Solution B, 12µl of 20% SDS (sodium dodecyl sulphate) and 5µl of proteinase K.
- Then mixture was subjected to incubator for overnight at 37 degrees Celsius.

Day 2 of DNA extraction Protocol

- Take 250 µl of solution C and 250 µl of solution D in a separate tube and add this in the tube containing sample (that was incubated overnight).
- The tubes were centrifuged at 13000 rpm for 10 min. aqueous phase collected into newly labeled Eppendorf tubes.
- 55µl of sodium acetate and 500µl of iso-propanol were added to precipitate DNA.

- Invert the tube several times and then centrifuge it at 13,000 rpm for 10 minutes.
- 200µl of 100% ethanol was added in tubes and centrifuge for 8 minutes at 13,000 rpm. Ethanol was discarded.
- DNA was dried by inverting the tubes on a clean surface for 30 minutes at room temperature. Once the tubes are dried add 200µl of TE buffer or PCR water to dissolve the DNA pallet.

3.11. Amplification Refractory Mutation System Polymerase Chain Reaction (ARMS-PCR)

The type of PCR which is used to identify presence of point mutations or SNPs in the genomic DNA is Amplification Refractory Mutation System Polymerase Chain Reaction (ARMS PCR). For this kind of PCR specific tetra primers, two internal primers and two external primers, against the selected SNP were used. For each sample 25µl of reaction mixture was prepared. This mixture consisted of 9µl of Solis Bio-Dyne Master Mix (12.5mM $MgCl_2$), 1µl of each of the four primers, 1µl of the DNA sample and 11µl of PCR water. After preparation, the reaction mixture was vortexed/centrifuged for a few seconds to mix it well and remove any bubble formed. PCR tubes containing reaction mixture were then placed in the Thermocycler. During the optimization process of the primers, multiple reactions were carried out together using the Gradient Thermocycler which was already programmed with a range of temperatures. Each PCR was composed of 3 steps and 35 cycles. Thermocycling conditions and steps are mentioned below.

Step 1: In stage1 the DNA template denaturation was allowed to occur at 95 degrees Celsius for 5 minutes. It was done to break the double stranded DNA into single strand. In stage 2 the denaturation of the daughter DNA took place. This stage had 35 cycles and each cycle took 30 seconds.

Step 2: Multiple annealing temperatures were provided in the gradient PCR and 30 seconds were allocated for this step too.

Step 3: Annealing of the primers was followed by their extension for 30 seconds at 72 degrees Celsius. The final extension was allowed to occur at 72 degrees Celsius for 7 minutes.

After completion of PCR, the tubes were stored in the freezer until running the PCR product on gel electrophoresis.

3.12. Gel Electrophoresis

Electrophoresis is a technique used to separate and sometimes purify macromolecules – especially proteins and nucleic acids – that differ in size, charge, or conformation. It is one of the most widely used techniques in biochemistry and molecular biology.

Principle:

The basic principle of this technique is to separate the DNA based on their charge to mass ratio. We ran our DNA sample/ PCR products in agarose gel. DNA is negatively charged and moves, under the force of an electric current, through the matrix of an agarose gel.

3.12.1. 1% (W/V) Agarose gel protocol

For visualization of genomic DNA 1% (w/v) Agarose gel was used. It was prepared by adding 1 gram agarose in 100ml of 1X TBE or TAE buffer. This mixture was then microwaved until the agarose completely dissolved in the buffer. 10µl of Ethidium Bromide was added in the mixture after cooling it down a little, for staining the DNA. Mixture was then poured in the gel casting tray and was kept at room temperature for solidification. After the gel was set, it was placed in buffer tank of electrophoresis apparatus. The 5 – 7µl of DNA sample combined with 1 – 2µl of loading dye were loaded in each well. Electrophoresis was performed at 80 V and 300 A for 30 – 40 minutes in 1X TBE or TAE buffer. DNA bands were analyzed by placing the gel in UV transilluminator or gel doc.

3.12.2. 2% (W/V) Agarose gel protocol

For visualization of PCR product 2% (w/v) agarose gel was prepared. The preparation procedure was the same as mentioned above except this time 2g of Agarose was dissolved in 100ml of buffer. When the gel was prepared, 10 – 12µl of the PCR product from individual tube was combined with 1 – 2µl of loading dye and then loaded in a well. Electrophoresis was performed at 80 V for 20 minutes in 1X TBE or TAE buffer. The gel was again analyzed by placing the gel in UV transilluminator or gel doc.

3.13. ALT/SGPT Test

ALT levels of the control and disease samples both were measured to get an idea about the change in liver function of the diseased individuals. Kit method was used for this purpose. The kit contained two solutions named R1 and R2. R1 contained Tris-buffer and Alanine amino acid while R2 contained Alpha ketoglutarate as a substrate, Lactate dehydrogenase enzyme and NADH cofactor. Blood samples were first centrifuged at 3000rpm for about 2 minutes to separate the serum from them. The working reagent was prepared by adding R1 and R2 in 4:1. For 1 sample 400µl of R1 and 100µl of R2 were taken and 50µl of sample was added in it. The reactions which take place after addition of samples are as follows:

$$\text{Alpha-ketoglutarate + Alanine} \xrightarrow{\text{ALT}} \text{Glutamate + Pyruvate}$$

$$\text{Pyruvate + NADH} \xrightarrow{\text{LDH}} \text{NAD}^{+} + \text{Lactate}$$

Lactate is the colorless end product, and its absorbance was measured at 340nm using a (UV-visible) spectrophotometer (Microlab 300 – Semi-automated chemistry analyzer). The probe of spectrophotometer dipped in the reaction solution takes up nearly 400µl of it and checks its absorbance/optical density and gives the concentration of ALT present. The formula for concentration is:

Concentration = ΔAbs/min * kit factor

Kit factor is 1746 for any absorbance to be checked at room temperature using 340nm wavelength. Normal ALT values lie below 40-41 IU/L.

3.14. Viral RNA Extraction

As the samples used were of HCV induced HCC patients, therefore the viral load of these HCV induced HCC blood samples was also measured to compare for comparison and association of SNPs with the severity of disease. For that, RNA of hepatitis C virus was extracted from the blood serum of the patients using kit method. For viral RNA extraction from one serum sample, 200µl of serum was taken in an Eppendorf and 500µl of VNE buffer was added. This mixture was vortexed for 5-7 seconds and 500µl of 75% ethanol was added and again vortexed for 5-7 seconds. It was then poured in spin column and centrifuged at 8000rpm for 1minute. The spin column was then taken out and the collection tube was discarded. RNA was retained on the filter of spin column which was then placed in a new collection tube and 500µl of wash buffer 1 was added into the spin column. The spin column was again centrifuged at 8000rpm for 1 minute and the collection tube was discarded. Second washing was done with 750µl of wash buffer 2 two times. In second washing, both times spin column was centrifuged at 14000rpm for 1 minute and collection tube was discarded. After washing the filter with trapped RNA was placed in a new collection tube and centrifuged for 1 minute at 14000rpm for complete drying of wash buffers from the filter and the collection tube discarded. Lastly, the filter was placed in a new Eppendorf and 50µl of RNAse free water was added in it. The Eppendorf with filter was centrifuged at 8000rpm for 1 minute leading to the dissolution of RNA in the RNAse free water. The dissolved RNA

was collected in the Eppendorf below. The filter was discarded and Eppendorf with RNA was saved at -20 °C for qRTPCR.

3.15. Quantitative Reverse Transcriptase Polymerase Chain Reaction (qRT-PCR)

This kind of PCR is used for the quantification of RNA. The RNA in sample is first transcribed by reverse transcriptase into cDNA which is then amplified and quantified at the same time through real time PCR. For real time quantification of the reaction SYBR® Green fluorescent dye is used which does not bind to ssDNA and only binds in the minor groove of dsDNA. When excited by the light source, the dye produces signal which is almost a hundred times stronger when it bound to the DNA molecule than in the unbound state. The change in fluorescence over the time is recorded against the cycle number and an amplification plot is generated which shows the progress of PCR. 44µl of reaction mixture which contained 43µl master mix and 1.2µl RT enzyme, was used for 6µl of each sample. At least two standards with known concentrations were used. The standards are used to determine the upper and lower limit of the concentrations obtained from the samples. The sample concentrations are compared with the standard curve and a linear graph is plotted which gives the starting quantity/copy number of the template molecule on x-axis against cycle threshold (CT) on y-axis. The sample which smaller CT value has more viral load than the one with larger CT value. Viral load is generated at the end of the reaction.

Chapter 4: Results

4.1. Predicted structure of PKCι

3D structures of the protein were predicted via I-TASSER (J. Yang et al., 2015) which is the most advance and reliable tool. It predicts protein structure based on multiple threading approach. The structure was then visualized through PYMOL (Lu, 2020). Different domains of the protein were identified from INTERPRO and highlighted by PYMOL. Protein PKCι contains around 4 important domains (Figure 14). Limon color represent PB1 domain (25-108 AA), PE/DAG domain (140-192 AA) is highlighted in magenta color, wheat color shows Protein kinase domain (254-522 AA) while blue color represents AGC kinase domain (523-594 AA).

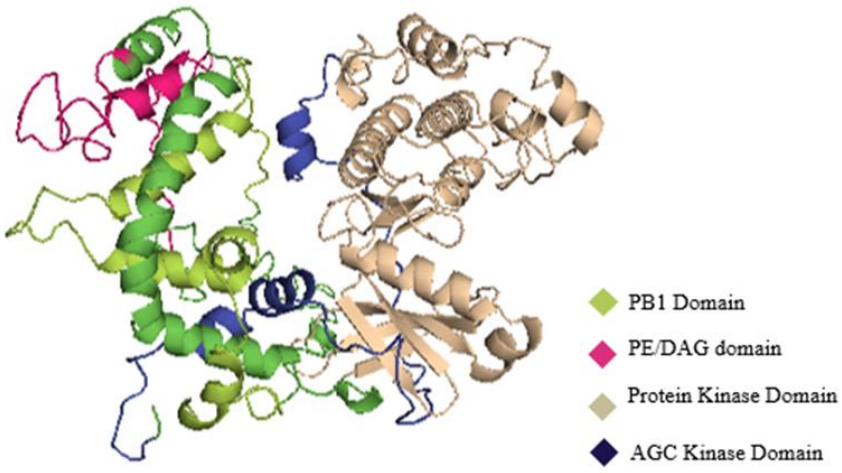

Figure 13. Predicted protein model of PRKCI. (Limon color: PB1 domain (25-108 AA), magenta color: PE/DAG domain (140-192 AA), wheat color: Protein kinase domain (254-522 AA), blue color: AGC kinase domain (523-594 AA))

4.2. Phylogenetic Tree and Subcellular Localization

The evolutionary relationship of the PKCs is depicted through the phylogenetic tree given in figure 14. It shows that all the members of this family of proteins originated

from a common ancestor protein. The branches originate from common branch points which join the closely linked members of the protein family. The scores on each branch represents the substitution per site, showing the extent of evolution of each protein from its other family members.

