Cathepsin D, an Aspartic Protease as a Potent Therapeutic Option for Breast Cancer: Design, Synthesis and Bioevaluation against Triple-Positive and Triple-Negative Breast Cancers

THESIS

Submitted in partial fulfillment

Of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

ANANTARAJU HASITHASHILPA

ID No 2013PHXF010H

Under the Supervision of

Prof. P. YOGEESWARI

BITS Pilani Pilani | Dubai | Goa | Hyderabad

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

2014

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

CERTIFICATE

This is to certify that the thesis entitled "**Cathepsin D, an Aspartic Protease as a Potent Therapeutic Option for Breast Cancer: Design, Synthesis and Bioevaluation Against Triple-Positive and Triple-Negative Breast Cancers**" and submitted by **ANANTARAJU HASITHASHILPA** ID No. **2013PHXF010H** for award of Ph.D. of the Institute embodies original work done by her under my supervision.

Signature of the Supervisor:

Name in capital letters: **P.YOGEESWARI**

Designation: **Professor**

Date:

ACKNOWLEDGEMENT

First and foremost I would like to thank God Almighty who has given me this opportunity.

I would like to give heart full thanks to my supervisor, Prof.P.Yogeeswari, Professor and Associate Dean (Sponsored Research and Consultancy Division), Department of Pharmacy, BITS-Pilani Hyderabad Campus for her continuous guidance, suggestions and support. She was always an inspiration to me. I am thankful for her enthusiasm, patience and love for research. She always encourage new thoughts and give constant support throughout the project. The experiences gained under her guidance are memorable and will be cherished throughout my life.

I deeply acknowledge and heartfelt thanks to Prof.D.Sriram, Professor, Department of Pharmacy, BITS-Pilani Hyderabad Campus for his valuable suggestions, guidance offered to me during this period.

I am thankful to acknowledge my DAC member Dr. Vamsi Krishna Venuganti, for his support and encouragement during this period.

I am grateful to Prof.Bijendra N. Jain, Vice-Chancellor, BITS-Pilani Campus and Prof.V.S.Rao, Director, BITS-Pilani Hyderabad Campus for allowing me to carry out my doctoral research work in the institute.

I am thankful to Prof.M.M.S.Anand, Registrar and Prof.S.K.Verma, Dean, Acamedic Research (Ph.D. Programme), BITS-Pilani for their support to do my research work.

I would like to Prof.M.B.Srinivas, Dean, Administration and Prof.Vidya Rajesh, Associate Dean, Academic Research (Ph.D Programme), BITS-Pilani Hyderabad Campus for their continuous support and encouragement during my research work.

I would like to express my gratitude to Dr.Srikanth Charde, Assistant Professor and Head, Department of Pharmacy, BITS-Pilani Hyderabad Campus for providing me with all necessary laboratory facilities and for having helped me at various stages of my research work.

I sincerely acknowledge the help rendered by Dr.Punna Rao, Dr.Sajeli Begum, and other Faculty members, Department of Pharmacy, BITS-Pilani Hyderabad Campus.

I am grateful to express my sincere thanks to all my friends Madhu Babu, Brindha, Ganesh.S, Saketh, Gangadhar, Poorna, Ganesh.P, Brahmam, Priyanka.S, Radhika, Suman, Bobesh, Anup, Mary Priyanka, Prakruthi, Shubham, Preeti, Vishnu kiran, Omkar for the time they had spent for me and helped me to complete my work.

I express my thanks to our laboratory attenders and demonstrators Mr. Rajesh, Mr.Venkat, Mr.Seenu, Ms.Saritha and Ms.Rekha for all their support.

I deeply acknowledge the Department of Science and Technology, New Delhi,, India for DST-INSPIRE award [IF130633].

At the end, but from bottom of my heart, I thank and deeply appreciate my family A.V.N. Ramesh (Father), K. Aruna Kumari (Mother), A. Sravana Neela (Sister) and A. Yashwanth Raj (Brother) for their faith in me and for continuous support. It was under their influence I gained so much drive and an ability to tackle challenges head on. No words can express how grateful I am for their love and support. Last but not the least, I would like to thank my husband (YHNB SAILESH) for his constant support at various phase of my life, he is my every hope and every dream I have ever had.

Hasitha Shilpa Anantaraju

ABSTRACT

Cathepsin D (CatD) (EC.3.4.23.5) is an aspartic protease expressed in all tissues. Most of the metastatic breast adenocarcinoma cells over express CatD. CatD is well-known as a prognostic biomarker in detecting advance stages of mammary cancers. The main aim of this study was to discover small molecule inhibitors against CatD, a clinically proven prognostic marker for breast cancer and to explore the mechanisms by which CatD could be a useful therapeutic target for triple-positive and triple-negative breast cancers (TPBC & TNBC). The crystal structure of CatD at 2.5 Å resolution (PDB: 1LYB), which was complexed with Pepstatin A, was selected for computer-aided molecular modelling. The methods used in our study were pharmacophore modelling and molecular docking. Virtual screening was performed to identify small molecules from an in-house database and a large commercial chemical library. Cytotoxicity studies were performed on human normal cell line HEK293T and growth inhibition studies on breast adenocarcinoma cell lines, namely MCF-7, MDA-MB-231, SK-BR-3, ZR-75-1 and MDA-MB-468. Furthermore, RT-PCR analysis, invitro enzyme assay, and cell cycle analysis ascertained the validity of the selected molecules. A set of 29 molecules was subjected to an *in vitro* fluorescence-based inhibitory activity assay, and among them seven molecules exhibited >50% inhibition at 25µM. These molecules also exhibited good growth inhibition against TPBC and TNBC cancer types. Among them, molecules1 and 29 showed single-digit micromolar GI_{50} values against MCF-7 and MDA-MB-231 cell lines.

We synthesized 17 analogues for M29 to achieve more promising lead molecule. Among the analogues synthesized, compound CIS-1 was found to the most active in enzymatic as well as anti-cancer assays. CIS-1 was found to inhibit ROS and reduce Acridine orange relative fluorescence which is an indicative of lysosomal degradation in cellular compartments. At its GI_{50} value of

1.02±0.05 μM, CIS-1 was found to down regulate EGFR, VEGF, MMP-9 which are the important growth factors and matrix metallo proteases involved in Extracellular Matrix (ECM) construction and invasion mechanisms. This was further demonstrated using angiogenesis assays like scratch assay, Boyden chamber assay and tube formation assays. Gelatin zymography studies also indicated the role of CIS-1 in inhibiting active forms of MMP-2 and MMP-9 proteases using cell culture supernatants. We performed DNA methylation studies using COBRA assay which evaluated the role of CatD inhibition on gene silencing.

Certain RTPCR studies were performed in hypoxic conditions developed using cobalt chloride on cancer cells. These studies resulted in huge gene expressions in relative to invasion behaviour of breast cancer cells. The effect of CIS-1 was further evaluated using *invivo* zebrafish embryo angiogenesis model which was alkaline phosphatase dependent assay. Olaparib, a PARP-1 inhibitor was tested in combination with CIS-1 and evaluated for their effective growth inhibition against MDA-MB-231 cells.

In conclusion, novel non-peptidic small molecule inhibitors were optimized identified using molecular docking studies. Analogues were synthesized and characterized for the most active M29 molecule identified through docking study. The potent analogue was further studied for their growth proliferative and mechanism- based activity on TPBC and TNBC cell types.

Contents

Table of Figures

List of Tables

List of Abbreviations

α	:	Alpha
β	:	Beta
μM	:	Micro molar
Å	:	Angstrom
ADME	:	Absorption, Distribution, Metabolism, Elimination
AO	:	Acridine Orange
CatD	:	Cathepsin D
CSC	:	Cancer stem cell
DCFHDA	:	Dichloro,dihydro-Fluorescein DiAcetate
DMSO	:	DiMethyl Sulfoxide
ECM	:	Extracellular Matrix
EGFR	:	Epidermal Growth Factor Receptor
EMT	:	Epithelial mesenchymal transition
ER	:	Estrogen Receptor
ERBB2	:	Epidermal growth factor receptor 2
FBS	:	Fetal Bovine Serum
PBS	:	Phosphate Buffered Saline
GLIDE	:	Grid based Ligand Docking and Energetics
H1F	:	Hypoxia inducing Factor
IC50	:	Half maximal Inhibitory Concentration
MMP	:	Matrix Metallo Proteases
MTT	:	(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
PAGE	:	Polyacrylamide Gel Electrophoresis
PARP	:	Poly ADP ribose polymerase
pCD	:	Pro Cathepsin D
PCR	:	Polymerase Chain Reaction
PDB	:	Protein Data Bank
PR	:	Progesterone Receptor
QPlogBB	:	QuickProp predicted Blood Brain partition coefficient

QPlogCaco	:	QuickProp predicted Caco-2 cell permeability
QPlogHERG	:	QuickProp predicted the human Ether-a-go-go-Related gene
QPlogP	:	QuickProp predicted Partition coefficient
RTPCR	:	Real Time Polymerase Chain Reaction
SAR	:	Structure-Activity Relationship
SP	:	Standard Precision
TPBC	:	Triple Positive Breast Cancer
TNBC	:	Triple Negative Breast Cancer
VEGF	:	Vasculo Endothelial Growth Factor
XP	:	Extra Precision

Chapter 1. INTRODUCTION

1.1. Cancer

Cancer is a complex disease in which cells in a specific tissue are no longer fully responsive to the signals within the tissue that regulate cellular differentiation, survival, proliferation and death. As a result, these cells accumulate within the tissue, causing local damage and inflammation. There are over 200 different types of cancer. After a quarter century of rapid advances, cancer research has generated a rich and complex body of knowledge, revealing cancer to be a disease involving dynamic changes in the genome. The barriers to development of cancer are embodied in a teleology that states cancer cells have defects in regulatory circuits that govern normal cell proliferation and homeostasis [Hanahan D. and Weinberg A.R., 2000].

In an attempt to battle this disease the World Health Organization have developed three principles on cancer prevention: they have estimated that current prevention strategies could prevent up to one-third of new cancers; they have suggested that improved early screening could result in the detection of one-third of cancers at an early stage; and they have proposed that a comprehensive treatment strategy could improve survival and quality of life for another one-third of patients with advanced cancer. These strategies offer the most cost-effective, long-term control of cancer. According to WHO India office, the estimated cancer incidence and cumulative risk by 2015 was reported to be one in 9. Amongst all, lung, breast, stomach and oral cavity cancers were reported to be the common occurrences [Nair M.K., *et al.*, 2015]. Sooner or later during the development of most types of human cancer, primary tumour masses spawn pioneer cells that move out, invade adjacent tissues, and hence travel to distant sites where they may succeed in founding new colonies. According to Sporn, these distant settlements of tumour cells—metastases—are the cause

of 90% of human cancer deaths [Hanahan D. and Weinberg A.R., 2000].Over the past decade, tumours have increasingly been recognised as organs whose complexity approaches may even exceed that of normal healthy tissues. Many human tumours are histopathologically diverse, containing regions demarcated by various degrees of differentiation, proliferation, vascularity, inflammation and invasiveness. Recent evidences pointed out a new dimension of intratumor heterogeneity and a sub-class of neoplastic cells within tumours, known as cancer stem cells (CSCs). The mechanism-based targeted therapies to treat human cancers have been one of the challenging and remarkable progress of research in to mechanisms of cancer pathogenesis [Nair M.K., *et al.,* 2015].

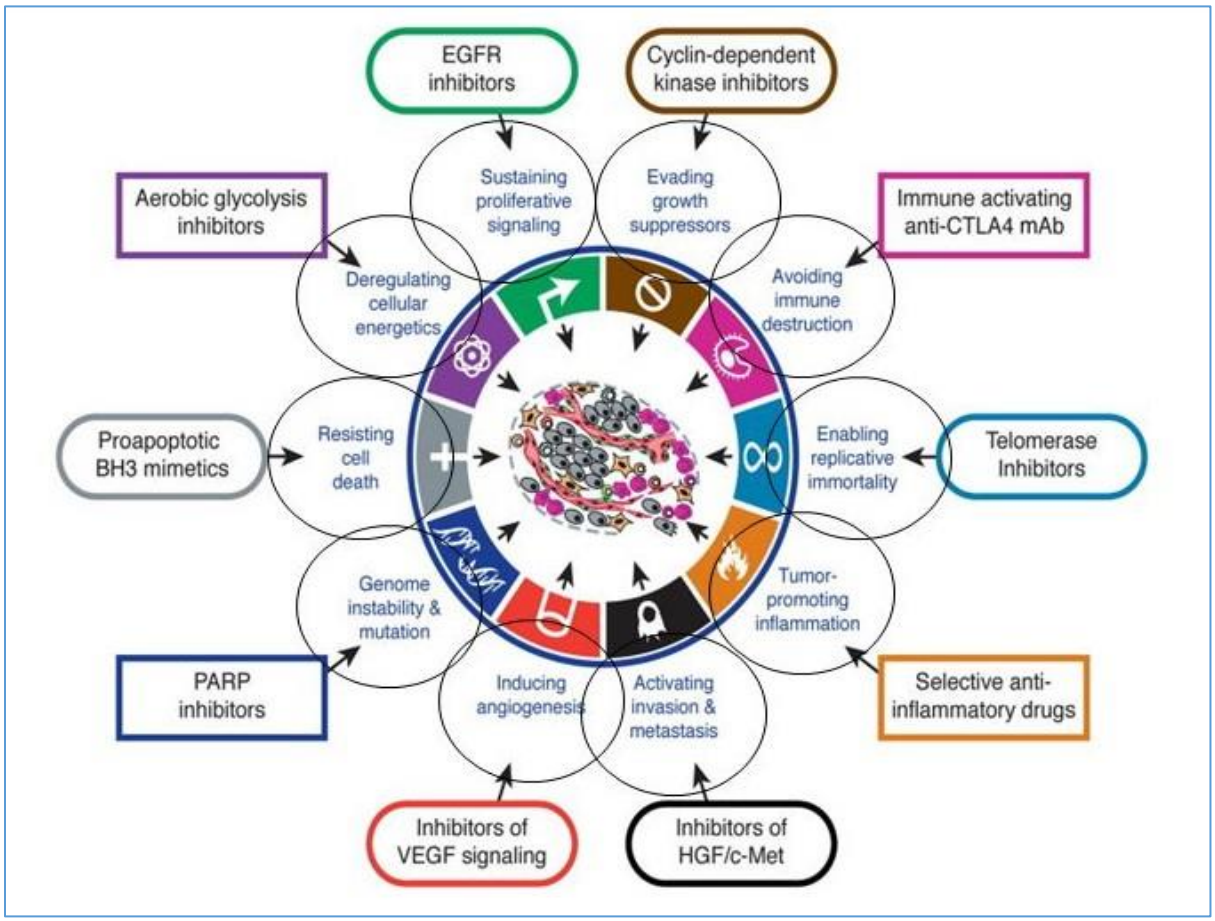


Figure 1: Therapeutic targeting of the Hallmarks of cancer [Nair M.K., *et al.,* 2015].

1.2. Breast cancer and its causes

Breast cancer is the second leading cancer type for the estimated cases and deaths in females all over the world [Siegel R.L., *et al.,* 2015]. Breast cancer is a complex and intrinsically heterogeneous disease that has been typically defined and classified using numerous clinical and pathologic features to predict outcome and treatment response. Classically, these clinical features include: age, tumour size, axillary node involvement, angio-lymphatic invasion, histologic grade, hormonal receptor status and ERBB2 (also known as HER-2/neu) amplification [Bosch A., *et al.,* 2010].

The increased incidences are due to increase in number of risk factors including lower age of menarche, late age of first pregnancy, shorter or no periods of breast feeding, fewer pregnancies, and a later menopause [Howell A., *et al.,* 2014]. . Apart from these, obesity, alcohol consumption, inactivity, and role of endogenous and exogenous hormones add to the burden of breast cancer. The impact of hereditary breast cancer for instance, BRCA1 and BRCA2 mutations has also increased over the time period. The new approaches to prevention is more into understanding the biological basis of breast development and the potential targets for therapeutic interventions [Breast cancer risk factors; 2009].

1.3. Types of breast cancers

Breast cancers can be broadly classified based on histology, molecular and functional subtypes. Histologically classification is based on architectural features and growth patterns of breast cancers namely insitu carcinoma and invasive carcinoma. Molecular classification is based on the intrinsic molecular subtypes of breast cancer identified by microarray analysis of patient tumour specimens namely luminal, basal and HER2 enriched. Functional classification is based on the tumour initiating cells [Malhotra G.K. *et al.,* 2010; Alison K.H., *et al.,* 2012] Breast cancer world in 2014

has come up with a new class subtypes that lead to targeted therapies. The subtypes are ER+ (Estrogen receptor positive), HER+ and triple-negative breast cancers. The year 2014 appeared to be a transitional year between receptor-based approaches that have dominated the past decades and newer approaches that brought us closer to a cure [Butcher L., *et al.*, 2015].

1.4. Role of biomarkers in breast cancer

Breast cancer treatment has experienced several changes in the past decades due to discovery of specific prognostic and predictive biomarkers that enabled research to develop individualized therapies considering different molecular subgroups. The way towards earlier diagnosis of breast cancers due to sophisticated improved imaging methodology and screening programs depicts the need of new factors and combinations of biomarkers to quantify the residual risk of patients [Weigel M.T. and Dowsett M., 2010]. The well-established molecular biomarkers i.e., estrogen and progesterone receptors have played a vital role in treatment of patients with endocrine therapy for many years. In recent years, human epidermal growth factor receptor (HER2) has been validated to be prognostic as well as predictor of response to HER2 targeting therapy. The emerging biomarkers are Ki67 (marker of proliferation), cyclin D1 (overexpressed at the mRNA and protein level) in 50% of breast cancer cases, Cyclin E (positive regulator of cell cycle), Cathepsins importantly Cathepsin B, D and L2 (responsible for invasion and metastasis), other biomarkers such as surviving, CD68, stromelysin 3 etc. Current research is involved in identifying circulating tumour cells, circulating miRNAs and tumour-specific DNA as biomarkers in the diagnosis of invasive and metastatic breast cancers [Weigel M.T. and Dowsett M., 2010; Heneghan H.M., *et al.*, 2010; Chirurgica M., 2002].

Imaging technologies (mammography, digital mammography, and magnetic resonance imaging (MRI)) have been adopted clinically for mass screening purposes, but there is resistance for

seeking such services on a yearly basis, given the relative complexity and high cost-to-benefit ratio of these imaging methodologies [Misek D.E. and Evelyn K.H., 2011]. The increasing pace of large and randomized clinical trials helps research to exploit neoadjuvant study design to increase the efficiency of correlative biomarker evaluation. Result of some of the targeted therapies intended to cure early breast cancers were tamoxifen (selective estrogen receptor modulator), trastuzumab (humanized monoclonal antibody), tyrosine kinase inhibitors, heat shock protein inhibitors, microtubule inhibitors and various immunotherapies [McArthur H.L. and Dickler M.N., 2008]. A recent review has listed all possible biomarkers that can be used for diagnosis, prognosis and predictive drug responses of breast cancer. The biomarkers are importantly needed to be validated in clinical setting to be translated into clinically useful test. The main factor that leads to breast cancer mortality is the presence of metastasis. The potential biomarkers till 2014 were listed as BRCA1/2, plasminogen activator, osteopontin, FGFR2, PTEN, PARP, sirtuins, Snail 1, PIK3CA, STAT3, TIMP-1, micro RNAs [Pultz B.D.A., *et al.,* 2014]. These biomarkers can further be helpful in stratification of patients as well as prevention of under or over treatments.

1.5. Clinical findings and current research strategies

According to National cancer institute-NIH, the possible types of treatment include surgery, radiotherapy, chemotherapy, targeted therapy, immune therapy, hormone therapy, stem cell transplant, precision medicine (selected based on the genetic profile of the person's cancer). Many clinical findings revealed the increasing resistance to chemotherapeutics. For almost three decades, tamoxifen has been the mainstay of endocrine therapy for the treatment of hormone-receptor positive breast cancer in pre-menopausal women and is also frequently used by post-menopausal women, both in early or advanced breast cancer. It has been estimated that approximately 50% of breast cancer patients undergoing chemotherapy develop tamoxifen resistance. Recent study

showed that about 20% of post-menopausal breast cancer patients who received tamoxifen in first line of therapy develop resistance and was considered non-responders to treatment. Some studies have shown that more than one-third of patients with metastatic breast cancer do not respond to first-line anthracyclines or taxanes, and disease progression occurs in less than 1 year, leading to the death of more than 90% of patients with metastatic breast cancer [Pultz B.D.A., *et al.*, 2014]. So, the current breast cancer research is more focused on to 'targeted' therapies to increase the chances of early detection and treatment of a particular mutation. Several clinical trials are currently assessing the efficacy of adding either a monoclonal antibody (cetuximab) or a tyrosine kinase inhibitor (erlotinib; Roche). EGLN1 (also known as PHD2) is a member of the Phd family of oxygen-sensitive prolyl hydroxylases that regulate the protein levels of the hypoxia-inducible factor-α might provide an excellent target for anti-metastatic drug discovery as per the recent nature reviews in cancer.

There are many gaps identified in current breast cancer research due to lack of financial and practical resources. Critical gaps include knowledge of genetic changes, initiation of breast cancer developmental signalling pathways, view on progression of breast cancer, updates on therapies and targets, intelligence on the new treatments and patient groups stratified using biomarkers, strategies to prevent estrogen receptor negative tumours and also the psychosocial aspects [Thompson A., *et al.*, 2008].

1.6. Breast cancer invasion and metastasis

1.6.1. Metastasis prognostic markers

Breast cancer is a clinically heterogeneous disease. Approximately 10–15% of patients with breast cancer have an aggressive disease and develop distant metastases within 3 years after the initial detection of the primary tumour. Breast cancer generally starts as a local disease and metastasizes

to the lymph nodes and distant organs. Metastasis is a complex multistep process involving complex molecular and genetic changes in the malignant cells that enable them to interact with surrounding structures including basement membranes structure, extracellular matrix, stromal cells, immune cells, endothelial cells and specialised cells at the metastatic sites. There are metastasis prognostic markers to identify the occurrence of breast cancer invasion. New prognostic markers are urgently needed to identify patients who are at the highest risk for developing metastases, which might enable oncologists to begin tailoring treatment strategies to individual patients. The markers are namely tumour size, axillary lymph node status, histological grade relative to tumour size, angio invasion in patients with lymph node negative tumours, uPA/PAI1 protein levels, steroid-receptor expression levels, ERBB2 (epidermal growth factor receptor 2) gene amplification and protein expression and gene expression profiling. New molecular technologies, such as DNA microarrays, support the idea that metastatic capacity might be an inherent feature of breast tumours. Improving our understanding of the molecular mechanisms of the metastatic process might also improve clinical management of the disease. According to the widely held model of metastasis, rare subpopulations of cells within the primary tumour acquire advantageous genetic alterations over time, which enable these cells to metastasize and form new solid tumours at distant sites. Recently, two studies also reported novel markers to predict distant metastasis risk and clinical outcome in patients with oestrogen-receptor-positive breast tumours who have been treated with adjuvant tamoxifen. Ma et al. showed that the expression ratio of the two genes HIXB13 and IL17BL accurately predicts metastasis development.

1.6.2. Models of the metastatic process

The metastatic properties of tumour cells were extensively investigated in the late 1970s and early 1980s by means of ‘experimental metastasis’ assays. Traditional model of metastasis suggests that

only subpopulations of tumour cells acquire metastatic capacity, late in tumorigenesis. Spontaneous metastasis assays indicate that all tumour cells have the capability to develop a metastasis. The 'dynamic heterogeneity' model proposes that the frequency with which metastatic variants arise within the primary tumour determines its metastatic potential. The 'clonal dominance' theory proposes that metastatic sub clones within a primary tumour can overgrow and dominate the tumour mass itself. The 'Geno metastasis hypothesis' proposes that metastasis occurs through transfection of susceptible cells in distant organs with circulating oncogenes.

New models of the metastatic process in breast cancer suggests the following:

1. Gene-expression profiling of human primary breast tumours can predict metastasis risk, which indicates that the capacity to metastasize might be acquired early during tumorigenesis.
2. Primary tumours with metastasizing capacity display the poor-prognosis signature and an additional tissue-specific expression profile can help in predicting the site of metastasis.
3. The parallel evolution model proposes that the dissemination of metastatic cancer cells occurs early in oncogenesis and independently from tumour cells at the primary site.
4. Another model states that only breast cancer stem cells, not the non-tumorigenic bulk of the tumour, have the ability to metastasize and form new tumours.

Whereas, the integrative model of breast cancer metastasis states that the oncogenic mutations occurring in a breast stem cell can cause the transformation to a breast cancer stem cell, generating 'poor-prognosis' tumours. In the metastatic poor prognosis tumours, under the influence of stromal fibroblasts, only the population of breast cancer stem cells has the ability to metastasize. There might be variant cancer stem cells that differ in their tissue selectivity for metastasis, expressing an additional tissue-specific profile. Secondly, mutations occurring in differentiated progenitor cells might form a non-metastatic 'good-prognosis' breast carcinoma [Weigelt B., *et al.*, 2005].

1.6.3. Pathology of metastasis

The important fundamental prerequisites for metastasis are unlimited cellular proliferation, evasion of cell intrinsic and environmental constraints, angiogenesis and gaining the capacity to detach, migrate and move away from the original locations. The pathology involved for the tumour cells to metastasize includes metastatic initiation and dissemination of the aggressive tumour cells leading to its entry into the circulation. Followed by the epithelial mesenchymal transition (EMT) that comprises a sequence of events in which the epithelial cells lose many of their epithelial characteristics and transits into mesenchymal cells. This phenomenon causes many complex changes in cell architecture and encompasses a wide spectrum of inter-cellular and intra-cellular changes. Recent insights into molecular EMT-controlling mechanisms indicate that several complex signalling pathways are involved in the initiation and execution of EMT in the contexts of development and cancer progression. A number of specific molecular pathways and transcription factors have been reported as ‘‘EMT triggers’’ including TGFß, Twist1/2, PI3 K/Akt pathway, CTEN, Snail, Slug, and Zeb1.

The tumour microenvironment in invasive breast carcinoma includes different types and subtypes of cells contributing as pro-metastatic or anti-metastatic regulators. Fibroblasts, macrophages and lymphocytes play a vital role in the tumour microenvironment. Cancer associated fibroblasts (CAFs) stimulate the adjacent malignant cells to secrete growth factors, chemokines, cytokine, and ECM remodelling enzymes. Whereas, the presence of T lymphocytes is reported to be associated with good prognosis in the case of Cytotoxic CD8+ memory T cells and CD4+ T helper 1 cells. Cytotoxic CD8+ memory T cells are able to eradicate malignant cells at the invasive front of the tumour. Another important component of metastasis pathology is the invasion of the extracellular matrix (ECM). Following detachment from the primary tumour cell mass, the breast carcinoma

cells first adhere to the matrix components. Within the interstitial ECM, breast carcinoma cells must create their passageway for migration through active proteolytic degradation of the ECM components. Such enzymatic degradation is carried out by proteases at the invading edge of the tumour. Different classes of proteases have been evidenced as pro-metastatic proteins including the serine, cysteine, aspartic and the matrix metalloproteinases (MMPs). Proteolysis of ECM facilitates breast cancer cell invasion through their propulsion, or simply, locomotion.

1.6.4. Structure and composition of extracellular matrix (ECM)

The extracellular matrix (ECM) is the non-cellular component present within all tissues and organs, and provides not only essential physical scaffolding for the cellular constituents but also initiates crucial biochemical and biomechanical cues that are required for tissue morphogenesis, differentiation and homeostasis. The ECM is composed of two main classes of macromolecules: proteoglycans (PGs) and fibrous proteins. The main fibrous ECM proteins are collagens, elastins, fibronectins and laminins. PGs fill the majority of the extracellular interstitial space within the tissue in the form of a hydrated gel. Collagen is the most abundant fibrous protein within the interstitial ECM and constitutes up to 30% of the total protein mass of a multicellular animal. Tissue homeostasis is mediated by the coordinated secretion of fibroblast metalloproteinases (MMPs) and. this is counterbalanced by tissue inhibitors of metalloproteinases (TIMPs). The growth factors (GFs) are ECM-bound which differentially modulate cell growth and migration and, when released, comprise part of a tightly controlled feedback circuit that is essential for normal tissue homeostasis.

Upon binding to ECM, degradation products and cytokines, monocytes rapidly differentiate into macrophages. These activated macrophages, in turn, secrete and release multiple GFs, MMPs and cytokines that promote angiogenesis and stimulate fibroblast migration and proliferation.

1.6.5. Signalling pathways in breast cancer metastasis

The stagnated research of breast cancer metastasis is due to lack of effective tools to widely understand the complex network of signalling pathways that drives the multistep process of the metastatic cascade. Many studies have used powerful methodologies to define a gene expression as a part of signal transduction pathway, and have tested its ability in metastatic progression.

Under conditions of low oxygen (hypoxia), which are common in large tumors, the hypoxia inducible factor-1α transcription factor is a target of pathway which activates a pro-angiogenic gene expression. Many other signalling genes were chemically and genetically activated to assess their pathway activity. Two main signalling pathways that are highly targeted namely the JUN/MAPK pathway and the PI3K/AKT pathway [Guille A., *et al.,* 2013]. Some of the pathway-based analyses genes were Src, H-Ras, E2F3, Myc, β-catenin, TCF/Wnt, and transforming growth factor (TGF)-β. Several other studies have recently explored MTDH (Metadherin) signalling, with the NF-κB, phosphoinositide 3-kinase-AKT, Ha-Ras, FOXO3a, and Myc pathways [Blanco M.A. and Kang Y., 2011].

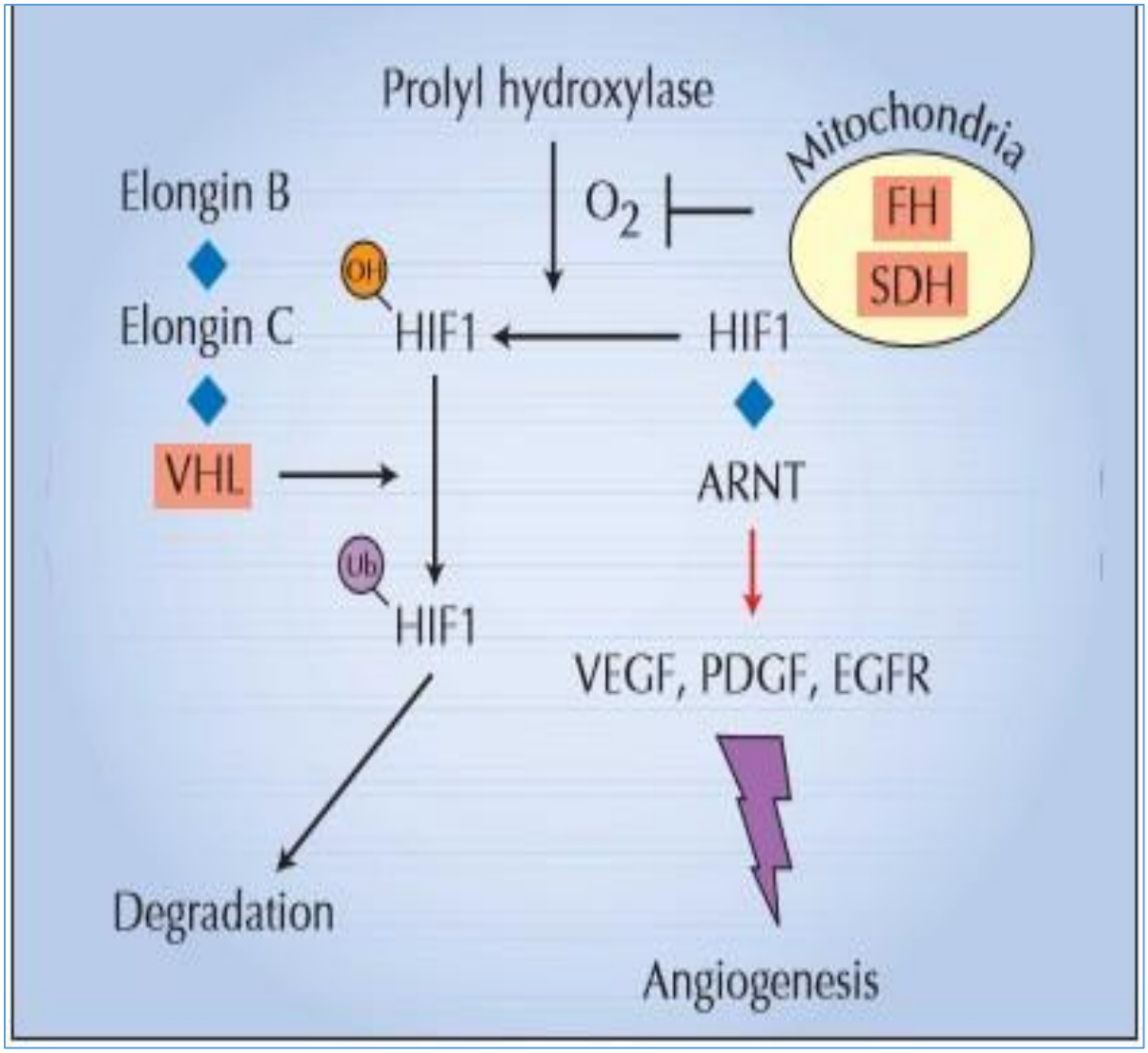


Figure 2: H1F pathway [Frantz C., *et al.*, 2010].

Multiple genes and signalling pathways have been shown to have the ability to influence metastatic progression, but few universal signalling events have been established as truly essential to the metastatic program. The heterogeneity of breast cancers makes them both a fascinating and challenging solid tumor to diagnose and treat. Alan Ashworth and colleagues have discovered the involvement of new mutational genes i.e., PARP and BRCA and their synthetic-lethal approach in tumor treatment

Chapter 2. Literature review

2.1. Cathepsins

As discussed in the introduction part, tumor invasion and metastasis play a vital role in the progression of malignant diseases. The background processes that allow the progression are importantly by the up-regulation of different proteases that degrade the ECM and basement membrane. Proteases are a large group of enzymes that catalyse the cleavage of peptide bonds in proteins. They are subdivided into five categories: metalloproteases, including matrix metalloproteases (MMPs), cysteine proteases, serine proteases, aspartic proteases and threonine proteases. This further releases metallo and serine proteases like MMPs outside the cells, thereby increasing lysosomal proteases, cathepsin B, L, D. Cathepsins belong to papain family and are a class of globular proteases. They are categorized into three groups based on the active amino acid residue i.e., cysteine proteases (B, C, H, F, K, L, O, S, V, W), aspartic (D and E) and serine (G) cathepsins [Nomura T. and Katunuma N., 2005].

Cathepsins A, B, D, E, F, G, H, L, K, S, X and Z were known to be elevated in cancer. Few cathepsins are highly expressed in metastatic tumors and found in the invasion edges of cancers, thereby causing tumor invasion. Cathepsins mediate cancer metastasis, they are highly expressed in invasive tumor, and they increase the motility and invasion of cancer cells. Cathepsins are synthesized as inactive proenzymes and processed to become mature and active enzymes. Some cathepsins are ubiquitously expressed, such as cathepsin B, L, H, and C, whereas the newly found cathepsins K, W, and X are expressed by specific cells and tissues. Furthermore, they mediate dissemination of cancer cell, mediate degradation of ECM and collagen, and also mediate angiogenesis. Cathepsins can activate other proteases, thereby indirectly affecting invasion by

participating in proteolytic cascades. Cathepsins directly cleave components of the ECM, such as laminin, fibronectin, tenascin-C, and type Ⅳ collagen, which leads to limited proteolysis of the ECM. Cathepsins can inactivate cell adhesion proteins by cleaving the cell surface protein E-cadherin, which is the principal component of adherens junctions [Tan G.J., *et al.*, 2013]. Of all cathepsins, CatB, L and D have mitogenic activity in tumor cells and act intracellularly resulting in cancer invasion or matrix degradation. In cancer tissues, cathepsins promote malignancy and so can be a good target for cancer therapy [Tan G.J., *et al.*, 2013]. There is a huge scope in developing proteinase inhibitors those may be good candidates for cancer therapy.

Cathepsin	Elevated in cancer	Location of cancer
A	Yes	Malignant melanoma
B	Yes	Breast carcinomas, melanoma, gastric cancer, lung cancer, colon cancer, ovarian cancer, cervical cancer pancreatic carcinomas, glioblastoma thyroid carcinoma,cholangiocarcinomas, hepatocellular carcinomas, bladder caner
C	Unclear	
D	Yes	Thyroid carcinomas, squamous cell carcinoma, renal cell, carcinoma, glioma brain tumors, laryngeal carcinoma, breast cancer, lung cancer, ovarian carcinoma
E	Yes	Pancreatic ductal adenocarcinoma, gastric cancer
F	Yes	cervical carcinoma
G	Yes	Breast cancer
H	Yes	Breast carcinoma, colorectal cancer, melanoma, head and neck carcinoma, glioma, prostate cancer
L	Yes	Breast cancer, lung cancer, gastric cancer, colon cancer, head and neck carcinomas, melanomas, gliomas, ovarian cancer, pancreatic cancer
K	Yes	Gastric cancer, squamous cell carcinoma, basal cell carcinoma, breast tumor, lung cancer, melanomas, prostate tumors, renal tumor
O	Unclear	
S	Yes	Astrocytoma, gastric cancer, hepatocellular carcinomas, glioblastomas, melanoma, gastric cancer, pancreatic islet cell cancer
V	Unclear	
W	Unclear	
X	Yes	Prostate cancer, gastric cancer, malignant melanomas, prostate cancer, lung tumors, breast cancer, colonrectal cancer
Z	Yes	Melanomas, gastric cancer, hepatocellular carcinomas, pancreatic carcinomas

Figure 3: Cathepsin family members expressed in cancers [Tan G.J., *et al.*, 2013].

2.2. Cathepsin D

CatD and CatE are aspartic proteases. The human CatD gene is located on chromosome 11 in the region p15 of H-*ras* oncogene. There are three molecular forms of CatD found in the cell namely the precursor (procathepsin D), the intermediate single-chain and the mature double-chain. Procathepsin D is found in Golgi complex and is enzymatically inactive. The other two forms are enzymatically active and found in lysosomes and endosomes. Procathepsin D secreted by the tumor cells increases progression of breast and lung cancer. Parallelly, CatD intense expression in high-grade carcinomas acts as important marker for invasive potential and aggressive behaviour of tumors [Saraswat-Ohri, S. and Vetvicka, V., 2011].

2.2.1. Human CatD: Lysosomal targeting and drug design

CatD (3.4.23.5) have a total of 6 reported crystal structures (**1LYA, 1LYB, 1LYW, 4OBZ, 4OC6**). The mature form of CatD exists in two-chain form due to post-translational cleavage event. The 2.5 Å resolution structures of native CatD and its complex with Pepstatin (aspartic protease inhibitor) were made into crystal structure with PDB (Protein Data Bank) code: 1LYB. CatD consists of three distinct regions that are typical for aspartic proteases namely an N-terminal domain (residues 1-188), a C-terminal domain (residues 189-346), and an interdomain, anti-parallel β-sheet composed of N-terminus (residues 1-7), the C-terminus (residues 330-346), and the interdomain-linking residues. Several structural features distinguish CatD from other aspartic proteases. The N and C domains contribute to the catalytic active site of CatD i.e., Asp-33 and Asp-231 [Saraswat-Ohri, S. and Vetvicka, V., 2011]. The central statine hydroxyl group occupies the position of water molecule and interacts with the two active-site aspartate residues in the enzyme. In the Pepstatin complex, the subsites S3, S2, S1 and S2’ are partially filled and the active

site pocket can accommodate bulkier inhibitor substituents. The Pepstatin-cathepsin D complex is stabilized by numerous hydrogen bonds between backbone atoms of the inhibitor and both main chain and side chain atoms of the enzyme [Majer P., *et al.*, 1997]. The main amino-acid interactions of the complex were Asp 33, Asp 231, Ser 80, Ser 235, Gly 35, Gly 79, and Gly 233.

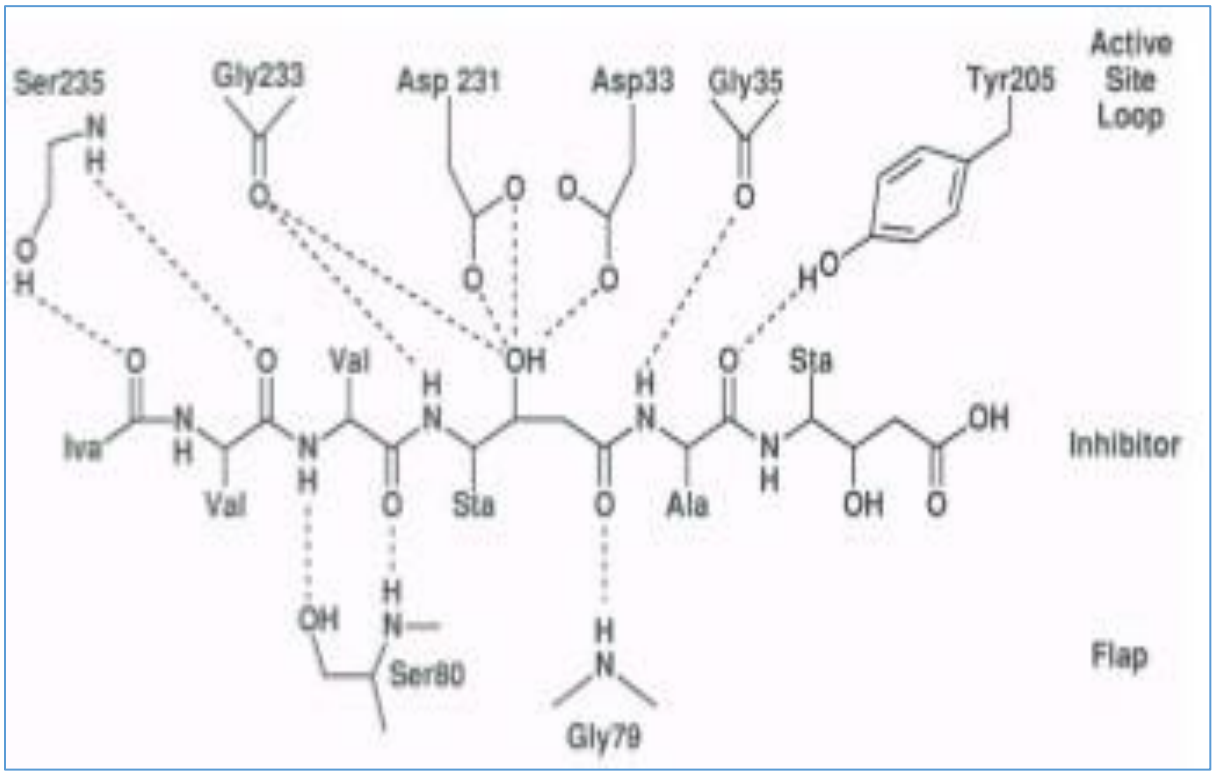


Figure 4: Schematic hydrogen-bonding diagram for Pepstatin bound to CatD [Majer P., *et al.*, 1997].

Residues around active site are well defined in CatD structures and the active site is similar to many other aspartic proteases according to the homology studies. The active site cleft has similarities with renin, pepsin and chymosin. There are 30 lysine and arginine residues on the surface of CatD, 20 on renin and only three on pepsin [Metcalf P. and Fusek M., 1993].

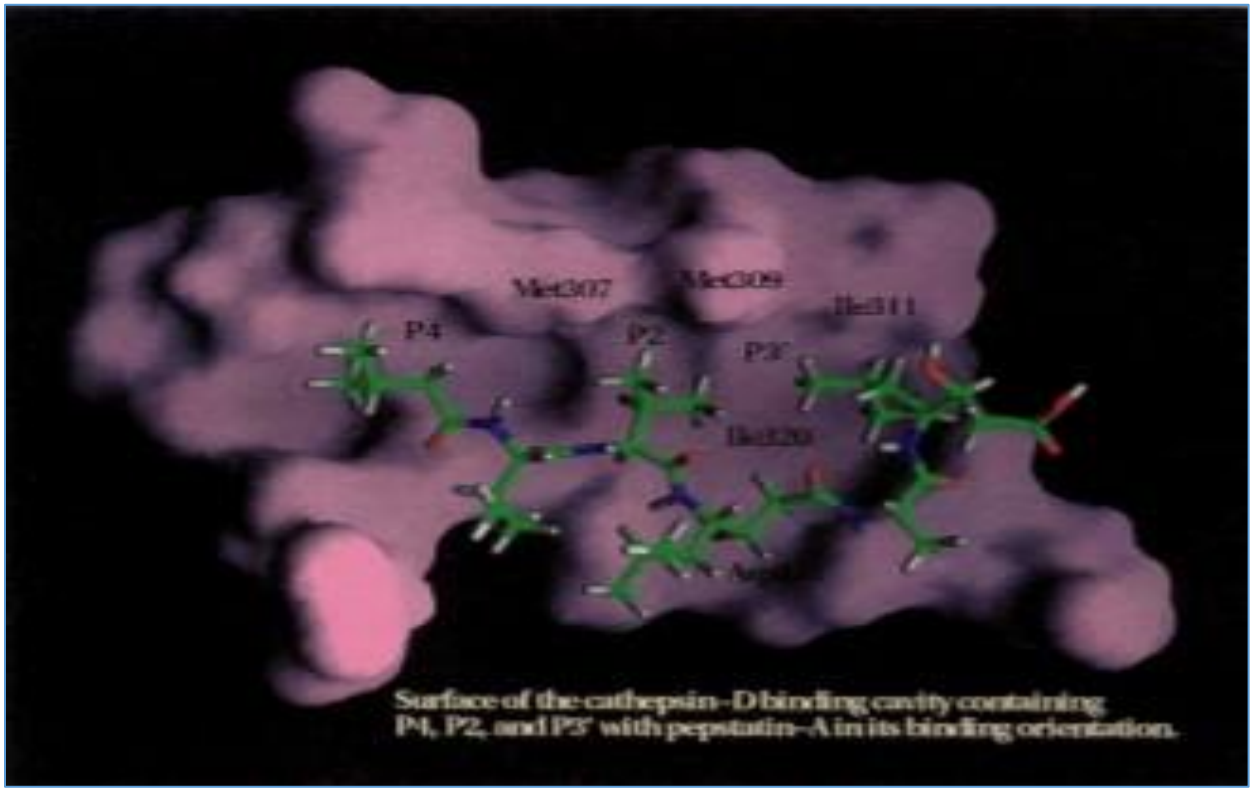


Figure 5: CatD-Pepstatin complex [[Metcalf P. and Fusek M., 1993].