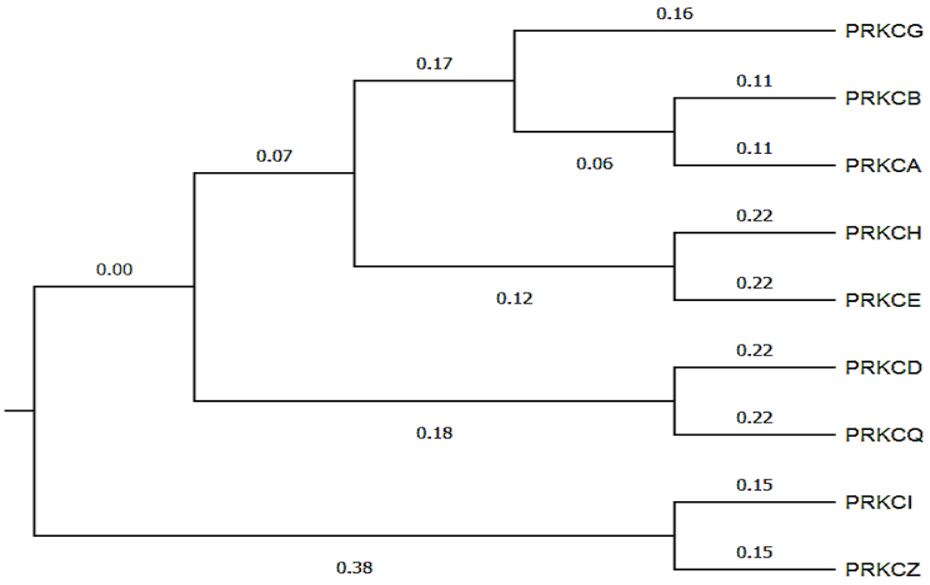

Figure 14. Phylogenetic tree of PKC protein family

The route of PKC gamma localization is given in figure 15. The red line represents the path taken by the protein to localize to its compartment within the cell. The score shows the probability/likelihood of the event. So, it is suggested based on the score that PKC Iota is localized in the cytoplasm.

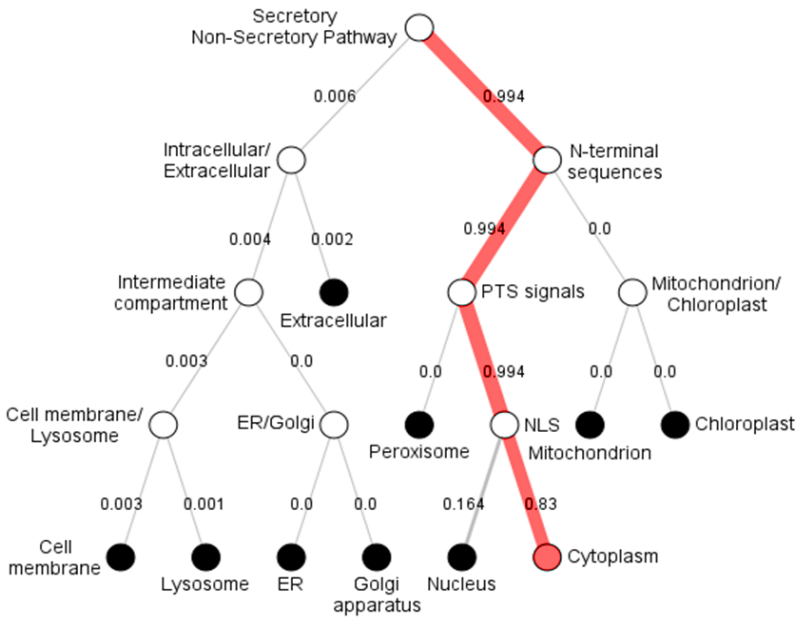

Figure 13. Subcellular localization of PKCι

4.3. Identified missense SNPs of PKCι

The Ensemble genome browser provides a total of 300 non-synonymous SNPs which were subjected to six different tools to analyze their impact on protein structure and function. Among the 300 nsSNPs 18 SNPs were detected to be pathogenic/deleterious after analyzing on five different tools i.e., SIFT, PROVEAN, PHD-SNP, SNAP, and Mutation Assessor. Table 2 represent the detail information of all the 18 SNPs that were predicted to be disease causing. Besides non-synonymous SNPs Ensemble genome browser also provides SNPs in intronic region,3 and 5 prime UTRs and synonymous SNPs etc. and their data are show in figure 16.

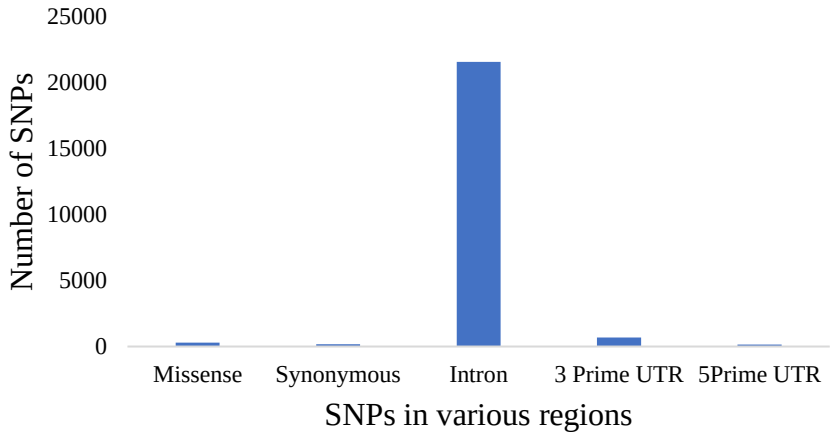

Figure 16. Graphical representation of PRKCI SNPs

Table 2. Deleterious non-synonymous SNPs in PRKCI gene using different computational tools (SIFT, PROVEAN, PHD-SNP, SNAP, and Mutation Assessor.).

Variant ID	**SIFT**	**PolyPhen**	**PROVEAN**	**PHD-SNP**	**SNAP**	**Mutation assessor**
rs1199520604	deleterious	probably damaging	Deleterious	deleterious	deleterious	medium
rs202068509	deleterious	possibly damaging	Deleterious	deleterious	deleterious	medium
rs1197750201	deleterious	probably damaging	Deleterious	neutral	deleterious	medium
rs1279890266	deleterious	possibly damaging	Deleterious	neutral	deleterious	Medium
rs567380681	deleterious	probably damaging	Deleterious	deleterious	deleterious	Medium

rs146841636	deleterious	probably damaging	Nil	neutral	neutral	Medium
rs56154494	deleterious	probably damaging	Deleterious	deleterious	deleterious	Medium
rs369872734	deleterious	probably damaging	Deleterious	neutral	neutral	Medium
rs1361108822	deleterious	probably damaging	Deleterious	deleterious	deleterious	Medium
rs1050315708	deleterious	probably damaging	Deleterious	deleterious	deleterious	Medium
rs185526558	deleterious	probably damaging	Deleterious	deleterious	deleterious	Medium
rs1379694217	deleterious	probably damaging	Deleterious	deleterious	deleterious	Medium
rs1373137084	deleterious	probably damaging	Deleterious	deleterious	deleterious	High
rs773463648	deleterious	probably damaging	Deleterious	deleterious	deleterious	Medium
rs1469931815	deleterious	possibly damaging	Deleterious	deleterious	deleterious	Medium

rs1577377456	deleterious	probably damaging	Nil	deleterious	deleterious	Medium
rs1189963071	deleterious	possibly damaging	Deleterious	deleterious	deleterious	High
rs1475798615	deleterious	probably damaging	Deleterious	deleterious	neutral	Medium

The above SNPs were further filtered out by applying another tool PolyPhen and only those SNPs were selected which were predicted to be damaging by all the six tools that were used and their data are given in table 3.

Table 3. Deleterious non-synonymous SNPs in PRKCI gene using different computational tools (SIFT, PROVEAN, PolyPhen, PHD-SNP, SNAP, and Mutation Assessor.).

Variant ID	**SIFT**	**PolyPhen**	**PROVEAN**	**PHD-SNP**	**SNAP**	**mutation assessor**
rs1199520604	deleterious	probably damaging	Deleterious	deleterious	deleterious	medium
rs202068509	deleterious	possibly damaging	Deleterious	deleterious	deleterious	medium
rs567380681	deleterious	probably damaging	Deleterious	deleterious	deleterious	medium
rs56154494	deleterious	probably damaging	Deleterious	deleterious	deleterious	medium

rs1361108822	deleterious	probably damaging	Deleterious	deleterious	deleterious	medium
rs1050315708	deleterious	probably damaging	Deleterious	deleterious	deleterious	medium
rs185526558	deleterious	probably damaging	Deleterious	deleterious	deleterious	medium
rs1379694217	deleterious	probably damaging	Deleterious	deleterious	deleterious	medium
rs1373137084	deleterious	probably damaging	Deleterious	deleterious	deleterious	high
rs773463648	deleterious	probably damaging	Deleterious	deleterious	deleterious	medium
rs1469931815	deleterious	possibly damaging	Deleterious	deleterious	deleterious	medium
rs1189963071	deleterious	possibly damaging	Deleterious	deleterious	deleterious	high

The SNP with variant ID rs ID1199530606 was selected for further *in silico* analysis and wet lab validation process.

4.4. Functional analysis of the selected SNP

The effect of nsSNPs on protein stability, structure and functions were analyze by three different tools i.e. I Mutant, MutPred and HOPE.

4.4.1. Analysis by I-Mutant

After submitting the protein sequence and specifying the above-mentioned SNP I-Mutant generated the data which demonstrated its effect on the stability of

overall protein structure. According to the data, RI of the given SNP ID is 6 and its DDG value is -0.30 Kcal/mol which shows decreased stability of the protein structure with the induction of this mutation (Capriotti, Fariselli, & Casadio, 2005).

4.4.2. Analysis by MutPred

MutPred analysis of Glycine with Tryptophan substitution at position 34 in PRKCI predict the following results:

a) Gain of catalytic residue at M37 (P = 0.0049)
b) Gain of sheet (P = 0.0149)
c) Gain of MoRF binding (P = 0.0416)
d) Loss of loop (P = 0.0512)
e) Gain of ubiquitination at K29 (P = 0.0635).

4.4.3. Analysis by HOPE

HOPE report gave the analysis of the given mutation as follows:

i. The mutant residue is bigger than the wild-type residue which might lead to bumps.
ii. Mutation is in the PB1 domain of the protein near a highly conserved position which is usually damaging for the protein (Venselaar, Te Beek, Kuipers, Hekkelman, & Vriend, 2010).

4.5. Molecular Dynamic Simulations Results analysis

Molecular dynamics simulations generated highly complex data after thousands to millions of steps. So, to extract useful information further analysis was required.

Some of the analytical tools used in this study for analysis of a typical short protein simulation of PRKCI are discussed below.

4.5.1. Root Mean Square deviation (RMSD)

RMSD stands for root mean square deviation. It numerically signifies the variation between a target structure and a reference structure. In molecular dynamics, we are less interested in the starting points and more in how structures and parts of structures change over time. For example, if we compute RMSD between two sets of atomic coordinates, the result will tell us how much the protein conformation has changed. RMSD values of both the wild PRKCI and mutated PRKCI were plotted against time in the given graph (Fig. 17). The increase in the RMSD plot with time shows the protein steadily deviates from its original conformation. Upon comparison, it shows that mutation causes slight decrease in stability of protein structure over time. The peaks evident in the graph indicates the occurrence of number of conformations which are accessed at some point in the trajectory

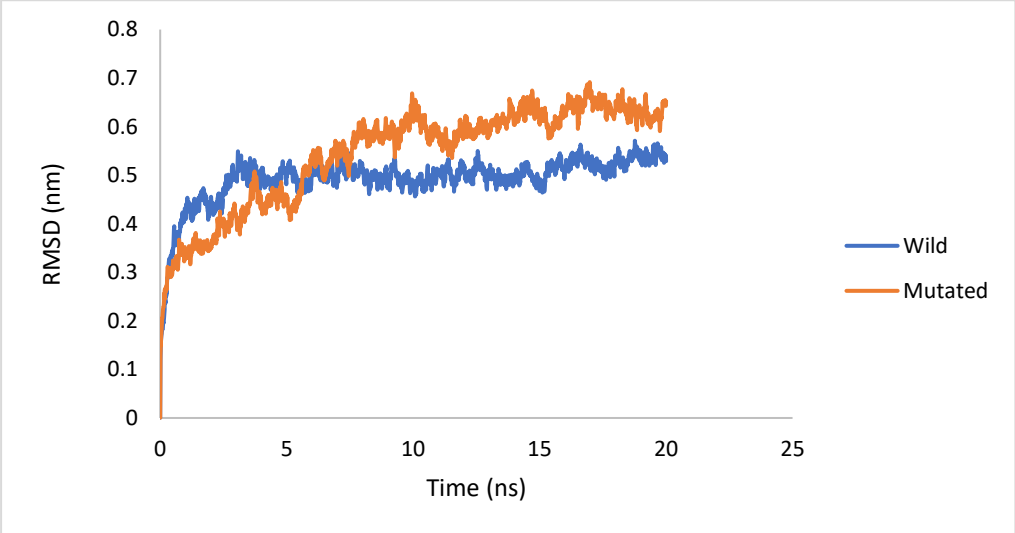

Figure 17. RMSD plot for PRKCI simulation

4.5.2. Root Mean Square Fluctuation (RMSF)

RMSF stands for root mean square fluctuation. It is a numerical measurement like RMSD but instead of showing positional difference in the entire structure over time, it gives the fluctuation of the individual residue over time during simulations. A graph of

RMSF values against residues for both the wild and mutated structure was plotted (Fig. 18). Comparison of the peaks shows that wild is more stable than mutated protein. Higher RMSF values at the end most likely are loop regions with more conformational flexibility, where the structure is not as well defined. Experimental spectroscopic techniques can be applied to detect the secondary structure of the protein.

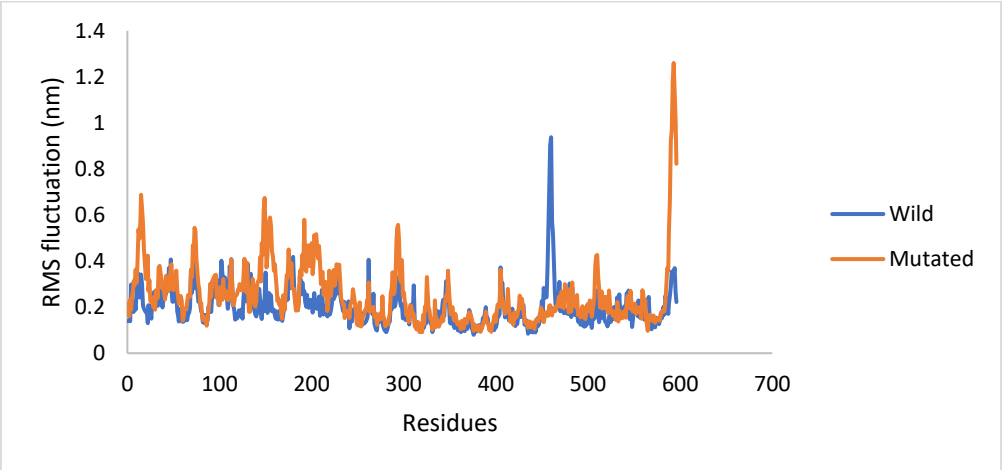

Figure 18. RMSF plot for PRKCI simulation

4.5.3. Number of Intra-protein Hydrogen Bonds

Hydrogen bond is the most important force and plays a critical role in maintaining the 2D or 3D structure of a protein. The increase or decrease in number of Hydrogen bonds over the time give an idea about weather a protein structure is stable or not. In figure 19, number of hydrogen bonds are plotted against time for 20 ns simulation. As we are analyzing a single mutation so there should not be a significant change in the number of hydrogen bonds. And as depicted in the graph there is no significant change in number of hydrogen bonds over the time which should be the case.

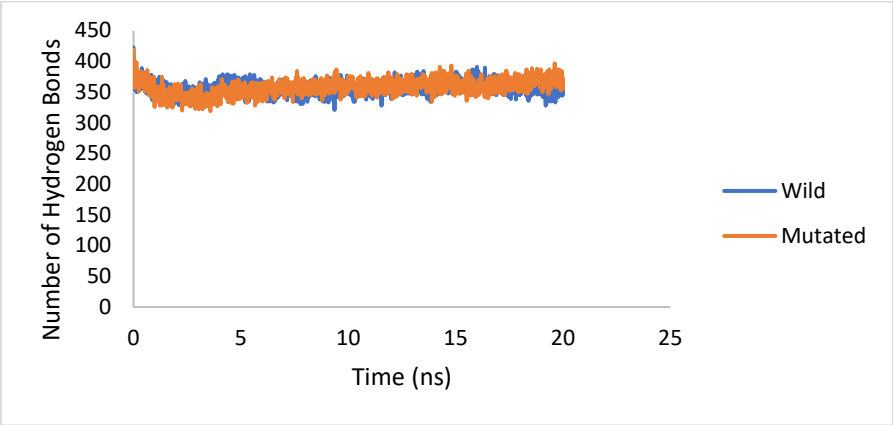

Figure 19. Intra-protein Hydrogen Bonds plot for PRKCI simulation

4.5.4. Radius of gyration

Radius of gyration is the distribution of atoms of a protein around its axis. Rg is calculated by measuring the distance from the point of rotation to the point where the effect of energy transfer is maximum. Radius of gyration gives us insight into the compactness of a protein structure (Garg et al.). In the current study radius of gyration was plotted as a function of time for 20ns simulation. From the figure 20 it is revealed that the Rg for both the wild and mutated structures was 2.64nm at 0ns. The Rg for both the structure changes between 0-5ns and 15-20ns only.

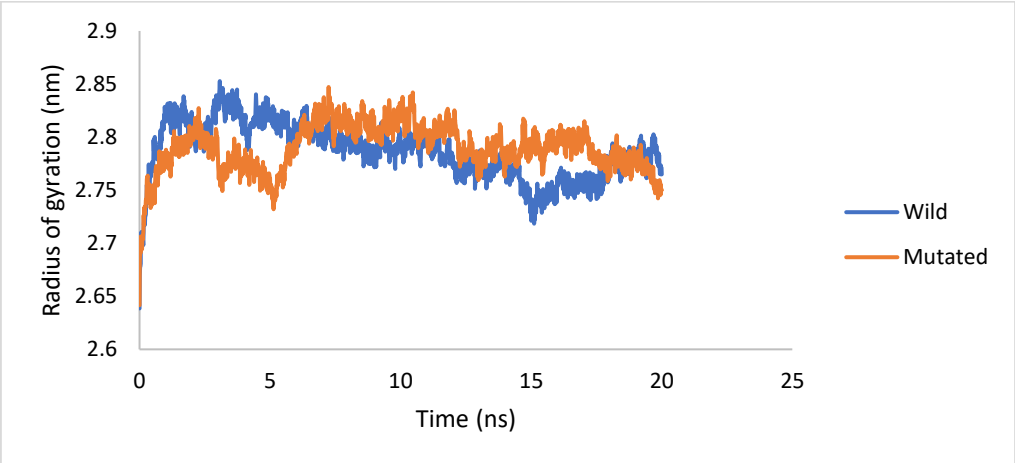

Figure 20. Radius of gyration plot for PRKCI simulation

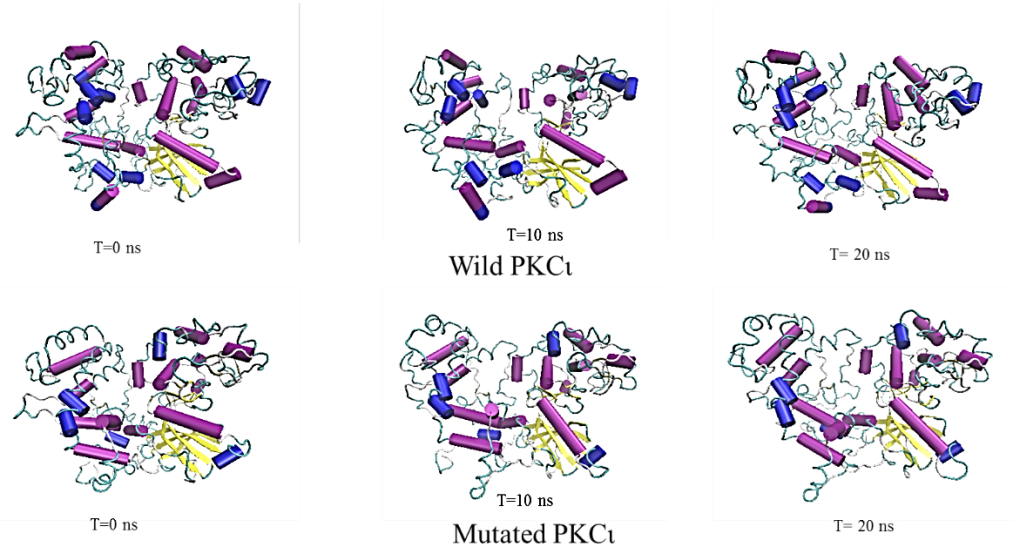

Figure 21. Snapshots of wild and Mutant PKCι Simulations at different time points

4.6. Decahydrocyclopentachrysenol

The ligand selected for docking is Decahydrocyclopentachrysenol (SID: 103580173) and its structure obtained from PubChem is given in the figure 22.