Differences in the subsite structures of human cathepsin D and related human aspartic proteases may be exploited for drug design. Comparison of the crystal structures of cathepsin D with human renin reveals that the substrate binding cleft in cathepsin D is wider overall than in renin. Cathepsin D has smaller residues in the S2 (Gly-79 vs. Ser-76) and S4 (Leu-236 vs. Tyr-220) subsites, which leads to larger subsite volumes, contribute more to the potency of inhibitors for cathepsin D than for pepsin and renin. Aspartic peptidases can accommodate up to 9 amino acids long substrate into their active site binding cleft.

2.2.2. Cathepsin D as a prognostic biomarker

Breast cancer can be a terminal diagnosis due to the fact that, despite best attempts, early detection has failed and a diagnosis is made at a late stage. As such, there is intense interest in identifying serum (or plasma) biomarkers that may delineate the presence, absence or extent of disease when tumors cannot be palpated or visualized. These biomarkers could be potentially useful for prognosticating, but they are conceivably valuable in surveillance [Abott D.E., *et al.*, 2010].

CatD is a well-characterized lysosomal hydrolase, serving critical functions in intra-cellular protein degradation. Under normal conditions, less than 10% of CatD is secreted as pCD into the extracellular milieu and is also detected in the serum. In mammary malignancies, however, CatD is aberrantly over-produced and hyper-secreted by both malignant and extra-tumoral cells such as macrophages and fibroblasts [Abott D.E., *et al.*, 2010; Duffy M.J., et al., 1991]. The overexpression of CatD in breast cancer cells leads to the hypersecretion of the 52-kDa pCD into the extracellular environment. In addition, the increased pCD affects tumor growth (mitogenic effect), increases angiogenesis and tumor metastases. Large studies and one meta-analysis found that pCD/ CatD level in tumor homogenate represents an independent prognostic factor [Vetvicka V. and Fusek M., 2012]. Intracellular CatD in breast cancer cells has been shown to have 8–16 times more activity than normal mammary cells.

CatD expression became of clinical interest in breast cancer as a potential marker of estrogen responsiveness. The radiometric immunoassays were used to measure the amount of CatD in tumor cytosols. The clinical findings suggest the usefulness of CatD as a prognostic marker to identify women with a high risk of relapse. Foekens et al stated that cytosolic CatD levels are predictive of early relapse and survival not only in node-negative breast cancer patients but also in node-positive, premenopausal and post-menopausal subgroups of breast cancer patients. The clinical

findings states that CatD is partially or wholly responsible for the aggressive nature of tumors which depicts the increasingly urgent development of its specific inhibitors [Westley B.R. and May F., 1999].

A clinical study diagnosed 226 early breast carcinoma patients and determined the concentration of three CatD forms (52-, 48- and 34-kDa proteins) by a radioimmunoassay. On the basis of differences in cathepsin D levels either within an ER (-)/PR(-) phenotype or between this and either ER(+)/PR(+) or ER(+)/PR(-) phenotypes, a concentration of 39 pmol/mg was determined as the cutoff value for distinguishing estrogen-regulated cathepsin D expression. Estrogen-regulated cathepsin D expression was recognized as a high-risk biomarker for low-risk (histological grade I) breast cancer patients and as a low-risk biomarker for high-risk patients [Markicevic M., *et al.*, 2013]. Other study summarizes that serum from patients with invasive disease possessed accentuated serum CD activity. Furthermore, patients with metastatic disease have serum CD activity that exceeds those of patients with limited, local disease [Abott D.E., *et al.*, 2010] CatD is unique in association with breast cancer which has a high risk of metastasis. Almost all of the clinical findings has shown that CatD concentration is an independent factor for identifying breast cancer with a poor prognosis. High CatD concentration is associated with shorter metastasis-free survival. The predictive value of CatD was much greater in axillary lymph node-negative patients than in node-positive patients [Spyratos F., *et al.*, 1989].

2.2.3. CatD and its post translational processing

A wide variety of proteases have been implicated throughout the development of the mammary gland, but misregulated expression of some of the same proteases has been implicated in the development and progression of breast cancer. Understanding the relationship between the physiological functions of particular proteases during development and their pro-tumorigenic

functions during cancer is critical for developing appropriately targeted cancer therapeutics. CatD is constitutively expressed in nearly all cells, where it is trafficked to endosomes and lysosomes, and serves a role as a major acid hydrolase in intracellular protein turnover. CatD has been functionally linked to a number of mechanisms involved in cancer progression, including proliferation, invasion, metastasis and angiogenesis. In various breast cancer cell lines, the trafficking of pCD through alternative pathways has been observed, resulting in differing degrees of extracellular secretion vs. lysosomal accumulation.

Glycosylation occurs only in cathepsin D, cathepsin E. and memapsin. Important feature of aspartic peptidases is the structural changes accompanying the transformation of newly translated proteins into zymogens and subsequently into active enzymes. All human aspartic peptidases are synthesized as pre-pro-enzymes. In the first proteolytic posttranslational modification, the signal "pre" sequence is cleaved off. This signal sequence is usually 20 amino acids long and is removed within the rough endoplasmic reticulum. Zymogens, formed by this first posttranslational processing, then travel to a place of their action and only when they reach the particular compartment or tissue are they activated into the fully active enzymes. The activation involves cleavage off of the pro-sequence, which is also called the propeptide or the activation peptide. CatD needs the catalysis of other lysosomal peptidases to reach full maturation.

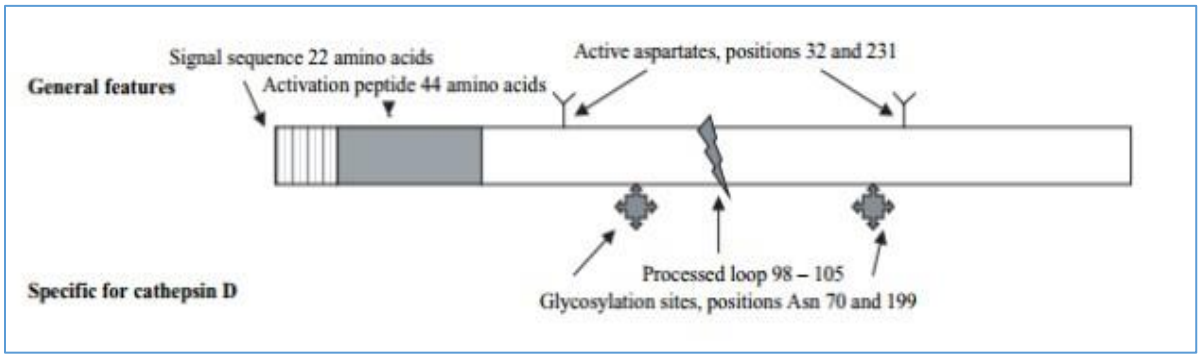


Figure 6: The catalytic active site positions of CatD [Tedone T., *et al.*, 1997]**.**

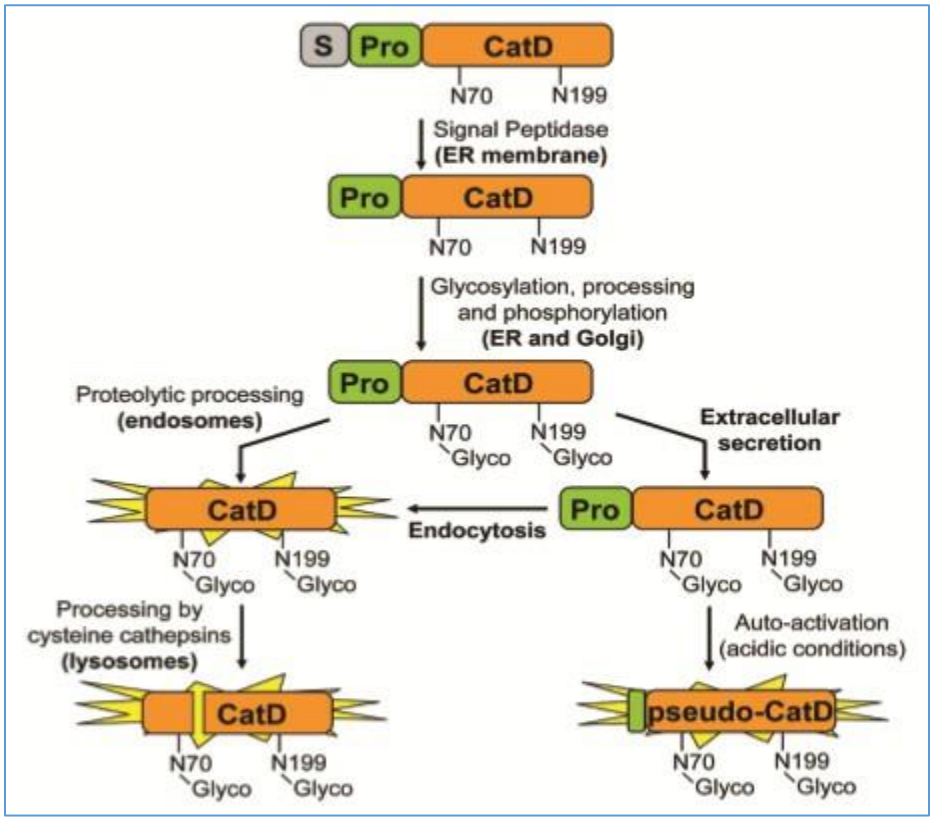


Figure 7: Structure n Processing and activation of CatD [Fusek M. and Vetvicka V., 2005].

Similar to other aspartic peptidases, mature CatD is formed after several proteolytic changes occurring in different cellular compartments. pCD originates after the first proteolytic cleavage involving the removal of the signal peptide by signal peptidases. The molecule of pCD has a molecular weight of 52 kDa. It consists of 392 amino acids and is, under normal conditions, glycosylated at two Asp residues. The activation peptide is a 44 amino acid long peptide which is folded in and above the active site cleft. pCD is transported to lysosomes and in the acidic milieu undergoes further proteolytic processing forming CatD consisting of one chain of 348 amino acids.

2.2.4. Regulation of CatD in breast cancer invasion and its therapeutic importance

Many clinical studies have correlated high cytosolic CatD levels in the primary tumor to high probability of metastasis and poor prognosis, suggesting that its cytosolic concentration could be

useful as a prognostic marker of aggressive breast tumors. Increased CatD activity in the cytosol of malignant breast carcinomas suggests that it may play an active role in metastatic spread. A study using a hormone-independent varient of MCF-7 cells having an increased metastatic potential in nude mice concluded that the transition from low to high invasive phenotype was most strongly associated with an increased expression of CatD in the highly invasive varient. Many hypotheses have been raised concerning the mechanism by which CatD may function. These include direct stimulation of the tumor cells, facilitated release of growth factors, and the degradation and reworking of the extracellular matrix to permit transmigration of the tumor cells. Tumors are known to have an even more acidic intercellular microenvironment than normal tissues that is probably created and maintained by specific transporters on the plasma membrane. The molecular size intercellular microenvironment where the cell is in contact with the extracellular matrix probably reaches extremely low pH levels and thus facilitates the action of the acidic proteases. The lowering of extracellular pH or the inhibition of transporters that alkalinize the cell has been demonstrated to increase the secretion of CatD in tumor cells [Tedone T., *et al.,* 1997].

It is notable that all the studies that have found CatD to be a marker of poor prognosis have measured the protein levels in tumor cytosols. This means that the stromal component of the tumor contributes to the overall level of CatD measured. Two immunohistochemical studies that examined CatD levels in breast tumors showed that if the tumor cells alone were scored for cathepsin staining then high levels of the enzyme were found to be a marker of good prognosis, with increased overall and disease-free survival [Johnson M.D., *et al.,* 1993].

Cell line	Invasion[a]	Nude mouse[a]	Cathepsin D secretion[b]
MCF-7-2	++	P	100
MCF-7-ADR	+++	P	0
MDA-MD-231	+++++	LI	2
MDA-MB-435	+++++	M[c]	0
MDA-MB-435s	++++[d]		25
MDA-MB-468	+	P	48
SK-Br-3	+	N	98

Figure 8: The invasiveness, tumorigenicity and CatD secretion of 7 breast cancer cell lines [Johnson M.D., *et al.*, 1993].

Several approaches, such as immunohistochemistry, in situ hybridization, cytosolic immunoassay and Northern and Western blot analyses have indicated that in most breast cancer tumors CatD is over expressed 2- to 50-fold compared to its concentration in other cells such as fibroblasts or normal mammary glands. The mechanism of CatD action on metastasis could involve either its interaction with the plasma membrane M6P/IGF-II receptor or its catalytic activity. CatD in the acidic vesicles is able to intracellularly digest many types of proteins, including proteins of the engulfed ECM, and thus provide food, amino acid supply and space for invasive breast cancer cells. CatD also has the potential to degrade or activate many important molecules able to play a role in one of the steps leading to metastasis. CatD can thus trigger the proteolytic cascade leading to ECM degradation by activating pro-cathepsin B and/or degrading cystatins. CatD facilitates the production of nutrients such as amino acids from the ECM and could also favour cancer cell proliferation via specific inactivation of growth inhibitors and activation (or liberation) of growth factors. CatD liberates several growth factors and cytokines are entrapped in an inactive form in the ECM, including pro-TGFβ, which may facilitate tumor cell invasion, and factors of the fibroblast growth factor (FGF) family which display a high angiogenic activity [Garcia M., *et al.*,

1996]. There are many other invasive factors whose malfunction contributes to several types of malignancies including breast cancer. Recent research in implications for matrix-dependent breast cancer cell invasion and metastasis stated the role of c-Myb in stimulating cell proliferation, suppressing differentiation, and apoptosis. c-Myb is involved in the regulation of epithelial-to-mesenchymal transition (EMT) and invasion in neuroblastoma, colon carcinoma, and embryonic kidney cells through the upregulation of the transcription repressor Slug. It is also reported to regulate the invasive behaviour of breast cancer cells in a matrix-dependent manner, possibly via a novel signalling axis causing the deregulation of MMP1, MMP9, and cathepsin D expression [Knopfova L., *et al.*, 2012].

In case of involuting mammary gland, Cathepsins B, D and L are elevated at the reversible stage of involution and remain high until 96 hrs post weaning. At the reversible phase of involution, CatD's cleavage does not proceed beyond the generation of the single chain active enzyme. These precise and timely post-translational modifications prompted us to speculate on CatD's significance in the involution process and re-population of the mammary tissue with adipocytes. Furthermore, CatD was found to preferentially cleave Histone 3(H3) at its amino terminal tail. It is likely that this post-translational modification or ''histone code'' is the signal for its cleavage by CatD and initiation of irreversible stage of involution. From functional perspective, CatD is critically involved in breast cancer progression and metastasis, deregulated synthesis and elevated secretion of CatD are hallmarks of cancer [Khalkhali-eliis Z., *et al.*, 2014].

2.2.5. Importance of CatD in Triple Negative Breast Cancer

Triple-negative breast cancer (TNBC), an aggressive variant of breast cancer characterized by lack of expression of the estrogen receptor (ER) , progesterone receptor (PR) and the human epidermal growth factor receptor 2 (HER-2) accounts for 15-26% of breast cancer. Currently there is no

molecular-based targeted therapy for TNBC, and unfortunately only approximately 20% of these tumours respond well to standard chemotherapy. Thus, developing improved treatments for TNBC is one of the highest priorities of current breast cancer research. Numerous agents are in various phases of clinical development, including several different poly (ADP-ribose) polymerase inhibitors, JAK kinase, and EGFR inhibitors as well as "revived" classical chemotherapeutic agents such as platinum salts [Polyak K., 2011]. The breast cancer patients with higher histologic grade, larger size, high ERK protein expression, low E-cadherin expression and Ki-67 staining may have a tendency toward local and visceral metastases. CatD and Ki-67 index were reported factors strongly associated with outcome of TNBC patients. Analysis demonstrated that the combination of 3 prognostic factors, including positive for ≥4 lymph nodes , high CatD expression and high Ki-67 index, provided a strong prognostic power to differentiate the patients with worse outcome [Huang L., *et al.,* 2013]. Radioimmunoassay was used demonstrate that high CatD levels in node-negative breast cancer patients are related to a worse prognosis [Kute T.E., *et al.*, 1992].

MDA-MB-231, a triple negative breast cancer cell line were able to survive and secrete proteolytically active CatD. In addition to its mitogenic activity as a ligand at a neutral pH, CatD can also be activated as a protease when breast cancer cells are exposed to an extracellular pH of 6.6. The MDA-MB231 cells are more aggressive and more efficient than MCF7 cells in spontaneously acidifying an extracellular milieu. They constitutively overexpress CatD corresponding to a basal-type breast cancer [Maynadier M., *et al.,* 2013].

2.2.6. Reported Peptide CatD inhibitors till date

The inhibitor is a compound showing chemical affinity toward an enzyme or bearing structural similarity to a substrate which binds to the catalytic site of an enzyme and forms an inactive enzyme-inhibitor complex. Most cathepsin D inhibitors are synthetic, peptide and polypeptide

produced by microorganisms, plants and animals. Specific anti-cathepsin D antibodies also have an inhibitory effect.

2.2.6.1. Synthetic inhibitors

Synthetic inhibitors of CatD are micromolecular organic compounds esterifying the carboxyl group of the Asp33 or Asp231 residue in the catalytic site. The derivatives include diazoacetyl-glycine ethyl ester, diazoacetyl-phenyloalanin methyl ester, diazoacetyl-2,4-dinitrophenyl-ethylenediamine and other diazol compounds. Other moieties which can also act as CatD inhibitors are ditiophosgen and 2,2 dichloro-1,3-ditiocyclobutanone, and methyl blue and tetranitrometane. These compounds were known to inactivate pepsin as well.

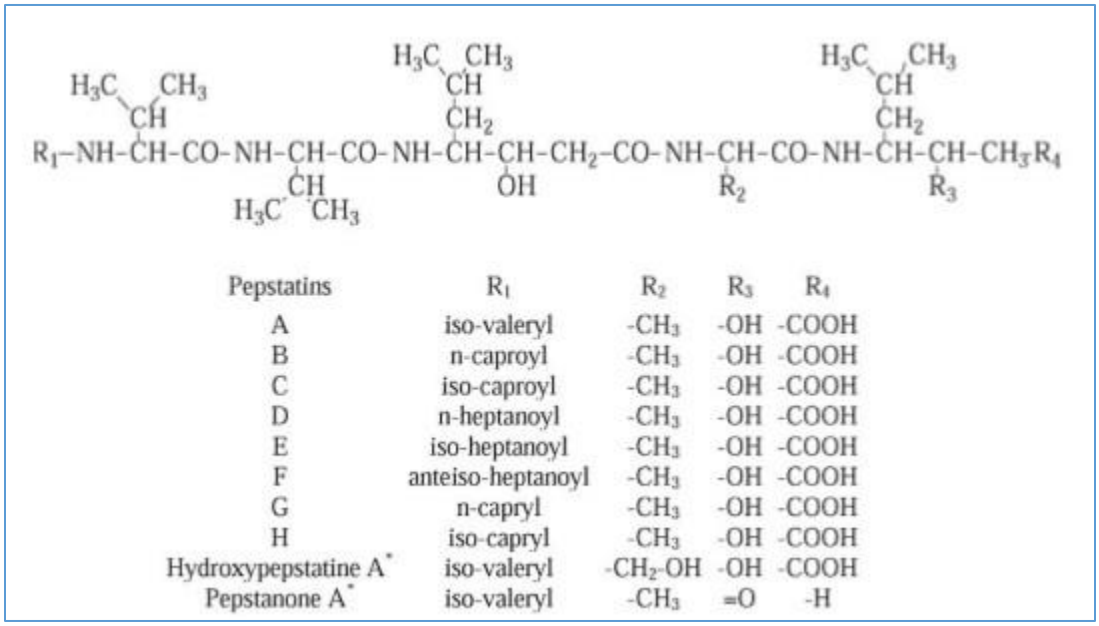
2.2.6.2. Synthetic substrate analogues

CatD activity is inhibited by structural analogues of synthetic substrates. These are oligopeptides containing at least five amino acid residues in the molecule and having L-amino acid replaced by D-amino acid. CatD vulnerability to hydrolysis and degree of inhibition by these analogues depend on the number and location of D-amino acids in the molecule.

2.2.6.3. Inhibitors derived from microorganisms

Inhibitors of aspartyl proteases are synthesized by *Streptomyces testaceus, Streptomyces argenteolus* and other species of the genus *Streptomyces*. These inhibitors are called pepstatins as they inhibit the activity of pepsin. They also inhibit the activities of CatD, CatE, renin, pseudorenin, aspartyl proteases produced by microorganisms and plants [Gacko M., *et al.*, 2007]. Pepstatins are pentapeptides containing non-amino acid constituents, typical amino acids and atypical amino acid. They are composed of a sequence of 1 acyl radical, 2 valin residues, 1 statin

residue, 1 alanine residue and 1 statin residue, and are shortly called AcL-Val-L-Val-L-Sta-L-Ala-L-Sta.



Pepstatins	R_1	R_2	R_3	R_4
A	iso-valeryl	$-CH_3$	-OH	-COOH
B	n-caproyl	$-CH_3$	-OH	-COOH
C	iso-caproyl	$-CH_3$	-OH	-COOH
D	n-heptanoyl	$-CH_3$	-OH	-COOH
E	iso-heptanoyl	$-CH_3$	-OH	-COOH
F	anteiso-heptanoyl	$-CH_3$	-OH	-COOH
G	n-capryl	$-CH_3$	-OH	-COOH
H	iso-capryl	$-CH_3$	-OH	-COOH
Hydroxypepstatine A*	iso-valeryl	$-CH_2$-OH	-OH	-COOH
Pepstanone A*	iso-valeryl	$-CH_3$	=O	-H

Figure 9: Structure of pepstatins, hydroxypepstatin and pepstanones [Gacko M., *et al.,* 2007].

The inhibition of aspartyl protease action by pepstatins depends to a large extent on the presence of acid residue in their structure. Pepstatin also binds to pCD and inhibits its auto activation [Agarwal N.S and Rich D.H., 1986].

2.2.6.4. Polypeptide plant inhibitors

Polypeptide inhibitors of CatD were collected from plant organs like seeds, bulbs, fruits and at times from leaves, roots and flowers. Well defined amino acid residues constituting the inhibitory reactive site, complementary to the amino acid residues of the protease catalytic site, are found on the surface of the polypeptide inhibitor molecule. The inhibitor forms a permanent complex with an in the 1:1 ratio and blocks the action of cathepsin due to hydrophobic and ionic hydrogen bonds of the reactive site of the inhibitor, the catalytic site of cathepsin and their closest environment.

2.2.6.4.1. Potato inhibitors of CatD

Potato bulbs contain six or eight polypeptide inhibitors of CatD and show a considerable similarity in the amino acid composition, sequence. Potato inhibitors possess two inhibitory sites: one binding CatD and the other binding trypsin, and they do not inhibit the activity of pepsin. Transcription of the potato inhibitor genes is up-regulated by the plant hormone abscisic acid and the fatty acid derivative jasmonic acid. Both of these have been implicated in signal transduction pathways mediating responses to wounding and to invasion [Cater S.A., *et al.*, 2002]. The CatD inhibitor from seeds of *Vicia sativa* L. does not penetrate into fibroblasts cells, but it penetrates inside breast cancer cells [Roszkowske J.W., *et al.*, 2004]. JIP21, is another defensive protein characterized from tomato with a novel function as CatD inhibitor. The N-glycosylation site has been found to be common in potato and tomato inhibitors. Glycosylation participates in important processes, such as maintenance of protein conformation and solubility, stabilization of the polypeptide against uncontrolled proteolysis, intracellular sorting and externalization of glycoproteins, or mediation of its biological activity [Lison P., *et al.*, 2006].

2.2.6.4.2. CatD inhibitors produced from pumpkin fruit

Cathepsin D inhibitor, which also inhibits the activity of pepsin and acid protease from fungus *Glomerella cingulata* has been isolated from pumpkin fruit. It is composed of 96 amino acid residues and occurs in three molecular forms differing slightly in the amino acid sequence.

2.2.6.4.3. Inhibitors derived from other plants

Cathepsin D inhibitors can be found in the seeds of lentil, vicia, wheat, barley, tomato leaves, tobacco leaves, aubergine fruit (*Solanum melongea*) and from trailing nightshade leaves (*Solanum dulcamara*) and black nightshade (*Solanum nigrum*) [Gacko M., *et al.*, 2007].

2.2.6.5. Inhibitors synthesized by lower animals

Cathepsin D activity is inhibited by equistatin, a protein inhibitor isolated from Actinia equina. Equistatin inhibits the activity of cystein proteases: papain, cathepsin B, cathepsin Land cathepsin D, but not the activity of pepsin, chymotrypsin or HIV protease.

2.2.7. Reported non-peptidic CatD inhibitors till date

Structure-based design coupled with combinatorial chemistry has been successfully used to design nonpeptide inhibitors of CatD [Huo S., *et al.,* 2002].

2.2.7.1. Acyl guanidine derivatives as potent CatD inhibitors

The acylguanidine moiety has been reported as specific binder forming H-bond interactions to the two catalytic aspartic acids of the structurally related aspartic protease.

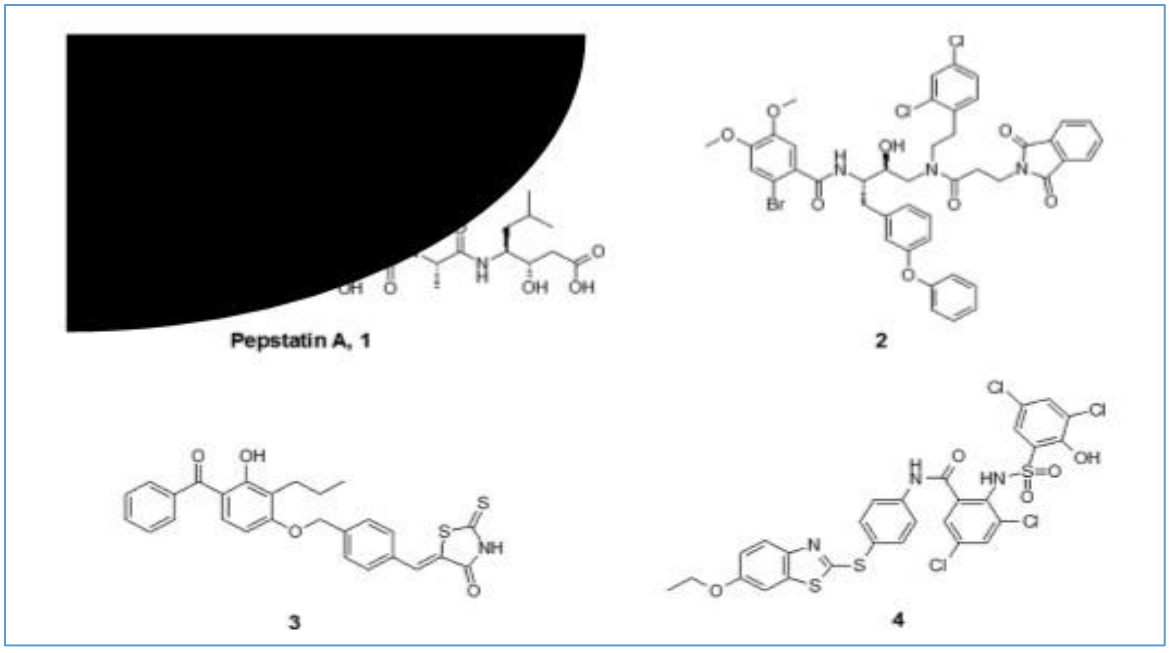


Figure 10: Reported Cathepsin D inhibitors [Gradter U., *et al.,* 2014].

The hydroxylamine inhibitor 2 formed several H-bonds within CatD binding cleft between the residues i.e., statine OH-group and the catalytic Asp 231/Asp 33, with both amide carbonyl O-atoms and Gly 79/ Ser 80 of the flap region and both methoxy O-atoms and Ser 235. Many other compounds were synthesized as additional substituents to inhibitor 2 such as derivatives of phthalimide or 3, 4-dimethoxybenzyl. More than 40 analogues were synthesized based on the H-bonding of the compounds with CatD at the catalytic active site. This lead to the successful derivatization of CatD inhibitors with nanomolar range activity [Gradter U., *et al.,* 2014].

2.2.7.2. Novel statine motifs as CatD inhibitors

Plasmepsin I and II are malarial aspartic proteases which are structurally similar to CatD enzyme. The statine motifs synthesized against them were also tested for inhibitory activity and selectivity over CatD.

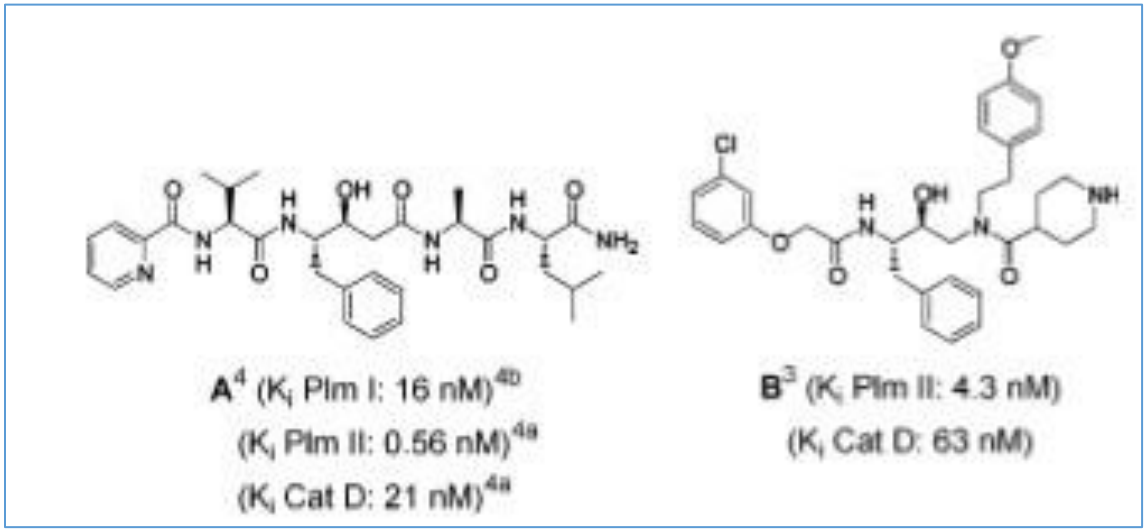


Figure 11: Statine motifs as CatD inhibitors.

This inhibitor also shows good selectivity toward CatD (Ki > 4000 nM). High selectivity toward CatD is also seen for other synthesized inhibitors. It is apparent that CatD selective inhibitors can be designed by reframing form utilizing the S3 binding pockets and by use of appropriate

arylbenzoic acids or bulky alkyl acids as capping groups for the N-terminus of the statine-like core. From examination of the X-ray crystal structures of CatD and Plasmepsin II, both in complex with pepstatin A, it can be seen that amino acid Ile 290 in plasmepsin II has a larger amino acid Met 307 in the corresponding position in CatD and that Leu 292 in proximity to the S2 pocket in plasmepsin II has the amino acid Met 309 in the corresponding position in CatD.

2.3. Novel strategies for Triple Negative Breast Cancer and scope of CatD inhibitors

The overall poor prognosis of patients with TNBC and their tendency to relapse with distant metastases indicate a definite need for effective systemic therapies for this disease. Chemotherapy is currently the mainstay of systemic treatment for these patients because hormonal and Her2-directed therapies are not effective. Despite high proliferation rates and TP53 mutation rates, TNBC seems not to show increased sensitivity to taxanes (compared with non-TNBC) in the metastatic setting [Carey L., *et al.,* 2015]. It has been reported that TNBC had the highest response with combination of cyclophosphamide, methotrexate and 5-fluorouracil.

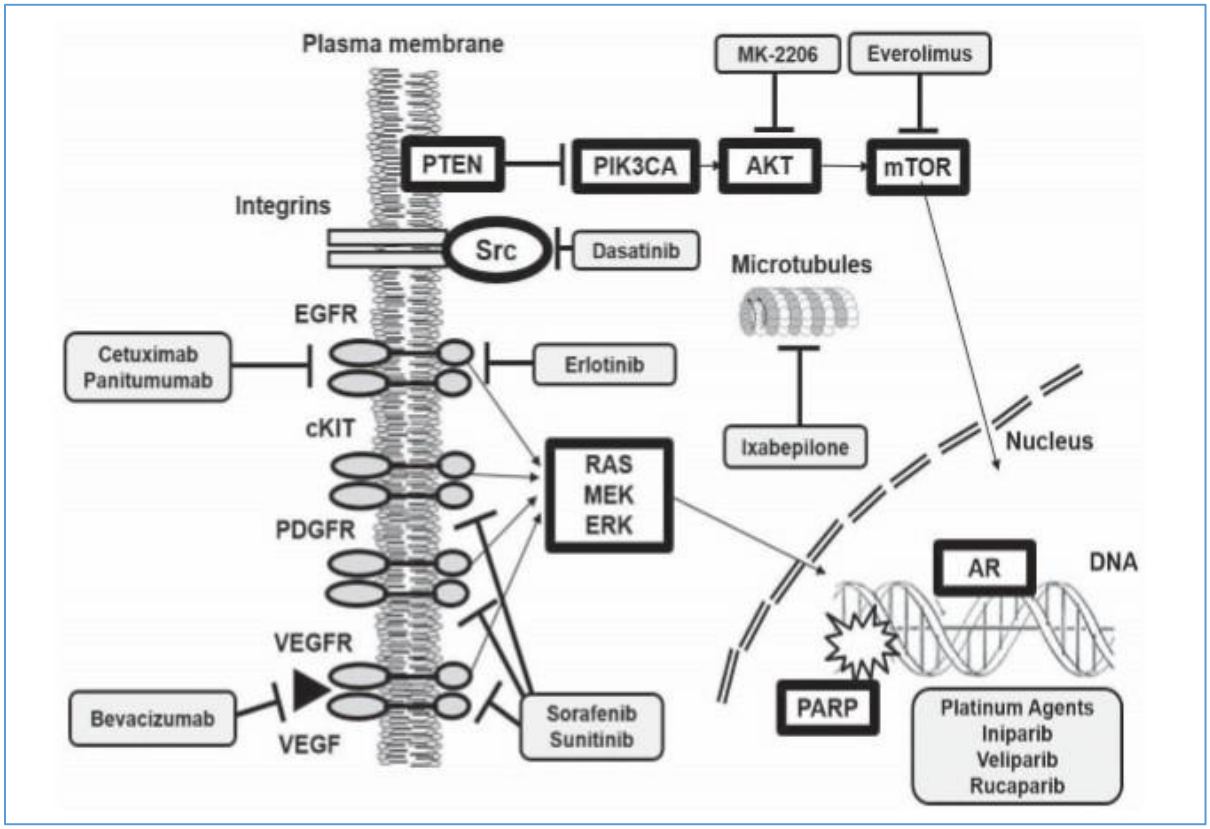


Figure 12: Neoadjuvant therapeutic strategies and drugs under investigation for triple negative breast cancer [Carbognin L., *et al.*, 2015].

There are many trials on neoadjuvant therapies to treat TNBC patients. Among them, PARP inhibitors are producing encouraging results. PARPs are a family of nuclear enzymes involved in the detection and repair of DNA damage. PARPs are critical enzymes in cell proliferation and are differentially upregulated in many cancers including TNBC and BRCA1-associated and BRCA2-associated tumors. Few PARP inhibitors are Olaparib, Iniparib, Veliparib and Rucaparib. Second set of inhibitors are angiogenesis inhibitors, which are actively investigated in patients with TNBC following positive results from the E2100 and AVADO trials of bevacizumab (VEGF inhibitor) in combination with a taxane in patients with metastatic breast cancer [Carey L., *et al.*, 2010]. Other antiangiogenic agents are currently being evaluated in clinical trials in the neoadjuvant setting of

TNBC. These include agents targeting VEGF and platelet derived growth factor (PDGF) receptor tyrosine kinase, including sunitinib and sorafenib. Epidermal growth factor receptor is overexpressed in approximately 60-70% of TNBC and thus it theoretically represents one of the possible targets for its treatment. In this regard cetuximab, a chimeric monoclonal antibody binding the extracellular domain of the EGFR, is well evaluated as single agent or in combination with other targeted therapies or chemotherapy in several trial in metastatic TNBC [Carbognin L., *et al.,* 2015]. A recent study demonstrated the efficacy and acceptable toxicity profile of panitumumab, an anti-EGFR antibody, combined with an anthracycline-taxane-based chemotherapy for operable TNBC. Activity and expression of the tyrosine kinase c-Src are frequently increased in breast cancers, and is associated with increased motility and invasiveness. The anti tumor activity of the dual abl/src kinase inhibitor dasatinib, was modest when given as monotherapy to heavily pre-treated patients with TNBC [Carey L., *et al.*, 2010].

There are many other pathway strategies in treating TNBC. The intracellular signalling PIK3CA/Akt/mTOR pathway, common activated in breast cancer, plays a key role in regulating multiple cellular processes including proliferation, survival, and differentiation [Clark O., *et al.,* 2014]. This strategy is being evaluated in TNBC, in which the PTEN loss-of function mutation or reduced expression has been observed in around 30% and these have been found to be associated with activation of Akt. Preclinical data suggest that combination treatment with rapamycin and cyclophosphamide produced a synergistic effect between the two drugs in term of response in nude mice bearing tumor xenograft of the TN MDA-MB-231 breast cancer cells. Currently, other agents that target one or more components of PIK3CA signalling pathway are being evaluated in the neoadjuvant setting of TNBC such as the AKT inhibitor MK-2206. The range of genetic abnormalities seen in TNBCs has opened the door to other therapeutic strategies, many of which

are being evaluated in clinical trials. These include agents targeting kinases, including mTOR, androgen receptor, TGF-β, and the TRAIL receptor [Carey L., *et al.*, 2010].

2.4. New functional aspects of CatD

CatD has been studied over last three decades, mainly for its role in cancer development and as an independent tumor marker. Numerous physiological functions of CatD have been suggested, based on its ability to cleave structural and functional proteins and peptides. The other new functional aspects are metabolic degradation of polypeptide hormones and growth factors, activation of enzymatic precursors, processing of enzyme activators and inhibitors, brain antigen processing and regulation of programmed cell death [Benes P., *et al.*, 2008].

2.4.1. Role of CatD in Neuronal ceroid lipofuscinosis (NCL)

Inactivation of CatD homologue in drosophila or naturally occurring mutation of CatD gene in sheep and American bulldogs, resulted in progressive neuronal accumulation of auto fluorescent storage material lipofuscin and neurodegeneration. The mutations in CatD significantly reduced CatD activity and/or affect pCD stability, post translational processing and intracellular targeting resulting in NCL development.

2.4.2. Role of CatD in skin development and function

Altered expression of CatD in the skin has been reported during wound healing and psoriasis. CatD protein expression and enzymatic activity increases during epidermal differentiation. Morphology of stratum corneum in CatD- deficient mouse was impaired resembling to the human skin diseases. In addition, pCD enhanced proliferation and regeneration of keratinocyte cell line HaCaT *invitro*. CatD enzymatic activity was also known to be involved in the processes of epithelial differentiation, epidermal barrier function and wound healing.

2.4.3. CatD in association with neurodegenerative disorders

Human CatD deficiency is responsible for an autosomal recessive neurodegenerative disease that can manifest in early childhood. Several animal models indicate that CatD deficiency leads to fatal neurodegeneration. CatD-deficient mice and sheep exhibit cerebral atrophy and show neuronal accumulation of auto fluorescent storage material mainly granular osmiophilic deposits. The CatD deficient patient described here has similar granular deposits in Schwann cells derived from skin biopsy material and shows cerebral and cerebellar atrophy on cranial MRI scans [Steinfeld R., *et al.,* 2006]. CatD has comparable functions in humans and other mammals and seems to play a key role in the homeostasis of neuronal structures. However, the variation in phenotypic presentation of CatD deficiency between mice, sheep, dogs, and humans points to additional species-specific differences in CatD function and thus disease mechanism.

2.4.3.1. Alzheimer's disease (AD)

And in neuro related aspects, CatD is noted to play vital role in Alzheimer's disease (AD). In AD, senile plaques and tangles show abundant CatD immunoreactivity. The CatD amounts were also found to be increased in the cerebrospinal fluid (CSF). CatD has been implicated in the processing of amyloid precursor protein (APP), apolipoprotein E (apoE) and the tau protein which are all important pathological factors of AD. According to the results of two recent meta-analyses, it is concluded that CD polymorphism on the risk of AD development is rather small on population level [Benes P., *et al.,* 2008]. Novel CatD inhibitors were reported to block the formation of hyper phosphorylated tau fragments in hippocampus. The degradation of b-amyloid precursor protein (APP) by the acid protease (cathepsin D) clearly indicates that APP is a good substrate for this protease. A number of studies report the cleavage of APP by cathepsin D and suggest that cathepsin

D could generate the pathogenic bA4 amyloid peptides from its precursor *invitro* [Malik M., *et al.,* 2011a].

2.4.3.2. Autism

Autism is a severe neurodevelopmental disorder characterized by problems in communication, social skills and repetitive behaviour. CatD is the prominent lysosomal aspartic acid protease abundantly expressed in brain and hydrolyses selective peptide bonds of target proteins with high specificity. The alterations of CatD activities in autistic subjects may be due to altered inflammatory and apoptotic processes. CatD protein expression was notably higher in the pyramidal and granule cells of hippocampus, in the neurons of the cerebellum, and in the frontal cortices of autistic subjects. The autistic brains showed a loss of pyramidal neurons and granule cells in the hippocampus and a significant loss of purkinje cells in the cerebellum. A number of studies demonstrated elevated levels of several inflammatory cytokines like TGF-β1, MCP-1, IL-6, IL-8, IL-10, GM-CSF, TNF-α, IFN-γ. The increased CatD in both cerebellum and frontal cortex could be a consequence of increased release of CatD, in response to apoptotic stimuli [Malik M., *et al.,* 2011b; Sheikh A.M., *et al.,* 2010].

2.4.3.3. Parkinson's disease

Parkinson's disease (PD) is the second most common neurodegenerative disease. The upregulation of Cat D was shown to occur at an early stage in experimental models of AD leading to slow apoptosis of neurons. The same was studied in the progressive neuronal damage as seen in PD, focusing on the caudate nucleus (CN). CatD can act as a molecular trigger for neuronal damage in CN. Using a nonhuman primate model of PD, it was reported that lysosomal instability and alterations in a lysosomal protease, CatD, in CN can lead to degeneration and dysfunction of

neurons. In the chronic MPTP model of PD, there was a significant upregulation of CatD in the CN. There is an increased staining for CatD in neurons, and an increase in the number of lysosomes in neurons in MPTP-induced PD. CatD could be used as a therapeutic strategy in PD as it clears toxic α-synuclein aggregates. Thus, CatD can be involved in the long-term neuronal damage in PD in regions remote from the site of primary insult, and underscores the need to study brain regions outside the sites of initial damage in neurodegeneration to help understand the basis of disease progression [Yelamanchili S.V., *et al.,* 2011].

2.4.4. Role of CatD in atherosclerosis

Atherosclerosis, a phenomenon characterized by the accumulation of low-density lipoprotein (LDL)-derived lipids in the ECM of arterial intima, recruitment of macrophages. Macrophage pericellular environment is sufficiently acidic for CatD activation and can further induce hydrolytic modification of LDL. Recent findings showed that CatD expression is reduced in macrophages of low HDL-cholesterol subjects and has its role in intracellular metabolism and transport of phospholipids and cholesterol.

2.5. CatD similarities in zebrafish and human

Zebrafish (*Danio rerio*) currently represents the most powerful '*in vivo'* vertebrate model for the study of several human diseases. A study demonstrated the processing and lysosomal targeting of zebrafish CatD expressed in human cells in comparison to human CatD expressed in zebrafish cells. The amino acid sequences of zebrafish and of human CatD revealed a high level of identity (64%) especially in the regions surrounding the catalytic aspartates of the active site. The predicted size for the carbohydrate-free mature polypeptides amounts approximately to 37kDa and to 36.5, for the human and the zebrafish CatD, respectively. Zebrafish CatD lacks the cysteine residue

corresponding to C91 in human sequence, which is responsible for one intra-chain Disulphide Bridge at the level of light chain. This loss in Disulphide Bridge differentiates zebrafish CatD from human CatD for its structural rigidity [Follo C., *et al.,* 2013].

The teleost zebrafish has exceptional utility as a human disease model system and represents a promising alternative model in cancer research. Zebrafish embryo allows disease-driven drug target identification and in vivo validation, thus representing an interesting bioassay tool for small molecule testing and dissection of biological pathways alternative to other vertebrate models. The basic vascular plan of the developing zebrafish embryo shows strong similarity to that of other vertebrates. In particular, angiogenesis occurs in the formation of the intersegmental vessels (ISVs) of the trunk that will sprout from the dorsal aorta at 20 hpf (hours post fertilization). Thus, advantage was taken in screening CatD inhibitors to measure their effect in inhibiting angiogenesis [Tobia C., *et al.,* 2011].

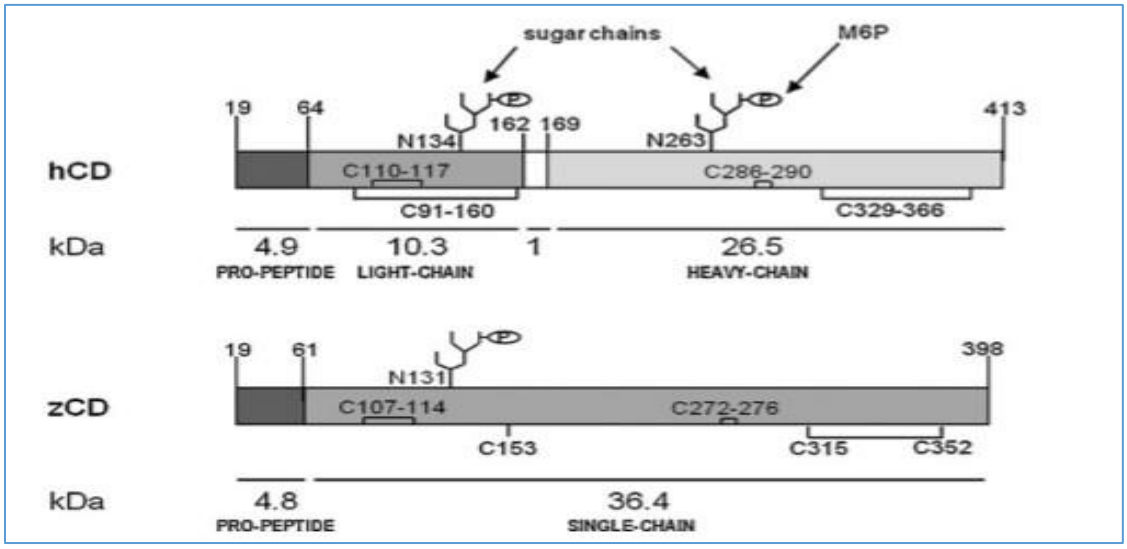


Figure 13: Structural similarities between Human and Zebrafish CatD [Follo C., *et al.,* 2013].