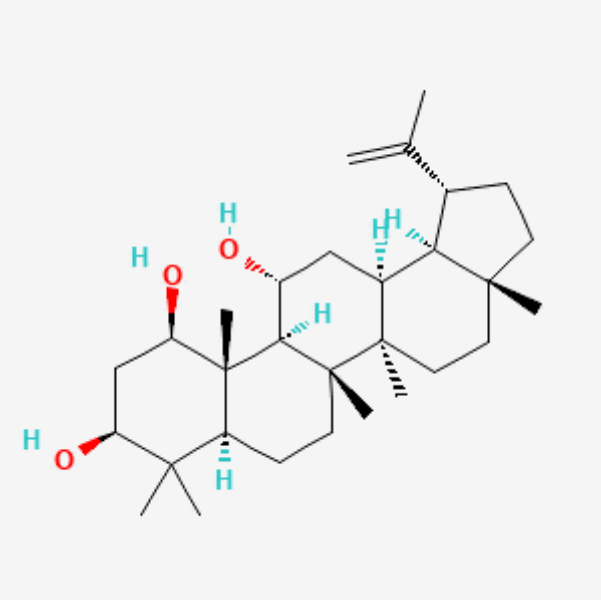

Figure 22. Chemical structure of Decahydrocyclopentachrysenol obtained from PubChem (Source: National Center for Biotechnology Information (2021). PubChem Compound Summary for CID 44559187. Retrieved July 31, 2021, from https://pubchem.ncbi.nlm.nih.gov/compound/44559187).

4.6.1. AdmetSAR Analysis of Decahydrocyclopentachrysenol

AdmetSAR analysis of the ligand (Decahydrocyclopentachrysenol) shows the physio-chemical properties of the ligand. The AdmetSAR report also provides regression properties, and ADMET (Absorption, Distribution, Metabolism, Excretion & Toxicity) properties of the drug

4.6.2. Lipinski's Rule of Five

The Lipinski's Rule of five shows the physio-chemical properties of the drug i.e., molecular weight, lipophilicity, hydrogen bond acceptor, hydrogen bond donor and rotatable bonds present in the drug. The drug is considered an optimum if its molecular weight is less than 500 Dalton, hydrogen bonds acceptor is less than 10, less than 5 hydrogen bond acceptor and AlogP less than 5 (C. A. Lipinski, Lombardo, Dominy, & Feeney, 2001). The properties predicted by AdmetSAR for Decahydrocyclopentachrysenol, and their values are given in table 4.

Table 4. Physio-chemical parameters of Decahydrocyclopentachrysenol estimated through admetSAR

Lipinski's Rule of Five and Veber's rule		
S.NO	**Characteristics**	**Values**
1	Molecular Weight	458.73
2	AlogP	5.97
3	H-Bond Acceptor	3
4	H-Bond Donor	3
5	Rotatable Bonds	1

4.6.3. Regression Analysis

Regression analysis shows that how the drug interacts with biomolecules of the body. AdmetSAR report shows the predicted properties i.e., solubility, binding to plasma membrane oral toxicity, etc. of the drug along with their values and units. Table 5 shows the detail of regression properties.

Table 5. Regression Properties of Decahydrocyclopentachrysenol molecule estimated through AdmetSAR

Admet predicted profile --- Regression Properties			
S.NO	**Properties**	**Values**	**Unit**
1	Water Solubility	-3.998	logS
2	Plasma Protein Binding	1.081	100%
3	Acute Oral Toxicity	3.278	mol/kg
4.	Tetrahymena pyriformis	1.034	pIGC50 (ug/L)

4.6.4. ADMET Properties

AdmetSAR provides the pharmacokinetic properties i.e., absorption, distribution, metabolism, excretion, and toxicity of Decahydrocyclopentachrysenol and predicted that it does not cross blood brain barrier, is not carcinogenic, and its oral toxicity falls in category I etc. The detail of these properties is given in table 6.

Table 6. Pharmacokinetic Properties of Decahydrocyclopentachrysenol molecule estimated through AdmetSAR

Absorption		
Blood-Brain Barrier	BBB-	0.2861
Human Intestinal Absorption	HIA+	0.9929

P-glycoprotein Substrate	Substrate	0.7709
P-glycoprotein Inhibitor	Non-inhibitor	0.6215
	Non-inhibitor	0.8559
Distribution		
Subcellular localization	Lysosome	0.6469
Metabolism		
CYP450 2C9 Substrate	Non-substrate	0.8253
CYP450 2D6 Substrate	Non-substrate	0.8926
CYP450 1A2 Inhibitor	Non-inhibitor	0.8120
CYP450 2C9 Inhibitor	Non-inhibitor	0.8284
CYP Inhibitory Promiscuity	Low CYP Inhibitory Promiscuity	0.6948
Toxicity		
Human Ether-a-go-go-Related Gene Inhibition	Weak inhibitor	0.9069
Carcinogens	Non-carcinogens	0.9054
Acute Oral Toxicity	I	0.8567

4.7. Interpretation of Docked Protein-ligand interaction

Among the five different protein-ligand interactions predicted by CB dock (given in table), the most stable one with the least vina score i.e., -9.5 (highlighted in red) was selected and the structure is given in figure 23.

Table 7. Vina scores and cavity sizes of CB Dock predicted possible Protein-ligand interactions

S. No	Vina Score	Cavity size
01.	-9.6	4377
02.	-8.4	1338
03.	-7.2	677
04.	-6.9	1570
05.	-6.4	754

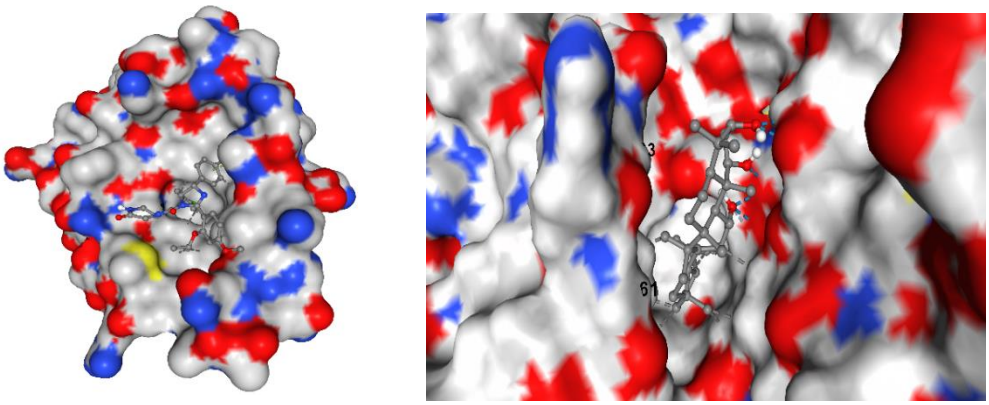

Figure 23. Surface view of PKCι and Decahydrocyclopentachrysenol docked inside its binding cavity of size 4377 (left) Close up picture of docked Nepeticin inside the binding cavity of PKCι (right).

The different kinds of molecular interactions of amino acids in the protein binding cavity with the ligand were visualized through LigPlot+ (fig.24). These interactions showed that the ligand was embedded deep in the hydrophobic pocket of the protein and formed strong hydrophobic interactions with it. The residues which make hydrophobic interactions include Arg 262, Leu385, Val268, Val335, Ala281, Ile332, Lys283, Thr395, val316 and Asn396. These residues are making hydrophobic interactions with carbon or oxygen atoms. On the other hand, Lys380, Asp 382 and Asn

383 are the residues which interacts through Hydrogen bonds with the ligand. The bond lengths of Hydrogen bonds are shown in dotted lines and the values are given in Angstrom(A). Fig.24 and fig.25 represent the 2D and 3D interaction of receptor and ligand.

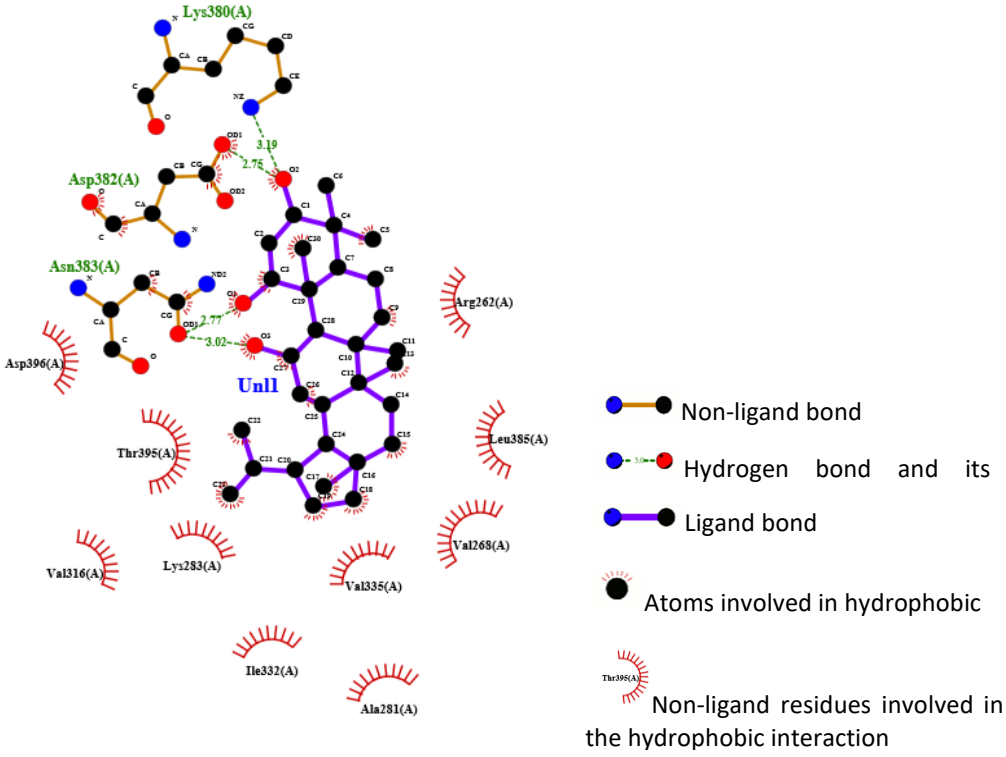

Figure 24. LigPlot visualization of docked Decahydrocyclopentachrysenol in the binding pocket of PKCι. The black and purple ball and stick structure is the ligand, Decahydrocyclopentachrysenol, and the red spiked semi-circles surrounding it are the residues present in the binding pocket forming hydrophobic interactions with the ligand while the green dotted line gives the H-bond formed between the binding pocket with the ligand.

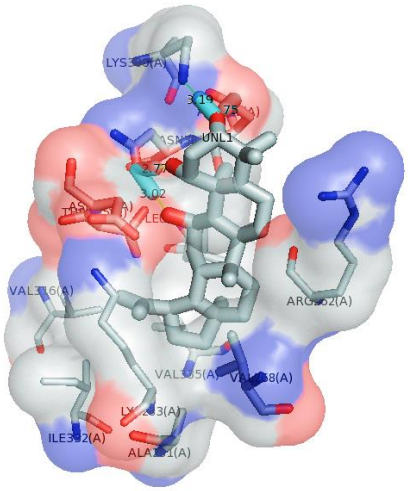

Figure 25. 3D representation of protein-ligand interaction via PyMOL

4.8. Analysis of Genotype Data of HCC and Control Samples

The analysis of PRKCI polymorphism in HCC patients and control were done based on odd ratio and relative risk. The data revealed that the SNP in homozygous mutated form i.e., TT with P-value <0.0001 and odd ratio and relative risk value 4.134 and 2.012 respectively, could be significantly associated with the disease than the heterozygous GT genotype and homozygous wild genotype GG. This shows that polymorphism in G allele increase the risk of disease occurrence.

Table 8: Comparison of PRKCI polymorphism in HCC patients and control

Genotype	Frequency distribution		Odds Ratio		Relative risk		P-value
	Patients%	Control%	Value	95% CI	Value	95% CI	
GG	23.00%	36.00%	0.5310	0.2834 to 1.000	0.7138	0.4912 to 0.9914	0.0623
TT	65.00%	31.00%	4.134	2.247 to 7.278	2.012	1.501 to 2.750	<0.0001

GT	12.00%	33.00%	0.2769	0.1349 to 0.5600	0.4697	0.2767 to 0.7418	0.0006

4.8.1. Data Based on gender

A similar Comparison of PRKCI polymorphism but with respect to gender in HCC patients and controls were also carried out (Table 9). The data that was obtained support the results mentioned in table 1. In both male and females, the mutated homozygous allele TT were found to be associated with the disease. The P-value for male and female having allele TT was 0.0006 and 0.0015 respectively depicting that the risk in males is slightly greater than in females. This difference could also be due to the relatively greater incidence rate of liver cancer, more specifically HCC in males than in females

Table 9: Comparison of PRKCI polymorphism in HCC patients and control with respect to gender

Genotype	**Frequency distribution**		**Odds Ratio**		**Relative risk**		**P-value**
	Patients%	**Control%**	**Value**	**95%CI**	**Value**	**95%CI**	
GG (F)	26.00%	40.74%	0.5111	0.2130 to 1.223	0.6927	0.4145 to 1.082	0.1466
TT (F)	64.00%	31.48%	3.869	1.701 to 8.831	1.995	1.319 to 3.116	0.0015
GT (F)	10.00%	27.78%	0.2889	0.1088 to 0.8573	0.4667	0.2046 to 0.9125	0.0260

GG (M)	20.00%	30.43%	0.5714	0.2282 to 1.398	0.7500	0.4262 to 1.182	0.2505
TT (M)	66.00%	30.43%	4.437	1.802 to 10.01	2.024	1.350 to 3.164	0.0006
GT (M)	14.00%	39.13%	0.2532	0.1011 to 0.6524	0.4623	0.2307 to 0.8207	0.0058

4.9. Analysis Of ALT In Patient vs Control

The ALT level of patients were significantly higher when compared with the control sample. The average concentration of ALT in patients were 106.84 IU/L which were significantly higher from the normal ALT range which is less than 41 IU/L. In the control sample ALT level was under 41 IU/L. Figure 26 represent the comparison of ALT concentration in patient vs control.

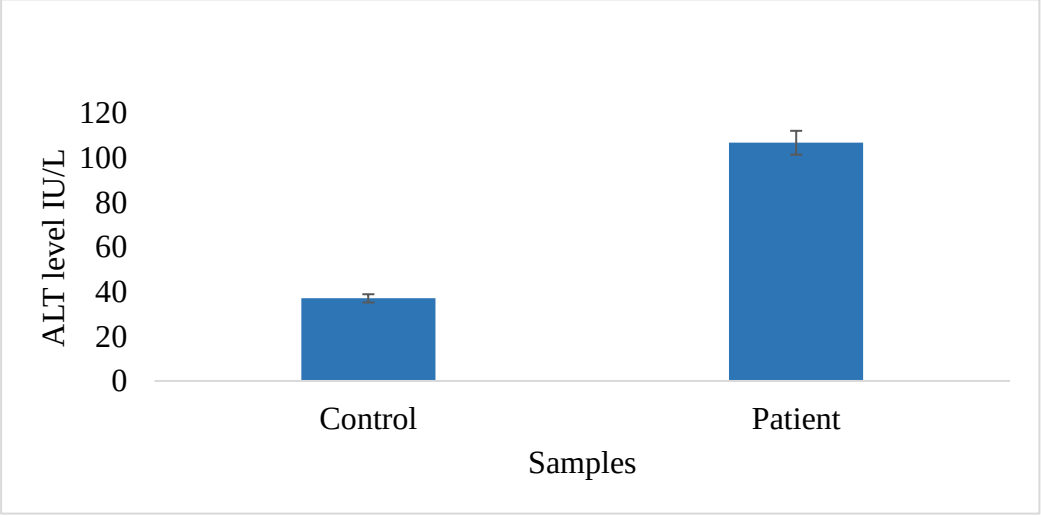

Figure 26: Comparison of ALT concentration in HCC patients and control

4.10. Viral Load Analysis of Patients with Different Alleles

Viral Load in patients was measured by qRT-PCR. The analysis of viral load against genotype shows significant difference. Average of viral load for genotype GG, TT and GT was taken and then analyzed which shows that patients with GG allele has significantly high (560095306.7 copies/ml) viral load compared to patients with TT and GT alleles given in figure 27 which suggests that there may be a correlation of viral load and genotype**.**

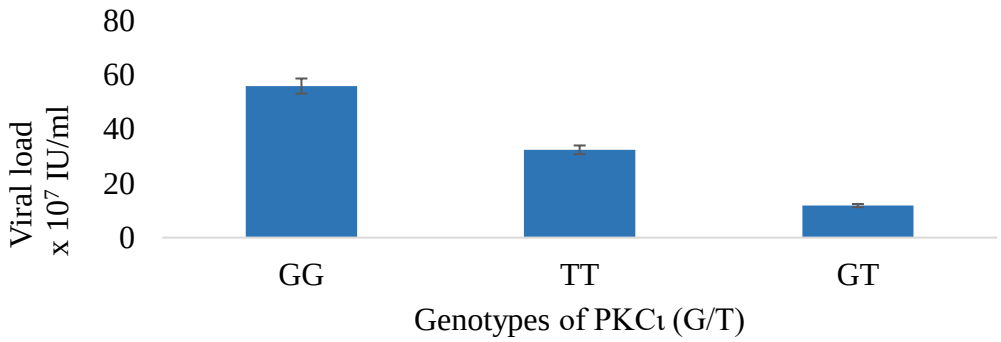

Figure 27: HCV Viral load plotted against homozygous wild (GG), heterozygous (GT) and homozygous mutated (TT) genotypes of HCC patients.

Chapter 5: Discussion

This study predicted protein model for PRKCI gene, offers insight into its nsSNPs and their impact on the structure and function of this protein. For understanding the function of protein, it is important to have knowledge of protein three-dimensional structure or tertiary structure (3D). 3D structures of the protein PKCι was predicted via I-TASSER which is the most advance and reliable tool and has extensively been used previously (Zhang, 2009). It predicts protein structure based on multiple threading approach. In this study I-TASSER predicted five models and each model is assigned an individual confidence score (C-Score). The higher the C-Score, the greater the confidence and vice-versa. The best model among the five models was chosen based on c-score (-3.07) and InterPro prediction (Blum et al., 2021).

Major problem with most cancers is the detection of the disease in early stages and HCC is no more different. The change in protein expression and its structural and functional properties has often been associated to the presence of certain kinds of polymorphisms in the genetic sequence (Decock et al., 2007; Fu et al., 2010; Robert & Pelletier, 2018); however not much is known about the role of PRKCI variations in relation to natural polymorphism. It is, therefore, crucial to identify deleterious nsSNPs in the PRKCI gene, as nsSNPs cause the most significant damaging impact on the protein structure and function (Stehr et al., 2011). The current study only considered missense variants because of their direct involvement in disease pathologies and their impacts on the adopted treatment regimen So, in this study single nucleotide polymorphism in PRKCI was correlated with HCC for the purpose of

identification of new prognostic marker for HCC. To analyze the effect of a large number of nsSNPs, using bioinformatics tools is a cost-efficient option, which allows identification of damaging SNPs with functional importance and with detailed analysis of an individual's disease susceptibility followed by the investigation of the structural basis of disease-causing mutations The retrieved non-synonymous SNPs from Ensemble Genome Browser were subjected to different tools i.e., SIFT, PROVEAN, PolyPhen, PHD-SNP, SNAP, and Mutation Assessor to predict the impact of single nucleotide polymorphism on protein. The selected SNP (rs ID1199530606) when sifted through all these tools were the most deleterious SNP, supported by at least five out of six tools. This SNP (rs ID1199530606) falls within the PB1 domain (which is exclusive to atypical PKCs) (Parker et al., 2014) of PKCι. Glycine is changing to Tryptophan at this position. The PB1 domain facilitates protein-protein interactions between PKCι and other proteins having PB1 domain such as Par-6 (partitioning-defective 6) (Joberty, Petersen, Gao, & Macara, 2000; Lin et al., 2000; Noda et al., 2001; Qiu, Abo, & Steven Martin, 2000), ZIP/p62 (Hirano et al., 2004; Puls, Schmidt, Grawe, & Stabel, 1997), and MEK5 (MAPK (mitogen-activated protein kinase)/ERK (extracellular-signal-regulated kinase) kinase 5) (Diaz-Meco & Moscat, 2001; Hirano et al., 2004). PB1 domain is near a highly conserved region which makes it a striking therapeutic target for the treatment of cancer (Parker et al., 2014). Further structural and functional analysis of the selected SNP (rs ID1199530606) was performed via I-Mutant, HOPE and MutPred. The results from these tools suggested that with the induction of the selected point mutation, which is in the PB1 regulatory domain of PKCι, the stability is likely to decrease along with Gain of catalytic residue at M37, Gain of sheet, Gain of MoRF binding, Loss of loop and Gain of ubiquitination at K29. All these predictions implicate that the impact of

single amino acid change is not restricted to that residue but can also affect the function of other residues.

MD simulations were carried out to get insights into the magnitudes of the damaging effects as it has previously used in many studies for the functional analysis of SNPs (Chitrala & Yeguvapalli, 2014). The results from the MD simulation showed that G into W mutation in PKCι induced random motions, more coil and bending conformations and loss of a few helixes, indicating significant structural alternation. These results showed that the impact of single amino acid change may not be restricted to that specific position and may interfere with amino acids at other positions. Higher RMSF values at the end most likely are loop regions with more conformational flexibility, where the structure is not well defined. Experimental spectroscopic techniques can be applied to detect the secondary structure of the protein.

The wet lab results confirmed the insilico prediction. The analysis of PCR results after amplifying the SNP in nearly 200 samples (100 controls and 100 HCV induced HCC patients) and applying statistical tools using GraphPad Prism, it was observed that the homozygous mutant allele TT has significant correlation (odd ratio (4.134), relative risk (2.012) and P-value <0.0001) with HCC compared to homozygous GG and Heterozygous GT. The analysis of PCR data with respect to gender also indicated the association of the same genotype in both males and females, although its relative risk and odds ratio in males is slightly greater than in females. This difference could also be due to the relatively greater incidence rate of liver cancer, more specifically HCC in males than in females (Hafeez, Mahmood, Khan, & Malkani, 2020; Singal, Lampertico, & Nahon, 2020). The association of polymorphism with genetic disease give us an idea about susceptibility and can also be used for early diagnosis (Kamaraj

& Bogaerts, 2015; Kamaraj & Purohit, 2013; Karchin, 2009). So, this SNP (rs ID1199530606) could be potential biomarker for prognosis of HCC.