CHAPTER 3. Objectives and plan of work

3.1. Objectives

Targeting Cathepsins has given a new glimmer of hope for cancer therapies in developing new therapeutic strategies. Researchers have explored a dominant role of CatD in breast cancer tumor progression, invasion and angiogenesis. CatD is ubiquitously present in all the cells and performs proteolytic processing and degradative activities inside the acidic vesicles. In case of tumors, the levels of CatD expression was known to increase devastating the progression. The main pathway CatD taken in causing tumour invasion was reported to be ECM degradation and angiogenesis. CatD, in particular was highly quantified in the breast cancer cytosols during clinical findings and studied as a prognostic biomarker for breast cancer invasion and metastasis. Our study deals with deducing the mechanism through which CatD acts in breast cancer cells and in breast cancers, the triple negative breast cancers (TNBC). As, CatD has a crystal ligand Pepstatin A deposited in the protein bank, we have studied the catalytic active site of CatD in developing small molecule inhibitors. As CatD was known to be overexpressed in breast cancers, its inhibition should probably be a promising strategy of treatment. There is a huge scope in developing CatD inhibitors as a means of targeted therapy in treating TNBC. Hence our primary objective is to develop small molecule non-peptidic inhibitors against CatD and to evaluate their role in inhibiting TNBC invasion and metastasis mechanisms.

The main objectives of the proposed work was to:

I. Design and identification of small molecule inhibitors using e-Pharmacophore approach and structure based drug design. The crystal with lessee resolution deposited in the protein data bank was used for this step.

II. High-throughput virtual screening from large commercial asinex database and in-house BITS databases to achieve inhibitors of diversified scaffolds and with best interaction pattern.

III. *In vitro* enzyme based inhibition studies of the selected molecules for determining IC_{50} values.

IV. *In vitro* cytotoxic and growth inhibition studies of the selected molecules from the databases to shortlist potential leads.

V. Synthesis of analogues by making necessary modifications to the lead molecule.

VI. *In vitro* cell based studies to evaluate the anticancer activity of synthesized analogues on a panel of breast cancer cell lines.

VII. *In vitro* cell based biological evaluation of potent leads for their metastatic mechanisms

VIII. Gene expression studies using RT-PCR analyses to build pathway strategies.

IX. *In vivo* anti-angiogenic effect evaluation of the potent leads using zebrafish.

3.2. Plan of work

The plan of work was classified into following categories

3.2.1. Design and identification of novel Cathepsin D inhibitors using molecular modelling methods.

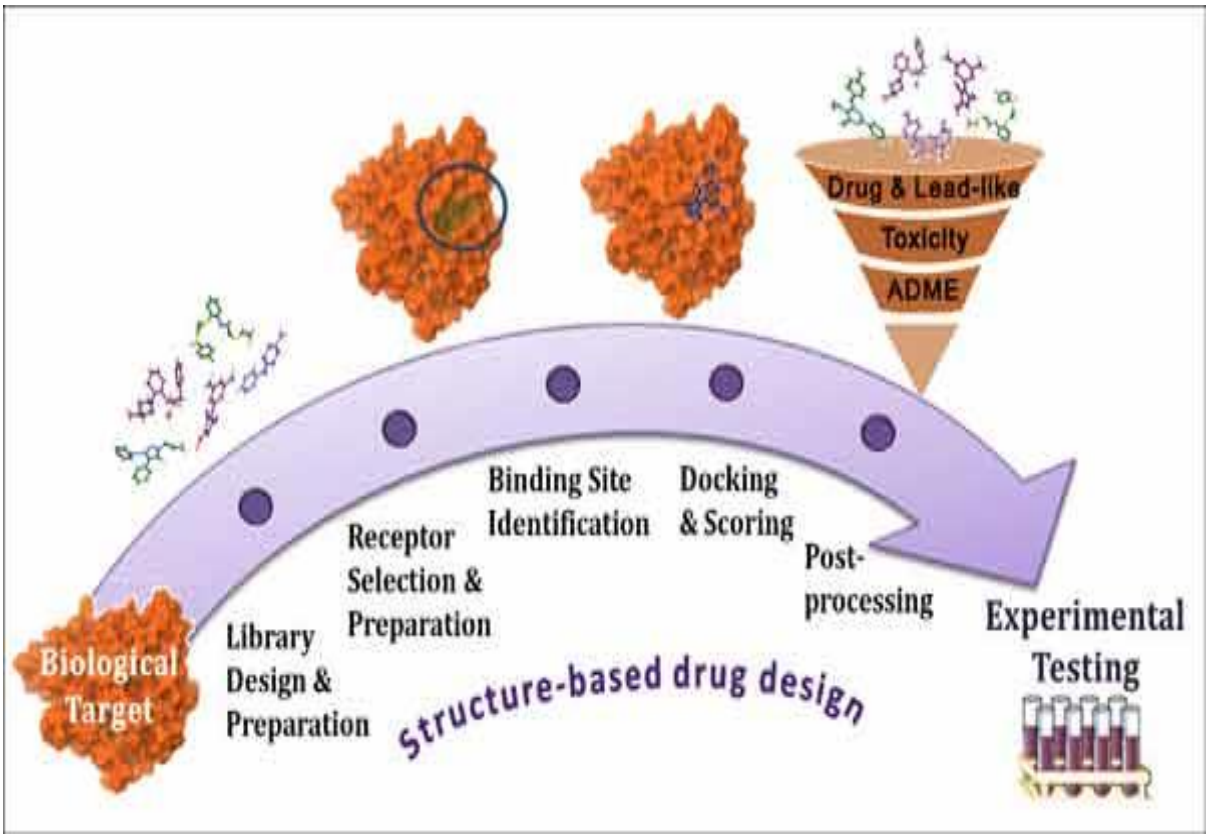


Figure 14: The flow of work in molecular modelling.

3.2.2. *In vitro* enzyme inhibition studies of Cathepsin D and renin enzymes

3.2.3. The effect of potent leads on cancer cell lines

3.2.3.1. *In-vitro* cytotoxicity and growth inhibition studies on various cancer cell lines including MCF-7, MDA-MB-231, SK-BR-3, MDA-MB-468 and ZR-75-1.

3.2.3.2 Cell cycle analysis and lysosomal permeability studies of the selected leads

3.2.4. In vitro biological pathway evaluation

3.2.4.1. Evaluation of gene expression studies involved in ECM degradation and angiogenesis mechanism

3.2.4.2. Evaluation of the role of leads in angiogenesis models

3.2.4.3. Effect of the leads on MMP 2 and 9 inhibition

3.2.5. *In-vivo* anti-angiogenic effect using zebrafish embryo

CHAPTER 4. MATERIALS AND METHODS

4.1. Molecular modelling software details

Computational work was carried in Intel Core 2 Duo E7400 2.80GHz capacity processor with internal memory of 2GB RAM, executing functions with RHEL 5.2 operating system. Schrodinger drug discovery software package consisting of modules for protein and ligand preparation and Glide for high throughput virtual screening (docking) was used in this study. Specialized applications like LigPrep 2.5 (Schrodinger 2011) (Schrodinger LLC, New York, USA) were executed for building and minimizing the structures, and was also used to add hydrogens and to generate various stereoisomers at neutral pH (pH 7). To identify and select structurally diverse compounds from different clusters, Canvas 1.4 software package was used. The pharmacophore models were generated using PHASE 3.4 implemented in Maestro 9.3 software package. All docking studies were performed using GLIDE (Grid Ligand Docking with Energetics) 5.8 (Maestro 9.3), and all ADME (Absorption, Distribution, Metabolism, and Excretion) properties were calculated using QikProp 3.5 (Maestro 9.3).

4.2. Identification of novel CatD small molecule inhibitors using structure based drug design approach

4.2.1. Design of CatD inhibitors

To design CatD inhibitors we have utilized structure based drug design approach (e-pharmacophore based) as described below. There are six crystal structures of CatD (Homo *sapiens*) deposited in the protein data bank. The crystal structures are namely, **1LYA, 1LYB, 1LYW, 4OBZ, 4OC6, 4OD9.** We have selected the one with lesser resolution of 2.5 Å i.e., 1LYB and coupled with Pepstatin A, specific inhibitor of CatD.

4.2.1.1.Structure based drug design strategy

Structure-based virtual screening plays an important role in drug discovery and complements other screening approaches. Structure-based drug design is the design and optimization of a chemical structure with the goal of identifying a compound suitable as a drug candidate. It is based on knowledge of the drug's three-dimensional structure and how its shape and charge cause it to interact with its biological target, ultimately eliciting a medical effect. The ideal target macromolecule for structure-based drug design is one that is closely linked to human disease and binds a small molecule in order to carry out a function. The target molecule usually has a defined binding pocket [Anderson A.C., 2003]. In present study, we have implemented a new protocol for generating energy-optimized pharmacophore (e-pharmacophore) based on the energies between the atoms. The drug design approach begins with the identification of a potential ligand binding site on the target molecule. The ligand binding site can be the active site, as in an enzymes. Once the structure and target site are identified, e-pharmacophore approach was followed to identify the potential entities [Anderson A.C., 2003].

4.2.1.2.Protein and ligand preparation

As discussed above, the crystal structure with lesser resolution PDB: 1LYB was selected and retrieved from the protein data bank (PDB). The protein preparation wizard was used from the maestro software package to add the hydrogen atoms, bond orders and formal charges during the protein preparation. The unnecessary unbounded water molecules were removed during this process, and the ones which were involved in the hydrogen bond interactions between protein and ligand were retained. Furthermore, OPLS_2005 force field was applied and the resulting structure was energetically minimized [Sastry G.M., 2013]. Interactions of the ligand with the protein

residues in the active site were visualized using ligand interaction diagram in Schrödinger suite version 9.3.

The ligand preparation using LigPrep includes addition of hydrogen atoms, removal of unwanted molecules such as water, small ions. The LigPrep neutralizes charged groups, then generate ionization and tautomeric states with Epik. Further, generates stereoisomers, low-energy ring confirmations and optimizes the geometries. Finally, the prepared ligand was docked XP using glide in both rigid as well as flexible modes. The root-mean-square deviation (**RMSD**) was measured in between rigid and flexible ligands using superposition. The input RMSD of the ligand was also ascertained.

4.2.1.3.Preparation of Commercial and in-house databases

Asinex is the commercial database available with us containing 5 lac unique structures used in the present study [Asinex, ASINEX Corp, USA]. All the molecules in asinex have diversified scaffolds, giving us a wide scope in filtering the best leads. Database molecules were all prepared using LigPrep and Epik to expand protonation and tautomeric states at pH 7.0. In-house database is our BITS database consisting of 2500 compounds which were synthesized in our laboratories. All these molecules were also optimized and prepared using LigPrep. Further, using Phase application in the maestro software package the database molecules were indexed with automatic creation of pharmacophore sites for each conformer to allow rapid database alignments and screening.

4.2.1.4.E-Pharmacophore generation and validation

Molecular docking studies were performed using an e-pharmacophore approach which helps in the rapid in silico screening of chemical databases. ‘Scripts’ option in Maestro was used to generate

pharmacophores, on the energies and features of the co-crystallized ligand. In addition, PHASE was used to identify the features in a molecule that match with the features in the generated hypothesis. The available pharmacophoric features provided by PHASE during the alignment were hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively ionisable group (N), positively ionisable group (P), and aromatic ring (R). Furthermore, different pharmacophore models were generated with minimum to maximum number of features to screen all the molecules from the ASINEX and BITS databases. Datasets were prepared using reported CatD inhibitors in the literature together with a decoy set. The decoy set, provided by Schrodinger, is a collection of 1000 ligands which exhibit "drug-like" properties, chosen from 1 million compound library. A total of 84 reported non-peptidic CatD inhibitors with different scaffolds (e.g., acyl guanidine, hydroxyl ethylamine) were drawn using Maestro's 2D sketcher and were energetically minimized using LigPrep. Stereo isomers were also generated by adding sufficient hydrogen molecules at neutral pH (pH 7) using LigPrep. On the basis of the 10 featured hypotheses generated using Pepstatin A, many 3–7 featured hypotheses were generated considering the positions of donors, acceptors and aromatic rings. PHASE generated scores were considered to evaluate the hypothesis ability in differentiating reported CatD inhibitors from the decoy set. Enrichment Factor (EF) and Goodness of Hit (GH) were calculated using Equations. 1 and 2:

$$\text{EF} = (Ha \times D)/\,(Ht \times A) \quad \ldots\ldots\ldots\ldots (1)$$

$$\text{GH} = \big((Ha\ /4\ Ht\ A) \times (3A\ +\ Ht)\big) \times \big(1\ -\ (Ht\ -\ Ha/\ D\ -\ A)\big) \ldots\ldots\ldots\ldots (2)$$

where '**Ht**' is the total set of molecules used from the hit list; '**Ha**' is the set of actives from the hit list; '**A**' denotes the actives from the decoy set; and '**D**' is the number of molecules from the decoy set. The best pharmacophores were further used to select molecules from the databases following molecular docking protocols.

4.2.1.5.Grid generation

The best pharmacophore models were used to screen diverse scaffolds from the ASINEX and BITS databases. GLIDE was used for the virtual screening of compound libraries. The crystal structure of CatD (PDB ID: 1LYB) was minimized using the protein preparation wizard (OPLS_2005 force field). The co-crystallized ligand, Pepstatin A, was selected during the receptor grid generation to define x, y and z axes of the grid. The sites for molecular docking were defined during the grid generation which allowed us to filter molecules from the selected active site.

4.2.1.6.Virtual screening by pharmacophore model and molecular docking

Virtual screening was implemented to retrieve the compounds that fit the chemical features present in the best hypothesis. For the e-pharmacophore approach, explicit matching was required for the most energetically favourable site (scoring better than 1.0 kcal/mol) that finds matching pharmacophores in the ligands. For filtering the database molecules, a minimum of 4-5 sites were required to match for hypotheses with 5-7 sites. The best pharmacophores with good EF and GH scores were selected and run through PHASE to find matches against the ASINEX and BITS databases. Molecules that matched the respective features of the selected pharmacophore were further considered for docking studies using GLIDE. Initially, the favourable interactions between screened ligand molecules (from pharmacophore screening) and CatD, in flexible mode of docking, were examined by means of GLIDE. The three different levels of docking precisions provided by GLIDE were utilized for virtual screening methodology, namely HTVS (high throughput virtual screening), SP (standard precision), and XP (extra precision). HTVS was first used to identify molecules binding to CatD followed by SP and then XP docking for refinement. Ligand interactions in relation to Pepstatin A, along with XP GLIDE scores, number of hydrogen

bonds, alignment of the molecule inside the active site pocket, and ADME properties were considered to select the final set of compounds for further in vitro studies.

4.2.1.7. ADME profile screening

The selected compounds after virtual screening were further selected for in-silico prediction of absorption, distribution, metabolism, excretion and toxicity using QikProp module [QikProp, v3.5, Schrödinger 2012]. QikProp efficiently evaluated pharmaceutically relevant properties for over half a million compounds per hour, making it an indispensable lead generation and lead optimization tool. Accurate prediction of absorption, distribution, metabolism, elimination (ADME) properties prior to expensive experimental procedures, such as HTS, could eliminate unnecessary testing on compounds that would ultimately fail; ADME prediction can also be used to focus lead optimization efforts to enhance the desired properties of a given compound. The QikProp module was used to calculate the ADME properties of the selected molecules. This provided an understanding of the solubility and absorption levels of the selected set of molecules from the database. Caco-permeability with a value of 500 as high permeability. QP log Po/w (octanol/water partition coefficient ratio) was calculated from a range of −2 to 6.5; QP log S (aqueous solubility) was considered good if the value was between −6.5 and 0.5. Furthermore, the percentage of human absorption was considered high if it was >80 % and poor if it was < 25%.

4.3. Selection of the right cell line for breast cancer research

Breast cancer is a complex and heterogeneous disease. Gene expression profiling has contributed significantly to our understanding of this heterogeneity at a molecular level, refining taxonomy based on simple measures such as histological type, tumour grade, lymph node status and the presence of predictive markers like oestrogen receptor and human epidermal growth factor

receptor 2 (HER2) to a more sophisticated classification comprising luminal A, luminal B, basal-like, HER2-positive and normal subgroups [Holliday D.I. and Speirs V., 2011; Stingl J. and Caldas C., 2007; Polyak K., 2011]. Each of these subgroups has different risk factors for incidence, response to treatment, risk of disease progression, and preferential organ sites of metastases. Luminal tumours are positive for estrogen and progesterone receptors, and the majority respond well to hormonal interventions (MCF-7 and ZR-75-1 cell lines), whereas HER2+ tumors have amplification and overexpression of the ERBB2 oncogene and can be effectively controlled with a diverse array of anti-HER2 therapies (ZR-75-1 cell line). Basal-like tumours in general lack hormone receptors and HER2 (MDA-MB-468 cell line); thus, the majority of these tumors are also called triple-negative breast cancer (TNBC). Claudin-low classification of breast carcinoma has intermediate response to chemotherapy (MDA-MB-231 cell line). A typical class with lack of hormone receptors but presence of HER2 are trastuzumab responsive (SKBR3 cell line). So, based on the above classification we have selected cell lines for our study on breast carcinomas.

4.3.1. Quantification of Cathepsin D gene expression in the selected panel of cell lines

A real-time polymerase chain reaction (RT-PCR) is a laboratory technique of molecular biology based on the polymerase chain reaction (PCR). It monitors the amplification of a targeted DNA molecule during the PCR. The amount of an expressed gene in a cell can be measured by the number of copies of an mRNA transcript of that gene present in a sample. The polymerase chain reaction (PCR) is a common method for amplifying DNA; for mRNA-based PCR the RNA sample is first reverse-transcribed to complementary DNA (cDNA) with reverse transcriptase. Real-time PCR is carried out in a thermal cycler with the capacity to illuminate each sample with a beam of

light of at least one specified wavelength and detect the fluorescence emitted by the excited fluorophore. We used SYBR green as a fluorescent dye in our study [Fraga F., *et al.,* 2008].

RT-PCR studies were carried out to bring out a logical selection of a set of breast cancer cell lines depending on their differential CatD gene expressions. The consideration was on the presence or the absence of differential gene expressions, e.g., estrogen receptor (ER+/−), progesterone receptor (PR+/−), and human epidermal growth factor receptor (HER/neu) which helped us study the activities of molecules considering both TNBC and TPBC types. Total mRNA was quantified from 5×10^6 cells MCF-7 (triple-positive), MDA-MB-231 (triple-negative), MDA-MB-468 (triple-negative), and SK-BR-3 (ER-, Her2 amplified). RNA was extracted using the TRI reagent according to the manufacturer's protocol (Sigma). PCR amplification, for converting mRNA to cDNA, was run using Verso cDNA kit, with optimized protocol for 45 cycles. Each cycle had a 15-second denaturation time at 95 ◦C, annealing for 1 min at 52 ºC-55 ºC, and extension for 1 min at 72 ºC. The mRNA quantities were amplified using a PCR master mix to optimize the annealing temperature of the designed CatD human primers. The cDNA entities formed were further confirmed by electrophoresis, running through 1 % agarose gel at 100 volts for 30 min. RT-PCR was performed following the manufacturer's instructions using a fluorescent dye, SYBR green (Kappa 268 Bio-systems), helpful in the detection of PCR products.

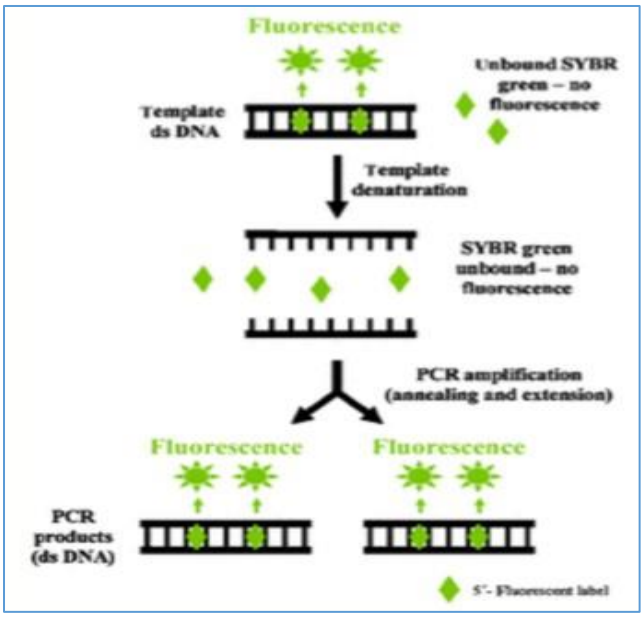


Figure 15: SYBR green detection [Smith C.J. and Osborn A.M., 2008].

We used CatD (human) primers procured from Sigma, i.e., forward primer 5′ CAT TGT GGA CAC AGG CAC TTC 3′ and reverse primer 5′ GAC ACC TTG AGC GTG TAGTCC 3 ′ to determine DNA fragment [Fukunda M.E., *et al.*, 2005]. The RT-PCR was initiated with denaturation at 94 ^{o}C for 3 minutes, followed by 45 cycles of denaturation at 95 ^{o}C for 20 s, annealing at 61 ^{o}C for 10 s, and elongation 72 ^{o}C for 2 min. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as the house-keeping gene to normalize the gene expressions. Normalized relative levels of CatD expressions in the selected breast cancer cell lines were the ratios of CatD and GAPDH gene expressions.

4.3.2. Quantification of CatD enzyme in the cell lysates

We have quantified the amounts of CatD present in the cell lysates to further validate target specific inhibitions. The clinical findings and the gene expression studies have already suggested the levels of CatD to be present and to act as an important prognostic marker. So, based on the reported data

and our studies we have estimated the quantity of CatD in the cell line panel selected. MCF-7, MDA-MB-231, MDA-MB-468, SKBR3 breast carcinoma cell lines were all cultured in the flasks. The Cathepsin-D Activity Assay kit is a fluorescence-based assay that utilizes the preferred cathepsin-D substrate sequence GKPILFFRLK(Dnp)-D-R-NH2) labelled with methyl coumarin (MCA). Cell lysates that contain cathepsin-D will cleave the synthetic substrate to release fluorescence, which can then easily be quantified using a fluorescence plate reader at Ex/Em = 328/460 nm (BioVision, Inc., Milpitas, CA).

The cells were harvested and collected (1×10^{6}) by means of centrifugation. The cells were lysed in 200 µL of chilled cell lysis buffer and incubated on ice for about 10 min. Later, the cells were centrifuged for 5 min at higher rpm and the supernatant which is the clear cell lysate was collected in to a fresh labelled tube. The lysate (50 µL) was added in to a 96-well plate along with the master mix which contains reaction buffer and substrate. The plate was allowed to be incubated at 37 ^{o}C for 1-2 hour. As a final step, the samples were read using a fluorescence plate reader at Ex/Em = 328/460 nm. A standard graph with different amounts of CatD (purified CatD) was plotted using the same procedure. The relative fluorescence units were calculated and extrapolated from the standard graph to deduce the amounts of CatD present in the cell lysates.

4.3.3. Survival of the breast cancer cell lines in acidic microenvironment

The acidic environment of tumour facilitates the maturation of the pro-cathepsin D and triggers downstream signalling pathways. Tumor cells are known to adapt in acidic microenvironment by preserving ATP levels. The effect of proteases and aspartyl protease inhibitor pepstatin A has been previously analysed by conditioning media of human MDA-MB-231 breast cancer cells at pH 6.6 and pH 7.4. CatD close to acidic pH and hypoxic solid regions of tumors was known to contribute

more in triggering a proteolytic cascade facilitating cancer cell invasion and metastasis [Maynadier M., *et al.,* 2013].

As a part of assessing lysosomal CatD activity in acidic conditions of tumor microenvironment, the metastatic breast cancer cell lines were plated using RPMI 1640 and MEM media with 10% FCS at pH 7.4 or 5.5, and the pH 5.5 was achieved by adding 2 N HCl to the culture medium. After growing the culture for 5 days, cells were stained with 0.4% trypan blue and analysed by flow cytometry [Matarrese P., et al., 2010]. The survival probabilities of the cell lines were calculated at different pH levels.

4.4. *In vitro* Enzyme inhibition assay studies

Aspartic proteases are a catalytic type of protease enzymes that use an activated water molecule bound to one or more aspartate residues for catalysis of their peptide substrates. In general, they have two highly conserved aspartates in the active site and are optimally active at acidic pH. Nearly all known aspartyl proteases are inhibited by pepstatin. Eukaryotic aspartic proteases include pepsins, cathepsins and renins. There is considerable structural as well as enzymatic action similarities between CatD and renin. So, we have tested our molecules on renin enzyme also, to estimate the off-target effect of our compounds.

4.4.1. Cathepsin D enzyme inhibition assay

CatD (EC 3.4.23.5) enzyme inhibition studies were carried out using Cathepsin D Inhibitor Screening Kit (Fluorometric) (BioVision, Inc., Milpitas, CA). This is a fluorescence-based assay which makes use of the CatD substrate sequence GKPIL-FFRLK (DNP)-D-RNH2 labelled with methyl coumarin (MCA). CatD acts by cleaving the synthetic substrate to release the quenched fluorescent group MCA, which will then be measured by a fluorescence plate reader at Ex/Em of

328/460 nm. The relative efficacy of test inhibitors was compared to the positive control inhibitor, Pepstatin A (IC_{50} < 0.1 nM).

The assay procedure implemented was followed as described by the manufacturer's protocol. The kit consists of CatD reaction buffer, CatD substrate (1mM), CatD, pure enzyme of human origin and pepstatin A. The positive control was prepared by mixing 10 µL of reconstituted CatD with 40 µL of CatD reaction buffer. The background control includes reaction buffer alone. The inhibitor reference control includes 10 µL of reconstituted CatD, 10 µL of pepstatin A and 30 µL CatD reaction buffer. The test inhibitor samples were prepared in the same way as inhibitor control using our compounds at different concentrations. Both controls and test samples were pre-incubated at 37 ^{o}C for 10 min. Followed by, addition of the substrate mix consisting of 2 µL of CatD substrate and 48 µL of reaction buffer into all the wells. An incubation period of 1–2 h at 37 ^{o}C was followed by measuring the fluorescence using a micro-plate reader.

The readings were further measured as per the standardized protocol:

$$\%INHIBITION = (RFUTest\ inhibitor - RFUBackground\ control) / (RFUPositive\ control - RFUBackground\ control) \times 100$$

4.4.2. Renin enzyme inhibition assay

Renin (EC 3.4.23.15), also known as an angiotensinogenase, is an enzyme that participates in the renin-angiotensin system (RAS) which mediates extracellular volume (i.e. blood plasma, lymph and interstitial fluid), and arterial vasoconstriction. BioVision's Renin inhibitor screening Kit uses a synthetic peptide substrate with a fluorophore (EDANS) at one end and a quencher (DABCYL) at the other end. Renin catalyses the cleavage of FRET substrate resulting in a product that is detected fluorometrically at Ex/Em = 328/552 nm. In the presence of a Renin inhibitor, the rate of hydrolysis of the substrate is decreased.

The assay procedure implemented was followed as described by the manufacturer's protocol. The kit consists of renin assay buffer, renin substrate, active human renin and inhibitor control (Aliskiren). Initially, renin enzyme solution was prepared as indicates. The stock solutions of the compounds were prepared at 1000x the highest test concentration and diluted to 4x test concentration with renin assay buffer just before use. A volume of 25 µL of the candidate compound working solution, Inhibitor Control (Aliskiren) working solution or Renin Assay Buffer was added into wells containing the Renin Enzyme Solution as candidate screen, Inhibitor Control, or Enzyme Activity Control, respectively. The contents are mixed well, and incubated for 5 minutes at 37 ºC. Followed by the addition of 25 µL substrate solution into each well. The plate is measured using a fluorescence plate reader in a kinetic mode for 30-60 min at Ex/Em of 328/552 nm. Any two time points (T1 and T2) were selected in the linear range of the plot and the corresponding RFU1 and RFU2 values were obtained. Furthermore, slopes were calculated for all the samples, including enzyme activity control, by dividing the net ΔRFU (RFU2-RFU1) values with the time ΔT (T2-T1).

$$\%\ Relative\ inhibition\ =$$

$$\frac{(Slope\ of\ enzyme\ activity\ control - Slope\ of\ enzyme\ activity\ with\ compound\ screen)}{Slope\ of\ enzyme\ activity\ control}\ x\ 100$$

4.5. *In vitro* cell based assay studies

Many biological assays require the measurement of surviving and/or proliferating mammalian cells. A colorimetric assay for living cells utilizes a colourless substrate and which is modified into a coloured product by any living cell and not by dead cells or tissue culture medium. Tetrazolium

salts are the attractive candidates for this purpose, as they help to measure the activity of various dehydrogenase enzymes [Mosmann T., 1983].

4.5.1. Cell culture and MTT cell proliferation assay

The breast cancer cell lines used in the present study were namely MCF-7, MDA-MB-231, MDA-MB-468, SKBR3 and ZR-75-1 of American Type Culture Collection origin and were kindly provided by Incozen Therapeutics Pvt. Ltd. The normal kidney cell line used in our study was HEK293T. MCF-7 was cultured in Minimum Essential Medium Eagle (MEM), MDA-MB-231 and MDA-MB-468 were cultured in Leibovitz's L-15 medium, SKBR3 was cultivated in Mc Coy's 5a medium, ZR-75-1 was processed in RPMI 1640 medium and HEK293T was cultured in Dulbecco's Modified Eagle's Medium (DMEM). The media was made complete growth medium by adding 10% Fetal Bovine Serum (FBS, heat inactivated), 1% penicillin/Streptomycin under conditions of 5% CO_2, 80% relative humidity (RH) at 37 ^{o}C.

Measurement of cell viability and proliferation forms the basis for numerous in vitro assays of a cell population's response to external factors. The reduction of tetrazolium salts is now widely accepted as a reliable way to examine cell proliferation. The yellow tetrazolium MTT (3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide) is reduced by metabolically active cells, in part by the action of dehydrogenase enzymes, to generate reducing equivalents such as NADH and NADPH [Bernas T. and Dobrucki J., 2002]. The resulting intracellular purple formazan can be solubilized and quantified by spectrophotometric means. The MTT Cell Proliferation Assay measures the cell proliferation rate and conversely, when metabolic events lead to apoptosis or necrosis, the reduction in cell viability.

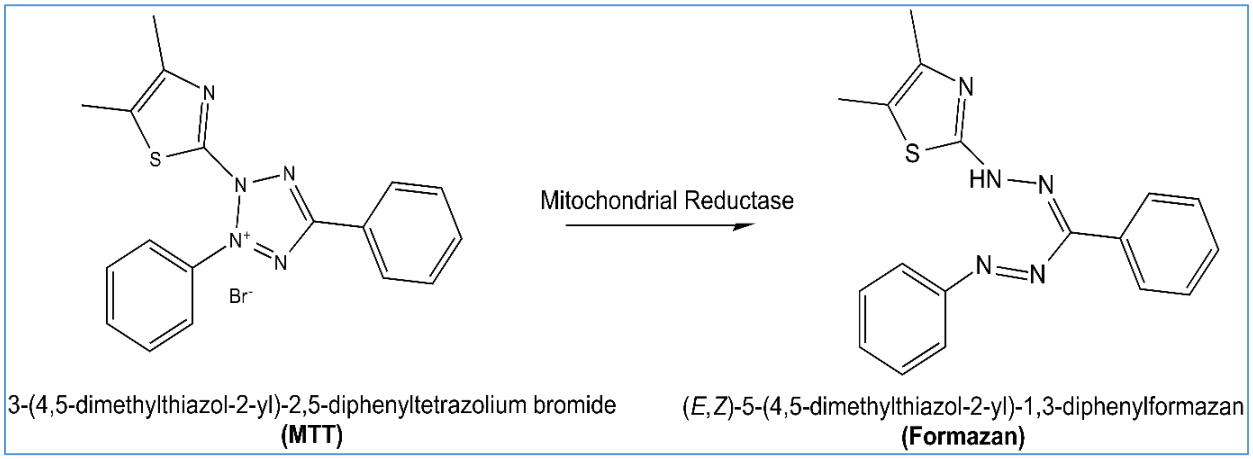


Figure 16: Principle involved in MTT assay [Bernas T. and Dobrucki J., 2002].

Cells were harvested and seeded into 96-well tissue culture plates with 5,000 cells per well and allowed to attach and gain morphology overnight. Cells were treated with test compounds at concentrations ranging from 100µM to 1 nM in triplicates for a drug-incubation period of 48–72 h. The incubation time varied depending upon the doubling time of individual cell lines. A stock solution of a test compound (10 mM) was prepared using DMSO and all the other concentrations were diluted with medium to achieve <1% DMSO while exposing to cells. A DMSO control was also placed to measure its toxic effects, if any, on the cancer cells [Meerloo J.V., *et al.,* 2011]. A zero reading was taken to note the absorbance of day 1 cells plated. After test compound incubation, MTT was added at a concentration of 5 mg/mL, and plates were incubated for 3–4 hrs at 37◦C. Cell viability was quantified by dissolving the formed formazan crystals in DMSO and measuring the absorbance spectrophotometrically at 550– 600 nm using a Vector X3 2030 multi-label reader (Perkin Elmer).

4.6. Synthesis and Characterisation

The top active lead compound from in-house BITS database found to be the most potent was taken up for synthesising further analogues in order to improve its activity. Several modifications were made in the parent test compound to further synthesise its derivatives using a proper synthetic protocol.

4.6.1. Synthetic scheme used to improve the activity of the potent lead compound against CatD obtained from virtual screening

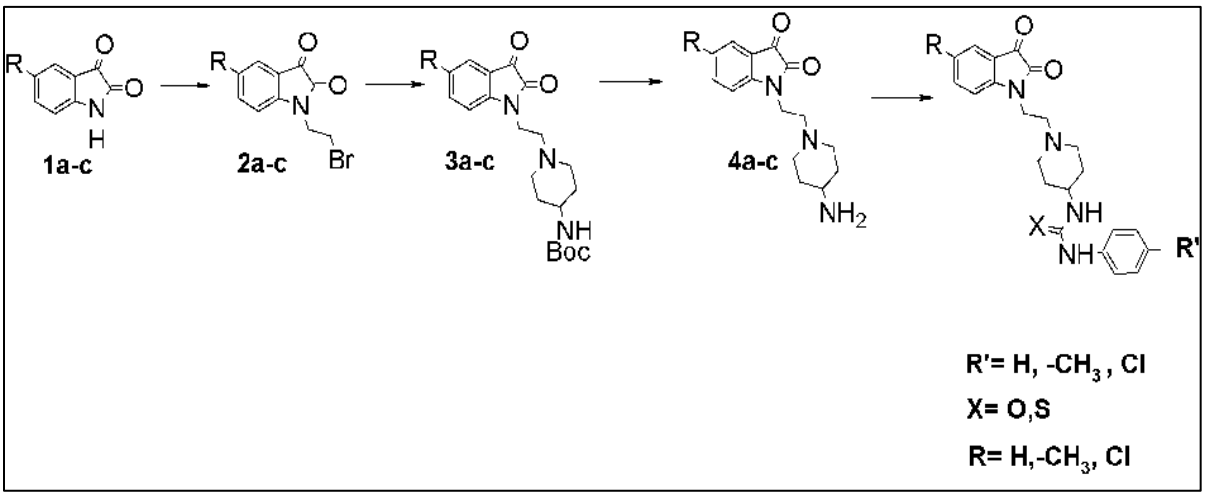


Figure 17: Synthetic scheme involved in the thesis

All commercially available chemicals and solvents were used without further purification. TLC experiments were performed on alumina-backed silica gel 60 F_{254} plates (Merck, Darmstadt, Germany). Homogeneity of the compounds was monitored by thin layer chromatography (TLC) on silica gel 60 F_{254} coated on aluminium plates, visualized by UV light and $KMnO_4$ treatment. All ^{1}H and ^{13}C NMR spectra were recorded on a Bruker AM-300 (300.12 MHz, 75.12 MHz) NMR spectrometer, BrukerBioSpin Corp, Germany. Chemical shifts were reported in ppm (δ) with reference to the internal standard TMS. The signals were designated as follows: s, singlet; d, doublet; t, triplet; m, multiplet. Molecular weights of the synthesized compounds were checked by

LCMS 6100B series Agilent Technology. Elemental analyses were carried out on an automatic Flash EA 1112 Series, CHN Analyzer (Thermo).

4.6.2. Molecular docking of all the synthesised analogues

All the synthesized analogues were drawn using 2D sketcher and were optimized using LigPrep application. They were further docked using glide XP and the grid generated for CatD (PDB: 1LYB). The analogues were cross verified for their interaction pattern and occupancy inside the active site pocket with respect to the parent compound. Using QikProp, ADME properties were also verified for the analogues in order to consider these analogues for our further studies.

4.7. Enzyme inhibition studies of the synthesized analogues

Similar procedure was followed as mentioned in 4.4

4.7.1. Cathepsin D enzyme inhibition assay

Similar procedure was followed as mentioned in 4.4.1

4.7.2. Renin enzyme inhibition assay

Similar procedure was followed as mentioned in 4.4.2.

4.8. MTT cell proliferation assay for the synthesized analogues

Similar procedure was followed as mentioned in 4.5.1.

4.9. Analyses using Flow sight

The parent test compound and the most actives from the synthesized analogues were further taken for mechanism-based studies. Flow sight is an instrument that includes flow cytometry along with cell imagery application. This technique helps in counting, examining and sorting microscopic particles suspended in a stream of fluid. It allows simultaneous multi parametric analysis of the physical and chemical characterisitics of single cells flowing through a laser. Cell components are fluorescently labelled and then excited by the laser to emit light at varying wavelengths.

4.9.1. Cell cycle analysis using propidium iodide

PI is a fluorogenic compound that binds stoichiometrically to nucleic acids so that fluorescence emission is proportional to the DNA (and RNA, which has to be removed if DNA is to be measured) content of a cell. This helps in simultaneous analysis of cell-cycle parameters of surviving cells. Selected molecules were evaluated for their effect on cancer cell cycle [Riccardi C. and Nicoletti I., 2006]. The DNA content was different in various phases of cell cycle and propidium iodide, which stains DNA, helped analyse the effect of a test compound on cell cycle regulation. Breast cancer cells MDA-MB-231 (1×10^6 cells/well) were seeded in six-well plates and treated with a test compound for 24 h. Cells were first harvested and washed, followed by fixation in 70 % ice-cold ethanol at 4 ^oC for 2 h. The cells were further centrifuged, washed well with cold PBS, and re-centrifuged. Finally, cells were evenly re-suspended in 250µL PBS, and stained with 10µL propidium iodide (PI; 1 mg/mL) and 10µL RNaseA (10mg/mL) for 30min at room temperature (RT). The samples were loaded to measure the forward scatter (FS) and side scatter (SS) to identify single cells. DNA content was observed using flow sight (Amnis, Millipore), and IDEAS analysis software was used to calculate percentage of cells in G0/G1, S and G2/M phases [Bakar M.F.A., et al., 2010].

4.9.2. Effect of the test compounds on Reactive Oxygen Species (ROS)

Reactive oxygen species (ROS) was measured using fluorescent dye called 2′, 7′-dichlorofluorescein diacetate (DCFHDA, Sigma Aldrich). The DCFHDA get deacetylated after penetrating into the cells because of cellular esterases, which was later oxidised by ROS into DCF (2', 7' dichlorofluorescein) fluorescent compound.

MDA-MB-231 cell lines were grown in media and seeded in a 6-well plate at cell density of 1 X 10^6 cells/well. The cells were treated with cobalt chloride at 100 µM concentration to induce hypoxia. The hypoxic condition is known to increase ROS levels inside the cells. The cells were then incubated for 24 h at 37 ^oC, once the cells were adhered to the bottom of the plate, the test compounds were incubated and cells were stained with DCFHDA for 40 min at 37 ^oC. Eventually all the wells were washed with PBS, to minimize the background signal, and then fresh media was added and the cells were harvested, trypsinised and made into pellets. The pellets were further re-suspended in PBS and immediately taken for examination using Flow sight (Amnis, Millipore). The channel 2 was used to measure fluorescence to acquire the data. The shift in the histograms obtained by plotting intensity Vs normalised frequency using IDEAS software, helps us to deduce the effect of our test compound on ROS [Himmelfarb J., *et al.,* 1992].

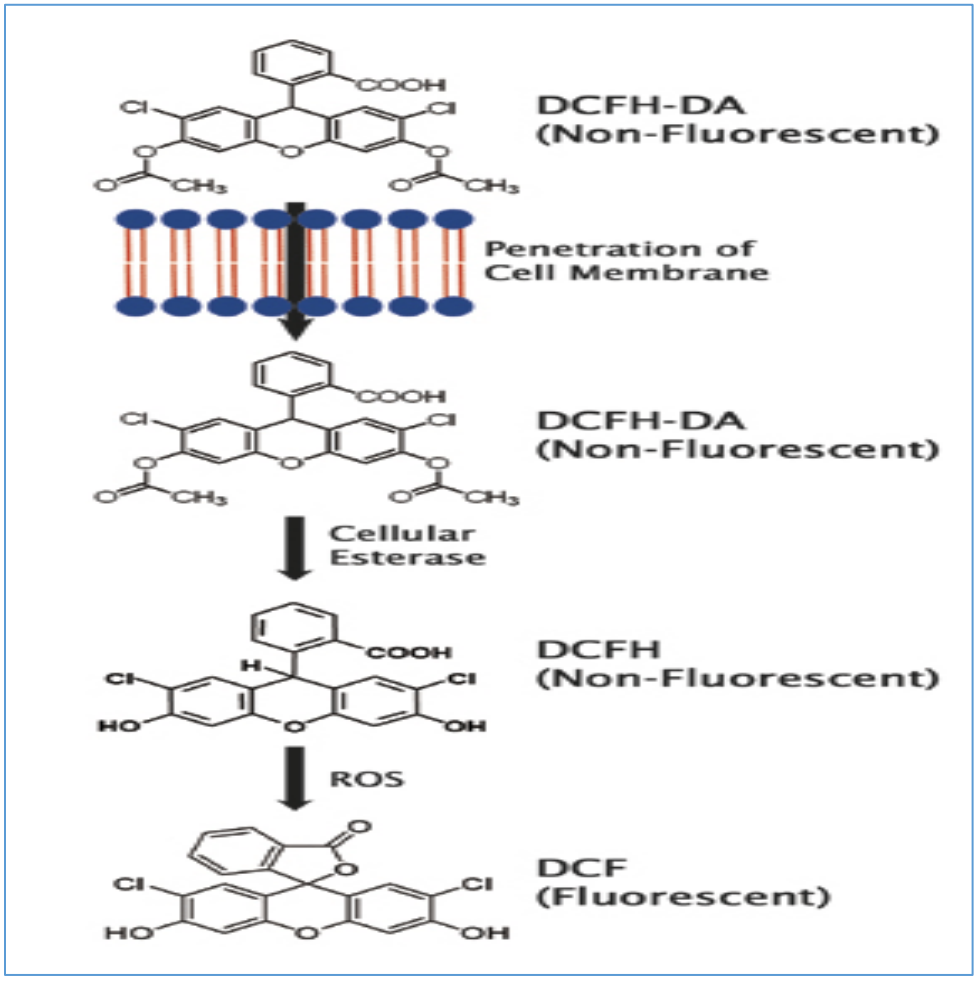


Figure 18: Principle involved in ROS generation [Cell Biolabs, Inc.,]

4.9.3. Lysosomal permeability estimation using Acridine Orange dye

A commonly used lysosomotropic compound is the lipophilic fluorochrome acridine orange (AO), which penetrates cellular membranes. AO is a metachromatic fluorochrome that when activated by blue light, gives a distinct red fluorescence at the high lysosomal concentration that results from an exposure of living cells to a low concentration of AO, and a weakly green fluorescence at the low obtained cytosolic concentration. AO is very useful for demonstrating intact lysosomes in living cells and for assaying lysosomal rupture under experimental conditions [Terman A., *et al.*,

2006]. Chloroquine is the most studied drug for its effect on the acidic organelles i.e., lysosomes using acridine orange treatment [Caporaso G.L., *et al.,* 1992].

MDA-MB-231 cell lines were grown in media and seeded in a 6-well plate at cell density of 1 X 10^6 cells/well. The cells were then incubated for 24 h at 37 ^oC, once the cells were adhered to the bottom of the plate, the test compounds were incubated and cells were stained with Acridine orange (10 µL from 0.5 µg/mL) for 30 min at 37 ^oC. Eventually all the wells were washed with PBS, to minimize the background signal, and then fresh media was added and the cells were harvested, trypsinised and made into pellets. The pellets were further re-suspended in PBS and immediately taken for examination using Flow sight (Amnis, Millipore). The channel 2 (505-506 nm emission) was used to measure fluorescence to acquire the data. The shift in the histograms obtained by plotting intensity Vs normalised frequency using IDEAS software, helps us to deduce the effect of our test compound on lysosomal content [Himmelfard J., *et al.*, 1992].

4.10. Angiogenesis assays performed for the most active test compound

As angiogenesis is essential for tumor growth and metastasis, controlling tumor-associated angiogenesis is a promising tactic in limiting cancer progression. Determining the migratory and invasive capacity of tumor and stromal cells and clarifying the underlying mechanisms is most relevant for novel strategies in cancer diagnosis, prognosis, drug development and treatment. There are different modes of cell travel *in vitro* methods to evaluate migration and invasion.

4.10.1.Migration chamber assay

The principle of this assay is based on two medium containing chambers separated by porous membrane through which cells transmigrate. The cells are seeded in the medium of upper part and can migrate in vertical direction through the pores of the membrane into the lower compartment,

in which medium containing an attractant (high serum) is present. The cells passed through the culture inserts are stained and quantified. The conventional cell culture inserts were purchased from HiMedia Laboratories Pvt. Ltd.

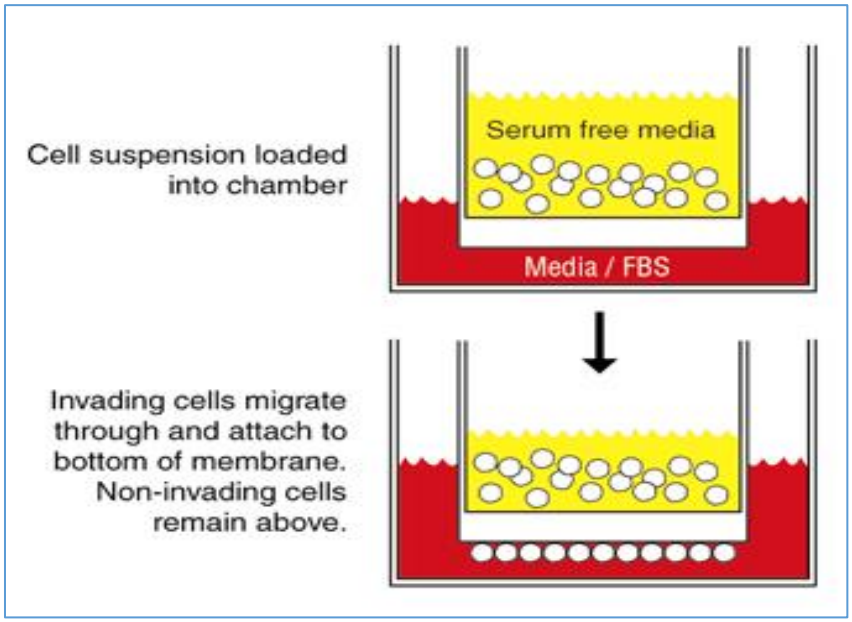


Figure 19: The mechanism involved in Migration chamber assay

The migration chambers were placed in a 24-well plate and 100 µL of serum-free medium was added in the chamber. Around 200 µL of $2.5x10^5$ cells/ mL in serum-free medium were added to the upper chamber along with the necessary drug concentrations. 750 µL of culture medium with serum was added to the lower chamber. The plate was incubated for 12-16 h at 37 ^oC. Later, the medium was removed from the upper chambers and were washed twice with PBS. The cells were allowed for fixation using formaldehyde (3.7% in PBS) at room temperature for 2 minutes [Kramer N., et al., 2013]. The cells were allowed for permeabilisation using 100 % methanol at room temperature for 20 min. The methanol was removed and the cells were washed twice with PBS. As a final step, the cells were stained with 0.5% of giemsa stain at room temperature for 15 min.

The stain was removed and the cells were washed twice with PBS. The non-migrated cells were scraped off using cotton swabs and the remaining migrated cells were counted under light microscope. Images were taken using phase contrast microscope [Chen H.C., 2005].

4.10.2.Tube formation assay

Over the past several decades, a tube formation assay using growth factor-reduced Matrigel has been typically employed to demonstrate the angiogenic activity of vascular endothelial cells *in vitro*. However, recently growing evidence has shown that this assay is not limited to test vascular behaviour for endothelial cells. Instead, it also has been used to test the ability of a number of tumor cells to develop a vascular phenotype. Some tumor cell lines such as melanoma B16F10 cells, glioblastoma U87 cells, and breast cancer MDA-MB-435 cells are able to form vascular tubules. This assay may serve as powerful utility to screen the vascular potential of a variety of cell types including vascular cells, tumor cells as well as other cells [Ponce M.L., 2009].

The melanoma cells B16F10 were grown in DMEM supplemented with 10% FBS and penicillin/streptomycin. Cells were washed with PBS three times and detached with 0.05% trypsin EDTA. After centrifugation with 2,000 rpm for 5 min, cell pellets were washed again with PBS. Then, the pelleted cells were counted with a haemocytometer. An aliquot of growth factor-reduced Matrigel (HiMedia) was warmed up at room temperature. Before completely thawed, it was transferred onto ice and its liquid was kept on ice for at least 10 min. Then 50 µL of Matrigel was plated to 96-well plates at a horizontal level that allows the Matrigel to distribute evenly, and incubated for 30 min at 37°C. Cells (1-2 x 10^4) were re-suspended with serum-free DMEM for tumor cells, and loaded on the top of the Matrigel. The test compounds were added to the serum-free medium. Following incubation at 37°C overnight, each well was analysed directly under a microscope. The tubules were proceeded for fixation using 100 µL of 10% formalin saline-based

solution [Francescone R.A., *et al.,* 2011]. Under a microscope with 10x phase contrast, tubules in each field were imaged and an average of tubules from 3-5 random fields in each well was counted [Ponce M.L., 2009; Arnaoutova I. and Kleinman H.K., 2010].

4.10.3.*In vitro* scratch assay

The *in vitro* scratch assay is a straightforward and economical method to study cell migration *in vitro*. The method helps us to study the regulation of cell migration by cell interaction with extracellular matrix (ECM) and cell-cell interactions. The cell movement can be calculated by measuring the decrease of uncovered regions at different time intervals. The assay is simple and has easy readout and analysis [Kramer N., *et al.,* 2013].

MDA-MB-231 cells were plated in a 6-well to get a monolayer of cells adhered by the next day. A scratch was made with a 200 µL pipette tip. The debris was removed and the edge of the scratch was made smooth by washing the cells with 1 mL of growth medium. All the wells were made these scratches uniformly with the same pipette tip passing through the middle of the 6-well plate. The test compounds were added to the well plates at the required concentrations. Depending on the incubations, the well plate was observed under the microscope for every 2, 4, 6, 12 and 24 h. The images required at each time point was taken and further analysed by scaling the distances between one side of the scratch and the other. The experiment was performed atleast three times to get large sample sizes that are easily quantified statistically [Liang C.C., 2007; Yarrow *et al.,* 2005].

4.11. Cell sensitivity Clonogenic assay

This assay measures the ability of the cells to proliferate and the ability of individual cells to form colonies. In this assay, the recovery time is comparable to the time for formation of colonies. These

assays were known as a 'gold standard' cellular-sensitivity assays. The cells in the exponential growth phase were exposed to the test compound. Following drug exposure, the cells are disaggregated to form a single-cell suspension and are plated out at low density to allow colony formation. The colonies are fixed, stained, and counted. Each colony is assumed to be derived from a single cell and thus the colony count is an estimate of the number of cells that survived the drug treatment. The number of cells within each colony depends on the number of cell doublings and can be used as an estimate of the effects, if any, of the drug on the cell doubling time.

The cells are trypsinised and collected in growth medium containing serum. The suspension is centrifuged (200g, 5 min) to pellet out cells and re-suspended in fresh growth medium. Cell counting was done using a haemocytometer and ensured that a single cell suspension is obtained. The cells were diluted to a density of 8×10^4 cells/ mL in a total volume of 10 mL. The cells were equilibrated with CO_2 and incubated at 37°C for 2–3 d to ensure that cells are in the exponential phase of growth for addition of test compound. Different concentrations of the test compounds were prepared and added to the well plate. The cells were evenly dispersed to ensure single-cell suspension after 24 h. The plate was periodically checked for the formation of colonies and changes in the media. After a period of 14 days, the medium was aspirated and the cells were washed twice with PBS. The cells were fixed using 4% paraformaldehyde for 2 min at room temperature. After fixation, crystal violet stain was added to the wells and allowed to incubate for 30 min at 37 ^oC. The excess dye was removed using normal water until the clear picture of the colonies were seen to be counted. The colonies formed in the control as well as treated were counted manually and the images were taken to observe the difference in formation of colonies if any. The percentage survival was calculated further from the data of number of colonies in each group.

4.12. Assessment of gelatinases (MMP-2 and MMP-9) by gelatin zymography

A major group of proteases that has been directly associated with tumor metastasis is the matrix metalloproteinases (MMPs), a family of endopeptidases known to cleave many ECM proteins. The gelatinases MMP-2 (gelatinase A, EC 3.4.24.24) and MMP-9 (gelatinase B, EC 3.4.24.35) are two members of the MMP family that have been extensively studied owing to their consistent association with tumor invasion and metastasis. The technique used to detect the presence or inhibition of these proteases is gelatin zymography [Toth M. and Fridman R., 2001].

4.12.1.Gel preparation

Initially all the stock solutions were prepared according to the recipes. The stocks to be prepared are acrylamide-bisacrylamide stock, separating gel and stacking gels buffer stocks, 1% w/v gelatin, 20% w/v SDS (sodium dodecyl sulphate), 10% w/v APS(ammonium per sulphate), running buffer stock, renaturing solution stock, developing buffer stock, staining and destaining solutions. The gel was prepared as an initial step and is 10 % polyacrylamide and 0.1% gelatin gel [Frankowski H., *et al.,* 2012]. The required volume of separating gel consisting of gelatin, acrylamide–bisacrylamide solution, SDS and APS was prepared and N, N, N , N -tetramethylethylenediamine (TEMED) was added in to the gel just before its addition into the gel cassette. On top of the separating gel, water was added to prevent entrapment of any air particles. The gel takes 30-40 min to get polymerised at room temperature. Polymerization is complete when a discrete line of separation can be noted between the gel and the water overlay. Decant the overlay water from the separating gel. Immediately stacking gel containing acrylamide–bisacrylamide solution, SDS, APS and TEMED was added on top of separating gel. The combs (10 wells) were rapidly inserted into the liquid stacking gel and is allowed to polymerise for another 30 min at room temperature.

4.12.2.Sample preparation

The protease samples were prepared meanwhile which were collected as cell lysates. MDA-MB-231 cells were plated in to a 6-well plate and treated with test compounds. After the incubation time, cells were washed twice with PBS and cold lysis buffer was added to each well. The cells were scraped into the lysis buffer using a rubber policeman, and the lysate is collected. The lysate was incubated on ice for atleast 15 min and vortex-mixed. The samples are further centrifuged (16,000g) for 20 min at 4 ^{o}C. The supernatant of the samples were collected and the protein concentration was measured. The sample to be loaded into the gel were prepared in the ratio of 3:1 with the supernatant and sample buffer respectively [Hawkes S.P., *et al.*, 2010a].

4.12.3.Development of zymogram

The gel prepared was placed inside the gel apparatus and filled with 1X running buffer. The samples were then loaded and the gel was allowed to run at constant voltage of 125 V until the bromophenol blue dye reaches the bottom of the gel which would approximately take 60-90 min. Later, the gel was carefully removed from the cassette and placed in a plastic tray containing 100 mL of renaturing solution. The gel was renatured for about 30 min with gentle agitation. The solution was decanted and the gel was washed twice with distilled water. Furthermore, the gel was incubated for an additional 30 min with developing buffer with gentle agitation. As a final step, the developing buffer was added freshly and the gel is allowed to be incubated overnight in a closed tray at 37 ^{o}C. The gel is stained for atleast 1h with staining solution and destained until the areas of gelatinolytic activity appear as clear sharp bands over the blue background. The pre-stained protein ladder with a range up to 200kDa was used to analyse the pro- and active forms of MMPs and their inhibition if any [Hawkes S.P., *et al.*, 2010b].

4.13. Quantification of gene expression using RT-PCR

The cellular processes involve a number of genes playing their role in regulating other downstream genes. So to study the downstream mechanistic pathway a set of genes thought to be interlinked were studied for their effects using a specific primer sequences. A series of steps are involved in determining gene expression.

4.13.1. Induction of hypoxia in tumor environment to study gene expressions

Hypoxic microenvironments and angiogenesis have been a focus of research from the past many years [Liu Q., et al., 2015]. The expression of numerous genes has been found to be regulated by hypoxic microenvironments, which plays a crucial role in biological characteristics. In the present study, $CoCl_2$ was utilized to mimic hypoxia. Ferrochelatase is inactivated by a substitution performed by cobalt ions under normoxic conditions, which blocks oxygen absorptivity in cytoblasts. A hypoxic microenvironment is subsequently formed [Sanchez L.M.L. *et al.*, 2014]. HIF-1α is rapidly degraded by the proteasome under normal conditions but is stabilized by hypoxic conditions. The process of angiogenesis possesses multiple steps, which is fundamental for tumor growth, invasion and metastasis. Growth factors are over-expressed and released, resulting in tumor over-growth. This promotes the vascular endothelial cells to migrate from the host vessel to tumor tissue [Okail M.S.A., 2010; Dai Z.J., *et al.*, 2012].

MDA-MB-231 cells were plated in to a 6-well plate $1x10^6$ cells/well and allowed to adhere incubating them overnight. $CoCl_2$ at the final concentration of 100 µM was added to the culture and incubated for 24 h at 37 ^oC [Wu D. and Yotnda P., 2011]. The test compounds were added to the wells as per the experimental protocol and followed by RT-PCR analyses.

4.13.2. Quantification of mRNA using TRI reagent

The cells were centrifuged and made pellets to start with mRNA extraction process. TRI reagent is a quick and convenient reagent used to isolate RNA. The pellet was added with 1 ml of TRI reagent and mixed evenly using pipette tip. A thick dense homogenous lysate was formed. Followed by phase separation, the samples were added with 0.2 mL of chloroform and vortexed vigorously for 15 seconds, and allowed to stand for 15 min at room temperature. The samples were then centrifuged at 12,000g for 15 min at 2-8 ^{o}C. This step separates the mixture into a red organic phase (containing protein), an interphase (containing DNA), and a colourless upper aqueous phase (containing RNA). The top most aqueous phase was collected into a fresh tube and 0.5 mL of isopropanol was added and mixed well. The samples were kept a side for 5-10 min at room temperature followed by centrifugation at 12000g for 10 min at 2-8 ^{o}C. This step results in RNA pellet formed towards the side of the tube. The isopropanol was removed and 75 % ethanol was added to wash the RNA pellet. The samples are further vortexed and centrifuged at 7,500 g for 5 min at 2-8 ^{o}C. Finally, the ethanol was removed and the tubes were allowed to air dry to remove any traces of ethanol. Biological grade water (50 µL) was added to the dried RNA pellet and dissolved by repeated pipetting with a micropipette at 55-60 ^{o}C for 10-15 minutes. The isolated mRNA was further measured and checked for any DNA contamination.

4.13.3. Conversion of mRNA to cDNA using PCR

The isolated mRNA was converted to cDNA using Thermo scientific verso cDNA kit. 20 µL sample of cDNA sample was freshly prepared before proceeding for RT-PCR analysis. The mRNA measured was calculated to be uniformly added from all the samples resulting in 200 ng/ µL concentration. The other components used for the conversion added were 5XcDNA synthesis buffer, dNTP mix, RNA random primers, RT enhancer and enzyme mix. To remove any

secondary structures the samples were heated at 70 ^{o}C for 5 min prior to enzyme addition. All the components added to PCR tubes were exposed to a cycling program of 42 ^{o}C, 30 min. the samples were stored at -20 ^{o}C.

4.13.4.Optimization of annealing temperature

The genes selected for the study were in relation to the effects of CatD in tumour cells. As CatD is more interlinked to ECM degradation, invasion and metastasis the following genes were examined: VEGF, EGFR, MMP-2, MMP-9, CXCR4, AKT and mTOR. The tumor inflammation related genes IL-1β, IL-6, IL-8, TNF-α, and NFκβ were selected. GAPDH was used as a house keeping gene for all the studies. The hypoxic conditions created were reported to upregulate the gene expressions of all the above selected genes along with H1F- α. The effect of newly targeted PPAR-γ, PARP-1, c-Myc and CDK1 genes were also studied for their role in relation to CatD inhibitors. All the primers were designed using primer blast and ordered from SIGMA Laboratories Pvt. Ltd. The oligonucleotides supplied were optimized for their annealing temperatures in our lab setup using 1 % agarose gel electrophoresis. The synthesised cDNA samples were taken and added with the forward and reverse primers of the above genes. PCR master mix purchased from Merck was used to run this experiment. A set of gradient temperatures was used to estimate the annealing capacity. The same cycling program that has to be used for RT-PCR was used. The samples were further run through 1% agarose gel at 100 V and observed for band intensities.

4.13.5.Performing RT-PCR and its analysis

The RT-PCR was initiated with denaturation at 94 ^{o}C for 3 minutes, followed by 45 cycles of denaturation at 95 ^{o}C for 20 s, annealing at 61 ^{o}C for 10 s, and elongation 72 ^{o}C for 2 min.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as the house-keeping gene to normalize the gene expressions. Normalized relative levels of gene expressions in the selected groups were the ratios of the gene and GAPDH expressions. The primer sequences used were:

IL-1β: Forward primer 5' GCAAGGGCTTCAGGCAGGCCGCG 3'

Reverse primer 5' GGTCATTCTCCTGGAAGGTCTGTGGGC 3'

IL-6: Forward primer 5' TTCGGTCCAGTTGCCTTCTC 3'

Reverse primer 5' GAGGTGAGTGGCTGTCTGTG 3'

TNF-α: Forward primer 5' CTCCAGGCGGTGCCTTGTTC 3'

Reverse primer 5' CAGGCAGAAGAGCGTGGTG 3'

NFκβ: Forward primer 5' GCGCTTCTCTGCCTTCCTTA 3'

Reverse primer 5' TCTTCAGGTTTGATGCCCCC 3'

IL-8: Forward primer 5' TGGACCCCAAGGAAAACTGG 3'

Reverse primer 5' TCTGGTCATGAGTACAACAAACT 3'

EGFR: Forward primer 5' CCTAAGATCCCGTCCATCGC 3'

Reverse primer 5' GGAGCCCAGCACTTTGATCT 3'

VEGF: Forward primer 5' CTACCTCCACCATGCCAAGT 3'

Reverse primer 5’ GCAGTAGCTGCGCTGATAGA 3’

MMP-2: Forward primer 5’ TCTCCTGACATTGACCTTGGC 3’

Reverse primer 5’ CAAGGTGCTGGCTGAGTAGATC 3’

MMP-9: Forward primer 5’ TTGACAGCGACAAGAAGTGG 3’

Reverse primer 5’ GCCATTCACGTCGTCCTTAT 3’

H1F-α: Forward primer 5’ TTCACCTGAGCCTAATAGTCC 3’

Reverse primer 5’ CAAGTCTAAATCTGTGTCCTG 3’

CXCR4: Forward primer 5’ CACTTCAGATAACTACACCG 3’

Reverse primer 5’ ATCCAGACGCCAACATAGAC 3’

AKT: Forward primer 5’ TCTATGGCGCTGAGATTGTG 3’

Reverse primer 5’ CTTAATGTGCCCGTCCTTGT 3’

mTOR: Forward primer 5’ CGCGAACCTCAGGGCAA 3’

Reverse primer 5’ TGGTTTCCTCATTCCGGCTC 3’

c-Myc: Forward primer 5’ TACCCTCTCAACGACAGCAG 3’

Reverse primer 5’ TCTTGACATTCTCCTCGGTG 3’

CDK1: Forward primer 5’ TTCAGGATGTGCTTATGC 3’

Reverse primer 5’ AGAGCAATTCCAAGCCAT 3’

E-Cadherin: Forward primer 5’ GAGAAACAGGATGGCTGAAGG 3’

Reverse primer 5’ TGAGGATGGTGTAAGCGATGG 3’

PPAR-γ: Forward primer 5’ GGCTTCATGACAAGGGAGTTTC 3’

Reverse primer 5’ AACTCAAACTTGGGCTCCATAAAG 3’

PARP-1: Forward primer 5’ CTACTCGGTCCAAGATCGCC 3’

Reverse primer 5’ TTGAAAAAGCCCTAAAGGCTCA 3’

4.14. In vivo zebrafish angiogenesis assay

Zebrafish embryos are transparent in which blood vessel development pattern is clearly visible by staining process and the angiogenesis can be quantified easily and accurately. This model is used to minimize the cost, reduce the time consumption and fasten the screening process thus it act as high throughput screening system.

24hpf (hour post fertilization) embryos were selected from stocks by observing them in stereo microscope and by comparing them with various developmental stages Zebrafish chart. The separated embryos were taken in Petri dish containing E_3 medium. E3 medium was replaced with Protease or Proteinase K solution (2µg/ml in E3 medium) and incubated for 5-10 mins for de-chorination. The embryos were flushed gently and washed with E3 3-4 times by pasture pipette to remove chorions completely. The 24hrs dechorionated embryos were taken in a petri plate containing vehicle. 6 embryos in a volume of 250µL of vehicle were distributed in 24-well

plate and 250µL of working concentrations of standard and test samples were added to respective wells as shown below. Various concentrations of compounds (1µM, 3µM, 10µM & 30µM) were added and embryos are then incubated at 28˚C for 48hrs (i.e., 72hpf). The embryos were washed with PBS for 2-3 times & then embryos were fixed with 4% Paraformaldehyde.

Fixed embryos were washed with PBS for 3-4 times for permeabilisation and the embryos were washed with 1ml Nano pure water for 5 minutes. Then washed with 1ml of ice cold acetone and kept for 7min at -20ºC.Later the embryos were again washed with Nano pure water for 5 mins. Finally, embryos were washed with PBS for 5 mins. After permeabilisation process the embryos were incubated with 1mL of NTMT buffer which is used for equilibration for 30 mins. Later eppendorf tubes containing embryos were wrapped with aluminium foil 300µL of NTMT+200µL of NBT/BCIP premix solution was added for staining endothelial cells of blood vessels and incubate for 10mins. After 10 mins 1ml of stopping buffer (0.25mM of EDTA+PBST) was added & 5-6 washings were done and incubated for 5 min. Later 1mL of solution containing 5% formamide + 10% hydrogen peroxide was added and incubated for 30 mins. Embryos were observed under stereo microscope and images were captured using a Zeiss AxioCam MR camera attached to a Zeiss Axioplan microscope under 5X and 10X using Zeiss fluorescent microscope. The extent of vessel formation in sub-intestinal vessel and inter-segmental region was observed.

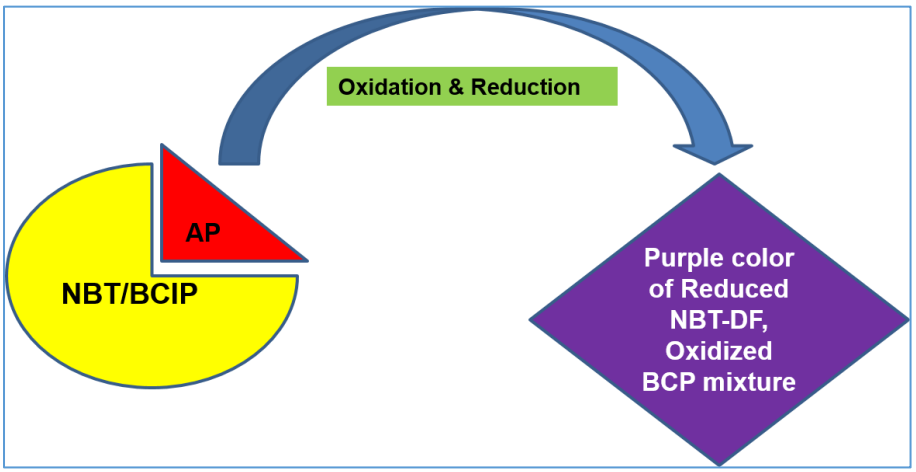


Figure 20: Qualitative estimation of blood vasculature Inhibition using NBT/BCIP.

4.15. COBRA assay

MDA-MB-231 cells were cultured using standard procedures and treated with CIS-1 and MIS-1 at IC_{50} concentration for two weeks. Genomic DNA was isolated from treated as well as untreated cells using Purelink genomic DNA isolation kit (Invitrogen). The genomic DNA was quantified and subjected to bisulphite conversion using Epitect Bisulfite Conversion Kit (Qiagen). Bisulphite converted DNAs from normal female and male blood were used as controls. These DNAs were then used to amplify a portion of LINE1 elements, HPRT promoter and SNRPN imprinting control region, using specific primers. PCR products for LINE1 and HPRT were digested with Taq1 restriction enzyme and that for SNRPN was digested with BstU1 restriction enzyme. The reaction products were resolved by polyacrylamide gel electrophoresis to identify digested fragments that reflect methylation. The primers used for the assay were listed below:

Primer Name	Sequence
HPRT F	GAGTTTTGGGAAAAGAGGATTG
HPRT R	AAACCCCTCCAAAAATACCC
LlNE 1 F	CCGTAAGGGGTTAGGGAGTTTTT
LlNE 1 R	RTAAAACCCTCCRAACCAAATATAAA
SNRPN F	ATAGTGTTGTGGGGTTTTAGGGGTTTAGTA
SNRPN R	CCCAAACTATCTCTTAAAAAAAACCACC

4.16. Combination effect of PARP-1 inhibitor and Cathepsin D inhibitor

PARP expression is thought to be increased both in triple-negative breast cancer and in breast cancer cells exposed to chemotherapy. The benefits of using PARP inhibitors in combination with other chemotherapy was found to be effective in treating TNBC. Olaparib is an oral PARP inhibitor in advanced-phase clinical development, which has been investigated as a single agent against BRCA1- and BRCA2-mutated advanced breast cancers in Phase II trial. We tested the combination of Olaparib and the one potent lead CatD inhibitor identified from our study [Crown J., *et al.,* 2012].

Stock solutions of 1 mM Olaparib and further dilutions (0.1, 1and 2 μM) were prepared in RPMI 1640. CatD inhibitor stock solution of 1 mM was prepared in DMSO, and other diluted solutions (0.25, 0.5, 1, 2, 5 and 10 μM) were prepared in RPMI1640. The final concentration of DMSO in all experiments was less than 0.5%. The stock solutions were sterilized using 0.22 μ microfilters under laminar flow hood and stored frozen. All dilutions were prepared fresh before addition to the cells, using phosphate buffered solution (PBS) as diluents. The cytotoxic effects of Olaparib and/or CatD inhibitor *in vitro* on MDA-MB-231 cells were tested with a rapid colorimetric assay

using MTT and compared with the untreated controls. This assay is based on the metabolic reduction of soluble MTT by mitochondrial enzyme activity of viable tumor cells, into an insoluble coloured formazan product, which can be measured spectrophotometrically after dissolving in DMSO.

Chapter 5. Results and Discussion

5.1. Selection and characterization of breast adenocarcinoma cell lines

A panel of breast cancer cell lines were selected and studied for their behavioural aspects in different tumor micro environments and target protein and gene expressions in our laboratory setup. As discussed in the "methods" column a wide range of selection was implemented in choosing the cell lines. All the cell lines were from the origin of ATCC and a kind gift from Incozen Therapeutics Pvt. Ltd.

5.1.1. CatD gene expression study

The cell lines were all grown in their respective media and plated into a 6-well plate as triplicates. The cell lines processed were namely MCF-7. MDA-MB-231, MDA-MB-468, and SKBR3. The mRNA quantifications of MCF-7, MDA-MB-231, MDA-MB-468, and SK-BR-3 were measured by reading absorbance at 260 nm (A_{260}). The relative CatD and GAPDH expression levels in all cell lines are shown in the **Figure 5.1.**

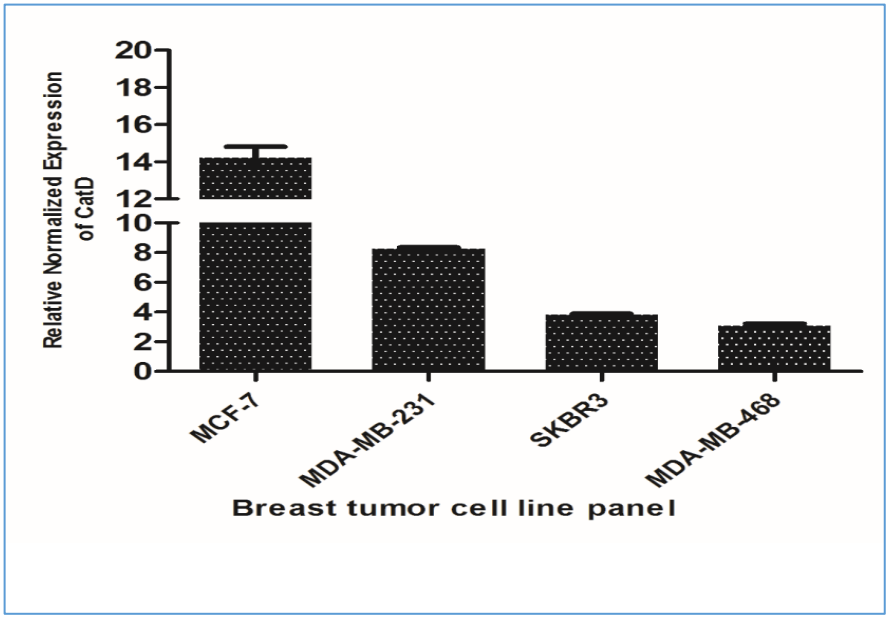


Figure 21: RT-PCR analyses for the expression of Cathepsin D (human) expressions in different human breast cancer cell lines that include triple-positive and triple-negative types of cancers.

CatD gene expression was higher in MCF-7 followed by triple-negative breast cancer cell lines MDA-MB-231 and MDA-MB-468. Among TNBCs, MDA-MB-231 is a highly metastatic tumor cell line and showed high levels of CatD gene expression. Thus, CatD expression was determined in the entire cell line panel selected and used to validate the inhibitory activities of the selected molecules.

We have further selected MCF-7 and MDA-MB-231 cell lines as main cell lines to study our compounds as they were known to show higher levels of CatD. Logically, MCF-7 is an estrogen positive cell line and MDA-MB-231 is an estrogen negative cell line. So, the test compounds inhibitions can be clearly understood for their effects on these two cell lines. MCF-7 was also reported to show increased levels of CatD expression by 17-β estradiol induction. Furthermore, many studies in the recent era have claimed the importance of developing small molecule inhibitors for the treatment of TNBCs. The TNBC cell line that we have studied, MDA-MB-231, is a highly metastatic and invasive cell line. We have further estimated the levels of CatD in the cell lysates.

5.1.2. Amount of CatD protein expression in cell lysates

The tumor cell lines were grown, seeded and were allowed to adhere by incubating them overnight. The cell lines were lysed and the supernatants were collected. The supernatants were measured for the levels of CatD using the substrate supplied in the Biovision kit. This is a fluorescence base assay that measures the concentrations of CatD in nano grams / 50 µL of cell lysate. We have found detectable ranges of 5-20 ng / 50 µL in all the selected cell lines.

The results are on par to gene expression studies with higher protein expression in MCF-7 followed by MDA-MB-231. This helps us in a better understanding of the cell lines in relation to our target of interest.

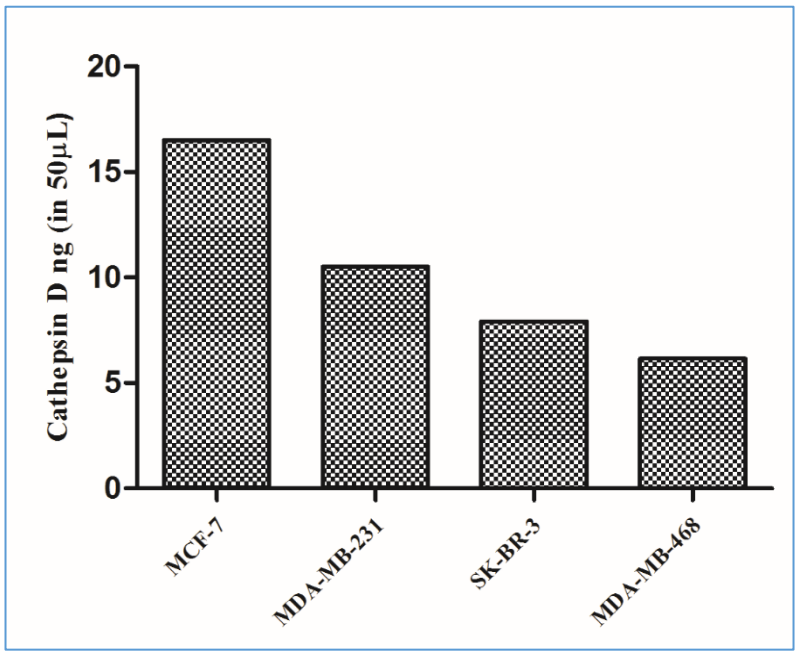


Figure 22: Levels of CatD protein expressions in different human breast cancer cell lines that include triple-positive and triple-negative types of cancers determined using fluorescence assay. The number of cells used for the assay 1x 10^5 cells/mL.

Based on these two studies, we confirmed the expression levels of CatD to be significantly higher in these tumor cell lines as seen in **Figure 5.2.** As known before, CatD is highly active at acidic pH i.e., tumor microenvironment. It is localised in the lysosomal components of the tumor cells. The tumor cell lines were hence grown in different pH levels to account their rate of survival.

5.1.3. Estimation of tumor cells survival in acidic microenvironment

The cell lines were studied for their survival in acidic microenvironment as per the protocol mentioned in "Methods" section. The media at pH 7.4 showed considerable equal cell viability when compared to media at pH 5.4 indicating adaptation of lysosomal volume at pH 7.4 analysed through flow sight (AMNIS, Millipore). The graph **Figure 5.3** was analysed on the basis of intensity, an indicative of percentage of cell survival**.**

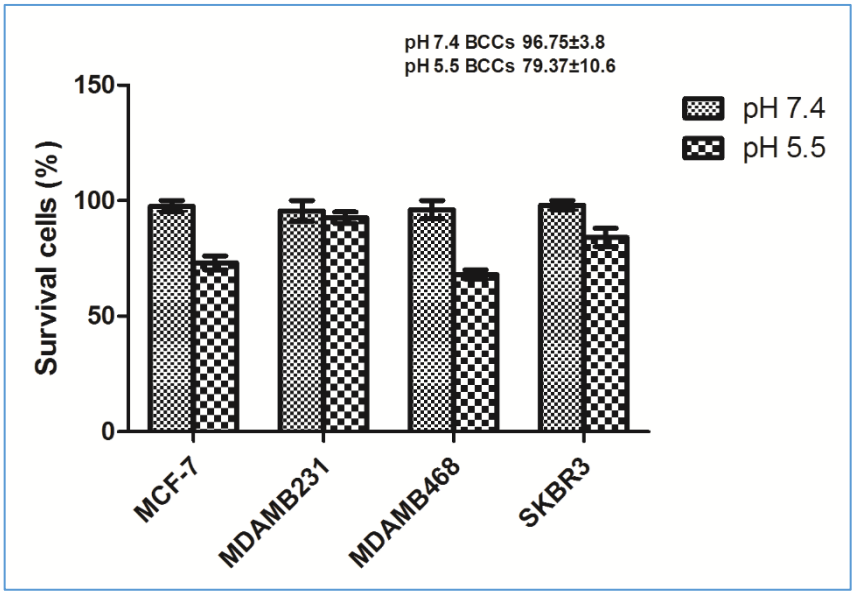


Figure 23: Cell survival analysis at different pH values of the growth medium was performed by Tryphan blue test. Data are reported as mean ±SEM of the percentage of surviving cells obtained in three separate experiments performed in triplicate. BCCs breast cancer cell lines.

The cell line MDA-MB-231 has equal survivability in both the pH ranges. This shows, there need not be any changes in the regular pH of media to study the role of CatD.

5.1.4. Conclusion

The whole idea in performing the above assays as a preliminary step is to confirm the suitability of breast adenocarcinoma cell lines in studying CatD as a target. There are many reports suggesting CatD as a prognostic biomarker in detecting breast cancers. Many clinical findings showed significantly higher expression rates of CatD in breast cancer cytosols of patients. The recent treatment therapeutic strategies were also looking forward in developing targeted therapies like the previous HER2 inhibitor Herceptin (trastuzumab). As CatD was known to be important even in triple negative cancers, developing protease inhibitors against TNBC is also a challenging research in supporting chemotherapy aspects of treatment. There are many other breast adenocarcinoma cell lines that can be considered with the above characterisitics. We have selected one cell line from each category which is considerably easy-to-handle and are with good proliferation rates. The invasion and metastatic features of MDA-MB-231 have further allowed us to develop different assay systems to understand the mechanisms through which the test compounds are showing inhibition. Taking these cell lines as basis, we have designed, synthesized and evaluated many small molecule inhibitors filtered from Asinex and BITS databases.

5.2. Design and bioevaluation of CatD inhibitors from Asinex database

A drug discovery process is to get the best hit in the closest possible way. Increasing insight into the genetics and molecular biology of cancer has resulted in the identification of an increasing number of potential molecular targets for anticancer drug discovery and development. These targets can be approached through exploitation of emerging structural biology, "rational" drug design, screening of chemical libraries, or a combination of these methods. The result is the rapid discovery of new anticancer drugs. Nowadays, a number of recent successful drugs have in part or in whole emerged from a structure-based research approach. As the structures of more and more proteins and nucleic acids become available, molecular docking is increasingly considered for lead discovery. Recent studies consider the hit-rate enhancement of docking screens and the accuracy of docking structure predictions. With more docking studies being undertaken, the "drug-likeness" and specificity of docking hits is also being examined. This process will ultimately help in discovering a new class of anti-cancer drugs that will help millions of cancer patients. CatD is also one such biomarker which has been studied from 1980s to know its role and importance in cancers. There are six crystal structures for CatD of human (Homo *sapiens*) origin deposited in the protein data bank till date.

In the present study, we employed e-pharmacophore based screening approach in the identification of novel non-peptidic small molecule inhibitors which were subjected to molecular docking studies. The identified hits were checked for their binding energies, ADME properties and *in vitro* studies based on which further optimizations were carried out.

5.2.1. Design of CatD inhibitors using structure based drug design

5.2.1.1. Selection of crystal structure and binding site analysis

There are many crystal structures of CatD of different origins till date deposited the protein data bank. The crystal structures of human origin were considered which were six co-crystallized with many substrates and chemical moieties. We have selected a crystal structure with PDB code 1LYB complexed with its specific inhibitor pepstatin A. The crystal structure selected has lesser resolution when compared to all other and is 2.5 Å resolution. The binding site of CatD is bilobed and with a deep active site pocket which can accommodate larger structural moieties. The active site pocket is lined with residues Gly35, Gly79, Gly 233, Ser80, Ser235, Asp231 and Asp33. The interactions Asp 231 and Asp 33 act as important catalytic active site. There are many other hydrophobic and π-π interactions between crystal ligand and the active site pocket. The crystal ligand of pepstatin A with interaction pattern was shown in **Figure 5.4**. Hence, these amino acid residues were taken as one of the basis in filtering the best hits from huge databases.

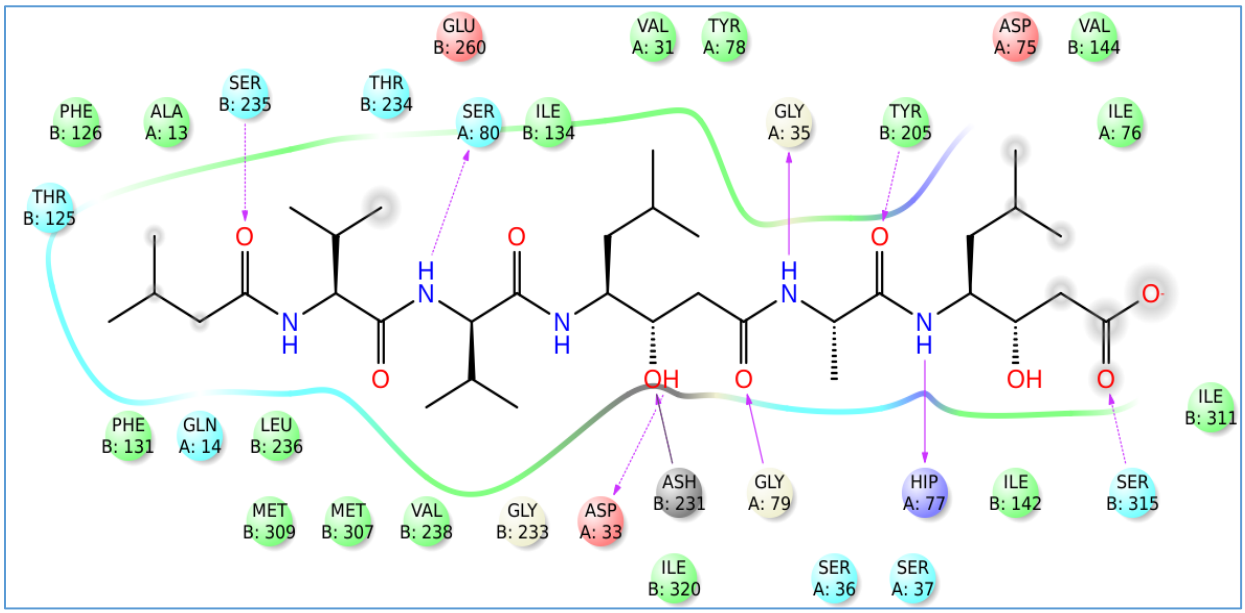


Figure 24: Interactive pattern of crystal ligand at the active site pocket of CatD 1LYB.

5.2.1.2. Binding active site pocket validation and e-pharmacophore generation

In order to validate the docking program utilized for e-pharmacophore generation, the crystal ligand was prepared and re-docked on to the grid of the active site. The crystal ligand was found to be oriented in the similar manner as if in the crystal structure with an RMSD of 0.179 Å. The ligand was found to be retaining the active site interactions as of crystal structure with XP glide score of -12.419 kcal.mol-1.

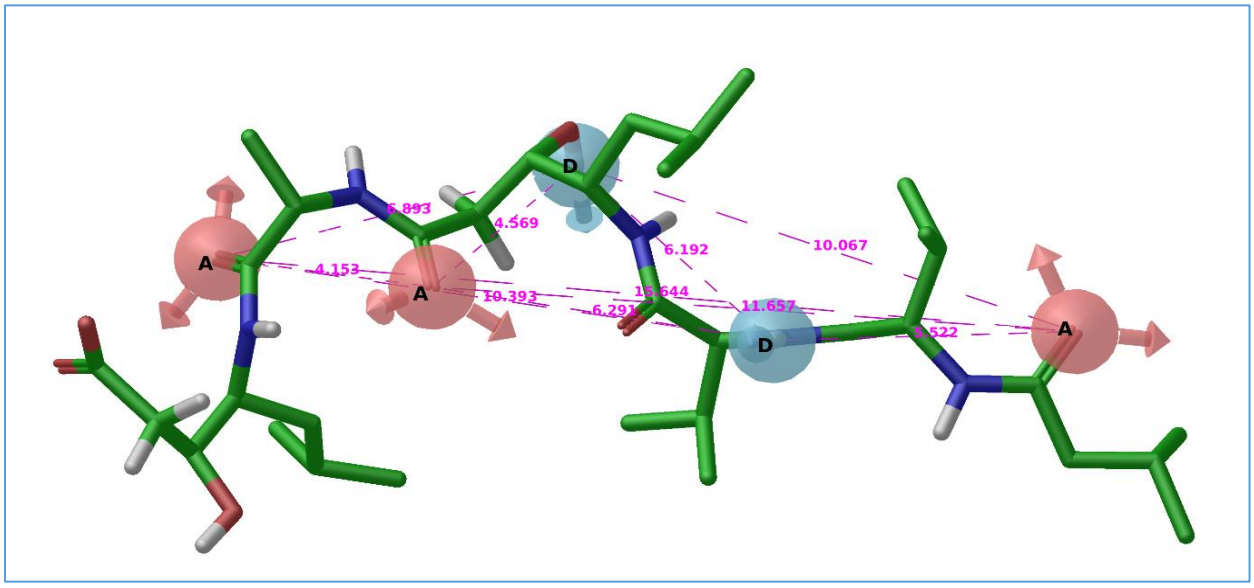


Figure 25: e-Pharmacophore generated for the crystal ligand from the crystal structure 1LYB.

Pharmacophore hypothesis was generated using the Glide descriptors which consist of polar, non-polar, van der Waals and electrostatic binding energies of the ligand with the protein which were mapped for the identification of pharmacophoric features. The pharmacophore generation protocol applied for the ligand yielded a hypothesis with ten features as shown in **Figure 5.5.** There were six acceptors and four donor groups in the generated ten-featured e-pharmacophore. On the basis of energies between the hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively ionisable (N), positively ionisable (P), and aromatic ring (R) moieties, we generated 3–7-featured pharmacophore models through combinations. The combinations were made by selecting different features each time generating respective pharmacophores. This further

helped us attain best-fit small molecules filtered on the basis of specific pharmacophore features. A set of 1084 compounds were screened against all the generated pharmacophores using PHASE, and the pharmacophore efficiencies were further calculated as discussed in the "Methods" section. Out of many hypotheses generated, we selected the five best hypotheses on EF and GH scores. Pharmacophores AAADD1, AADDD, AAADD2, AADD, and AAD displayed good EF and GH scores **Table 5.1,** and so were further selected for molecular docking to identify best hits from the databases. The binding affinities of the features were considered in preparing different e-pharmacophore **Table 5.2.**

Table 1: Energy scores of each feature in the generated e-pharmacophore.

Rank	Features	Energy values	Type
1	A6	-0.7	Aromatic ring
2	A7	-0.7	Aromatic ring
3	D11	-0.7	H-bond donor
4	A4	-0.59	Aromatic ring
5	D13	-0.59	H-bond donor
6	A1	-0.54	Aromatic ring
7	A3	-0.54	Aromatic ring
8	A5	-0.51	Aromatic ring
9	D8	-0.33	H-bond donor

Table 2: Statistical parameters followed on best five hypotheses after screening of the decoy sets of molecules.

Top 5 Hypotheses	AAADD.1	AADDD	AAADD.2	AADD	AAD
H_t[a]	223	78	124	7	99
H_a[b]	52	44	49	4	37
%Y[c]	23.32	56.41	39.52	57.14	37.37
%A[d]	61.90	52.38	58.33	4.76	44.05
EF[e]	3.01	7.28*	5.10	7.37	4.82
Fn[f]	32	40	35	80	47
Fp[g]	171	34	75	3	62
GH[h]	0.27	0.54*	0.41	0.44	0.37

[a] Total number of hit molecules from the database; [b] Total number of active molecules in hit list; [c] % yield of actives [(Ha/Ht) × 100]; [d] % Ratio of actives [(Ha/A) × 100]; [e] Enrichment Factor; f False negatives [A - Ha]; [g] False Positives [Ht - Ha]; [h] Goodness of fit score; * best EF and GH scores among five hypothesis

5.2.1.3. E-pharmacophore based screening of Asinex database

The top 5 e-pharmacophores were employed for screening asinex database (a commercial database) in search of compounds with similar pharmacophore features. The search criterion was put to match 3-5 features of the pharmacophore. The aim of virtual screening workflow operated in this work is to reduce the enormous chemical library into a manageable number of compounds which would be comfortable for identification of inhibitors thereby increasing the chances of

identification of drug candidate. The pharmacophore based screening of asinex database resulted in compounds with significant fitness value, which is a measure of how well the compound fits the pharmacophore. In the present study, 25,994 compounds out of 5 lac from asinex database were filtered after HTVS glide docking. Of which, 9454 compounds were filtered after SP docking and a total of 1785 compounds were shortlisted after final XP docking.