It has been established that high alanine aminotransferase (ALT) level is linked with hepatitis C virus (HCV) induced hepatocellular carcinoma (HCC) and can lead to the disease rapidly (Tarao et al., 1999). So, the ALT test was performed for both patients and control and the results revealed significantly increased ALT level of HCC patients (106 IU/L).

The association of viral load with genotype was performed to analyze the link of genotype with viral load. Our results demonstrated that patients with genotype GG have high viral load followed by TT and then GT. It might be due to clearance of viral load over time as literature shows clearance of HCV viral RNA in patients coinfected with HCV/HIC having rs12979860 polymorphism CC genotype (Lapiński, Pogorzelska, Kowalczuk, Nikliński, & Flisiak, 2013).

The association of PKCι variant with hepatocellular carcinoma open new direction for therapeutics. A potential drug, Decahydrocyclopentachrysenediol, against PKCι was selected on the basis of admetSAR predictions. The predictions for Lipinski's rule of 5 indicated Decahydrocyclopentachrysenediol as promising drug candidate. Lipinski's rule of 5 gives standard values of molecular mass less than 500 Da, H-bond donors no more than 5, H-bond acceptors no more than 10 and logP (octanal-water partition coefficient) up to 5. Decahydrocyclopentachrysenediol has all these values within the given range. Besides that, Admet profile of drugs also include prediction of drug behavior against 4 regression models which include logS (water solubility), plasma protein binding, acute oral toxicity and PIGC50 in Tetrahymena pyriformis

model organism. Water solubility of Decahydrocyclopentachrysenediol is -3.998 which is less than 1, thus indicating its hydrophobic properties. Besides that, 1.081 out of 100% molecules of Decahydrocyclopentachrysenediol can bind to plasma proteins and it can cause oral toxicity at the dose of 3.278 mol/kg. The concentration which can cause 50% growth inhibition of Tetrahymena pyriformis is 1.034 mol/kg. The pharmacokinetic properties of Decahydrocyclopentachrysenediol predicted by admetSAR indicate its subcellular localization in lysosomes and human intestinal absorption with 0.9930 probability. Apart from that, it is not hepatotoxic, biodegradable, not orally available and does not cross blood brain barrier. All these properties indicate that Decahydrocyclopentachrysenediol is a promising drug candidate. The crude extract of the plant from which Decahydrocyclopentachrysenediol is extracted has already been studied and tested for its antitumor, apoptosis inducing properties against HCC and other cancers (Wang et al., 2019; Yin et al., 2019). Therefore, molecular docking of Decahydrocyclopentachrysenediol with PKCι was performed to see its interaction with it. It showed good interaction with multiple hydrophobic interactions and a H-bond in the binding pocket of PKCι with vina score of -9.6. Docking of this molecule in PKCι was not specific for the mutated or wild type structure which keeps an open window for potential drug against this protein irrespective of its mutations.

Conclusion

The identified pathogenic SNP (rs1199520604) of PKCι through online tools showed strong association with hepatocellular carcinoma. So, it may be used for prognosis of hepatocellular carcinoma. The structure of PKCι was predicted, and structural and functional analysis shows that mutation at position 34 (G34W) caused the overall stability of PKCι to decrease. Correlation of genotypes with viral load give insight that there may be an association of genotype with viral load. Decahydrocyclopentachrysenediol-PKCι docking was performed, and the results proposed that Decahydrocyclopentachrysenediol can be a potential inhibitor of PKCι, and the results needs further invitro and in-vivo validation. The expression profile of PKCι upon G34W mutation needs to be explored further that might open new ways in cancer therapeutic field. Further Invitro and in-vivo studies are required to ascertain effects of the variants in native protein.

References

Yang, J., & Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. *Nucleic acids research, 43*(W1), W174-W181.

Zhang, C., Freddolino, P. L., & Zhang, Y. (2017). COFACTOR: improved protein function prediction by combining structure, sequence and protein–protein interaction information. *Nucleic acids research, 45*(W1), W291-W299.

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet, 46(3), 310-315. https://doi.org/10.1038/ng.2892

Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Research, 40(W1), W452-W457. https://doi.org/10.1093/nar/gks539

Choi, Y., & Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics, 31(16), 2745-2747. https://doi.org/10.1093/bioinformatics/btv195

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., Cannon-Albright, L. A., Teerlink, C. C., Stanford, J. L., Isaacs, W. B., Xu, J., Cooney, K. A., Lange, E. M., Schleutker, J., Carpten, J. D., Powell, I. J., Cussenot, O., Cancel-Tassin, G., Giles, G. G., MacInnis, R. J., Maier, C., Hsieh, C. L., Wiklund, F., Catalona, W. J., Foulkes, W. D., Mandal, D., Eeles, R. A., Kote-Jarai, Z., Bustamante, C. D., Schaid, D. J., Hastie, T., Ostrander, E. A., Bailey-Wilson, J. E., Radivojac, P.,

Thibodeau, S. N., Whittemore, A. S., & Sieh, W. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet, 99(4), 877-885. https://doi.org/10.1016/j.ajhg.2016.08.016

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. Nature Methods, 7(4), 248-249. https://doi.org/10.1038/nmeth0410-248

Reva, B., Antipin, Y., & Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res, 39(17), e118. https://doi.org/10.1093/nar/gkr407

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols, 5(4), 725-738. https://doi.org/10.1038/nprot.2010.5

Blum, M., Chang, H. Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., Richardson, L., Salazar, G. A., Williams, L., Bork, P., Bridge, A., Gough, J., Haft, D. H., Letunic, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Necci, M., Orengo, C. A., Pandurangan, A. P., Rivoire, C., Sigrist, C. J. A., Sillitoe, I., Thanki, N., Thomas, P. D., Tosatto, S. C. E., Wu, C. H., Bateman, A., & Finn, R. D. (2021). The InterPro protein families and domains database: 20 years on. Nucleic Acids Res, 49(D1), D344-d354. https://doi.org/10.1093/nar/gkaa977

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality

worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, *68*(6), 394-424. doi:10.3322/caac.21492

Capriotti, E., & Altman, R. B. (2011). Improving the prediction of disease-related variants using protein three-dimensional structure. *BMC Bioinformatics*, *12*(4), S3. doi:10.1186/1471-2105-12-S4-S3

Clements, O., Eliahoo, J., Kim, J. U., Taylor-Robinson, S. D., & Khan, S. A. (2020). Risk factors for intrahepatic and extrahepatic cholangiocarcinoma: A systematic review and meta-analysis. *J Hepatol*, *72*(1), 95-103. doi:10.1016/j.jhep.2019.09.007

Collins, F. S., Guyer, M. S., & Charkravarti, A. (1997). Variations on a theme: cataloging human DNA sequence variation. *Science*, *278*(5343), 1580-1581. doi:10.1126/science.278.5343.1580

Deng, N., Zhou, H., Fan, H., & Yuan, Y. (2017). Single nucleotide polymorphisms and cancer susceptibility. *Oncotarget*, *8*(66), 110635-110649. doi:10.18632/oncotarget.22372

Du, G. S., Wang, J. M., Lu, J. X., Li, Q., Ma, C. Q., Du, J. T., & Zou, S. Q. (2009). Expression of P-aPKC-iota, E-cadherin, and beta-catenin related to invasion and metastasis in hepatocellular carcinoma. *Ann Surg Oncol*, *16*(6), 1578-1586. doi:10.1245/s10434-009-0423-7

Ferlay, J., Ervik, M., Lam, F., Colombet, M., Mery, L., Piñeros, M., . . . Bray, F. (2018). Global cancer observatory: cancer today. *Lyon, France: International Agency for Research on Cancer*.

Joberty, G., Petersen, C., Gao, L., & Macara, I. G. (2000). The cell-polarity protein Par6 links Par3 and atypical protein kinase C to Cdc42. *Nat Cell Biol, 2*(8), 531-539. doi:10.1038/35019573

Kamaraj, B., & Bogaerts, A. (2015). Structure and Function of p53-DNA Complexes with Inactivation and Rescue Mutations: A Molecular Dynamics Simulation Study. *PLOS ONE, 10*(8), e0134638. doi:10.1371/journal.pone.0134638

Kamaraj, B., & Purohit, R. (2013). In Silico Screening and Molecular Dynamics Simulation of Disease-Associated nsSNP in TYRP1 Gene and Its Structural Consequences in OCA3. *BioMed Research International, 2013*, 697051. doi:10.1155/2013/697051

Karchin, R. (2009). Next generation tools for the annotation of human SNPs. *Brief Bioinform, 10*(1), 35-52. doi:10.1093/bib/bbn047

Khan, S. A., Tavolari, S., & Brandi, G. (2019). Cholangiocarcinoma: Epidemiology and risk factors. *Liver Int, 39 Suppl 1*, 19-31. doi:10.1111/liv.14095

Kulik, L., & El-Serag, H. B. (2019). Epidemiology and Management of Hepatocellular Carcinoma. *Gastroenterology, 156*(2), 477-491.e471. doi:10.1053/j.gastro.2018.08.065

Petrick, J. L., & McGlynn, K. A. (2019). The changing epidemiology of primary liver cancer. *Curr Epidemiol Rep, 6*(2), 104-111. doi:10.1007/s40471-019-00188-3

Ramadori, P., Li, X., & Heikenwalder, M. (2020). PKCλ/ι Loss Induces Metabolic Reprogramming in Liver Cancer: Hitting Two Birds with One Stone? *Cancer Cell, 38*(2), 152-154. doi:https://doi.org/10.1016/j.ccell.2020.07.003

Acevedo-Duncan, M., Patel, R., Whelan, S., & Bicaku, E. (2002). Human glioma PKC-ι and PKC-βII phosphorylate cyclin-dependent kinase activating kinase during the cell cycle. Cell proliferation, 35(1), 23-36.

Akimoto, K., Takahashi, R., Moriya, S., Nishioka, N., Takayanagi, J., Kimura, K., . . . Ohno, S. (1996). EGF or PDGF receptors activate atypical PKClambda through phosphatidylinositol 3-kinase. Embo j, 15(4), 788-798.

Baldwin, R. M., Parolin, D. A., & Lorimer, I. A. (2008). Regulation of glioblastoma cell invasion by PKC iota and RhoB. Oncogene, 27(25), 3587-3595. doi: 10.1038/sj.onc.1211027

Berra, E., Diaz-Meco, M. T., Dominguez, I., Municio, M. M., Sanz, L., Lozano, J., . . . Moscat, J. (1993). Protein kinase C zeta isoform is critical for mitogenic signal transduction. Cell, 74(3), 555-563. doi: 10.1016/0092-8674(93)80056-k

Berra, E., Díaz-Meco, M. T., Lozano, J., Frutos, S., Municio, M. M., Sánchez, P., . . . Moscat, J. (1995). Evidence for a role of MEK and MAPK during signal transduction by protein kinase C zeta. Embo j, 14(24), 6157-6163.