5.2.2. Molecular docking of the selected compounds from Asinex database

The XP glide score of pepstatin was -12.49 kcal.mol^{-1} which was initially docked. This was considered in shortlisting the 1785 selected compounds out of XP docking. The docking calculations were performed in accordance with the molecular docking protocols and interactions between the molecules and target protein. Molecules were ranked on the basis of docking score, interactions with amino acid residues, and numbers of hydrogen bonds in comparison with crystal ligand Pepstatin A and were further selected for SP and XP docking. The active site of CatD was found to be interacting with Pepstatin A with residues, namely Asp33, Asp231, Gly35, Gly79, Gly233, Ser80, and Ser235. The active site contains two aspartate residues, Asp33 and Asp 231, each of which plays a vital role in the catalytic function of CatD. The final XP docking resulted in 523, 139, and 124 molecules from 5-, 4-, and 3-featured e-pharmacophores, respectively. The active site has many areas that can accommodate bulkier substituents from the databases used. In addition, the docking score, ADME properties, and the alignment of the molecule in the active site were considered to select a total of 13 asinex molecules for in vitro activity studies.

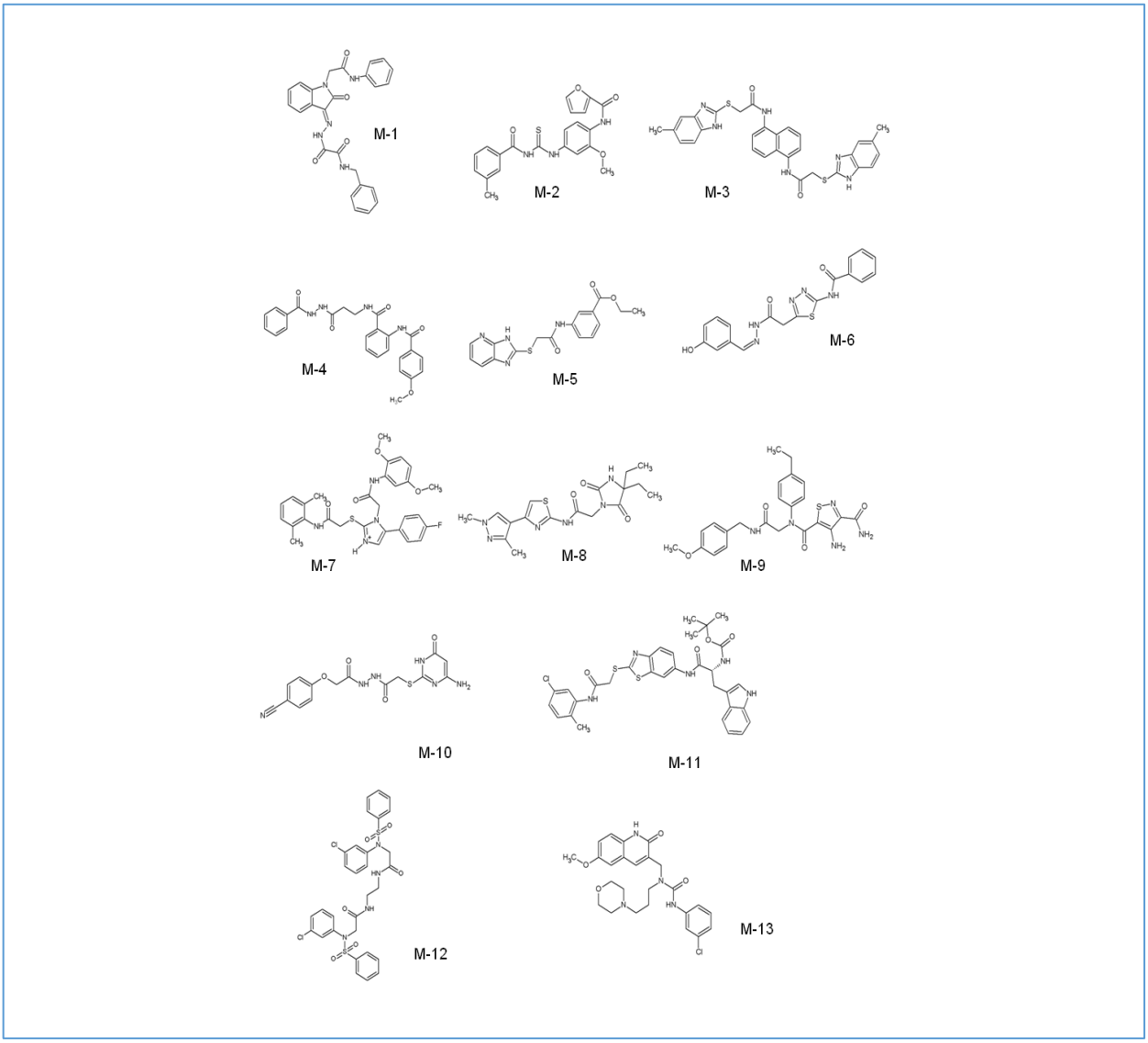


Figure 26: Structures of the selected 13 compounds from asinex database as CatD inhibitors

All the 13 selected compounds were found to retain the interaction pattern with the binding site and the XP glide score ranged between -7 kcal.mol^{-1} to -11 kcal.mol^{-1}. The π-π interactions were also considered while filtering the compounds from the database. The glide score energies of the compounds and the docking site interactions of the selected compounds were listed in **Table 5.3.**

Table 3: Glide score and amino acid interactions for the crystal ligand and selected 13 compounds

Asinex Molecules	**GLIDE score**	**Interacting residue** **H-Bond**
M-1	-9.34	Ser80,Gly35,Gly79,Ser235
M-2	-9.31	Ser80,Ser235,Gly233,Gly35
M-3	-9.25	Ser235,Ser80,Asp231,Tyr205
M-4	-8.98	Gly35,Asp231,Gly79(2),Ser80,Ser235
M-5	-8.41	Gly233,Gly79(2),Gly35,Tyr205
M-6	-8.42	Ser80,Asp231(2),Gly35,Gly233,Ala13
M-7	-10.6	Ser235,Ser80,Gly79,Gly35
M-8	-8.07	Ser80,Gly79,Gly233,Gly35
M-9	-8.9	Ser235,Ser80,Gly233,Gly79
M-10	-7.46	Gly233,Gly79,Asp33,Asp231,Thr125
M-11	-10.97	Ser235,Ser80(2),Gly35,Thr125
M-12	-9.82	Asp231,Gly79,Ser235,Ser80,Thr234
M-13	-8.78	Asp231,Gly35,Gly79,Ser235,Leu236
Pepstatin	-12.419	Asp33,Asp231,Gly35,Gly79,Gly233,Ser80,Ser235

The selected 13 compounds have diverse scaffolds of thiadiazole, thiazole, furan, oxazole and pyrazole moieties in their structures. The selected various scaffolds were taken for further enzymatic and *in vitro* studies to evaluate the effective CatD inhibitor.

5.2.3. ADME predictions of the selected asinex molecules

A total of 13 molecules from the ASINEX database were selected and examined for pharmaceutically relevant drug-like properties. The QikProp module was used for validating pharmaceutically relevant properties of the molecules in relation to Absorption, Distribution, Metabolism, and Excretion (ADME). Most of the hits had good calculated partition coefficients (QPlogPo/w), QPlogS, and QPPCaco, which are the key measures of solubility, absorption, and cell permeability. Lipinski's rule of five displayed zero violations of the molecules. The expected human oral absorption for the molecules ranged from 57 to 100 %. All the properties were listed out in the **Table 5.4**. All the 13 compounds were found to obey Lipinski rule of five, showing good percentage oral absorption thereby reflecting the drug likeliness of the compounds.

Table 4: *In silico* pharmacokinetic profile for the selected 13 compounds.

Asinex Molecules	**QPlogPo/w[a]**	**QPlogS[b]**	**QPPCaco[c]**	**Rule of 5[d]**	**% Human oral absorption[e]**
M-1	3.905	-6.033	193	0	90.72
M-2	5.102	-7.915	1007.1	1	100
M-3	5.487	-8.011	446.7	2	80.58
M-4	4.831	-6.417	464.5	0	100
M-5	2.787	-5.106	378.3	0	89.4
M-6	2.428	-5.658	52.1	0	71.89
M-7	6.402	-7.889	2989.2	2	100
M-8	2.457	-4.973	418.7	0	88.26
M-9	1.289	-2.715	20.5	0	57.97

M-10	0.037	-4.654	5.4	0	40.25
M-11	7.796	-11.424	416.7	2	93.57
M-12	3.336	-5.249	135.5	1	71.68
M-13	3.383	-4.26	312	0	91.4
Pepstatin	2.73	-4.645	3.559	3	13.92

[a] Predicted octanol/water partition co-efficient log p (acceptable range from −2.0 to 6.5). [b] Predicted aqueous solubility, log S. (acceptable range from –6.5 to 0.5) [c] Predicted Caco-2 cell permeability in nm/s (acceptable range: <25 is poor and >500 is great). [d] Number of violations of Lipinski's rule of five (maximum is 4) [e] Percentage of human oral absorption (<25% is poor and >80% is high).

5.2.4. *In vitro* enzyme inhibition studies

All the 13 asinex selected compounds were tested for CatD inhibition followed by selective renin inhibition assay using the kits procured from Biovision, Milpitas, U.S.A.

5.2.4.1. Cathepsin D enzyme inhibition study

Initially, Pepstatin standard graph was plotted considering concentrations from 100µM to 0.1 nM as per the protocol mentioned in the "methods" section. Furthermore, 13 molecules were selected for *in vitro* CatD inhibition using a flourometric assay (BIOVISION, U.S.A). The initial screening at 25µM revealed that three out of 13 molecules displayed >50 % inhibition as shown in **Figure 5.7.** These molecules were further screened at 10 µM, and fluorescence measurements (Ex/Em of 328/460 nm) showed an inhibition range from 40 to 50 %. Active molecules were correlated with their docking calculations. Molecules 1, 3, and 10 from the Asinex database showed strong docking interactions and docking poses inside the active pocket compared to all other database molecules. Among these, the docking poses of molecules 1 and 10 with different orientations

inside the active site pocket are shown in **Figure 5.8.** The active site pocket of CatD has two lobes and is connected by a cleft. The residues Gly79 and Ser80 forming the cleft are not displayed in the figure to provide a clear view of the alignment of molecules inside the active site pocket. These deselected residues are coloured green, as shown in **Figure 5.8.**

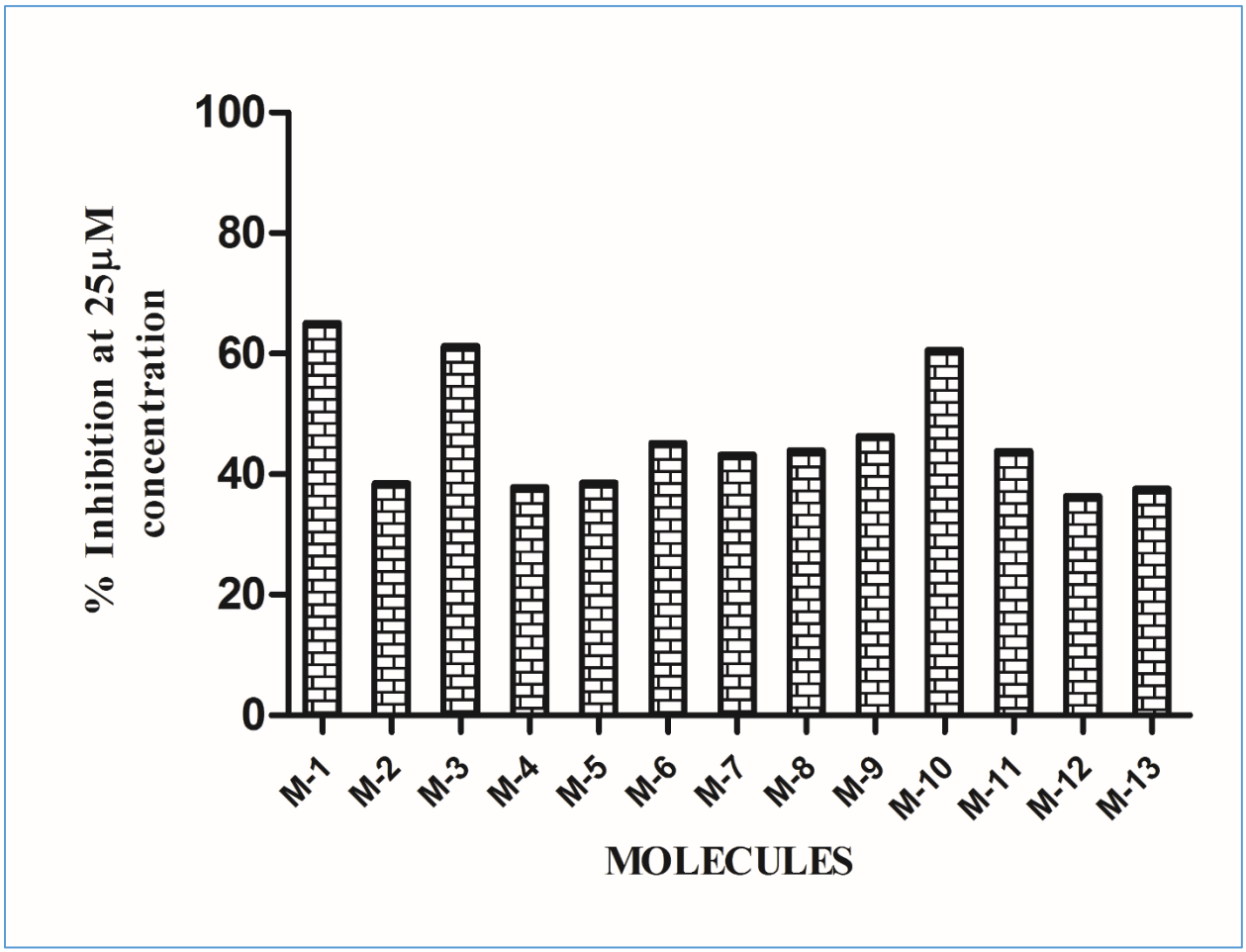


Figure 27: Percentage inhibitions of Asinex database molecules at 25µM concentrations analyzed using fluorescence plate reader at Ex/Em of 328/460 nm. Three out of 13 were showing >50 % inhibition profiles of CatD.

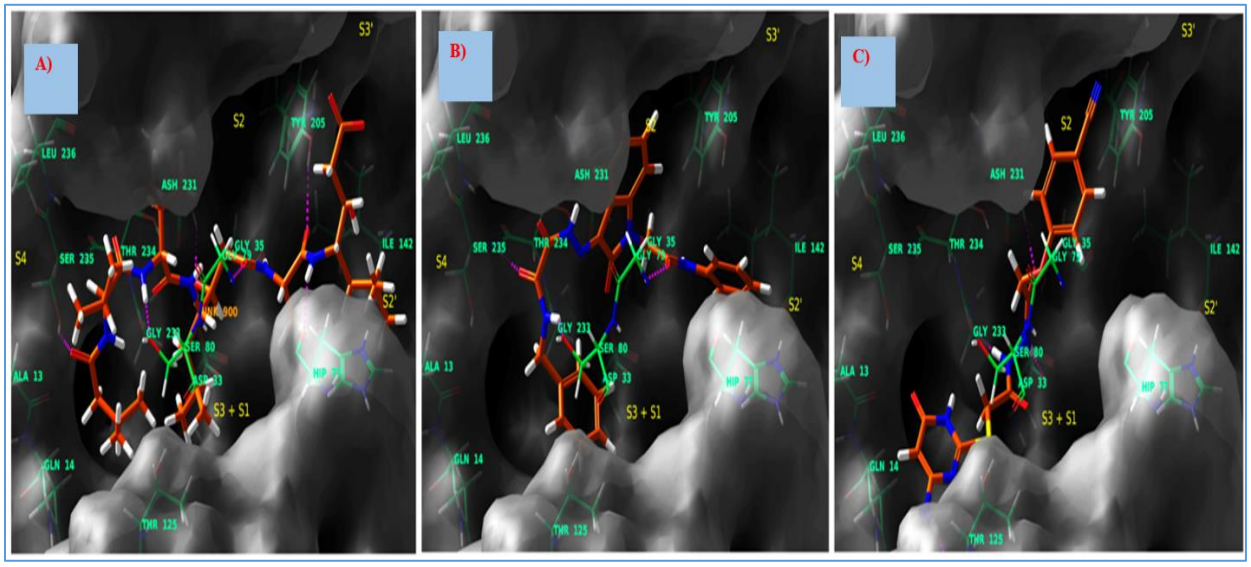


Figure 28: Binding poses displaying interactions A) Pepstatin A, crystal ligand (1LYB), B) most active Molecule 1 from docking, C) Active compound with good alignment in to the groove lead 10. All active site-forming crucial residues are shown in spring green and sites S1, S2, S3, S4, S2′, and S3 ′ are in yellow text.

This is a preliminary assay for the evaluation of selected compounds for their CatD inhibition. This further helps us to correlate docking studies with *in vitro* enzyme studies to construct a rationale in selecting the molecules for mechanistic based studies. Molecular docking process helps us to hasten the process in filtering the required compounds from a huge diversified scaffold library. The selected compounds may or may not possess inhibitory property on a laboratory scale-up. The molecules showing considerable CatD inhibition were alone taken for pathway analysis.

5.2.4.2. Renin enzyme inhibition study

Many studies reported the enzymological similarities between renin and CatD. The inhibitors against CatD were also known to inhibit renin due to the structural similarities which makes the molecules non-specific and less-effective in treating the actual ailment. Hence, the molecules which were known to have good CatD inhibition were also screened against renin. Molecules 1, 3 and 10 having good CatD enzyme inhibition were tested at 25 µM concentrations against renin. The fluorescence measurements were made at Ex/Em of 328/552 nm. This is a kinetic mode of study in which two timings were noted to show greater difference in inhibition i.e., 30 and 60 min. The relative inhibitions were calculated for these three molecules. Amongst all, molecule 1 was found to show lesser inhibition against renin which makes it more specific towards CatD. The percentage inhibitions of all the three molecules were shown in **Table 5.5.**

Table 5: The inhibition profile of the active Asinex molecules against CatD off-target "Renin", performed using fluorescence assay.

Asinex Molecules	**% Renin inhibition at 25 µM**
M-1	33.28
M-3	57.28
M-10	50.74

5.2.5. *In vitro* growth proliferation and cytotoxicity studies

CatD gene and protein expressions were studied in the above assays and we have selected a panel of breast cancer cell lines. The rank in order of CatD expression is MCF-7> MDA-MB-231 > SK-BR-3> MDA-MB-468. The therapeutic strategies have encouraged target specific research in developing inhibitors to treat breast carcinomas. There is also an increasing need in establishing chemotherapy approach for non-responsive TNBCs. Apart from enzymatic assays, all the molecules were studied for their role of inhibition in breast cancer cell lines. The immune profile and mutational status of the cell lines are listed in the **Table 5.6.**

We performed this assay to know the inhibitory activity of selected asinex compounds on all the cancer cell lines. We have selected breast adenocarcinoma cell lines (MCF-7, MDA-MB-231, MDA-MB-468 and SK-BR-3) and treated with molecules M1-M13. Importantly, MCF-7 and MDA-MB-231 cell lines were considered to know the effect of the molecules and were proceeded for their effect on other cell lines. The mutational data of the cell lines was also correlated with the growth proliferation effect. Half-maximal growth inhibition (GI_{50}) values were calculated based on the percentage inhibitions of all the molecules tested at 100 μM – 1 nM concentration ranges. Surprisingly, the active compounds have shown inhibitory activity in accordance to CatD gene expression and the GI_{50} values are as indicated in the **Table 5.7.**

In contrast to its effect on MCF-7 and MDA-MB-231, the active molecules did not affect the proliferation of HEK293T cells as seen in the **Table 5.7.** More than a fifty-fold change in GI_{50} was observed between the most sensitive cell line (MCF-7) versus normal HEK293T cells, indicating a better selectivity of M1, M3 and M10 towards cancerous cells with inherently high levels of CatD. We used paclitaxel as the reference standard as it is reported to inhibit all types of breast cancer cell lines effectively.

Table 6: Immuno profile and mutational status of the selected breast cancer cell line panel.

Sample No.	Cell line	Immuno profile	Mutation status
1	MCF-7	ER+, PR+,HER2-	CDKN2A, PIK3CA
2	MDA-MB-231	ER-,PR-,HER2-	CDKN2A, RAS, TP53, BRAF
3	MDA-MB-468	ER-,PR-,HER2-	PTEN, RB1, SMAD4, TP53
4	SK-BR-3	ER-,PR-,HER2+	TP53

Table 7: GI_{50} (μM) values of the promising ASINEX molecules tested on four different breast cancer cell lines and a normal cell line (HEK 293T). Molecules showing less than 10μM activity were studied completely.

Asinex molecules	**Half-Maximal concentration (GI_{50} μM)**				**HEK (CC_{50} μM)**
	MCF-7	**MDA-MB-231**	**SK-BR-3**	**MDA-MB-468**	
M-1	1.97±0.25	2.27±1.13	8.68±0.74	14.05±1.5	51.12±1.13
M-2	46.67	100	-	-	-
M-3	1.28±0.45	3.70±0.05	6.63±0.10	22.92±0.03	59.06±2.36
M-4	21.29±1.14	>100	-	-	-
M-5	64.25	97.82	-	-	-
M-6	>100	96.23	-	-	-
M-7	90.92	94.66	-	-	-
M-8	191.3	171.2	-	-	-
M-9	>100	1457	-	-	-
M-10	7.35±0.65	17±0.24	4.86±0.23	57.92±2.01	56.54±3.5
M-11	>100	>100	-	-	-
M-12	>100	>100	-	-	-
M-13	>100	>100	-	-	-
Paclitaxel[a]	0.0072	0.00241	0.005	0.0048	-

[a] Reference standard used

Based on the results obtained, the molecules 1 and 3 were found to show greater inhibition on MCF-7 and MDA-MB-231 and is in accordance to the levels of CatD gene expressions in the cell lines. In comparison to specificity towards CatD, M1 has more inhibition towards CatD

than renin where as M3 has almost equal levels of inhibition on CatD and renin. This confirms Molecule 1 to be potent CatD inhibitor out of the selected 13 molecules from the asinex database. Therefore, we performed cell cycle analysis of M1 on MDA-MB-231 cells, the TNBC type of cells. We have further focused our research on developing effective inhibitors to treat TNBC kind of cancers, as CatD was found to have a very important role in their invasion and metastasis.

5.2.6. CatD inhibition led to cell cycle arrest

One of the crucial features to examine the final arbiter of cell fate was cell cycle arrest. The effect of cell cycle arrest during CatD inhibition was determined. The effect of M1 on cell cycle in MDA-MB-231 cancer cell lines was determined by propidium iodide staining and analysed by flow sight (Amnis, Millipore). **Figure 5.9** displays the increase in cells in the G2/M phase ($p < 0.01$) when treated with 2.27 μM i.e., GI_{50} value of M1 on MDA-MB-231 cells compared to the control. There is a decrease in G0/G1 phase in treated cells when compared to control. Cell cycle data can be further correlated with the levels of CDK1 after treatment with M1 in these cells. The imaging system showed PI fluorescence differentiating various phases of cells with their DNA content. The samples were further analysed by IDEAS analysis software, and a histogram was plotted as Intensity in channel 4 (Propidium iodide) on X-axis and Normalized frequency on Y-axis.

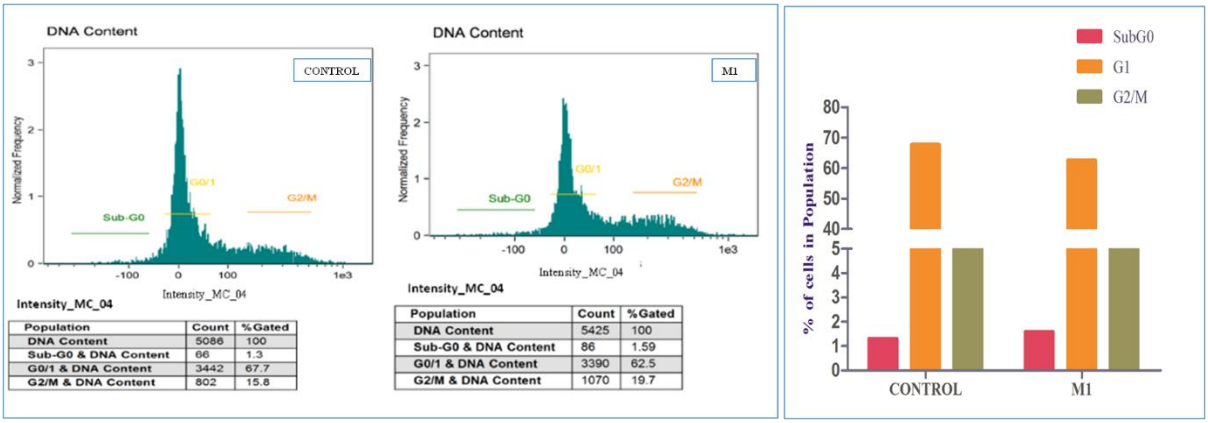


Figure 29: Cell cycle analysis of control, and Molecule (M) 1 with statistical parameters analyzed through IDEAS software gating the single cell population. Channels 1 & 9 are bright fields; Channel 4 (488 nm) for Propidium iodide. The same was represented by histogram using Graph Pad prism v.6.0.

5.2.7. Conclusion

CatD is a promising target whose inhibitors can effectively treat breast cancers and its invasion. In the present study, we have designed small molecule CatD inhibitors using molecular modelling studies and structure based drug design strategy. The inhibitors filtered were further tested against enzyme inhibition on an *in vitro* setup. We found 3 out of 13 to inhibit CatD at >50% and were studied on *in vitro* cancer cell lines. The cytotoxicity studies also resulted in the very three molecules i.e., M1, M3 and M10 to cause greater growth inhibition on the breast cancer cell lines. On whole, molecule 1 was found to cause greater CatD inhibition and lesser off-Target renin inhibition making it more specific to CatD. This was further confirmed by its inhibition on MCF-7 and MDA-MB-231 at GI_{50} values of 1.97±0.25 µM and 2.27±1.13 µM respectively. M1 further caused G2/M cell cycle phase arrest which can be correlated with many other kinases. Thus, we have achieved a new small molecule CatD inhibitor M1 from a huge commercial asinex database of 5 lac compounds. Further work is in progress to increase the possibility of affinity of these inhibitors through structural modification and developing structure-activity relationship.

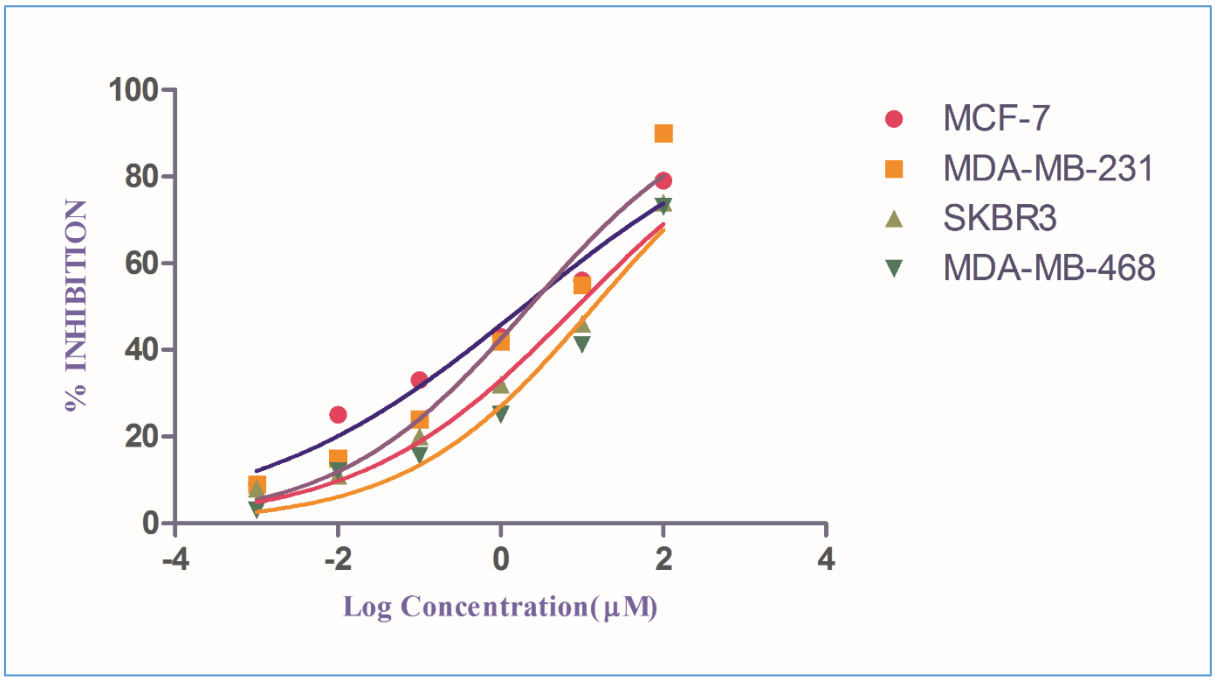


Figure 30: Growth inhibition curves representing cytotoxicity effect of M1 on different breast cancer cell lines as measured by MTT dye assay after different periods of cultivation in medium containing 10% FBS. The points represent mean of triplicates from a representative experiment.

5.3. Design and bioevaluation of CatD inhibitors from in-house BITS database

In the previous chapter, we have screened compounds from a huge commercially available asinex database. The present chapter deals with filtering of small molecule, non-peptidic inhibitors against CatD from our own in-house BITS database. We have a database of 930 molecules which were synthesized and characterized using mass spectroscopy, NMR spectroscopies in our synthesis laboratory. As we have already optimized all the schemes, the synthesis of the filtered molecules is going to be easy. So, we have screened the compounds using molecular docking studies against CatD crystal structure. As there are <1000 compounds we have employed direct glide XP docking using the grid generated from CatD 1LYB crystal structure.

5.3.1. Design of CatD inhibitors from BITS database

5.3.1.1. Selection of crystal structure and binding site analysis

The same procedure was followed as per section 5.2.1.1.

5.3.1.2. Grid generation

GLIDE was used for the virtual screening of compound libraries. The crystal structure of CatD (PDB ID: 1LYB) was minimized using the protein preparation wizard (OPLS_2005 force field). The co-crystallized ligand, Pepstatin A, was selected during the receptor grid generation to define x, y, and z axes of the grid. The sites for molecular docking were defined during the grid generation which allowed us to filter molecules from the selected active site. The grid generated was selected during the docking process which will generate the glide energy score based on the fitness of the molecules into the grid.

5.3.1.3. Molecular docking of in-house BITS database molecules

The XP glide score of pepstatin was -12.49 kcal.mol^{-1} which was initially docked. This was considered in shortlisting the 930 BITS compounds out of XP docking. The docking calculations were performed in accordance with the molecular docking protocols and interactions between the molecules and target protein. Molecules were ranked on the basis of docking score, interactions with amino acid residues, and numbers of hydrogen bonds in comparison with crystal ligand Pepstatin A and XP docking. The active site of CatD was found to be interacting with Pepstatin A with residues, namely Asp33, Asp231, Gly35, Gly79, Gly233, Ser80, and Ser235. The active site contains two aspartate residues, Asp33 and Asp 231, each of which plays a vital role in the catalytic function of CatD. After XP docking, 16 molecules were shortlisted out of 930 molecules from the database. The active site has many areas that can accommodate bulkier substituents from the databases used. In addition, the docking score, ADME properties, and the alignment of the molecule in the active site were also considered to select a total of 16 BITS molecules for in vitro activity studies. The selected molecules have diversified scaffolds with different interaction patterns.

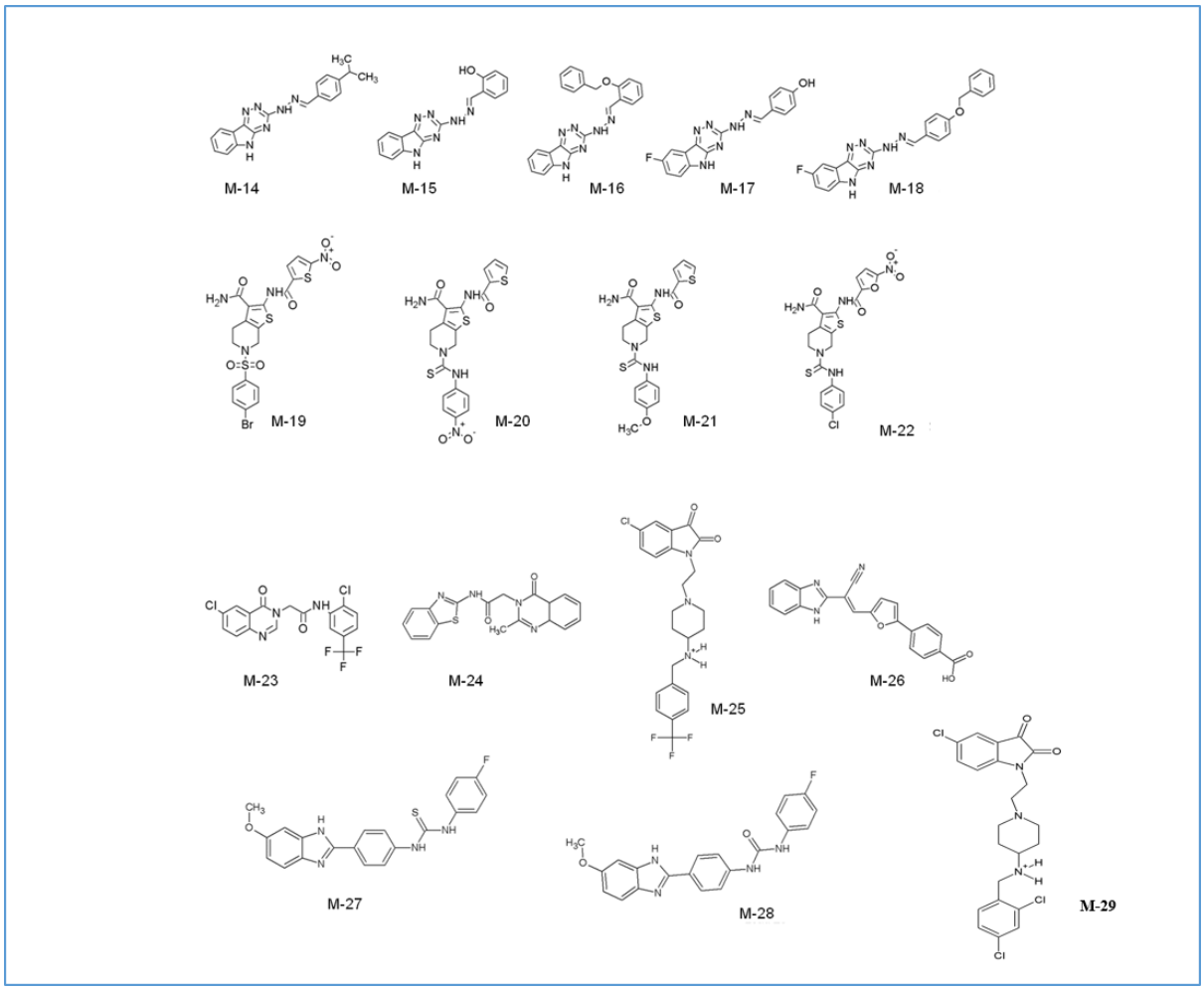


Figure 31: Structures of the molecules selected from in-house BITS database.

All the 16 selected compounds were found to retain the interaction pattern with the binding site and the XP glide score ranged between -6 kcal.mol^{-1} to -9 kcal.mol^{-1}. The π-π interactions were also considered while filtering the compounds from the database. The glide score energies of the compounds and the docking site interactions of the selected compounds were listed in **Table 5.8.**

Table 8: Interactions of the selected 16 compounds at CatD active site pocket.

BITS Molecules	**GLIDE score**	**Interacting residue H-Bond**
M-14	-7.86	Asp33,Gly35
M-15	-7.9	Gly233,Asp33,Gly35
M-16	-6.89	Asp231,Gly79
M-17	-7.21	Gly79,Asp33,Asp231,Thr125
M-18	-6.82	Asp33,Asp231,Gly79
M-19	-6.13	Asp231,Gly233,Ser235,Tyr205,Thr234
M-20	-7.08	Gly35,Gly233
M-21	-7.05	Gly233(2),Gly35
M-22	-6.13	Gly35,Gly233
M-23	-6.9	Asp231,Gly233
M-24	-6.62	Gly35,Gly79
M-25	-8.19	Gly233,Asp231,Gly79
M-26	-8.01	Gly35,Asp231.Ser235
M-27	-7.58	Ser235,Gly233,Gly35(2)
M-28	-7.74	Ser235,Gly233,Gly35(2)
M-29	-8.02	Gly 79,Gly 233, Asp 231
Pepstatin A	-12.419	Asp33,Asp231,Gly35,Gly79,Gly233,Ser80,Ser235

The selected 16 compounds have diverse scaffolds of triazole, isatin, furan and thiophene moieties in their structures. The selected various scaffolds were taken for further enzymatic and *in vitro* studies to evaluate the effective CatD inhibitor.

5.3.1.4. ADME predictions of the selected asinex molecules

A total of 16 molecules from the BITS database were selected and examined for pharmaceutically relevant drug-like properties. The QikProp module was used for validating pharmaceutically relevant properties of the molecules in relation to Absorption, Distribution, Metabolism, and Excretion (ADME). Most of the hits had good calculated partition coefficients (QPlogPo/w), QPlogS, and QPPCaco, which are the key measures of solubility, absorption, and cell permeability. Lipinski's rule of five displayed zero violations of the molecules. The expected human oral absorption for the molecules ranged from 57 to 100 %. All the properties were listed out in the **Table 5.9.** All the 16 compounds were found to obey Lipinski rule of five, showing good percentage oral absorption thereby reflecting the drug likeliness of the compounds.

Table 9: In silico pharmacokinetic profile for the selected 16 BITS database compounds.

BITS Molecules	**QPlogPo/w[a]**	**QPlogS[b]**	**QPPCaco[c]**	**Rule of 5[d]**	**% Human oral absorption[e]**
M-14	3.31	-5.24	611.07	0	96.21
M-15	1.70	-3.40	269.18	0	80.39
M-16	3.91	-4.74	915.43	0	100.00
M-17	1.84	-4.01	183.52	0	78.21
M-18	4.32	-6.18	624.39	0	100.00
M-19	2.22	-5.75	14.07	1	47.55
M-20	3.54	-7.35	45.01	0	77.28
M-21	4.11	-6.74	403.43	0	100.00
M-22	3.28	-7.03	36.47	1	61.13
M-23	3.98	-6.10	1087.93	0	100.00

M-24	2.43	-4.31	859.45	0	93.68
M-25	3.42	-4.12	73.95	0	80.44
M-26	3.51	-6.06	43.26	0	76.77
M-27	3.87	-6.03	813.58	0	100.00
M-28	4.88	-6.88	2308.28	0	100.00
M-29	3.39	-3.89	583.74	0	81.19
Pepstatin A	2.73	-4.65	3.56	3	13.92

[a] Predicted octanol/water partition co-efficient log p (acceptable range from −2.0 to 6.5). [b] Predicted aqueous solubility, log S. (acceptable range from –6.5 to 0.5) [c] Predicted Caco-2 cell permeability in nm/s (acceptable range: <25 is poor and >500 is great). [d] Number of violations of Lipinski's rule of five (maximum is 4) [e] Percentage of human oral absorption (<25% is poor and >80% is high).

5.3.2. *In vitro* enzyme inhibition studies

All the 16 BITS selected compounds were tested for CatD inhibition followed by selective renin inhibition assay using the kits procured from Biovision, Milpitas, U.S.A.

5.3.2.1. Cathepsin D enzyme inhibition study

Initially, Pepstatin standard graph was plotted considering concentrations from 100µM to 0.1 nM as per the protocol mentioned in the "methods" section. Furthermore, 16 molecules were selected for *in vitro* CatD inhibition using a flourometric assay (BIOVISION, U.S.A). The initial screening at 25µM revealed that four out of 16 molecules displayed >50 % inhibition as shown in **Figure 5.12**. These molecules were further screened at 10 µM, and fluorescence measurements (Ex/Em of 328/460 nm) showed an inhibition range from 40 to 50 %. Active molecules were correlated with their docking calculations. Among these, the docking poses of

molecule 29 with different orientations inside the active site pocket are shown in **Figure 5.13.** The active site pocket of CatD has two lobes and is connected by a cleft. The residues Gly79 and Ser80 forming the cleft are not displayed in the figure to provide a clear view of the alignment of molecules inside the active site pocket. These deselected residues are coloured green, as shown in **Figure 5.13.**

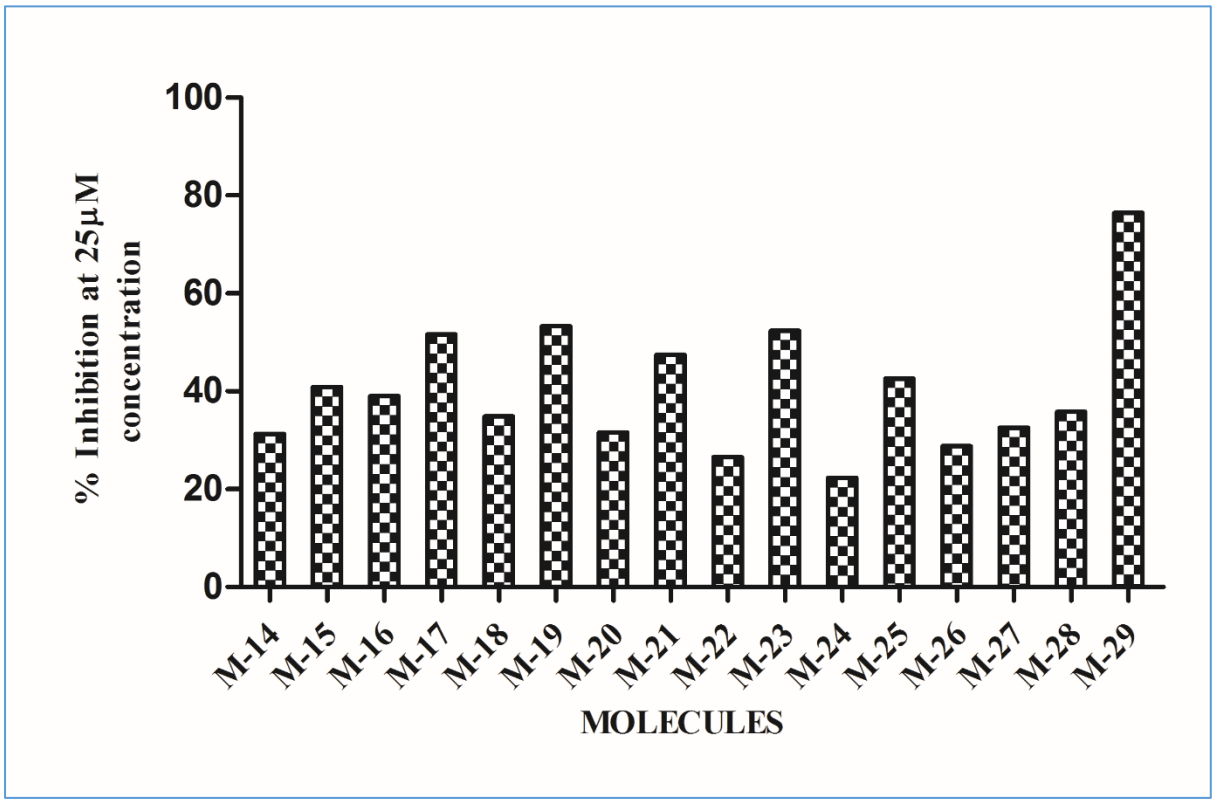


Figure 32: Percentage inhibitions of BITS database molecules at 25µM concentrations analyzed using fluorescence plate reader at Ex/Em of 328/460 nm. Four out of 16 were showing >50 % inhibition profiles of CatD.

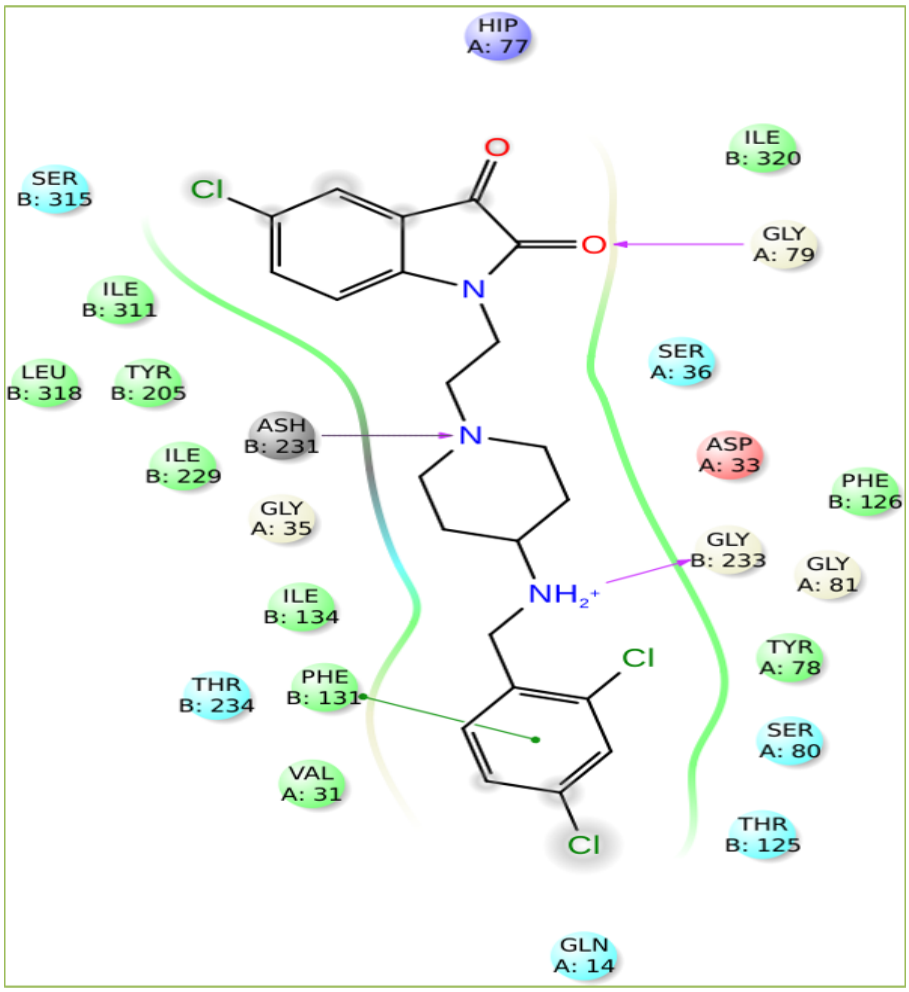


Figure 33: Ligand interaction pattern of the most active M29 among the analogues from docking. All active site-forming crucial residues are shown as hydrogen bond interactions.

This is a preliminary assay for the evaluation of selected compounds for their CatD inhibition. This further helps us to correlate docking studies with *in vitro* enzyme studies to construct a rationale in selecting the molecules for mechanistic based studies. The idea in selecting molecules from BITS database that are showing good docking scores helped us in synthesising further analogues of the potent lead molecule. Molecular docking process helps us to hasten the process in filtering the required compounds from a huge diversified scaffold library. The selected compounds may or may not possess inhibitory property on a laboratory scale-up. The molecules showing considerable CatD inhibition were alone taken for pathway analysis.

5.3.2.2. Renin enzyme inhibition study

Many studies reported the enzymological similarities between renin and CatD. The inhibitors against CatD were also known to inhibit renin due to the structural similarities which makes the molecules non-specific and less-effective in treating the actual ailment. Hence, the molecules which were known to have good CatD inhibition were also screened against renin. Molecule 29 having very high CatD enzyme inhibition were tested at 25, 10 and 1 µM concentrations against renin. The fluorescence measurements were made at Ex/Em of 328/552 nm. This is a kinetic mode of study in which two timings were noted to show greater difference in inhibition i.e., 30 and 60 min. The relative inhibitions were calculated for these three molecules. Amongst all, molecule 29 was found to show lesser inhibition against renin which makes it more specific towards CatD. The percentage inhibitions of M29 were shown in the graph **Figure 5.14.**

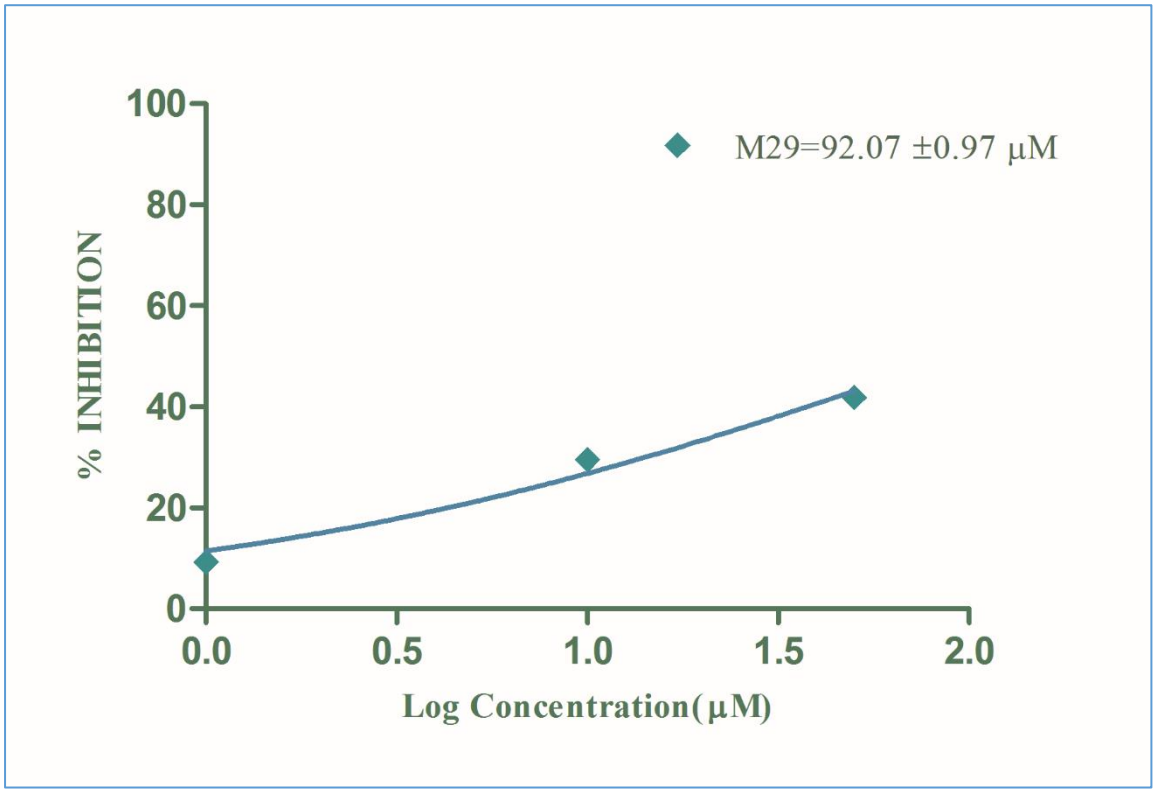


Figure 34: The off-target prediction of CatD active BITS molecule M29 against Renin enzyme using fluorescence based assay. We used three different concentrations 1, 10 and 25 μM and determined the IC_{50} value.

5.3.3. *In vitro* growth proliferation and cytotoxicity studies

CatD gene and protein expressions were studied in the above assays and we have selected a panel of breast cancer cell lines. The rank in order of CatD expression is MCF-7> MDA-MB-231 > SK-BR-3> MDA-MB-468. The therapeutic strategies have encouraged target specific research in developing inhibitors to treat breast carcinomas. There is also an increasing need in establishing chemotherapy approach for non-responsive TNBCs. Apart from enzymatic assays, all the molecules were studied for their role of inhibition in breast cancer cell lines.

We performed this assay to know the inhibitory activity of selected asinex compounds on all the cancer cell lines. We have selected breast adenocarcinoma cell lines (MCF-7, MDA-MB-231, MDA-MB-468 and SK-BR-3) and treated with molecules M14-M29. Importantly, MCF-7 and MDA-MB-231 cell lines were considered to know the effect of the molecules and were proceeded for their effect on other cell lines. The mutational data of the cell lines was also correlated with the growth proliferation effect. Half-maximal growth inhibition (GI_{50}) values were calculated based on the percentage inhibitions of all the molecules tested at 100 μM – 1 nM concentration ranges. Surprisingly, the active compounds have shown inhibitory activity in accordance to CatD gene expression and the GI_{50} values are as indicated in the **Table 5.10.**

In contrast to its effect on MCF-7 and MDA-MB-231, the active molecules did not affect the proliferation of HEK293T cells as seen in the **Table 5.10**. More than a fifty-fold change in GI_{50} was observed between the most sensitive cell line (MCF-7) versus normal HEK293T cells, indicating a better selectivity of M17, M19, M23 and M29 towards cancerous cells with inherently high levels of CatD. We used paclitaxel as the reference standard as it is reported to

inhibit all types of breast cancer cell lines effectively. Amongst all, M29 was known to effectively inhibit enzyme as well as growth of the cancer cell lines.

Table 10: GI_{50} (μM) values of the promising BITS database molecules tested on four different breast cancer cell lines and a normal cell line (HEK 293T). Molecules showing les than 10μM activity were studied completely.

BITS MOLECULES	**Half-Maximal concentration (GI 50 μM)**				**HEK (CC50 μM)**
	MCF-7	**MDA-MB-231**	**SK-BR-3**	**MDA-MB-468**	
M-14	20.05±1.18	42.35	-	-	23.5±1.12
M-15	79.49	90.45	-	-	100.2±0.07
M-16	73.25	100	-	-	29.5±0.78
M-17	3.4±2.12	5.7±0.005	5.96±2.13	8.42±2.32	44.82±2.01
M-18	350.6	ND	-	-	31.6±0.05
M-19	2.5±0.36	10.3±1.12	18.77±0.0056	17.24±1.56	19.2±2.32
M-20	97.61	>100	-	-	24.2±1.32
M-21	85.48	>100	-	-	13±0.03
M-22	66.67	>100	-	-	14.4±1.35
M-23	2.16±0.02	5.6±2.34	55.15±1.16	62.15±2.21	13.56±1.82
M-24	105.9	100	-	-	34.76±0.24
M-25	36.72	100	-	-	30.3±0.68
M-26	44.62	58.49	-	-	30.9±1.09

M-27	44.58	78.347	-	-	12.7±0.03
M-28	28.75	33.55	-	-	30.4±0.08
M-29	1.9±1.6	5.3±0.03	7.38±1.12	42.92±0.03	7.1±0.12
Paclitaxel[a]	0.0072	0.00241	0.005	0.0048	-

[a] Reference standard

Based on the results obtained, the molecules 17, 19, 23 and 29 were found to show greater inhibition on MCF-7 and MDA-MB-231 and is in accordance to the levels of CatD gene expressions in the cell lines. In comparison to specificity towards CatD, M29 has more inhibition towards CatD than renin and has greater growth inhibition on breast cancer cell lines. This confirms Molecule 29 to be potent CatD inhibitor out of the selected 16 molecules from the BITS database. The structural aspects of M29 was considered and taken for further synthetic scheme.

5.3.4. Conclusion

The approach in selecting lead molecule from BITS database resulted in identification of four molecules initially. The four molecules showing >50% enzyme inhibition are M17, M19, M23 and M29 molecules. All the molecules M14-M29 were further evaluated for their growth inhibition activities. In conclusion, by comparison of all the assay results M29 was found to have good interaction pattern and greater inhibitions. Thus, we achieved in obtaining M29, an isatin moiety from a database consisting 930 molecules. The molecule was further studied on structural activity relationships with respect to molecular docking studies. M1 from asinex database and M29 from BITS database were the two potent molecules that we have achieved during the first phase of study. As these molecules were determined to show promising effect in inhibiting TNBC cell line MDA-MB-231, we have made our study in specific to TNBC treatment. There is a huge scope in developing small molecule inhibitors to treat estrogen

negative cancers. CatD was also known to have greater importance in invasion and metastatic mechanisms of triple negative breast carcinomas. Hence, we took a step forward in synthesizing analogues to the identified potent lead molecules.

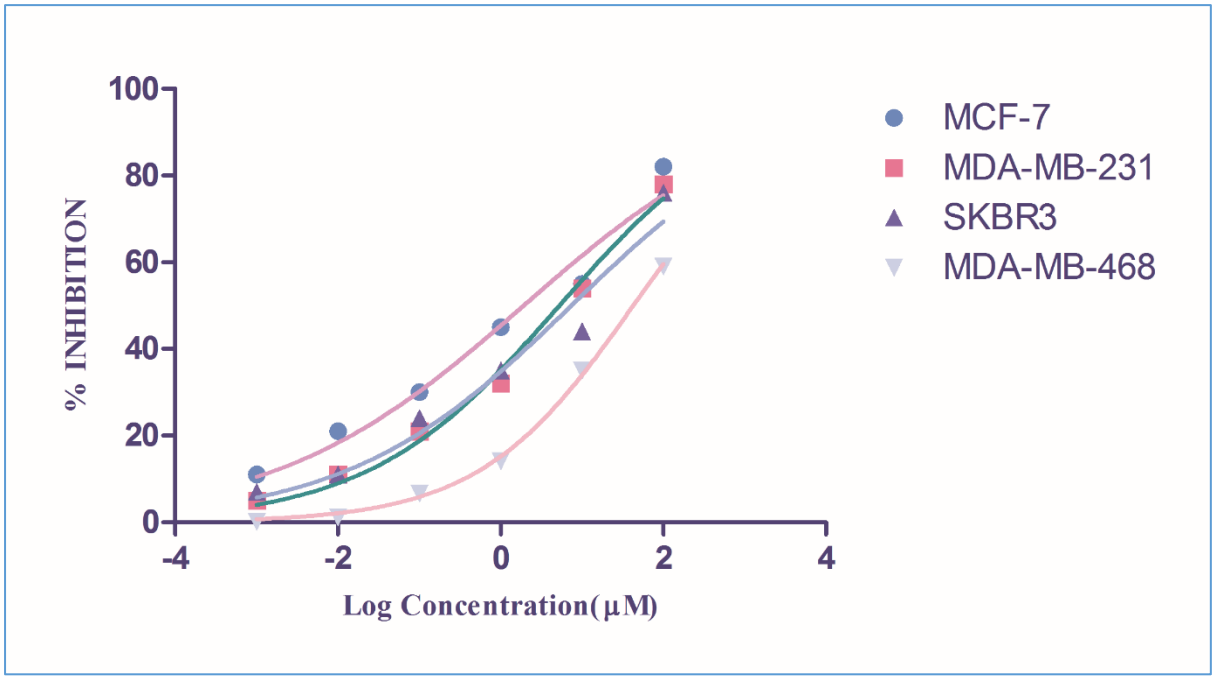


Figure 35: The growth inhibition curve of the most active BITS molecule M29 performed on different breast adenocarcinoma cell lines using MTT cell proliferation assay. The graphs were plotted using Graph Pad prism v.6.0.

5.4. Molecular docking, synthesis and lead identification from M29 analogues

We took molecules 1 and 29 from the above studies as potent lead molecules for further modifications. M1 was complex in structure and we could not achieve better interaction patterns of its analogues by changing side chains or aromatic rings of the molecule. We found it difficult in optimising the synthetic scheme to synthesize M1. The second molecule M29 is an isatin moiety and we have made changes in the connecting aliphatic chains as well as the aromatic ring in the molecule. This modifications resulted in good docking scores when compared to parent molecule M29. Thus, we have started our second phase of study by synthesizing analogues of M29 and in evaluating these analogues to achieve much promising molecule than M29. Initially, we made a set of 50-100 structures and started with molecular docking.

5.4.1. Molecular docking studies of M29 analogues

The grid generated from CatD PDB: 1LYB was used for molecular docking studies. All the structural modifications made were drawn using 2D sketcher and saved as .Mol files. In the latter step, all the files were imported and optimized using LigPrep application. As there are <100 molecules, we directly docked the molecules using glide XP run. To our surprise, we found a set of modifications to play very crucial role in interacting with the binding active site pocket of CatD. There were 18 molecules under this set which were considered for synthesis. The XP glide score of pepstatin was -12.49 kcal.mol^{-1} which was initially docked. This was considered in shortlisting the M29 analogues out of XP docking. The docking calculations were performed in accordance with the molecular docking protocols and interactions between the molecules and target protein. Molecules were ranked on the basis of docking score, interactions with amino acid residues, and numbers of hydrogen bonds in comparison with crystal ligand

Pepstatin A and XP docking. The docking scores and interacting residues for the set of 18 molecules were given in the **Table 5.11.**

Table 11: Glide scores and important interaction residues of the structurally modified M29 analogues.

Isatin analogues	**GLIDE score (Kcal/mol)**	**Interacting residue**	
		H-Bond	**π-π stacking**
IS-1	-5.748	Asp 33	PHE 131
IS-2	-4.836	Gly 35,Gly 233,Gly 79	PHE 131,TYR 205
IS-3	-6.174	Gly 233	PHE 131
IS-4	-3.697	Gly 233 (2), Gln 14	-
IS-5	-4.734	Gly 233,Thr 234	PHE 131
IS-6	-7.209	Gly 35,Gly 233,Gly 79	PHE 131
CIS-1	-5.303	Gly 233,Thr 234	PHE 131
CIS-2	-6.438	Gly 79	-
CIS-3	-5.148	Gly 233	-
CIS-4	-7.154	Gly 79,Thr 234	-
CIS-5	-5.402	Gly 233	PHE 131
CIS-6	-7.243	Gly 35,Gly 233,Gly 79	PHE 131
MIS-1	-5.75	Gly 233	-
MIS-2	-3.401	Gly 233 (2)	PHE 131
MIS-3	-6.207	Asp 231, Gly 233 (2)	-
MIS-4	-5.02	Gly 233,Ser 235	-

MIS-5	-5.734	Gly 233,Thr 234	PHE 131
MIS-6	-7.374	Gly 35,Gly 233,Gly 79	-
M29	-8.02	Gly 79,Gly 233, Asp 231	PHE 131

The set of 18 were named based on the substitutions made and IS being indicative of isatin moieties. The analogues were hence as above IS 1-6, CIS 1-6 and MIS 1-6. All the 16 selected compounds were found to retain the interaction pattern with the binding site and the XP glide score ranged between -3 kcal.mol^{-1} to -8 kcal.mol^{-1}. The π-π interactions were also considered while filtering the compounds from the database.

5.4.2. ADME predictions of M29 analogues

A total of 18 molecules as M29 analogues were selected and examined for pharmaceutically relevant drug-like properties. The QikProp module was used for validating pharmaceutically relevant properties of the molecules in relation to Absorption, Distribution, Metabolism, and Excretion (ADME). Most of the hits had good calculated partition coefficients (QPlogPo/w), QPlogS, and QPPCaco, which are the key measures of solubility, absorption, and cell permeability. Lipinski's rule of five displayed zero violations of the molecules. The expected human oral absorption for the molecules ranged from 57 to 100 %. All the properties were listed out in the **Table 5.12**. All the 18 compounds were found to obey Lipinski rule of five, showing good percentage oral absorption thereby reflecting the drug likeliness of the compounds.

Table 12: In silico pharmacokinetic profile of the selected M29 analogues from docking.

Isatin analogues	**QPlogPo/w[a]**	**QP logS**	**QPPCaco[c]**	**Rule of 5[d]**	**% Human oral absorption[e]**
IS-1	2.69	-4.31	212.79	0	84.34
IS-2	1.54	-3.10	68.50	0	68.81
IS-3	2.69	-4.27	174.80	0	82.85
IS-4	1.62	-3.60	46.54	0	66.23
IS-5	3.24	-4.81	202.88	0	87.23
IS-6	2.13	-4.01	89.62	0	74.35
CIS-1	3.51	-5.86	196.28	0	88.52
CIS-2	1.63	-3.51	49.77	0	66.87
CIS-4	2.11	-4.31	48.95	0	69.56
CIS-5	3.20	-5.12	143.35	0	84.29
CIS-6	2.64	-4.72	102.60	0	78.41
MIS-1	2.57	-4.02	169.78	0	81.90
MIS-2	1.87	-3.63	49.62	0	68.22
MIS-3	2.84	-3.72	187.10	0	84.21
MIS-4	2.20	-4.57	59.05	0	71.52
MIS-5	3.63	-5.73	206.95	0	89.65
MIS-6	2.42	-4.56	90.18	0	76.10
M29	3.39	-3.89	83.74	0	81.19

[a] Predicted octanol/water partition co-efficient log p (acceptable range from −2.0 to 6.5). [b] Predicted aqueous solubility, log S. (acceptable range from –6.5 to 0.5) [c] Predicted Caco-2 cell permeability in nm/s (acceptable range: <25 is poor and >500 is great). [d] Number of

violations of Lipinski's rule of five (maximum is 4) [e] Percentage of human oral absorption (<25% is poor and >80% is high).

5.4.3. Synthesis of the selected M29 analogues

Synthetic protocol employed for obtaining target molecule was shown in **Figure** ("Methods" section). Synthesis started with the alkylation of substituted isatins (1a-e) using 1, 2 dibromoethane as alkylating agent and K_2CO_3 as a base. The column chromatography technique was utilized to separate desired product from O-alkylated impurity. Also, excess usage of dibromoethane helped to reduce the di alkylated product. N-2-bromo ethyl isatin (2a-c) obtained from the previous step was treated with 4-N Boc aminopiperidine to get nitrogen linked isatin analogues (3a-c). Final scaffold for the synthesis of target molecule was obtained by treating (3a-e) with trifluoro acetic acid to get Boc cleaved amine (4a-c), followed by treatment with various substituted isocyanates and isothiocyanates resulted in urea and thiourea as final products. The substitutions are R, R' and X as seen in the positions of the **Figure 5.16.** The codes and modifications made accordingly were given in the **Table 5.13**. Among the designed molecules for synthesis, we could not obtain any yield for CIS-3 molecule and hence is not discussed in further procedures.

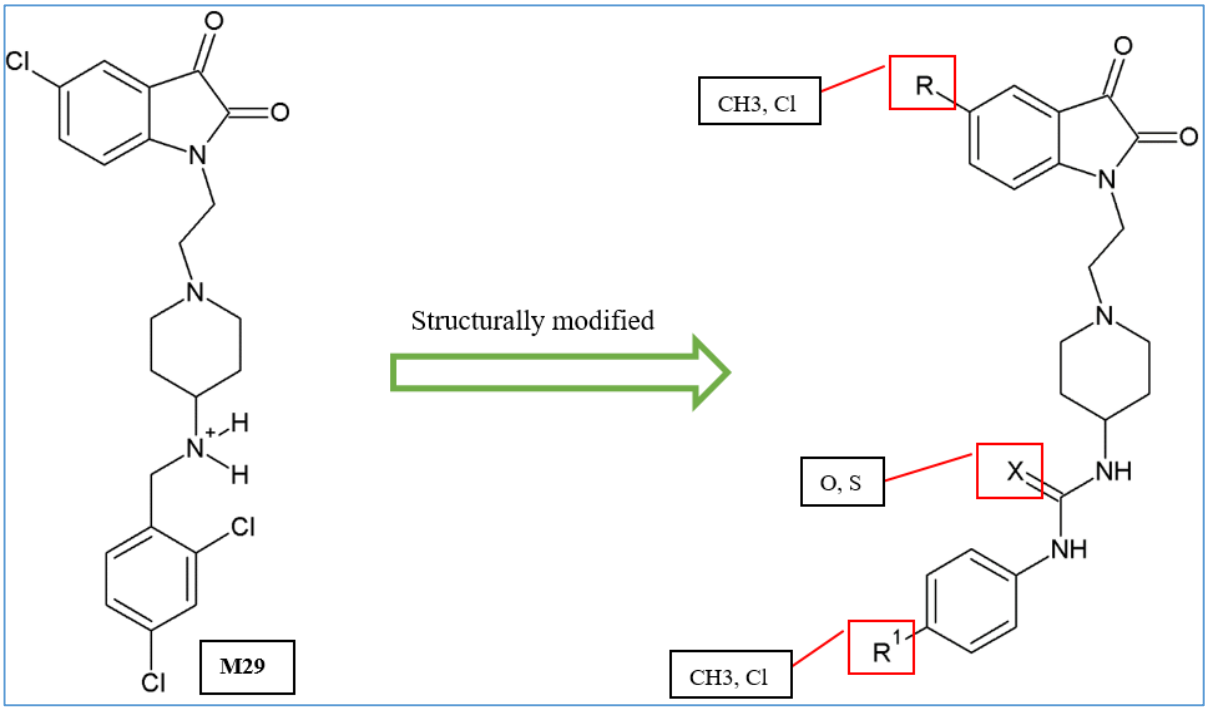


Figure 36: The structural modification of M29 made on the basis of docking and interaction pattern of the resulting analogues.

Table 13: The list of substitutions made contributing to the synthesis of 17 analogues, considering M29 as the parent lead molecule.

Compounds	**R**	**R'**	**X**
IS1	-	-	S
IS2	-	-	O
IS3	-	CH3	S
IS4	-	CH3	O
IS5	-	Cl	S
IS6	-	Cl	O
MIS1	CH3	-	S
MIS2	CH3	-	O

MIS3	CH3	CH3	S
MIS4	CH3	CH3	O
MIS5	CH3	Cl	S
MIS6	CH3	Cl	O
CIS1	Cl	-	S
CIS2	Cl	-	O
CIS4	Cl	CH3	O
CIS5	Cl	Cl	S
CIS6	Cl	Cl	O

We obtained 5-10 mg of each analogue and were well-characterized using NMR, IR and mass spectroscopies. The synthesised analogues were further taken for *in vitro* enzymatic and cytotoxicity evaluations.

5.4.4. *In vitro* Cathepsin D inhibition assay

Initially, Pepstatin standard graph was plotted considering concentrations from 100µM to 0.1 nM as per the protocol mentioned in the "methods" section. Furthermore, all the synthesized 17 analogue molecules were selected for *in vitro* CatD inhibition using a flourometric assay (BIOVISION, U.S.A). The initial screening at 25µM revealed that five out of 17 molecules displayed >50 % inhibition as shown in **Figure 5.17.** CIS-1 showed >70% enzyme inhibition and equal to its parent compound M29. Active molecules were correlated with their docking calculations. The active site pocket of CatD has two lobes and is connected by a cleft. The residues Gly79 and Ser80 forming the cleft are not displayed in the figure to provide a clear view of the alignment of molecules inside the active site pocket. These deselected residues are coloured green, as shown in **Figure 5.18.** The ligand interaction pattern of the most active analogues CIS-1 and MIS-1 were shown in **Figure 5.19.**

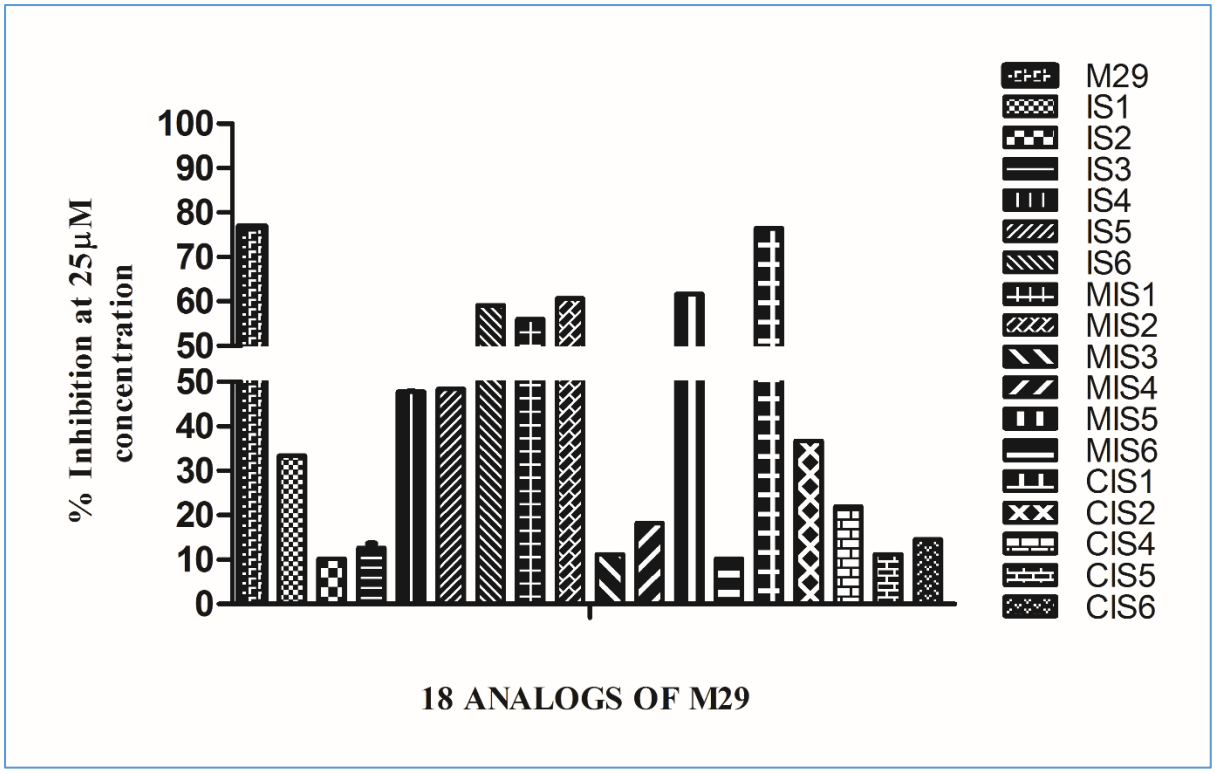


Figure 37: Percentage inhibitions of M29 analogues at 25µM concentrations analyzed using fluorescence plate reader at Ex/Em of 328/460 nm. Six out of 18 were showing >50 % inhibition profiles of CatD.

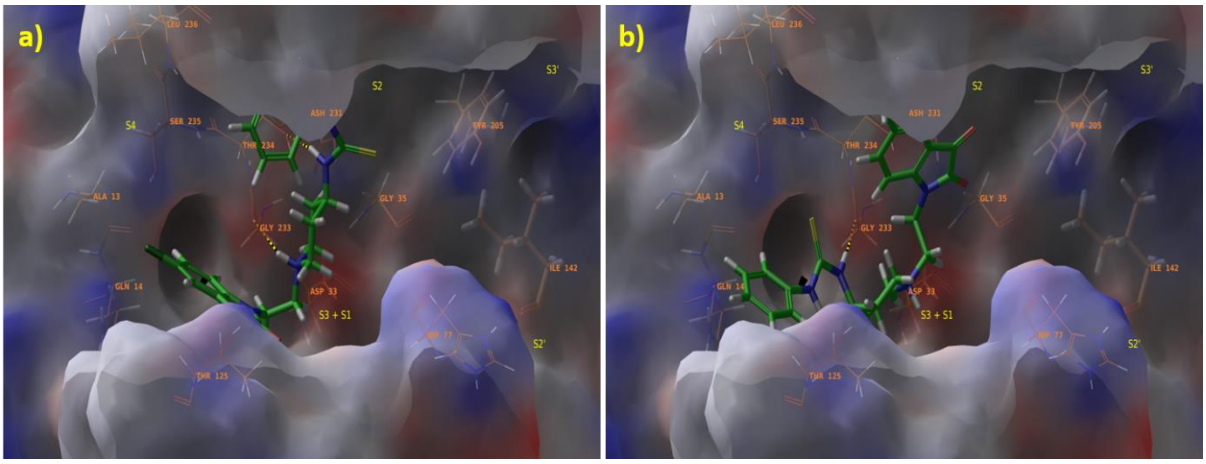


Figure 38: Binding poses of a) CIS-1 and b) MIS-1 aligned inside the CatD active site pocket.

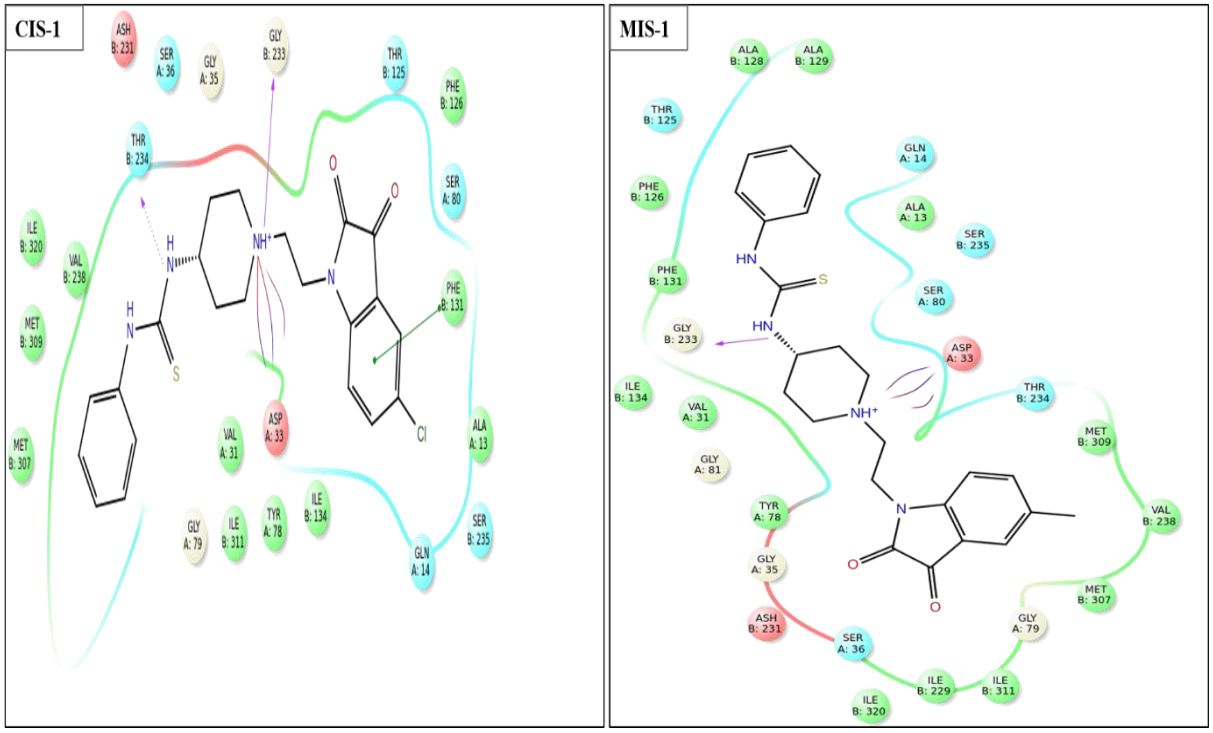


Figure 39: Ligand interaction plots of the active lead molecules from the synthesized analogues i.e., CIS-1 and MIS-1.

This is a preliminary assay for the evaluation of selected compounds for their CatD inhibition. This further helps us to correlate docking studies with *in vitro* enzyme studies to construct a rationale in selecting the molecules for mechanistic based studies. CIS-1 showed very good inhibition which was further concentrated for CatD related mechanistic studies.

5.4.5. *In vitro* Renin inhibition assay

This is a fluorescence based assay performed using kit procured from Biovision, U.S.A. Many studies reported the enzymological similarities between renin and CatD. The inhibitors against CatD were also known to inhibit renin due to the structural similarities which makes the molecules non-specific and less-effective in treating the actual ailment. Hence, the molecules which were known to have good CatD inhibition were also screened against renin. CIS-1 as well as MIS-1 having very high CatD enzyme inhibition and growth inhibitions were tested at 25, 10 and 1 µM concentrations against renin. The fluorescence measurements were made at

Ex/Em of 328/552 nm. This is a kinetic mode of study in which two timings were noted to show greater difference in inhibition i.e., 30 and 60 min. The relative inhibitions were calculated for these three molecules. CIS-1 was evaluated to show lesser renin inhibition when compared to MIS-1. The percentage inhibitions of all the three molecules i.e., M29, CIS-1 and MIS-1 were shown in the **Figure 5.20.**

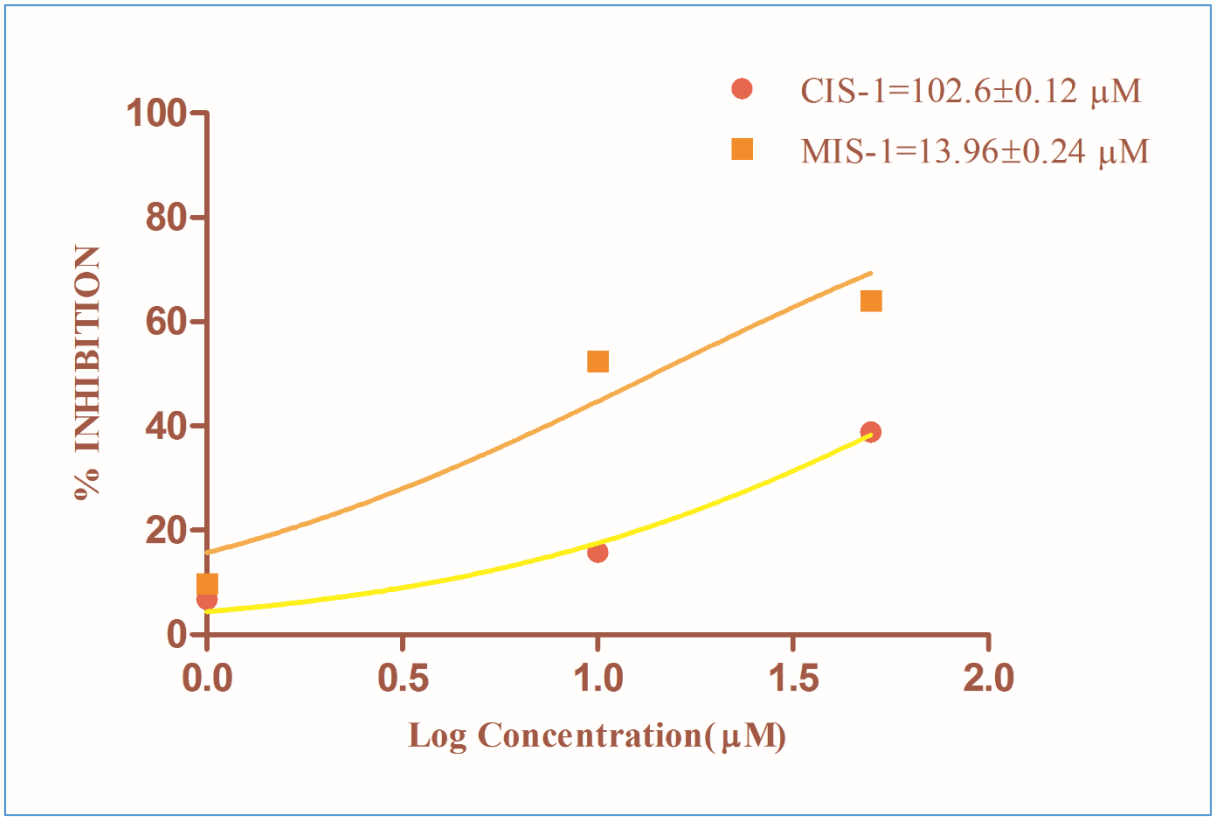


Figure 40: The off-target prediction of CatD active M29 analogues CIS-1 and MIS-1 against Renin enzyme using fluorescence based assay. We used three different concentrations 1, 10 and 25 µM and determined the IC_{50} value.

5.4.6. Cell proliferation assay-MTT

CatD gene and protein expressions were studied in the above assays and we have selected a panel of breast cancer cell lines. The rank in order of CatD expression is MCF-7> MDA-MB-231. In this study, we have included another new breast cancer cell line ZR-75-1 which has

immuno profiling of ER+, PR+ and HER2+. Apart from enzymatic assays, all the molecules were studied for their role of inhibition in breast cancer cell lines.

We performed this assay to know the inhibitory activity of selected asinex compounds on all the cancer cell lines. We have selected breast adenocarcinoma cell lines (MCF-7, MDA-MB-231, and ZR-75-1) and treated with all the synthesized analogue molecules. Importantly, MCF-7 and MDA-MB-231 cell lines were considered to know the effect of the molecules and were proceeded for their effect on other cell lines. The mutational data of the cell lines was also correlated with the growth proliferation effect. Half-maximal growth inhibition (GI_{50}) values were calculated based on the percentage inhibitions of all the molecules tested at 100 μM – 1 nM concentration ranges.

In contrast to its effect on MCF-7 and MDA-MB-231, the active molecules did not affect the proliferation of ZR-75-1 and HEK293T cells as seen in the **Table 5.14**. More than a forty-fold change in GI_{50} was observed between the most sensitive cell line (MCF-7) versus normal HEK293T cells, indicating a better selectivity of CIS-1 and MIS-1 towards cancerous cells with inherently high levels of CatD. In comparison to enzyme assay and cytotoxicity assays, CIS-1 was determined to have greater inhibitions. Later, CIS-1 was also compared to its parent molecule M29 and was found to be superior in inhibition.

Table 14: GI_{50} (μM) values of the promising M29 analogues tested on three different breast cancer cell lines and a normal cell line (HEK 293T). Molecules showing less than 10μM activity were studied completely.

Compounds	**GI50 (μM)**			**CC50(μM)**
	MCF-7[a]	**MDA-MB-231[b]**	**ZR-75-1[c]**	**HEK293T[d]**
IS1	6.4±0.98	5.52±1.2	71.88±1.3	4.89±0.25
IS2	14.9±0.01	15.26±0.22	85.73±0.02	5.83±0.3

IS3	16.31±2.2	11.54±2.21	41.5±0.08	7.43±0.21
IS4	2.52±2.23	8.5±1.15	45.34±1.13	3.31±1.1
IS5	1.14±2.03	8.84±1.26	39.92±0.02	4.53±0.49
IS6	4.58±3.33	11.16±0.05	82.52±0.05	2.98±0.71
MIS1	0.52±2.63	1.66±1.12	14.34±1.12	16.37±0.05
MIS2	6.2±2.5	9.25±0.03	44.33±1.03	27.12±0.21
MIS3	14.4±2.63	12.37±1.1	22.59±1.05	16.6±0.03
MIS4	11.13±1.12	15.46±1.23	31.16±1.1	27.13±1.12
MIS5	3.5±0.99	7.7±2.26	41.85±0.09	17.52±1.12
MIS6	13.05±1.6	23.09±0.99	62.68±0.07	13.52±1.5
CIS1	0.48±1.45	1.02±0.05	11.62±2.21	37.98±0.08
CIS2	1.48±2.2	1.15±0.56	>100	54.36±0.07
CIS4	7.19±0.99	9.9±1.19	53.21±1.45	71.95±0.12
CIS5	>100	78.22±0.65	78.81±1.13	76.54±0.23
CIS6	>100	78.4±1.12	80.33±1.12	95.74±0.21
M29	0.67±1.27	2.6±0.03	15.66±0.02	7.1±0.02
Paclitaxel[a]	0.0072±0.01	0.00241±0.02	-	-
Semaxanib [b](SU5416)	-	1.8±0.01	-	-
Olaparib[c]	-	0.55±0.12	-	-

[a] Reference standard for cancer cell lines

[b] Anti-angiogenic drug which is in Phase III clinical trails

[c] PARP-1 inhibitor, enzyme involved in DNA repair

5.4.7. Conclusion

We have selected molecules from databases and proceeded for synthesis in a motive to achieve more potent and specific CatD inhibitor. Hence, we have filtered and studied about 29

molecules and took one molecule M29 for further synthesis. We were able to synthesize 17 molecules which were again selected on the basis of molecular modelling studies. We achieved CIS-1 as the best molecule with greater enzymatic as well as growth inhibitions in comparison to M29. Thus, we took M29 and CIS-1 in all our further studies that include cell cycle analysis, ROS inhibition study, RTPCR analyses and zymography studies. This helped us to achieve best potent lead, non-peptidic, small molecule as CatD inhibitor out of our research.

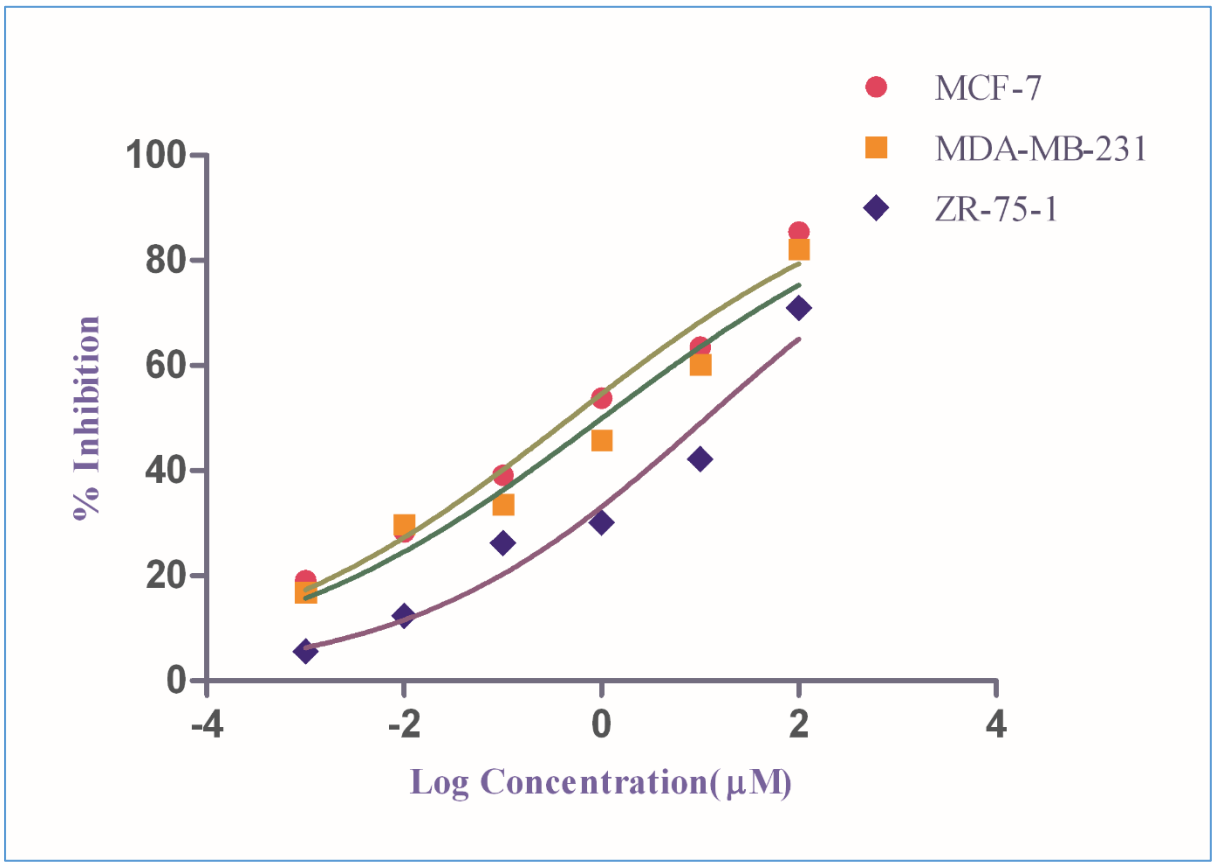


Figure 41: The growth inhibition curve of the most active M29 analogue CIS-1 performed on different breast adenocarcinoma cell lines using MTT cell proliferation assay. The graphs were plotted using Graph Pad prism v.6.0.

5.5. Lead bioevaluation for its role in causing CatD inhibition

The above assay studies concluded M29 from databases and CIS-1 and MIS-1 from the synthesized analogues to be the most active molecules. Apart from enzyme and growth inhibitions, the molecules were tested for their role in inhibiting CatD. CatD was known to cause invasion and metastasis by initiating extracellular matrix degradation. This is noticeably found in TNBC cell line MDA-MB-231. We have also found significant levels of CatD gene expression in MDA-MB-231 which allowed us to design and evaluate CatD inhibitors against TNBC. Our study further uses MDA-MB-231 cell line, and deduces cell cycle analysis, role of ROS, angiogenic effects and gene regulations that can be caused due to M29 and CIS-1.