Bjorkoy, G., Perander, M., Overvatn, A., & Johansen, T. (1997). Reversion of Ras- and phosphatidylcholine-hydrolyzing phospholipase C-mediated transformation of NIH 3T3 cells by a dominant interfering mutant of protein kinase C lambda is accompanied by the loss of constitutive nuclear mitogen-activated protein kinase/extracellular signal-regulated kinase activity. J Biol Chem, 272(17), 11557-11565. doi: 10.1074/jbc.272.17.11557

Desai, S. R., Pillai, P. P., Patel, R. S., McCray, A. N., Win-Piazza, H. Y., & Acevedo-Duncan, M. E. (2011). Regulation of Cdk7 activity through a phosphatidylinositol (3)-kinase/PKC-ι-mediated signaling cascade in glioblastoma. Carcinogenesis, 33(1), 10-19. doi: 10.1093/carcin/bgr231

Diaz-Meco, M. T., Berra, E., Municio, M. M., Sanz, L., Lozano, J., Dominguez, I., . . . et al. (1993). A dominant negative protein kinase C zeta subspecies blocks NF-kappa B activation. Mol Cell Biol, 13(8), 4770-4775. doi: 10.1128/mcb.13.8.4770

Diaz-Meco, M. T., & Moscat, J. (2001). MEK5, a new target of the atypical protein kinase C isoforms in mitogenic signaling. Mol Cell Biol, 21(4), 1218-1227. doi: 10.1128/mcb.21.4.1218-1227.2001

Díaz-Meco, M. T., Municio, M. M., Frutos, S., Sanchez, P., Lozano, J., Sanz, L., & Moscat, J. (1996). The product of par-4, a gene induced during apoptosis, interacts selectively with the atypical isoforms of protein kinase C. Cell, 86(5), 777-786. doi: 10.1016/s0092-8674(00)80152-x

Eder, A. M., Sui, X., Rosen, D. G., Nolden, L. K., Cheng, K. W., Lahad, J. P., . . . Atkinson, E. N. (2005). Atypical PKCι contributes to poor prognosis through loss of apical-basal polarity and cyclin E overexpression in ovarian cancer. Proceedings of the National Academy of Sciences, 102(35), 12519-12524.

Eder, A. M., Sui, X., Rosen, D. G., Nolden, L. K., Cheng, K. W., Lahad, J. P., . . . Mills, G. B. (2005). Atypical PKCiota contributes to poor prognosis through loss of apical-basal polarity and cyclin E overexpression in ovarian cancer. Proc Natl Acad Sci U S A, 102(35), 12519-12524. doi: 10.1073/pnas.0505641102

Erdogan, E., Klee, E. W., Thompson, E. A., & Fields, A. P. (2009). Meta-analysis of oncogenic protein kinase Ciota signaling in lung adenocarcinoma. Clin Cancer Res, 15(5), 1527-1533. doi: 10.1158/1078-0432.ccr-08-2459

Fields, A. P., & Regala, R. P. (2007). Protein kinase C iota: human oncogene, prognostic marker and therapeutic target. Pharmacol Res, 55(6), 487-497. doi: 10.1016/j.phrs.2007.04.015

Frederick, L., Matthews, J., Jamieson, L., Justilien, V., Thompson, E. A., Radisky, D. C., & Fields, A. P. (2008). Matrix metalloproteinase-10 is a critical effector of protein kinase Cι-Par6α-mediated lung cancer. Oncogene, 27(35), 4841-4853.

Frederick, L. A., Matthews, J. A., Jamieson, L., Justilien, V., Thompson, E. A., Radisky, D. C., & Fields, A. P. (2008). Matrix metalloproteinase-10 is a critical effector of protein kinase Ciota-Par6alpha-mediated lung cancer. Oncogene, 27(35), 4841-4853. doi: 10.1038/onc.2008.119

Greenman, C., Stephens, P., Smith, R., Dalgliesh, G. L., Hunter, C., Bignell, G., . . . Stratton, M. R. (2007). Patterns of somatic mutation in human cancer genomes. Nature, 446(7132), 153-158. doi: 10.1038/nature05610

Gustafson, W. C., Ray, S., Jamieson, L., Thompson, E. A., Brasier, A. R., & Fields, A. P. (2004). Bcr-Abl regulates protein kinase Ciota (PKCiota) transcription via an Elk1 site in the PKCiota promoter. J Biol Chem, 279(10), 9400-9408. doi: 10.1074/jbc.M312840200

Hirano, Y., Yoshinaga, S., Ogura, K., Yokochi, M., Noda, Y., Sumimoto, H., & Inagaki, F. (2004). Solution structure of atypical protein kinase C PB1 domain and its

mode of interaction with ZIP/p62 and MEK5. J Biol Chem, 279(30), 31883-31890. doi: 10.1074/jbc.M403092200

Huang, H. C., Huang, C. Y., Lin-Shiau, S. Y., & Lin, J. K. (2009). Ursolic acid inhibits IL-1β or TNF-α-induced C6 glioma invasion through suppressing the association ZIP/p62 with PKC-ζ and downregulating the MMP-9 expression. Molecular Carcinogenesis: Published in cooperation with the University of Texas MD Anderson Cancer Center, 48(6), 517-531.

Jin, Z., Xin, M., & Deng, X. (2005). Survival function of protein kinase C{iota} as a novel nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone-activated bad kinase. J Biol Chem, 280(16), 16045-16052. doi: 10.1074/jbc.M413488200

Joberty, G., Petersen, C., Gao, L., & Macara, I. G. (2000). The cell-polarity protein Par6 links Par3 and atypical protein kinase C to Cdc42. Nat Cell Biol, 2(8), 531-539. doi: 10.1038/35019573

Justilien, V., Jameison, L., Der, C. J., Rossman, K. L., & Fields, A. P. (2011). Oncogenic activity of Ect2 is regulated through protein kinase C iota-mediated phosphorylation. J Biol Chem, 286(10), 8149-8157. doi: 10.1074/jbc.M110.196113

Kampfer, S., Windegger, M., Hochholdinger, F., Schwaiger, W., Pestell, R. G., Baier, G., . . . Überall, F. (2001). Protein kinase C isoforms involved in the transcriptional activation of cyclin D1 by transforming Ha-Ras. Journal of Biological Chemistry, 276(46), 42834-42842.

Kjær, S., Linch, M., Purkiss, A., Kostelecky, B., Knowles, P. P., Rosse, C., . . . Patel, B. (2013). Adenosine-binding motif mimicry and cellular effects of a thieno [2, 3-d] pyrimidine-based chemical inhibitor of atypical protein kinase C isoenzymes. Biochemical Journal, 451(2), 329-342.

Liao, D. F., Monia, B., Dean, N., & Berk, B. C. (1997). Protein kinase C-zeta mediates angiotensin II activation of ERK1/2 in vascular smooth muscle cells. J Biol Chem, 272(10), 6146-6150. doi: 10.1074/jbc.272.10.6146

Lin, D., Edwards, A. S., Fawcett, J. P., Mbamalu, G., Scott, J. D., & Pawson, T. (2000). A mammalian PAR-3-PAR-6 complex implicated in Cdc42/Rac1 and aPKC signalling and cell polarity. Nat Cell Biol, 2(8), 540-547. doi: 10.1038/35019582

Linch, M., Sanz-Garcia, M., Rosse, C., Riou, P., Peel, N., Madsen, C. D., . . . Dillon, C. (2014). Regulation of polarized morphogenesis by protein kinase C iota in oncogenic epithelial spheroids. Carcinogenesis, 35(2), 396-406.

Linch, M., Sanz-Garcia, M., Soriano, E., Zhang, Y., Riou, P., Rosse, C., . . . Kjaer, S. (2013). A cancer-associated mutation in atypical protein kinase Cι occurs in a substrate-specific recruitment motif. Science signaling, 6(293), ra82-ra82.

Lopez-Garcia, L. A., Schulze, J. O., Fröhner, W., Zhang, H., Süß, E., Weber, N., . . . Zeuzem, S. (2011). Allosteric regulation of protein kinase PKCζ by the N-terminal C1 domain and small compounds to the PIF-pocket. Chemistry & biology, 18(11), 1463-1473.

Mah, I. K., Soloff, R., Hedrick, S. M., & Mariani, F. V. (2015). Atypical PKC-iota Controls Stem Cell Expansion via Regulation of the Notch Pathway. Stem cell reports, 5(5), 866-880. doi: 10.1016/j.stemcr.2015.09.021

Martiny-Baron, G., & Fabbro, D. (2007). Classical PKC isoforms in cancer. Pharmacological research, 55(6), 477-486.

Moscat, J., & Diaz-Meco, M. T. (2000). The atypical protein kinase Cs. Functional specificity mediated by specific protein adapters. EMBO Rep, 1(5), 399-403. doi: 10.1093/embo-reports/kvd098

Murray, N. R., & Fields, A. P. (1997). Atypical protein kinase C iota protects human leukemia cells against drug-induced apoptosis. J Biol Chem, 272(44), 27521-27524. doi: 10.1074/jbc.272.44.27521

Murray, N. R., Jamieson, L., Yu, W., Zhang, J., Gökmen-Polar, Y., Sier, D., . . . Fields, A. P. (2004). Protein kinase Ciota is required for Ras transformation and colon carcinogenesis in vivo. J Cell Biol, 164(6), 797-802. doi: 10.1083/jcb.200311011

Murray, N. R., Kalari, K. R., & Fields, A. P. (2011). Protein kinase Cι expression and oncogenic signaling mechanisms in cancer. Journal of cellular physiology, 226(4), 879-887.

Murray, N. R., Weems, J., Braun, U., Leitges, M., & Fields, A. P. (2009). Protein kinase C betaII and PKCiota/lambda: collaborating partners in colon cancer promotion and progression. Cancer Res, 69(2), 656-662. doi: 10.1158/0008-5472.can-08-3001

Noda, Y., Takeya, R., Ohno, S., Naito, S., Ito, T., & Sumimoto, H. (2001). Human homologues of the Caenorhabditis elegans cell polarity protein PAR6 as an adaptor that links the small GTPases Rac and Cdc42 to atypical protein kinase C. Genes Cells, 6(2), 107-119. doi: 10.1046/j.1365-2443.2001.00404.x

Paramio, J. M., Segrelles, C., Ruiz, S., & Jorcano, J. L. (2001). Inhibition of Protein Kinase B (PKB) and PKCζ Mediates Keratin K10-Induced Cell Cycle Arrest. Mol Cell Biol, 21(21), 7449-7459. doi: 10.1128/mcb.21.21.7449-7459.2001

Parker, P. J., Justilien, V., Riou, P., Linch, M., & Fields, A. P. (2014). Atypical protein kinase Cι as a human oncogene and therapeutic target. Biochem Pharmacol, 88(1), 1-11. doi: 10.1016/j.bcp.2013.10.023

Patel, R., Win, H., Desai, S., Patel, K., Matthews, J. A., & Acevedo-Duncan, M. (2008). Involvement of PKC-iota in glioma proliferation. Cell Prolif, 41(1), 122-135. doi: 10.1111/j.1365-2184.2007.00506.x

Pillai, P., Desai, S., Patel, R., Sajan, M., Farese, R., Ostrov, D., & Acevedo-Duncan, M. (2011). A novel PKC-ι inhibitor abrogates cell proliferation and induces apoptosis in neuroblastoma. The International Journal of Biochemistry & Cell Biology, 43(5), 784-794. doi: https://doi.org/10.1016/j.biocel.2011.02.002

National Center for Biotechnology Information (2021). PubChem Compound Summary for CID 44559187. Retrieved July 31, 2021 from https://pubchem.ncbi.nlm.nih.gov/compound/44559187.