5.5.1. Effect of M29 and its potent lead analogues on cell cycle arrest

One of the crucial features to examine the final arbiter of cell fate was cell cycle arrest. The effect of cell cycle arrest during CatD inhibition was determined. The effect of the molecules on cell cycle in MDA-MB-231 cancer cell lines was determined by propidium iodide staining and analysed by flow sight (Amnis, Millipore). **Figure 5.22** displays the increase in cells in the G2/M phase ($p < 0.01$) when treated with 1.8 μM i.e., GI_{50} value of Semaxanib and 2.6 μM i.e., GI_{50} value of M29 on MDA-MB-231 cells compared to the control. There is a decrease in G0/G1 phase in treated cells when compared to control. Whereas, there is a decrease in cells in the G2/M phase ($p < 0.01$) when treated with 1.02 μM i.e., GI_{50} value of CIS-1 and 1.66 μM i.e., GI_{50} value of MIS-1 on MDA-MB-231 cells compared to the control. There is an increase in G0/G1 phase in treated cells when compared to control. Cell cycle data can be further correlated with the levels of CDK1 after treatment with these molecules on MDA-MB-231. The imaging system showed PI fluorescence differentiating various phases of cells with their DNA content. The samples were further analysed by IDEAS analysis software, and a histogram

was plotted as Intensity in channel 4 (Propidium iodide) on X-axis and Normalized frequency on Y-axis.

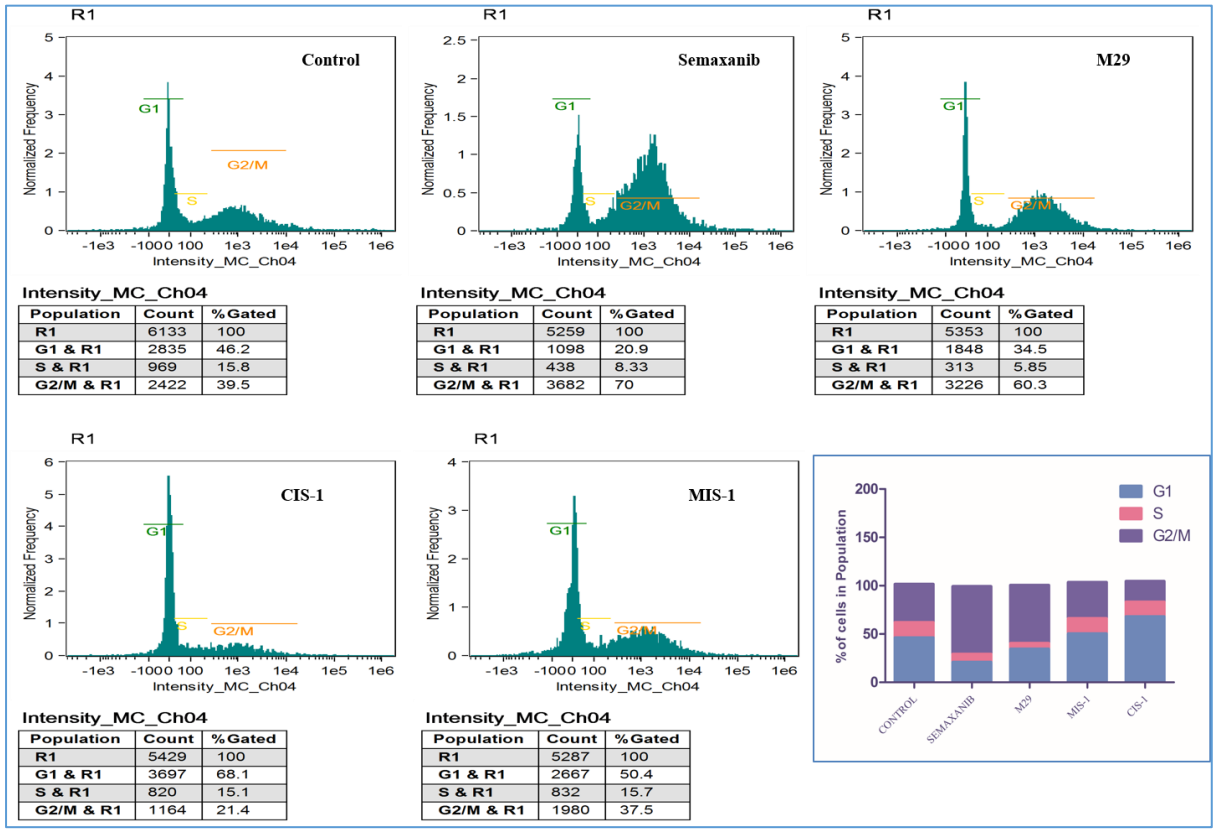


Figure 42: CIS-1 arrest the cell cycle in G1 phase and Semaxanib arrest the cell cycle in G2/M in MDA-MB-231 breast cancer cell lines. Cells were incubated with 1.02µM of CIS-1 and 1.8 µM of Semaxanib (SU5416) compound for 48 h before cell cycle analysis. The results were analysed using IDEAS flow sight software, AMNIS, Millipore. A graphical representation of the same was plotted using Graph pad prism v.6.0.

5.5.2. M29 and CIS-1 causes minimal ROS inhibitory effect

Reactive oxygen species (ROS) has a very crucial role in relation to cancer and its progression. Generally, the apoptosis-inducing molecules causes elevated ROS levels and thereby killing of cancer cells. The other way of mechanism is ROS inhibition is to act as antioxidant and prevent the worsening of cancer invasion and metastasis. As the designed CatD inhibitors are known to inhibit growth proliferation of cancer cells, they are evaluated for their role in ROS using

green fluorescent dye. The hypoxic conditions created using CoCl2 in MDA-MB-231 cells resulted in high levels of ROS generation which eventually helped us to evaluate the inhibitory profile of the molecules. The cells were treated with 1.8 µM i.e., GI_{50} value of Semaxanib, 2.6 µM i.e., GI_{50} value of M29 and 1.02 µM i.e., GI_{50} value of CIS-1. The cells which were analysed further using DCFHDA, a fluorescent dye which detects ROS showed a shift in the graphs. The shift in graph from the control towards left side indicates decrease in fluorescence which is ROS inhibition and vice-versa. We found a very less ROS detection in control cells and there is increased ROS in control cells treated with $CoCl_2$. We also observed more ROS inhibition in cells treated with CIS-1 as seen in **Figure 5.23** followed by M29 and SU5416. Overall, we determined minimal involvement of these molecules in ROS inhibition.

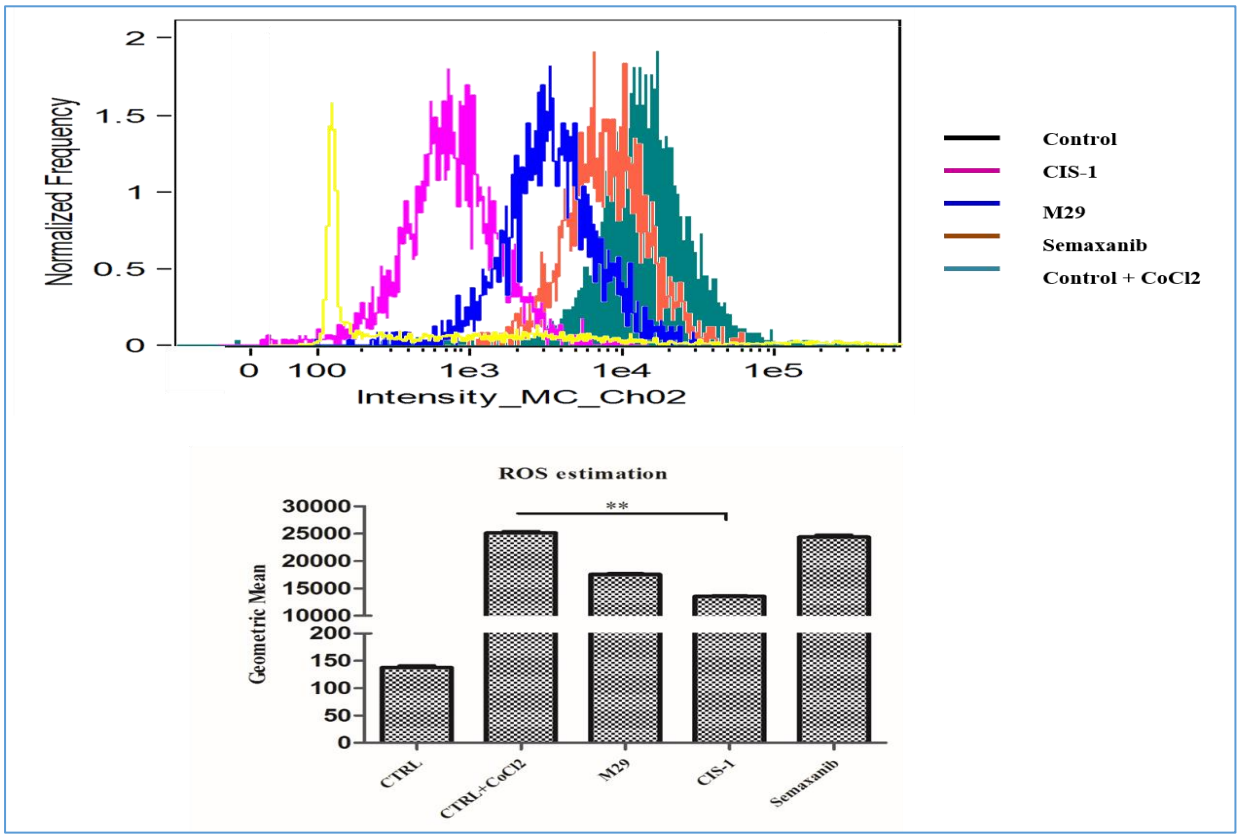


Figure 43: Generation of ROS in MDA-MB-231 cells using Cobalt chloride which mimics hypoxic tumor microenvironment. The control are without $CoCl_2$ and the drugs were incubated for 48 h. The cells were assessed by flow sight for ROS generation using DCFHDA green

fluorescent dye, and data are expressed graphically (**p<0.01 when compared with the $CoCl_2$ control) showing the percentage of DCF-DA-fluorescence intensity. The flow sight analysis showed a shift in ROS fluorescence towards left, which indicates ROS inhibition by the compounds in hypoxic conditions. Data are represented graphically using One-way ANOVA; Bonferroni's post-hoc analysis performed by Graph pad v.6.0.

5.5.3. Role of M29 and CIS-1 in lysosomal degradation

CatD is ubiquitously found in acidic compartments i.e., lysosomal parts of the cells. The molecules which act as CatD inhibitors were also predicted to cause lysosomal degradation resulting in inhibition. MDA-MB-231 cells were taken and stained with Acridine orange, a dye which permeabilizes into lysosomes. Previously, chloroquine was reported as a positive control for AO staining as chloroquine causes lysosomal membrane permeabilisation resulting in more red fluorescence. In the same way, these CatD inhibitors were also tested for their role in lysosomal degradation in breast cancer cells. The cells were treated with M29, CIS-1 and SU5416 (standard) at their GI_{50} values as the above flow sight analyses. The molecule CIS-1 showed a shift in fluorescence graph towards left in comparison to control. The scale on X-axis indicates the intensity of red fluorescence which says a shift in graph towards left as decrease in fluorescence intensity. Therefore, CIS-1 was determined to exhibit lysosomal degradation making the lysosomes unavailable for AO staining. CIS-1 effect is followed by M29, whereas SU5416 has no effect on lysosomal degradation as seen in **Figure 5.24.**

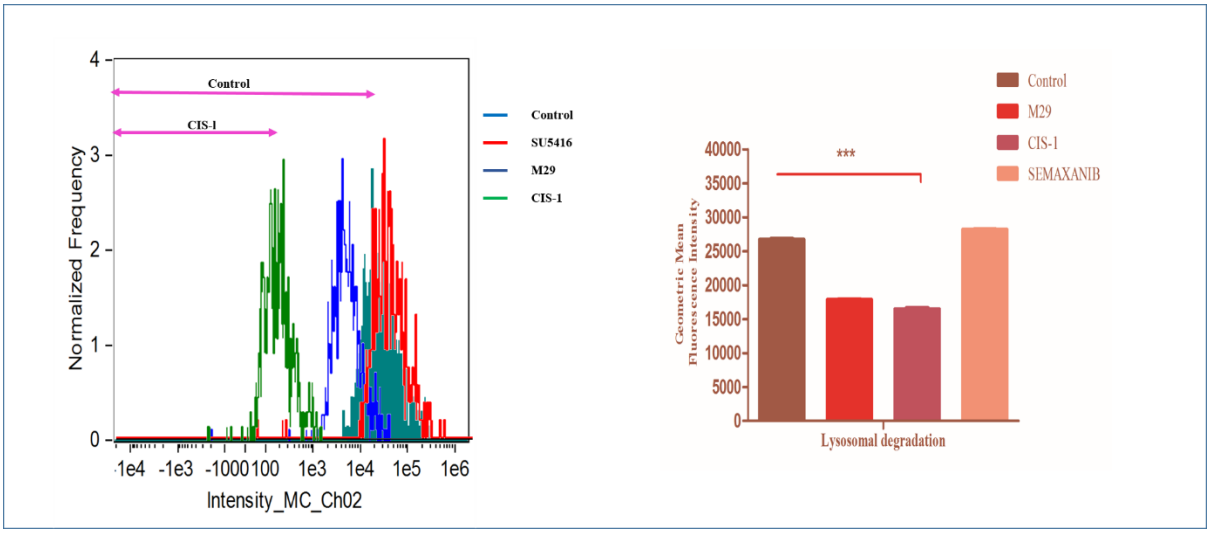


Figure 44: Detection of lysosomal fluorescence in MDA-MB-231 cells before and after treatment using Acridine orange (AO) dye. The control are without any treatment and the drugs were incubated for 48 h. The cells were assessed by flow sight for fluorescence generation using and data are expressed graphically (groups M29 and CIS-1 groups with ***p<0.001 when compared with the control) showing the percentage of AO-fluorescence intensity. The flow sight analysis showed a shift in fluorescence towards left, which indicates decreased fluorescence by the compounds due to lysosomal degradation. Data are represented graphically using One-way ANOVA; Bonferroni's post-hoc analysis performed by Graph pad v.6.0.

5.5.4. Evaluation of M29 and CIS-1 using angiogenesis assays

CatD is well studied and reported for its vital role in progression of breast cancer invasion and metastasis. In turn it causes ECM degradation releasing all the growth factors into the blood stream thereby aggravating metastatic mechanisms. Angiogenesis is the key mechanism involved in metastatic mechanism of cancer cell lines. An inhibitor known to cause prevent metastasis and invasion is referred to as anti-angiogenic drug. Some of the targeted anti-angiogenic drugs in the market were VEGF inhibitors, EGFR inhibitors, MMP inhibitors which were known to treat TNBCs. So, in view of these options we have tested our CatD inhibitors

for their role in effecting angiogenesis mechanism. We conducted *in vitro* scratch assay, tube formation assay and migration chamber assay to study the effect of the active molecules.

5.5.4.1. *In vitro* scratch assay

Molecules that are inhibiting the process of invasion are well detected using this assay. The scratch made in the middle of the adhered cells created a gap making the cells into two compartments. We treated the cells with M29, CIS-1 and SU5416 (standard) at their GI_{50} concentrations. The scratch dimension was measured initially at 0 h which was determined to be the same in both control as well as treated groups. The healing of the scratch was observed for every 4 h and we found a significant difference in 12 and 24h images. Two dark lines were drawn at the borders of the scratch to observe the decrease or increase in the scratch area. The rate of cell migration was determined by the available computing software that measures the distance travelled during the desired time frame. The distance travelled by the cells is minimum in the cells treated with CIS-1 followed by M29 and then SU5416 and this was clearly observed from the **Figure 5.25.**

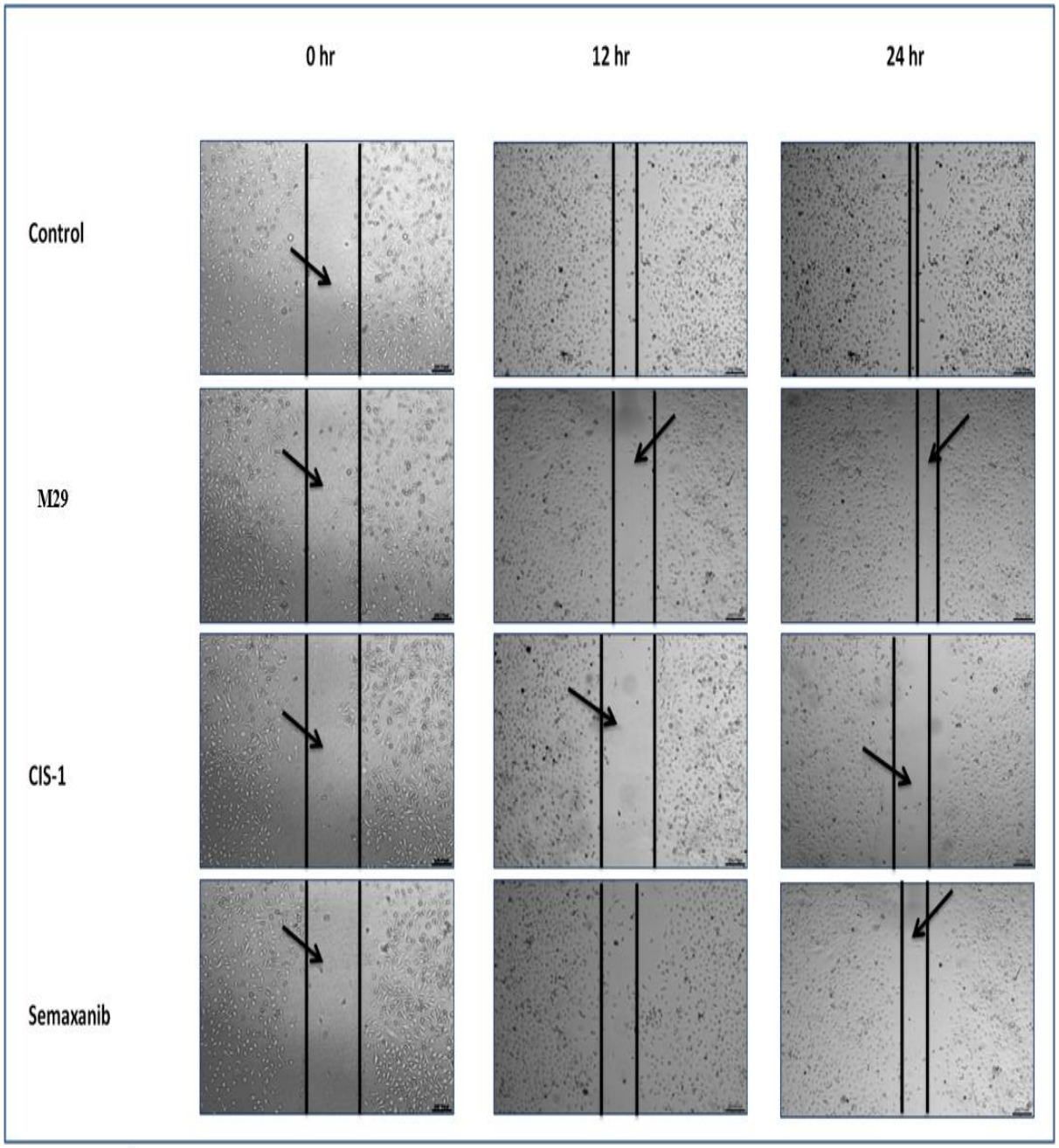


Figure 45: Effects of M29, CIS-1 and Semaxanib treatment on the migration of MDA-B-231 cells. Scratch wound healing assay was performed to examine their effects. Images are taken immediately after scratching the cultures 0 h, 12 h and 24 h.

5.5.4.2. Tube formation assay

We used matrigel to demonstrate the angiogenic activity of vascular endothelial cells using *in vitro* tube formation assay. B16F10, skin melanoma cells were used which has ability to form tubules when seeded on Matrigel. We observed the formation of tubules as seen in **Figure 5.26** from clear elongated cell bodies that connect to form polygon network. Once the cells were doubled, a solid vascular network was formed. The cells were seeded on to matrigel along with the molecules to be evaluated. The molecules used for the treatment were M29, CIS-1 and SU5416 at their GI_{50} concentrations. We found clear tubules formed in the control cells, and disruption of tubules in the treated groups as seen in the **Figure 5.25.**

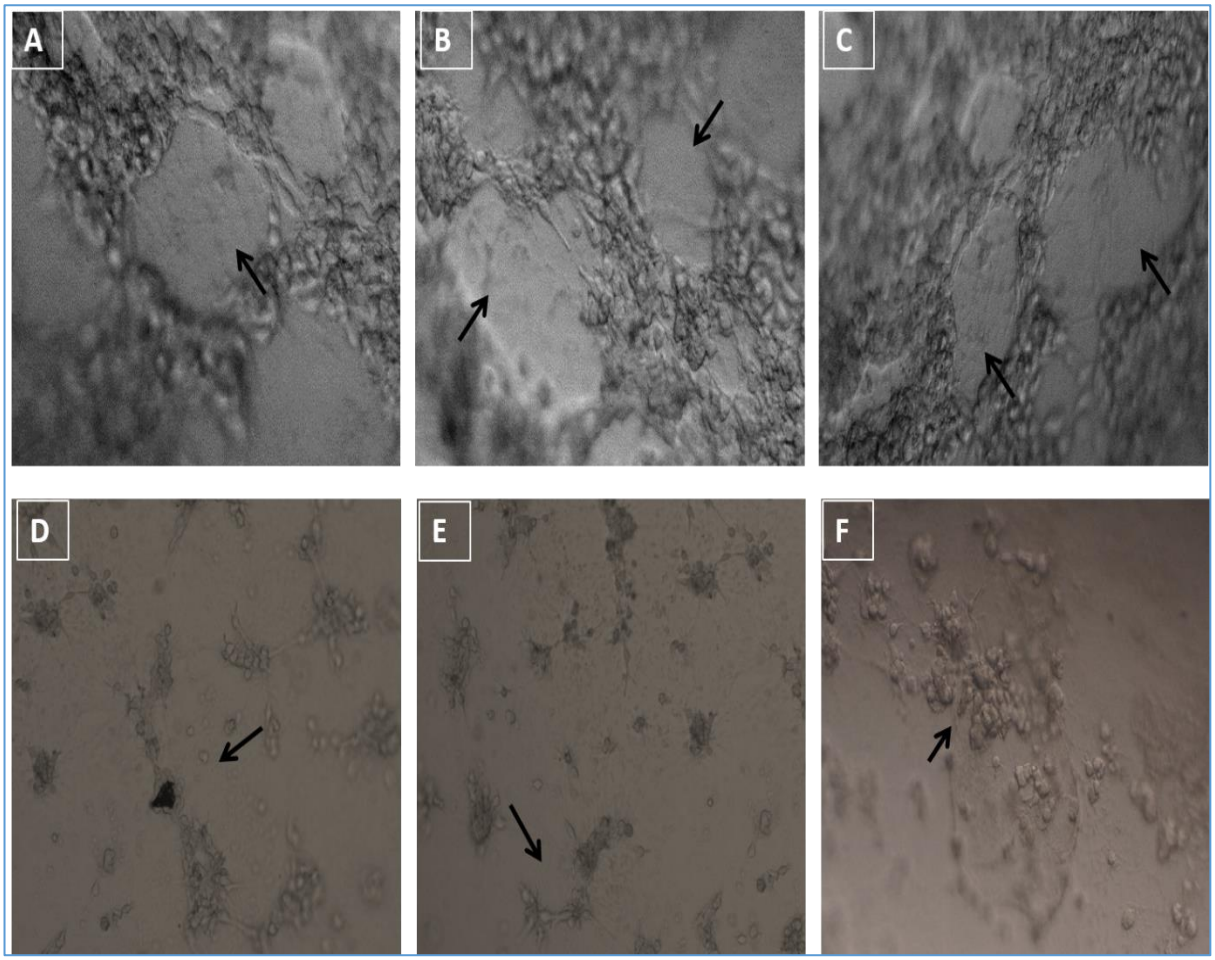


Figure 46: Tube formation assay: All the cells (2 x 10^4) were loaded on Matrigel and incubated overnight. Tubules were imaged using phase contrast. A, B, C images show B16F10 cells used as a positive control. A representative of 3-5 fields was shown. Bar: 100 µm. D) M29 E) CIS-1 and F) Semaxanib inhibits tube formation of tubules. The effect of the compounds on tube

formation of B16F10 was examined using an in vitro tube formation assay. Changes in cell morphology were observed using a phase contrast microscope (×40) and photographed.

The tested cells were maintained in a growth rate at log phase The analysis of the tubes was done under a microscope within 24h. We determined cell migration and cell-cell adhesion differences in the control and . treated cells. The vascular formation using matrigel was reported to have greater similarities with vascular property of tumor cell type.

5.5.4.3. Boyden chamber assay

This is yet another important type of angiogenesis assay which is based on chemotaxis phenomenon. We used MDA-MB-231 cells that has greater migration property suitable for this assay. The chambers procured are also known as cell-culture inserts which has two compartments. We maintained the two compartments at lower and higher levels of serum allowing the cells to migrate towards higher serum levels of medium on the basis of chemotaxis. We treated the cells with CIS-1, M29 and SU5416 (standard) as like in our previous assays at their GI_{50} concentrations. We observed the cell migration from the top (low serum) chamber to the bottom (high serum) chamber. The cells fixed were observed under microscope and were also counted. There were very few migrated cells in CIS-1 and M29 groups followed by SU5416-treated cells in comparison to control chambers as seen in the **Figure 5.27.**

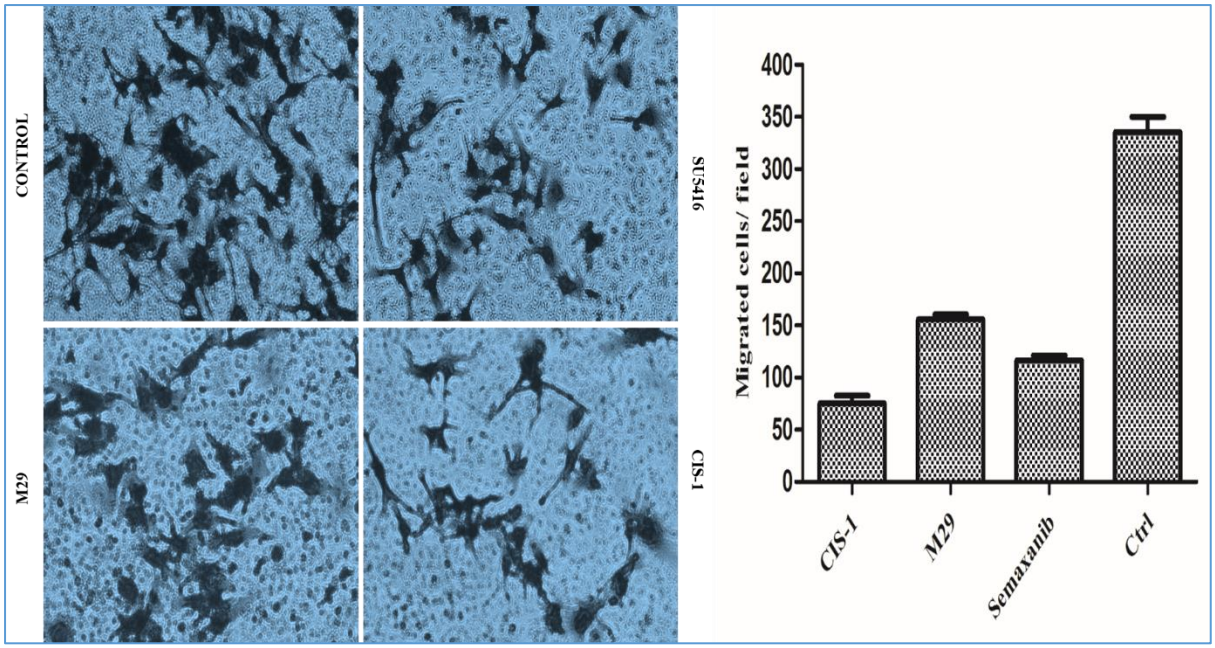


Figure 47: Cell migration analysis of MDA-MB-231 cells by Boyden chamber assay. Shown here are representative × 20 field microscope photos of Control MDA-MB-231 migrated cells (A), percentage of cells migrating after M29 treatment (B), CIS-1 treatment (C) and Semaxanib treatment (D). For invasion assay, Matrigel-coated Cell-culture inserts were used. For each panel (A-D), two independent experiments were performed in triplicates, bars indicate S.E.M.

5.5.5. Quantification of gene expression in relation to CatD inhibition

We have selected 18 different genes which were thought to be effected during CatD inhibition. The regulation of these genes in turn resulted in CatD modulation. Cancer which was known to be closely related to inflammation phenomenon was also taken in to consideration while studying the genes. The role of CatD in invasion and metastasis, the role of cancer in inflammation, the importance of new upcoming targeted genes as well as the ultimate downstream genes that result in angiogenesis were all given importance in estimating the pathway through which CatD inhibition can treat cancers. To find a difference in inhibition we have tried increasing the gene expression by inducing the cells with $CoCl_2$, thereby creating a hypoxic microenvironment. The expression levels of inflammatory mediators and growth

factors were significantly increased due to the chemical inducer $CoCl_2$ that mimics hypoxia. So, we performed gene expression studies of few genes in hypoxic conditions and few in normal conditions.

5.5.5.1. Role of inflammatory mediators and growth factors in CatD regulation under hypoxic conditions

The cells induced with the chemical inducer, CoCl2 were treated with our potent molecules M29, CIS-1 and the standard SU5416. mRNA was extracted from the cells after the incubation period and was converted into cDNA using PCR. The inflammatory mediators we used were IL-6, IL-1β, IL-8 along with cytokine TNF-α and transcription regulator NFκβ. As CatD is known to cause ECM degradation and release the growth factors out into the blood stream we considered the growth factors that are embedded in the matrix. The growth factors and proteases which were quantified namely, EGFR, VEGF, MMP-2 and MMP-9 genes. The hypoxic condition in the tumor microenvironment was known to induce H1F-α factor and CXCR4 chemokine in relation to metastatic mechanism of MDA-MB-231 cells. All the 11 genes were thus quantified and run through RT-PCR to evaluate the effect of CatD inhibitors in comparison to house-keeping gene GAPDH. Our potent analogue CIS-1 was found to significantly downregulate IL-8, NFκβ, VEGF, MMP-2, H1F-α and CXCR4 in comparison to M29 and SU5416 which were normalized with GAPDH as seen in **Figures 5.28 and 5.29**. Data were analysed using one-way ANOVA with post hoc Bonferroni's correction as seen in Figure 5.28 *P< 0.05, ***P< 0.001 versus. ANOVA, analysis of variance; RT–PCR, reverse transcription PCR.

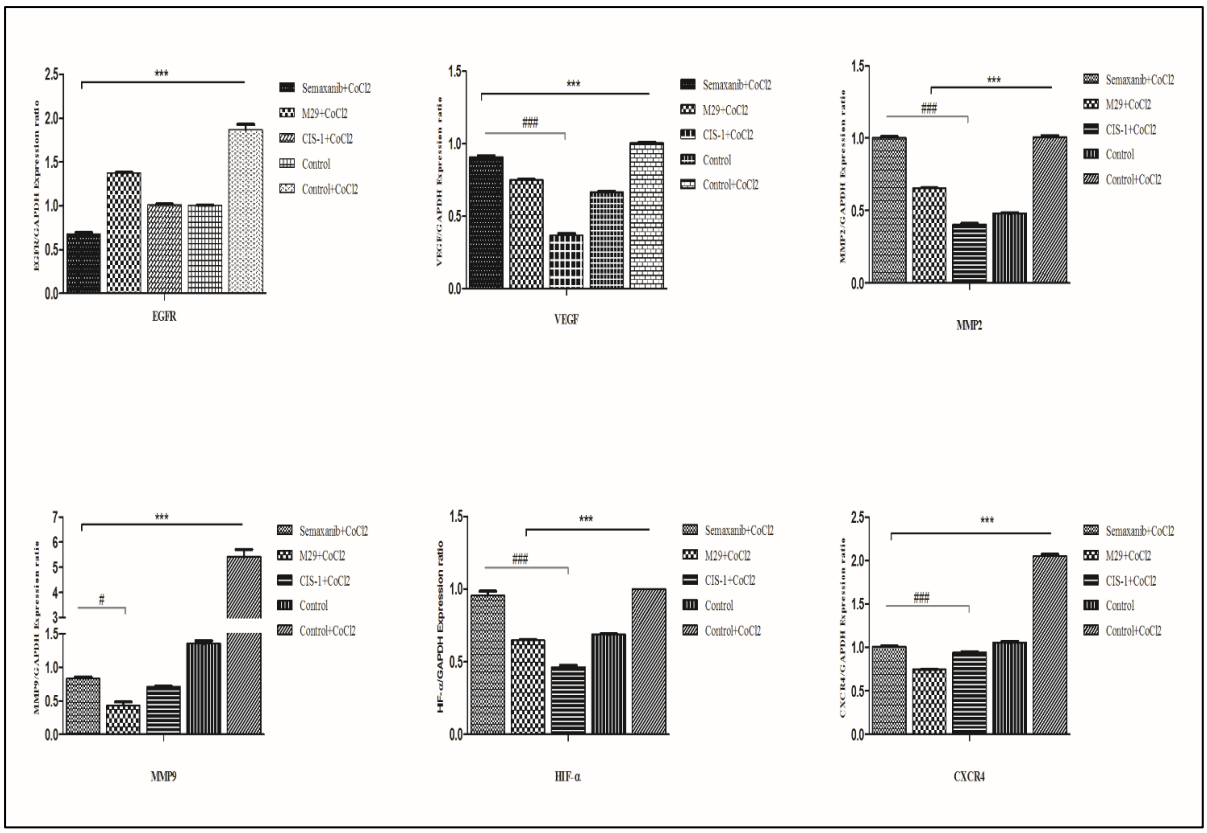


Figure 48: RT–PCR from cell culture mRNA, quantified EGFR, VEGF, MMP-2, MMP-9, H1F-α and CXCR4 gene expressions of control, control (induced hypoxia), M29, CIS-1 and SU5416 treatments. Data are mean±S.E.M. from three mRNA samples in each group. Data were analyzed using one-way ANOVA with post hoc Bonferroni's correction. *p<0.05, **p<0.01, ***p<0.001 versus control; $^{\#}$p<0.05, $^{\#\#}$p<0.01, $^{\#\#\#}$p<0.001 versus Semaxanib (SU5416). ANOVA, analysis of variance; RT–PCR, reverse transcription PCR.

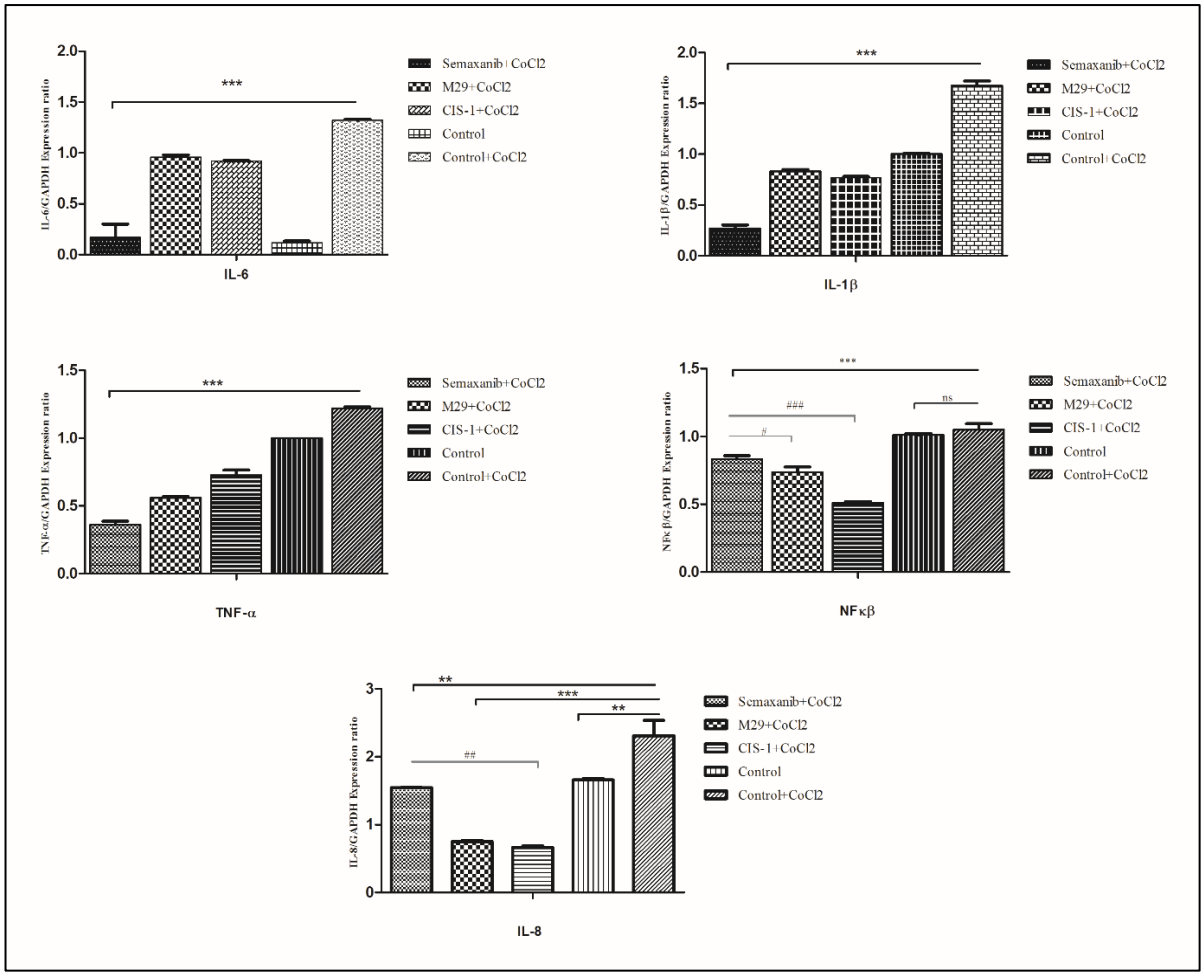


Figure 49: RT–PCR from cell culture mRNA, quantified IL-6, IL-1β, TNF-α, NFκβ and IL-8 gene expressions of control, control (induced hypoxia), M29, CIS-1 and SU5416 treatments. Data are mean±S.E.M. from three mRNA samples in each group. Data were analyzed using one-way ANOVA with post hoc Bonferroni's correction. *p<0.05, **p<0.01, ***p<0.001 versus control; $^{\#}$p<0.05, $^{\#\#}$p<0.01, $^{\#\#\#}$p<0.001 versus Semaxanib (SU5416). ANOVA, analysis of variance; RT–PCR, reverse transcription PCR.

5.5.5.2. Role of cell cycle regulators and novel target genes in CatD inhibition under normoxic conditions

The cell cycle regulators AKT/mTOR and CDK1 were generally the final downstream genes in any cell signalling pathway. Apart from these, PARP-1 a very novel target in developing TNBC inhibitors was also tested, as our, CatD inhibitors were found to act promisingly on MDA-MB-231 cells. The other targets were collected from the literature as they were reported to have close relation to CatD and also act as growth control regulators namely PPAR-γ, E-Cadherin (promising biomarker in breast cancers) and c-Myc. Among these genes, c-Myc, CDK1, AKT and mTOR were significantly downregulated more in M29 treated cells followed by CIS-1 group of cells. There was a very promising downregulation of PARP-1 by CIS-1 treated cells. But surprisingly, e-cadherin has been upregulated in all the three treated groups. This is further observed from the **Figure 5.30 and 5.31** in which the mRNA expressions are normalized with GAPDH.

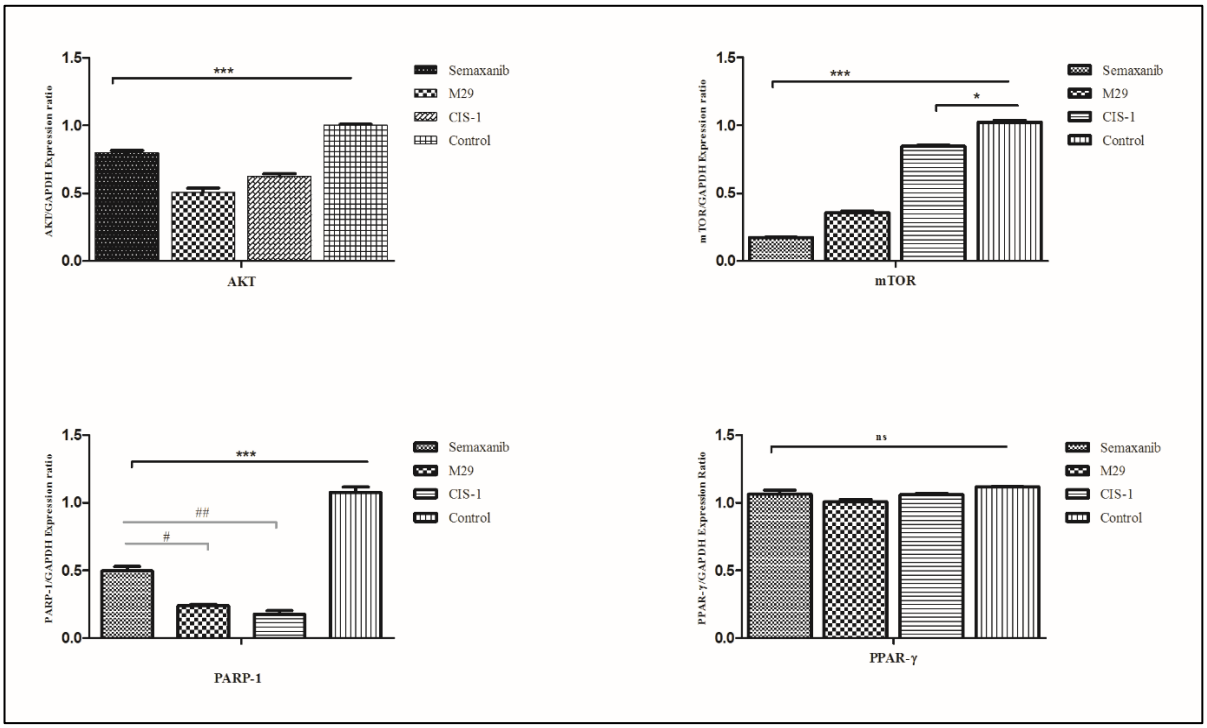


Figure 50: RT–PCR from cell culture mRNA, quantified AKT, mTOR, PARP-1 and PPAR-γ gene expressions of control, M29, CIS-1 and SU5416 treatments. Data are mean±S.E.M. from three mRNA samples in each group. Data were analyzed using one-way ANOVA with post hoc Bonferroni's correction. *p<0.05, **p<0.01, ***p<0.001 versus control; #p<0.05,

$^{\#\#}p<0.01$, $^{\#\#\#}p<0.001$ versus Semaxanib (SU5416). ANOVA, analysis of variance; RT–PCR, reverse transcription PCR.

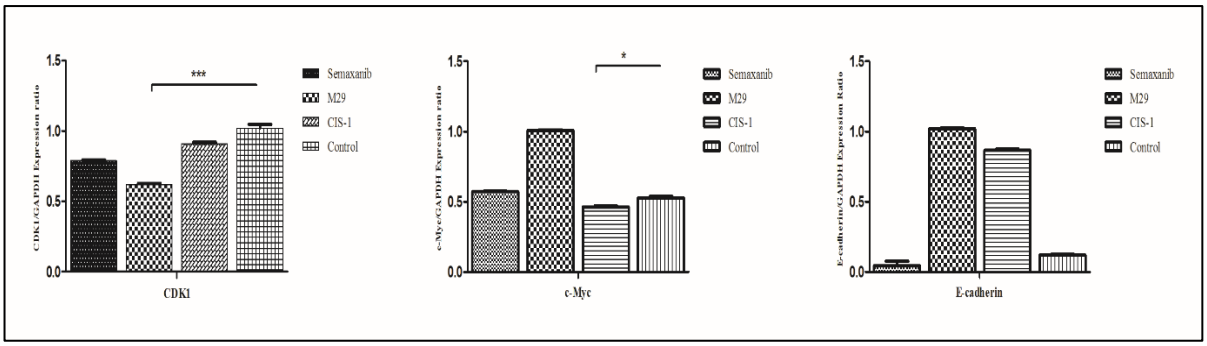


Figure 51: RT–PCR from cell culture mRNA, quantified CDK1, c-Myc and e-Cadherin gene expressions of control, M29, CIS-1 and SU5416 treatments. Data are mean±S.E.M. from three mRNA samples in each group. Data were analyzed using one-way ANOVA with post hoc Bonferroni's correction. $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$ versus control; $^{\#}p<0.05$, $^{\#\#}p<0.01$, $^{\#\#\#}p<0.001$ versus Semaxanib (SU5416). ANOVA, analysis of variance; RT–PCR, reverse transcription PCR.

5.5.6. Clonogenic assay for the evaluation of single cell proliferation effect caused by CatD inhibitors

We performed this assay to estimate the cell survival of treated and untreated MDA-MB-231 cells based on the ability of a single cell to grow into colony (1 colony = 50 cells). The cells were treated with GI_{50} concentrations of the molecules M29, CIS-1 and SU5416 (standard). The images were taken using microscope as seen in **Figure 5.32** and the cells were counted using Image J. The average colony count from three wells were used to calculate plating efficiency and surviving fraction as shown in the **Table 5.15**.