Puls, A., Schmidt, S., Grawe, F., & Stabel, S. (1997). Interaction of protein kinase C zeta with ZIP, a novel protein kinase C-binding protein. Proc Natl Acad Sci U S A, 94(12), 6191-6196. doi: 10.1073/pnas.94.12.6191

Qiu, R. G., Abo, A., & Steven Martin, G. (2000). A human homolog of the C. elegans polarity determinant Par-6 links Rac and Cdc42 to PKCzeta signaling and cell transformation. Curr Biol, 10(12), 697-707. doi: 10.1016/s0960-9822(00)00535-2

Regala, R. P., Weems, C., Jamieson, L., Khoor, A., Edell, E. S., Lohse, C. M., & Fields, A. P. (2005). Atypical protein kinase C iota is an oncogene in human non-small cell lung cancer. Cancer Res, 65(19), 8905-8911. doi: 10.1158/0008-5472.can-05-2372

Regala, R. P., Weems, C., Jamieson, L., Khoor, A., Edell, E. S., Lohse, C. M., & Fields, A. P. (2005). Atypical protein kinase Cι is an oncogene in human non–small cell lung cancer. Cancer Res, 65(19), 8905-8911.

Sanchez, P., De Carcer, G., Sandoval, I. V., Moscat, J., & Diaz-Meco, M. T. (1998). Localization of atypical protein kinase C isoforms into lysosome-targeted endosomes through interaction with p62. Mol Cell Biol, 18(5), 3069-3080. doi: 10.1128/mcb.18.5.3069

Sanz, L., Sanchez, P., Lallena, M. J., Diaz-Meco, M. T., & Moscat, J. (1999). The interaction of p62 with RIP links the atypical PKCs to NF-kappaB activation. Embo j, 18(11), 3044-3053. doi: 10.1093/emboj/18.11.3044

Scotti, M. L., Bamlet, W. R., Smyrk, T. C., Fields, A. P., & Murray, N. R. (2010). Protein kinase Ciota is required for pancreatic cancer cell transformed growth and tumorigenesis. Cancer Res, 70(5), 2064-2074. doi: 10.1158/0008-5472.can-09-2684

Selbie, L. A., Schmitz-Peiffer, C., Sheng, Y., & Biden, T. J. (1993). Molecular cloning and characterization of PKC iota, an atypical isoform of protein kinase C derived from insulin-secreting cells. J Biol Chem, 268(32), 24296-24302.

Sia, D., Villanueva, A., Friedman, S. L., & Llovet, J. M. (2017). Liver Cancer Cell of Origin, Molecular Class, and Effects on Patient Prognosis. Gastroenterology, 152(4), 745-761. doi:10.1053/j.gastro.2016.11.048

Sontag, E., Sontag, J. M., & Garcia, A. (1997). Protein phosphatase 2A is a critical regulator of protein kinase C zeta signaling targeted by SV40 small t to promote cell growth and NF-kappaB activation. Embo j, 16(18), 5662-5671. doi: 10.1093/emboj/16.18.5662

Stallings-Mann, M., Jamieson, L., Regala, R. P., Weems, C., Murray, N. R., & Fields, A. P. (2006). A novel small-molecule inhibitor of protein Kinase Cι blocks transformed growth of non–small-cell lung cancer cells. Cancer Res, 66(3), 1767-1774.

Takeda, H., Matozaki, T., Takada, T., Noguchi, T., Yamao, T., Tsuda, M., . . . Kasuga, M. (1999). PI 3-kinase gamma and protein kinase C-zeta mediate RAS-independent activation of MAP kinase by a Gi protein-coupled receptor. Embo j, 18(2), 386-395. doi: 10.1093/emboj/18.2.386

Weichert, W., Gekeler, V., Denkert, C., Dietel, M., & Hauptmann, S. (2003). Protein kinase C isoform expression in ovarian carcinoma correlates with indicators of poor prognosis. Int J Oncol, 23(3), 633-639.

White, W. O., Seibenhener, M. L., & Wooten, M. W. (2002). Phosphorylation of tyrosine 256 facilitates nuclear import of atypical protein kinase C. J Cell Biochem, 85(1), 42-53.

Win, H. Y., & Acevedo-Duncan, M. (2008). Atypical protein kinase C phosphorylates IKKalphabeta in transformed non-malignant and malignant prostate cell survival. Cancer Lett, 270(2), 302-311. doi: 10.1016/j.canlet.2008.05.023

Wodarz, A., Ramrath, A., Grimm, A., & Knust, E. (2000). Drosophila atypical asymmetric division in the Drosophila epithelium. Nature protein kinase C associates with Bazooka and controls 2001, 409: 522-525. polarity of epithelia and neuroblasts. J Cell Biol, 30, 1361-1374.

Wooten, M. W. (1999). Function for NF-kB in neuronal survival: regulation by atypical protein kinase C. J Neurosci Res, 58(5), 607-611. doi: 10.1002/(sici)1097-4547(19991201)58:5<607::aid-jnr1>3.0.co;2-m

Yang, Y. L., Chu, J. Y., Luo, M. L., Wu, Y. P., Zhang, Y., Feng, Y. B., . . . Cai, Y. (2008). Amplification of PRKCI, located in 3q26, is associated with lymph node metastasis in esophageal squamous cell carcinoma. Genes, Chromosomes and Cancer, 47(2), 127-136.

Yang, Y. L., Chu, J. Y., Luo, M. L., Wu, Y. P., Zhang, Y., Feng, Y. B., . . . Wang, M. R. (2008). Amplification of PRKCI, located in 3q26, is associated with lymph node

metastasis in esophageal squamous cell carcinoma. Genes Chromosomes Cancer, 47(2), 127-136. doi: 10.1002/gcc.20514

Zhang, L., Huang, J., Yang, N., Liang, S., Barchetti, A., Giannakakis, A., . . . Coukos, G. (2006). Integrative genomic analysis of protein kinase C (PKC) family identifies PKCiota as a biomarker and potential oncogene in ovarian carcinoma. Cancer Res, 66(9), 4627-4635. doi: 10.1158/0008-5472.can-05-4527

Zhang, P. W., Chen, L., Huang, T., Zhang, N., Kong, X. Y., & Cai, Y. D. (2015). Classifying ten types of major cancers based on reverse phase protein array profiles. PLoS One, 10(3), e0123147. doi: 10.1371/journal.pone.0123147

Sanz, L., Sanchez, P., Lallena, M. J., Diaz-Meco, M. T., & Moscat, J. (1999). The interaction of p62 with RIP links the atypical PKCs to NF-kappaB activation. *Embo j, 18*(11), 3044-3053. doi:10.1093/emboj/18.11.3044

Srivatanakul, P., Sriplung, H., & Deerasamee, S. (2004). Epidemiology of liver cancer: an overview. *Asian Pac J Cancer Prev, 5*(2), 118-125.

Tisdale, E. J. (2002). Glyceraldehyde-3-phosphate dehydrogenase is phosphorylated by protein kinase Ciota /lambda and plays a role in microtubule dynamics in the early secretory pathway. *J Biol Chem, 277*(5), 3334-3341. doi:10.1074/jbc.M109744200

Wooten, M. W., Seibenhener, M. L., Neidigh, K. B., & Vandenplas, M. L. (2000). Mapping of atypical protein kinase C within the nerve growth factor signaling cascade: relationship to differentiation and survival of PC12 cells. *Mol Cell Biol, 20*(13), 4494-4504. doi:10.1128/mcb.20.13.4494-4504.2000

Capriotti, E., Fariselli, P., & Casadio, R. (2005). I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic acids research, 33(Web Server issue), W306-W310. doi:10.1093/nar/gki375

Garg, R., Benedetti, L. G., Abera, M. B., Wang, H., Abba, M., & Kazanietz, M. G. (2014). Protein kinase C and cancer: what we know and what we do not. Oncogene, 33(45), 5225-5237. doi:10.1038/onc.2013.524

Venselaar, H., Te Beek, T. A., Kuipers, R. K., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics, 11, 548. doi:10.1186/1471-2105-11-548

Xie, J., Guo, Q., Zhu, H., Wooten, M. W., & Mattson, M. P. (2000). Protein kinase C iota protects neural cells against apoptosis induced by amyloid beta-peptide. *Brain Res Mol Brain Res, 82*(1-2), 107-113. doi:10.1016/s0169-328x(00)00187-x

Garg, R., Benedetti, L. G., Abera, M. B., Wang, H., Abba, M., & Kazanietz, M. G. (2014). Protein kinase C and cancer: what we know and what we do not. Oncogene, 33(45), 5225-5237. doi:10.1038/onc.2013.524

Murray, N. R., Kalari, K. R., & Fields, A. P. (2011). Protein kinase Ciota expression and oncogenic signaling mechanisms in cancer. J Cell Physiol, 226(4), 879-887. doi:10.1002/jcp.22463

Chitrala, K. N., & Yeguvapalli, S. (2014). Computational screening and molecular dynamic simulation of breast cancer associated deleterious non-synonymous

single nucleotide polymorphisms in TP53 gene. PLoS One, 9(8), e104242. doi:10.1371/journal.pone.0104242

Hafeez, S., Mahmood, A., Khan, R. U., & Malkani, N. (2020). Trends in Cancer Prevalence in Punjab, Pakistan: A Systematic Study from 2010 to 2016. Journal of Bioresource Management, 7(2), 8.

Singal, A. G., Lampertico, P., & Nahon, P. (2020). Epidemiology and surveillance for hepatocellular carcinoma: New trends. Journal of hepatology, 72(2), 250-261.

Wang, W., Wang, Y., Liu, M., Zhang, Y., Yang, T., Li, D., . . . Shi, L. (2019). Betulinic acid induces apoptosis and suppresses metastasis in hepatocellular carcinoma cell lines in vitro and in vivo. Journal of cellular and molecular medicine, 23(1), 586-595.

Yin, F., Feng, F., Wang, L., Wang, X., Li, Z., & Cao, Y. (2019). SREBP-1 inhibitor Betulin enhances the antitumor effect of Sorafenib on hepatocellular carcinoma via restricting cellular glycolytic activity. Cell death & disease, 10(9), 1-12.

Zhang, Y. (2009). I-TASSER: Fully automated protein structure prediction in CASP8. Proteins: Structure, Function, and Bioinformatics, 77(S9), 100-113.