Table 15: Calculation of plating efficiency and survival and comparison of colony counts

Test compounds	#CELLS PLATED	#COLONIES	Mean/X	%Plating Efficiency	% Survival
CONTROL	10,000	210,150,270	210	2.1	100
M29	10,000	110,130,160	133.3	1.33	63.33
CIS1	10,000	30,35,32	32.3	0.32	15.28
SU5416	10,000	38,41,35	38	0.38	18.09

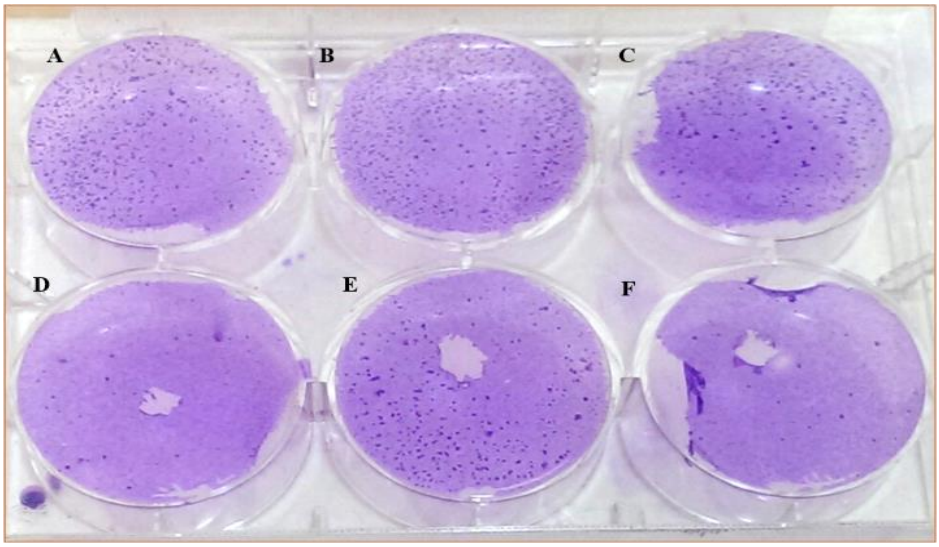


Figure 52: Digital image showing colonies (A, B & C) Control cells (D) CIS-1 treated cells, (E) M29 treated cells, (F) SU5416 treated cells. Cell survival of human cell lines MDA-MB-231 control and treated cells with M29, CIS-1 and Semaxanib were evaluated by the clonogenic assay. Colony formation in control and treated MDA-MB-231 cells is shown. A significant reduction in colony formation between treated and untreated cells was observed by photo micrographic analysis.

5.5.7. M29 and CIS-1 overcame the promising effects of CatD invasion on MDA-MB-231 cells

Gelatin zymography revealed the presence of matrix gelatinases MMP-2 and MMP-9 in the culture supernatants of MDA-MB-231. The approximate molecular weight of MMP-9 (92kDa) and MMP-2 (72kDa) bands in the culture supernatants were correlated with the pre-stained protein coloured ladder and the assay was run with the treated samples. We observed pro-MMP-9 and pro-MMP-2 bands, whereas the active forms were not found and were observed to be digested/ inhibited due to drug treatment after 24 h. We did not observe any gelatinolytic bands after 48h drug treatment as seen in the **Figure 5.33**. M29 and CIS-1 has been shown to prevent breast cancer cell invasion through various means by the above assays. Accordingly, we determined that they have abrogated the promoting effects of MMP-2 and MMP-9 role on MDA-MB-231 cells. Interestingly, we found that MMP-induced invasion was downregulated post-treatment. (**Figure 5.33**)

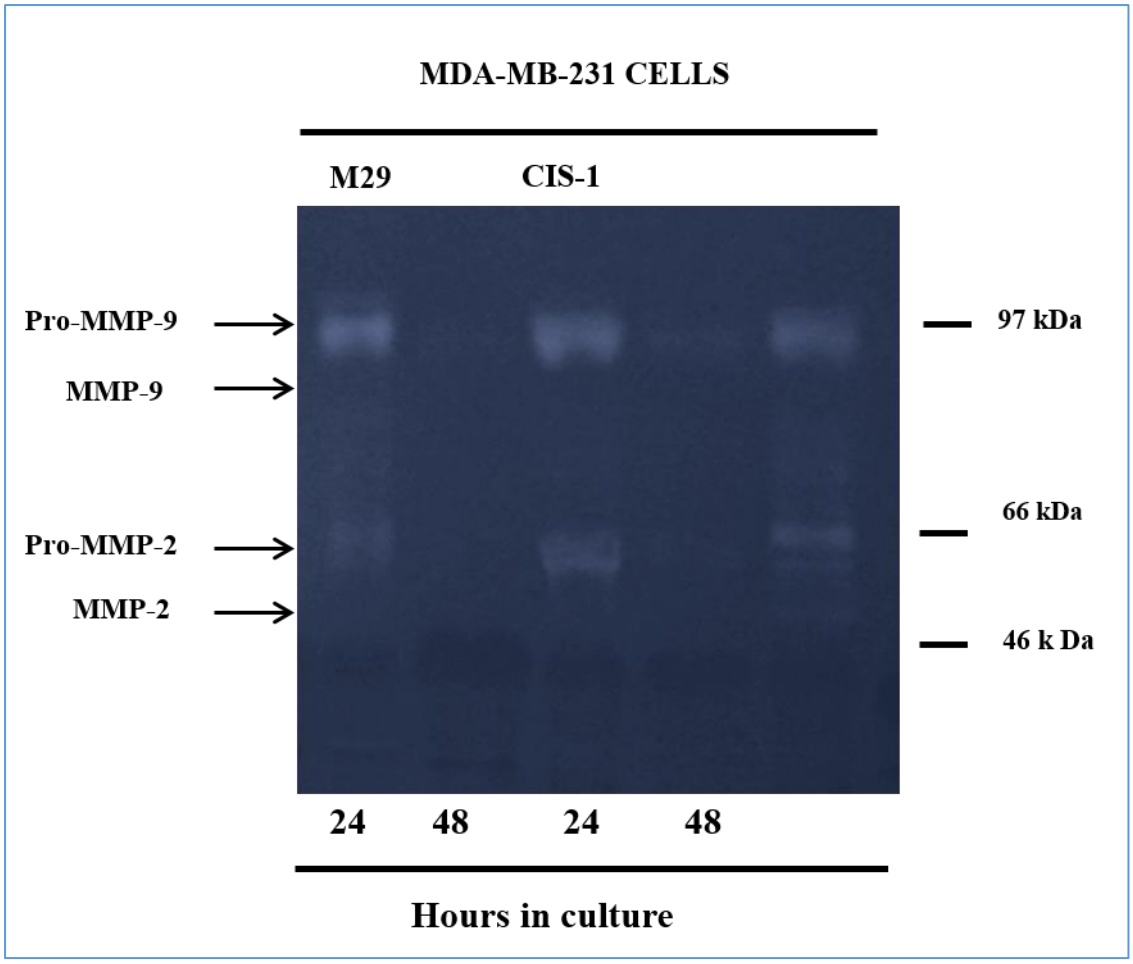


Figure 53: M29 and CIS-1 treatment reduces the release of MMP-2 and MMP-9 using gelatin zymography. Confluent MDA-MB-231 were treated with the lead molecules, and the media were collected and analyzed by gelatin zymography.

5.5.8. COBRA assay

Human genome consists of about half a million long interspersed nucleotide elements (LINE1 elements) that are heavily methylated repeat sequences. It has been reported that more than one-third of DNA methylation occurs in repetitive elements hence studying the methylation of these serves as a marker for global DNA methylation. Amplification of housekeeping gene hypoxanthine phosphoribosyl transferase (HPRT) located on X chromosome was used as a

positive control. Amplification of the imprinting control region of Small Nuclear Ribonucleoprotein-Associated Protein N (SNRPN) was also done.

It has previously been shown that certain drugs show epigenetic changes leading to increase or decrease in global DNA methylation levels in addition to effecting individual genes (109-112). To know whether CIS-1 and MIS-1 treatment shows any such DNA methylation changes, the treated and untreated MDA-MB-231 cells were subjected to COBRA assays (113) for LINE 1 elements, HPRT promoter and SNRPN imprinting control region. For all the sequences studied, the treated cells did not show any significant difference in methylation as compared to untreated cells (**Figure 5.34**). Since there are half a million LINE 1 repeat elements in the genome comprising of one third of the total methylation (114-116), it serves as a marker to study global DNA methylation (117). No difference was observed in the LINE 1 methylation after drug treatment suggesting that the drug does not alter gross methylation patterns in the cells.

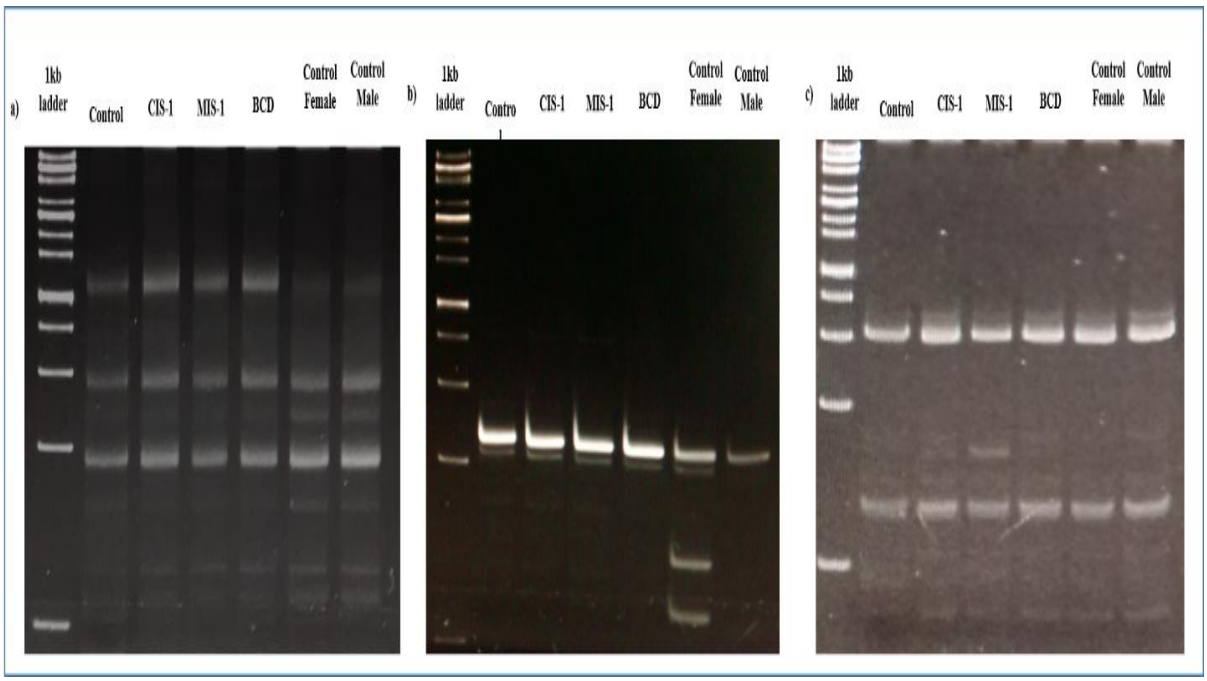


Figure 54: Analysis of the effect of treatment of CIS-1, MIS-1 and M29 (BCD) treatment on MDA-MB- 231 cells on DNA methylation pattern. (a) LINE1 PCR products digested with Taq1 restriction enzyme visualized by PAGE. (b) HPRT PCR products digested with Taq1 restriction enzyme visualized by PAGE (c) SNRPN PCR products digested with BstU1 restriction enzyme visualized by PAGE.

5.6. *In vivo* evaluation of angiogenic effect using Zebrafish embryos

Zebrafish (*Danio rerio*) has been a successful model system in both drug discovery and developmental biology studies and is increasingly being used in studies on angiogenesis. In particular, the embryos of zebrafish are used for this assay, due to its transparency in showing sub-intestinal blood vessel region. The angiogenesis pattern was increased by exposing the embryos to nicotine, which acted as pro-angiogenic agent. Many VEGF-pathway inhibitors were tested in their pre-clinical studies using this model.

Embryos were treated with CIS-1, M29 and Semaxanib as per the procedure described in the "Methods" section. Alkaline phosphatase inhibition is directly proportional to the anti-angiogenic effect of the molecule which was observed by the extent of vessel formation in sub-intestinal region of the embryo. Control embryos were stained and observed clearly for their SIV and ISV regions. Nicotine exposed embryos showed outgrowth in SIV region and was used as angiogenic model. Embryos were exposed to the concentrations of 1,3,10 and 30 µM for each molecule and total disruption of ISV and rupture of SIV regions was found at 30 µM for Semaxanib and at 10 µM concentration for M29 and CIS-1 as shown in Figure 5.35.

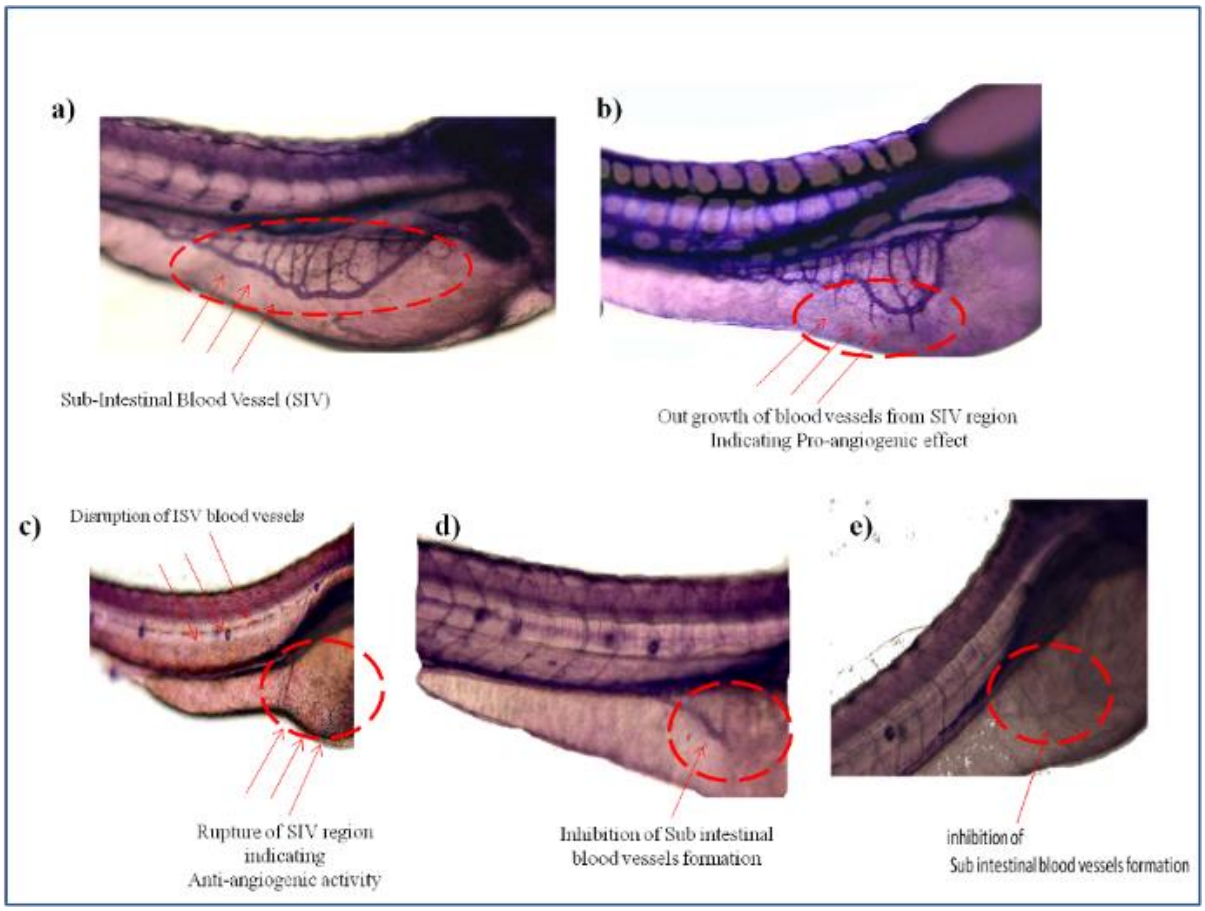


Figure 55: Zebrafish embryos treated with a) Treatment with 0.1% DMSO had no effect on vessel formation. The SIVs (arrow) were in the characteristic pattern. b) Treatment with nicotine to induce angiogenesis. SIV region showing out-growth of vessels. c) Semaxanib blocks both angiogenic and vasculogenic vessel formation in zebrafish embryos. d) Treatment with M29 at 10 µM concentration showing its effect on inter-segmental blood vessels and SIV region. e) Treatment with CIS-1 at 10 µM concentration showing growth inhibition of SIV region.

5.7. Combination effect of PARP-1 inhibitor and CIS-1

Olaparib is a poly ADP-ribose polymerase (PARP) inhibitor, which blocks enzymes involved in DNA repair. It's used among women suffering from ovarian cancer who have BRCA gene mutations identified through a genetic test. Because of phenotypic similarities between some "triple negative" breast cancers (TNBC) and the most prevalent type of breast cancer seen in BRCA1 mutation carriers, some have hypothesized that TNBC might also be specifically sensitive to PARP inhibition. There was a recent phase I trial of Olaparib in combination with Paclitaxel for the treatment of metastatic TNBC. The results of the present study showed that using fixed concentration of Olaparib (0.1, 0.5, 1 and 2 μM) in the presence of various concentrations of CIS-1, determined the effective concentration of synergism. Combination of Olaparib and CIS-1 in the highest tested concentrations (2 μM and 2 μM, respectively) killed 90% of the cells as seen in the **Figure 5.36**. So, in the chemotherapy of TNBC the combination strategy would serve to increase the efficiency in reducing tumors and as our compound CIS-1 was found to prevent invasive mechanism of the cancers, this is an added advantage in combination with Olaparib.

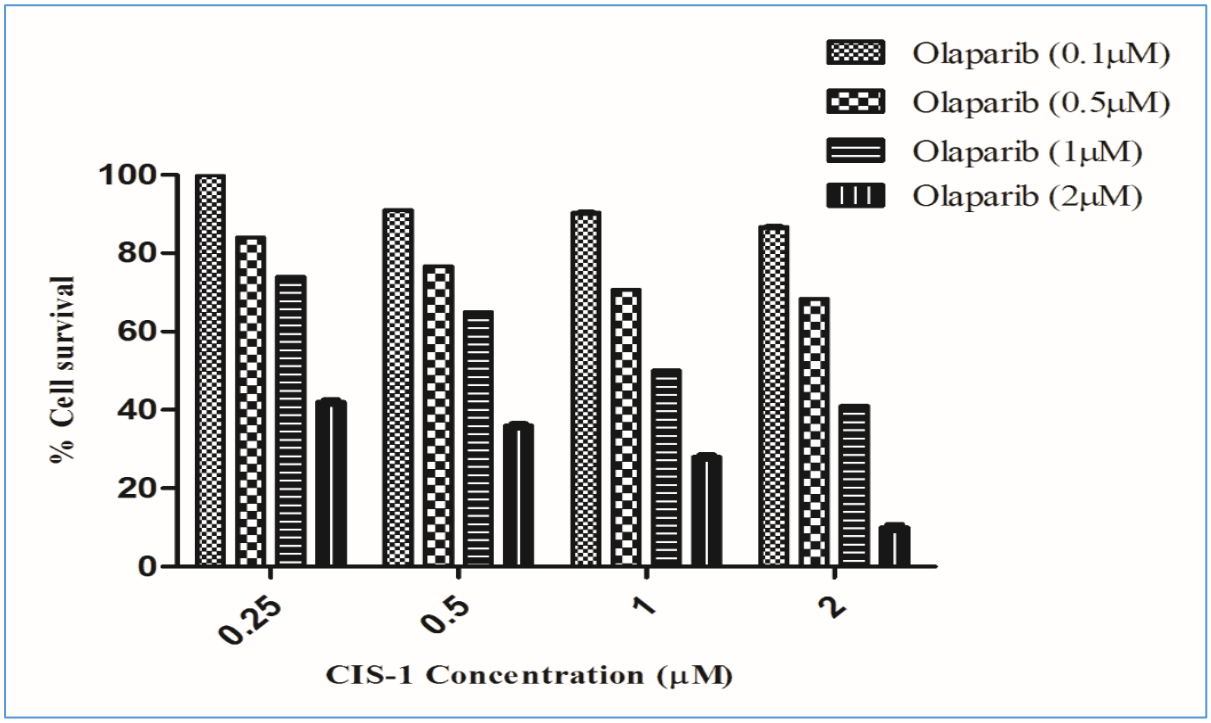


Figure 56: The effects of combination of Olaparib (0.1, 0.5, 1, 2 µM); and CIS-1 (0.25, 0.5, 1, 2 µM).Cells were incubated for 48 h at the presence of combination of Olaparib and CIS-1. Cells were incubated at the presence of different concentrations of Olaparib for 24 h, the media was then aspirated and CIS-1 was added in fresh media and incubated for another 48 h.

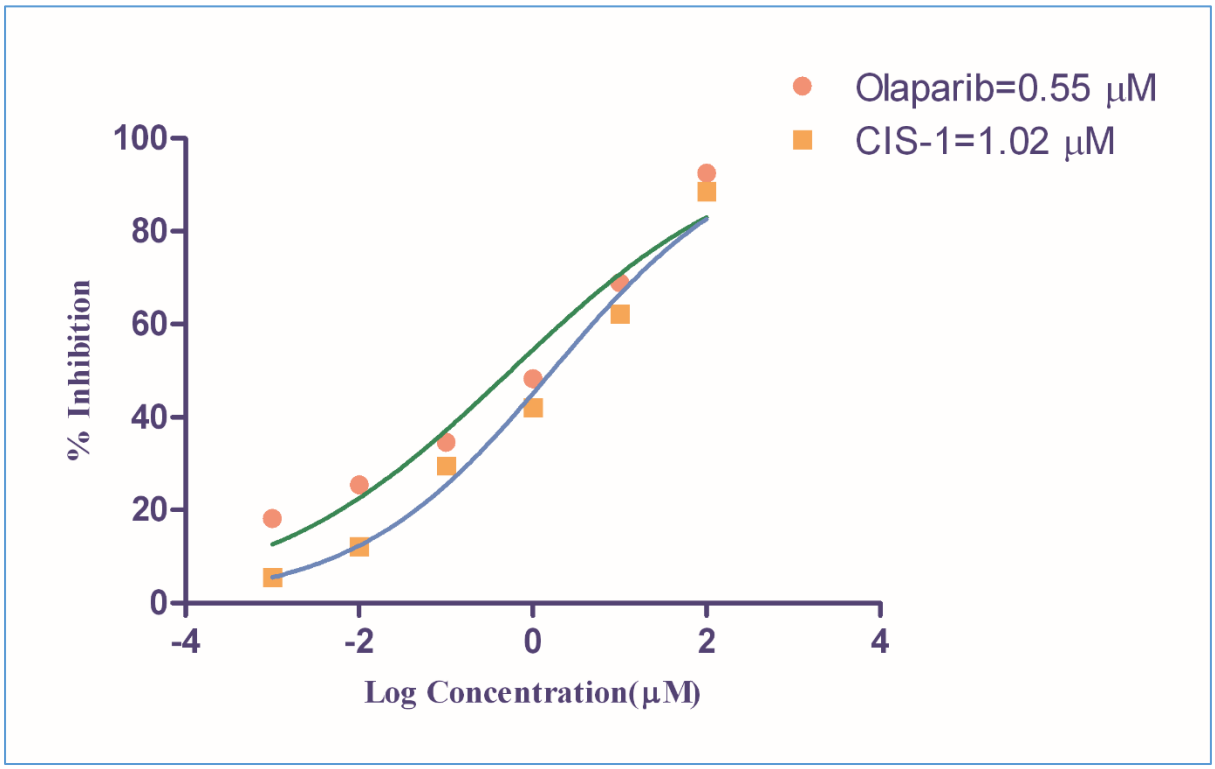


Figure 57: The effect of Olaparib and CIS-1 on MDA-MB-231 as individual drug compounds. The GI_{50} was calculated according to Graph Pad Prism v 6.0.

5.8. Conclusion

M29, the lead compound from BITS database was taken a step forward in synthesizing its analogues for achieving better CatD inhibition.

Among the synthesized 17 analogues, CIS-1 was identified to have >70% CatD inhibition enzymatically. Further cytotoxicity studies revealed the GI_{50} value of CIS-1 to be 1.02±0.05 µM on MDA-MB-231 cells. This lead molecule CIS-1 was taken for further mechanism based studies in comparison to M29. We created hypoxic tumor microenvironment by inducing the

cancer cells with cobalt chloride. This resulted in increased gene expressions of growth factors, metallo matrix proteases, hypoxic factors and interleukins. Since, EGFR, VEGF, MMP-2 and MMP-9 play a major role in invasion and metastasis we have tested M29 and its active analogue CIS-1 for their downstream effects on the above listed factors. Data suggested that, in addition to CatD inhibition, these molecules have a promising down regulating effect of these growth factors. We found the active molecules to also affect the gene expressions of c-Myc, CDK-1 and e-Cadherin which were all involved in the mechanism of ECM degradation, invasion and metastasis. In addition, PARP-1 inhibitor Olaparib was used to demonstrate the combination effect of the most active lead we identified from the whole study i.e., CIS-1. Data strongly suggested that CatD inhibition along with PARP-1 inhibition increased the effect of cancer growth cell inhibition. CIS-1 has itself been evaluated to show effective CatD inhibition along with its effective role in inhibiting invasion and metastasis. We used *in vivo* zebrafish embryo model to visualize the anti-angiogenic activity of CIS-1 and the results demonstrated CIS-1 to cause effective anti-angiogenic effect when compared to Semaxanib (Phase III), a standard reported anti-angiogenic drug.

Chapter 6. Recapitulation and conclusion

The current research in investigating biomarkers for early detection, prognosis and the prediction of treatment responses in breast cancer is rapidly expanding. On the other hand, no validated biomarker currently exists for use in routine clinical practice, and breast cancer detection and management remains dependent on invasive procedures. The efforts to find a cure for breast cancer are ongoing. Researchers develop and test new drugs and other treatments all the time. The focus on monoclonal antibodies, anti-angiogenesis drugs and immunotherapy has always been a demand in the lab. The promising example is Trastuzumab (Herceptin), a FDA- approved monoclonal antibody. Targeted therapeutic strategies had a mixed response in terms of successes and failures.

Our current study is to identify non-peptidic small molecule inhibitors against Cathepsin D, an aspartic protease which was found to over express in metastatic breast tumors. CatD came into lime light as an important prognostic biomarker found in the cytosol of breast cancer patients during the clinical trials. Furthermore, CatD was found to play a vital role in ECM degradation of tumor cells resulting in breast cancer invasion and metastasis. We performed a thorough literature search and found there is an every need to discover CatD inhibitors to treat metastatic breast cancer.

In summary,

- Using molecular modeling methods, in a high throughput screening against human CatD (PDB: 1LYB) and from Asinex and BITS databases, we successfully identified a novel scaffold of inhibitors of CatD.
- CatD enzyme activity and anti-cancer activity of the lead molecules was evaluated, both against the enzyme and in cell based assays.

- Based on the substructure analysis of the lead identified, we synthesized and characterized 17 compounds as analogues of lead molecule M29 to improve the potency.
- Of the molecules synthesized, compound CIS-1 was found to be the most active in cellular assays. Percentage inhibition of 76% using CatD enzyme assay for compound CIS- while the GI50 in MDA-MB-231 cells was 1.02±0.05μM.
- In addition to the anti-proliferative activity, compound CIS-1 inhibited ROS and increased lysosomal degradation detected using Acridine Orange in cancer cells.
- Cell cycle analysis indicated that compound CIS-1 caused a G0/G1 arrest in cells when treated at its GI_{50} concentration of 1.02 μM.
- Angiogenesis assays like wound healing assay, boyden chamber assay and tube formation assay also indicated the role of CIS-1 in preventing angiogenesis of cancer cells.
- An *invivo* zebrafish angiogenesis assay was further carried out to validate the anti-angiogenic effect of the most potent lead CIS-1.
- Prolonged treatment with CatD inhibitors resulted in downregulation of endothelial growth factors, matrix metallo proteases which has in turn effected the levels of AKT and mTOR.
- Combinations of CIS-1 with PARP-1 inhibitor Olaparib were found to be beneficial compared to that of single agents alone in the treatment of metastatic TNBC.

In conclusion, findings from this study have a strong molecular modeling backbone in optimizing and identifying CatD inhibitors and a validated logic in structurally modifying the identified leads to achieve more promising potent molecules. Our findings also suggested the mechanism through which the lead molecules were effective in causing CatD inhibition and the downstream effects involved in causing the inhibition. Our study also has a relevant *invivo*

evaluation and a combination of PARP-1 inhibitiors and CatD inhibitors which could provide a rationale treatment strategy in metastatic breast cancer patients.

Further studies in using *invivo* rodent tumor model mimicking human mammary cancer is warranted for obtaining insights into additional closer effect of identified leads. Our study illustrated the need in developing CatD inhibitors, the role of CatD inhibitors and the relationship between CatD and invasion factors and its combination effect with PARP-1 inhibitors. Thus, the use of CatD inhibitors could be a potential anti-tumor TNBC therapy. In addition, the CatD inhibitors would be important in other cancers such as colon, lung and prostate cancers.

Chapter 7. Future perspectives

Recent studies demonstrated the role of CatD in various cancers along with inflammatory disorders, Alzheimer's disease, lipid adipogenesis and many other neurological disorders. So, there is every scope to test the inhibitors identified in this study on all other ailments for their role in CatD inhibition. The inhibitors we achieved should be further tested in an *invivo* rodent tumor model setup that mimics the human metastatic breast cancer. The tumor model can be achieved by injecting 4T1 mouse mammary tumor cells into BALB/C strain of mice or by injecting MDA-MB-231 human breast cancer cells into nude mice. We procured 4T1 cells and analyzed the level of gene expressions that the lead CIS-1 down regulated during our study. Also, the correlation between PARP-1 inhibitors and CatD inhibitors can be extended for their combined effect. Apart from docking studies prediction of physicochemical properties, the lead CIS-1 can further be evaluated for its plasma and metabolism studies during the *invivo* study.

This study is a challenge and achievement in discovering small molecule non-peptidic CatD inhibitors and also to deduce the mechanism through which the lead molecules might act. This could serve as an important hope in scientists to combat cancer which is the "sword of hope".

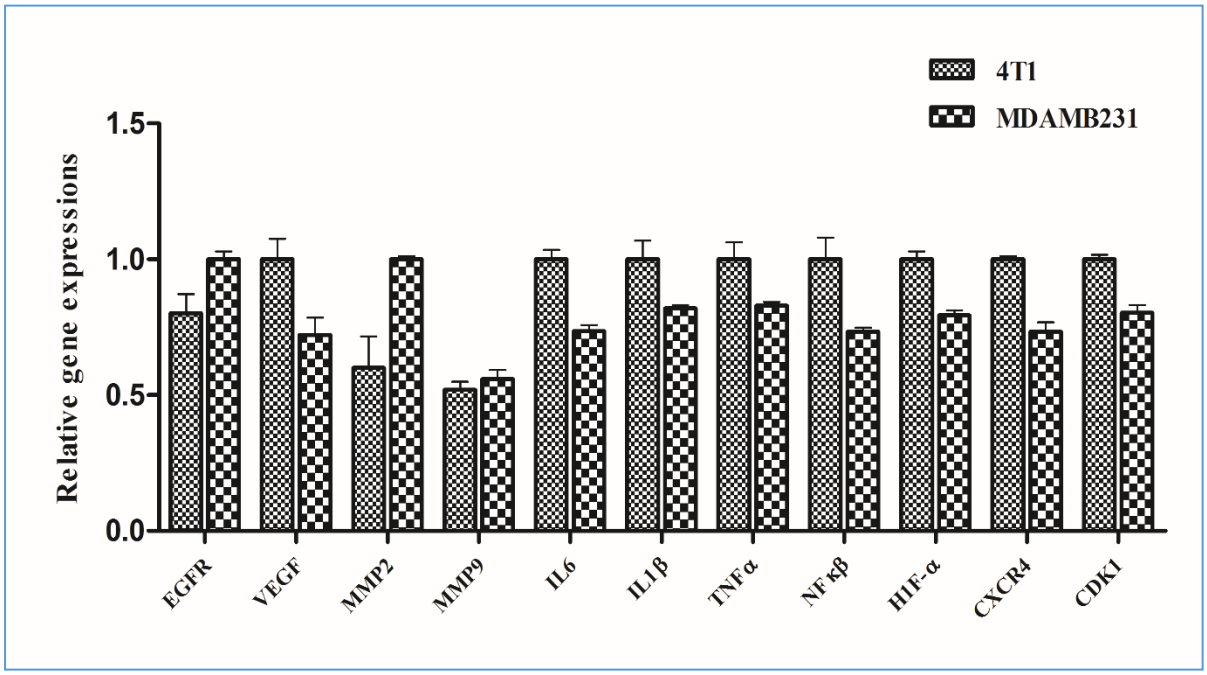


Figure 58: An RTPCR analysis which is a comparative study of levels of the listed gene expressions in 4T1 Vs MDA-MB-231 in order to develop *invivo* 4T1 mouse tumor model. 4T1: Mouse mammary tumor cells; MDA-MB-231: Human breast adenocarcinoma cells.

CHAPTER 9. Bibliography

Abbott D.E., Margaryan N.V., Jeruss J., Khan S., Kaklamani V., Winchester D.J., Hansen N., Rademaker A., Khalkhali-Ellis Z., Hendrix M.J. Reevaluating cathepsin D as a biomarker for breast cancer: serum activity levels versus histopathology. Cancer biology & therapy. 2010, 9, 23-30.

Agarwal N.S., Rich D.H. Inhibition of cathepsin D by substrate analogs containing statine and by analogs of pepstatin. Journal of medicinal chemistry. 1986, 29, 2519-2524.

Alderton G.K. Metastasis: Oxidizing abnormalities. Nature Reviews Cancer. 2009, 4, 227.

Alexov E. Calculating proton uptake/release and binding free energy taking into account ionization and conformation changes induced by protein–inhibitor association: application to plasmepsin, cathepsin D and endothiapepsin–pepstatin complexes. Proteins: Structure, Function, and Bioinformatics. 2004, 56, 572-584.

Allison K.H. Molecular Pathology of Breast Cancer. American journal of clinical pathology. 2012, 6, 770-780.

Al Okail M.S. Cobalt chloride, a chemical inducer of hypoxia-inducible factor-1α in U251 human glioblastoma cell line. Journal of Saudi Chemical Society. 2010, 14, 197-201.

Anderson A.C. The process of structure-based drug design. Chemistry & biology. 2003, 10, 787-797.

Arnaoutova I., Kleinman H.K. In vitro angiogenesis: endothelial cell tube formation on gelled basement membrane extract. Nature protocols. 2010, 5, 628-635.

Bakar M.F., Mohamad M., Rahmat A., Burr S.A., Fry J.R. Cytotoxicity, cell cycle arrest, and apoptosis in breast cancer cell lines exposed to an extract of the seed kernel of Mangifera pajang (bambangan). Food and chemical toxicology. 2010, 48, 1688-1697.

Beaujouin M., Prebois C., Derocq D., Laurent-Matha V., Masson O., Pattingre S., Coopman P., Bettache N., Grossfield J., Hollingsworth R.E., Zhang H. Pro-cathepsin D interacts with the extracellular domain of the β chain of LRP1 and promotes LRP1-dependent fibroblast outgrowth. Journal of cell science. 2010, 123, 3336-3346.

Beenken S.W., Bland K.I. Biomarkers for breast cancer. Minerva chirurgica. 2002, 4, 437-448.

Benes P., Vetvicka V., Fusek M. Cathepsin D—many functions of one aspartic protease. Critical reviews in oncology/hematology. 2008, 68, 12-28.

Bernas T., Dobrucki J. Mitochondrial and nonmitochondrial reduction of MTT: Interaction of MTT with TMRE, JC-1, and NAO mitochondrial fluorescent probes. Cytometry. 2002, 47, 236-242.

Blair R.J., Meng H., Marchese M.J., Ren S., Schwartz L.B., Tonnesen M.G., Gruber B.L. Human mast cells stimulate vascular tube formation. Tryptase is a novel, potent angiogenic factor. Journal of Clinical Investigation. 1997, 99, 2691.

Blanco M.A., Kang Y. Signaling pathways in breast cancer metastasis-novel insights from functional genomics. Breast Cancer Res. 2011, 2, 206.

Bosch A., Eroles P., Zaragoza R., Vina J.R., Lluch A. Triple-negative breast cancer: molecular features, pathogenesis, treatment and current lines of research. Cancer treatment reviews. 2010, 3, 206-215.

Butcher L. Payment Reforms Coming to Oncology. Oncology Times. 2015, 2, 1-26.

Cao Y., DePinho R.A., Ernst M., Vousden K. Cancer research: past, present and future. Nature Reviews Cancer. 2011, 10, 749-754.

Caporaso G.L., Gandy S.E., Buxbaum J.D., Greengard P. Chloroquine inhibits intracellular degradation but not secretion of Alzheimer beta/A4 amyloid precursor protein. Proceedings of the National Academy of Sciences. 1992, 89, 2252-2256.

Carbognin L., Furlanetto J., Vicentini C., Nortilli R., Pilotto S., Brunelli M., Pellini F., Paolo Pollini G., Bria E., Tortora G. Neoadjuvant Strategies for Triple Negative Breast Cancer:'State-of-the-art'and Future Perspectives. Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Cancer Agents). 2015, 15, 15-25.

Carey L., Winer E., Viale G., Cameron D., Gianni L. Triple-negative breast cancer: disease entity or title of convenience?. Nature reviews Clinical oncology. 2010, 7, 683-692.

Carew J.S., Espitia C.M., Esquivel J.A., Mahalingam D., Kelly K.R., Reddy G., Giles F.J., Nawrocki S.T. Lucanthone is a novel inhibitor of autophagy that induces cathepsin D-mediated apoptosis. Journal of Biological Chemistry. 2011, 286, 6602-6613.

Cater S.A., Lees W.E., Hill J., Brzin J., Kay J., Phylip L.H. Aspartic proteinase inhibitors from tomato and potato are more potent against yeast proteinase A than cathepsin D. Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology. 2002, 1596, 76-82.

Chen H.C. Boyden chamber assay. Cell Migration: Developmental Methods and Protocols. 2005, 15-22.

Clark O., Botrel T.E., Paladini L., Ferreira M.B. Targeted therapy in triple-negative metastatic breast cancer: a systematic review and meta-analysis. Core Evid. 2014, 9, 1-11.

Comen E.A., Robson M. Poly (ADP-ribose) polymerase inhibitors in triple-negative breast cancer. Cancer journal (Sudbury, Mass.). 2010, 16, 48.

Crown J., O'shaughnessy J., Gullo G. Emerging targeted therapies in triple-negative breast cancer. Annals of oncology. 2012, 23, 56-65.

Dai Z.J., Gao J., Ma X.B., Yan K., Liu X.X., Kang H.F., Ji Z.Z., Guan H.T., Wang X.J. Up-regulation of hypoxia inducible factor-1α by cobalt chloride correlates with proliferation and apoptosis in PC-2 cells. Journal of Experimental & Clinical Cancer Research. 2012, 31, 1.

Dent R.A., Lindeman G.J., Clemons M., Wildiers H., Chan A., McCarthy N.J., Singer C.F., Lowe E.S., Watkins C.L., Carmichael J. Phase I trial of the oral PARP inhibitor olaparib in combination with paclitaxel for first-or second-line treatment of patients with metastatic triple-negative breast cancer. Breast Cancer Res. 2013, 15, R88.

dos Anjos Pultz B., da Luz F.A., de Faria P.R., Lima A.P., Oliveira R.A., Silva M.J. Far Beyond the Usual Biomarkers in Breast Cancer: A. Journal of Cancer. 2014, 7, 559-571.

Duffy M.J., Brouillet J.P., Reilly D., McDermott E., O'Higgins N., Fennelly J.J., Maudelonde T., Rochefort H. Cathepsin D concentration in breast cancer cytosols: correlation with biochemical, histological, and clinical findings. Clinical chemistry. 1991, 37, 101-104.

Follo C., Ozzano M., Montalenti C., Ekkapongpisit M., Isidoro C. Similarities and differences in the biogenesis, processing and lysosomal targeting between zebrafish and human pro-Cathepsin D: functional implications. The international journal of biochemistry & cell biology. 2013, 45, 273-282.

Fraga D., Meulia T., Fenster S. Real-time PCR. Current Protocols Essential Laboratory Techniques. 2008, 10, 1-34.

Francescone R.A., Faibish M., Shao R. A Matrigel-based tube formation assay to assess the vasculogenic activity of tumor cells. Journal of visualized experiments: JoVE. 2011, 55.

Franken N.A., Rodermond H.M., Stap J., Haveman J., Van Bree C. Clonogenic assay of cells in vitro. Nature protocols. 2006, 1, 2315-2319.

Frankowski H., Gu Y.H., Heo J.H., Milner R., del Zoppo G.J. Use of gel zymography to examine matrix metalloproteinase (gelatinase) expression in brain tissue or in primary glial cultures. Astrocytes: Methods and Protocols. 2012, 221-233.

Frantz C., Stewart K.M., Weaver V.M. The extracellular matrix at a glance. Journal of cell science. 2010, 123, 4195-4200.

Fukuda M.E., Iwadate Y., Machida T., Hiwasa T., Nimura Y., Nagai Y., Takiguchi M., Tanzawa H., Yamaura A., Seki N. Cathepsin D is a potential serum marker for poor prognosis in glioma patients. Cancer research. 2005, 65, 5190-5194.

Fuseka M., Vetvicka V. Dual role of cathepsin D: ligand and protease. Biomed Papers. 2005, 149, 43-50.

Gacko M., Minarowska A., Karwowska A., Minarowski L. Cathepsin D inhibitors. Folia Histochemica et Cytobiologica. 2007, 45, 291-313.

Garcia M., Platet N., Liaudet E., Laurent V., Derocq D., Brouillet J.P., Rochefort H. Biological and clinical significance of cathepsin D in breast cancer metastasis. Stem Cells. 1996, 14, 642-50.

Gradler U., Czodrowski P., Tsaklakidis C., Klein M., Werkmann D., Lindemann S., Maskos K., Leuthner B. Structure-based optimization of non-peptidic Cathepsin D inhibitors. Bioorganic & medicinal chemistry letters. 2014, 24, 4141-4150.

Guille A., Chaffanet M., Birnbaum D. Signaling pathway switch in breast cancer. Cancer cell international. 2013, 13, 66-73.

Hanahan D., Weinberg R.A. The hallmarks of cancer. cell. 2000, 1, 57-70.

Hanahan D., Weinberg R.A. Hallmarks of cancer: the next generation. cell. 2011, 5, 646-674.

Hattori S., Fujisaki H., Kiriyama T., Yokoyama T., Irie S. Real-time zymography and reverse zymography: a method for detecting activities of matrix metalloproteinases and their inhibitors using FITC-labeled collagen and casein as substrates. Analytical biochemistry. 2002, 301, 27-34.

Hawkes S.P., Li H., Taniguchi G.T. Zymography and Reverse Zymography for Detecting MMPs, andTIMPs. Matrix metalloproteinase protocols. 2001, 399-410.

Heneghan H.M., Miller N., Lowery A.J., Sweeney K.J., Newell J., Kerin M.J. Circulating microRNAs as novel minimally invasive biomarkers for breast cancer. Annals of surgery. 2010, 3, 499-505.

Himmelfarb J., Hakim R.M., Holbrook D.G., Leeber D.A., Ault K.A. Detection of granulocyte reactive oxygen species formation in whole blood using flow cytometry. Cytometry. 1992, 13, 83-89.

Holliday D.L., Speirs V. Choosing the right cell line for breast cancer research. Breast Cancer Res. 2011, 13, 215.

Howell A., Anderson A.S., Clarke R.B., Duffy S.W., Evans D.G., Garcia-Closas M., Gescher A.J., Key T.J., Saxton J.M., Harvie M.N. Risk determination and prevention of breast cancer. Breast Cancer Res. 2014, 5, 446.

Huang L., Liu Z., Chen S., Liu Y., Shao Z. A prognostic model for triple-negative breast cancer patients based on node status, cathepsin-D and Ki-67 index. PloS one. 2013, 8, e83081.

Huo S., Wang J., Cieplak P., Kollman P.A., Kuntz I.D. Molecular dynamics and free energy analyses of cathepsin D-inhibitor interactions: insight into structure-based ligand design. Journal of medicinal chemistry. 2002, 45, 1412-1419.

Johnson M.D., Torri J.A., Lippman M.E., Dickson R.B. The role of cathepsin D in the invasiveness of human breast cancer cells. Cancer research. 1993, 53, 873-877.

Johansson P.O., Chen Y., Belfrage A.K., Blackman M.J., Kvarnstrom I., Jansson K., Vrang L., Hamelink E., Hallberg A., Rosenquist A, Samuelsson B. Design and synthesis of potent inhibitors of the malaria aspartyl proteases plasmepsin I and II. Use of solid-phase synthesis to explore novel statine motifs. Journal of medicinal chemistry. 2004, 47, 3353-3366.

Khalkhali-Ellis Z., Goossens W., Margaryan N.V., Hendrix M.J. Cleavage of histone 3 by Cathepsin D in the involuting mammary gland. PloS one. 2014, 9, e103230.

Knopfova L., Benes P., Pekarcikova L., Hermanova M., Masarik M, Pernicova Z, Soucek K., Smarda J. c-Myb regulates matrix metalloproteinases 1/9, and cathepsin D: implications for matrix-dependent breast cancer cell invasion and metastasis. Molecular cancer. 2012, 11, 1-15.

Kramer N., Walzl A., Unger C., Rosner M., Krupitza G., Hengstschlager M., Dolznig H. In vitro cell migration and invasion assays. Mutation Research/Reviews in Mutation Research. 2013, 752, 10-24.

Kute T.E., Shao Z.M., Sugg N.K., Long R.T., Russell G.B., Case L.D. Cathepsin D as a prognostic indicator for node-negative breast cancer patients using both immunoassays and enzymatic assays. Cancer research. 1992, 52, 5198-5203.

Liang C.C., Park A.Y., Guan J.L. In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. Nature protocols. 2007, 2, 329-333.

Lison P., Rodrigo I., Conejero V. A novel function for the cathepsin D inhibitor in tomato. Plant physiology. 2006, 142, 1329-1339.

Liu Q., Xu Z., Mao S., Chen W., Zeng R., Zhou S., Liu J. Effect of hypoxia on hypoxia inducible factor-1α, insulin-like growth factor I and vascular endothelial growth factor expression in hepatocellular carcinoma HepG2 cells. Oncology letters. 2015, 9, 142-1148.

Lopez-Sanchez L.M., Jimenez C., Valverde A., Hernandez V., Penarando J., Martinez A., Lopez-Pedrera C., Munoz-Castaneda J.R., Juan R., Aranda E., Rodriguez-Ariza A. $CoCl_2$, a mimic of hypoxia, induces formation of polyploid giant cells with stem characteristics in colon cancer. PloS one. 2014, 9, e99143.

Majer P., Collins J.R., Gulnik S.V., Erickson J.W. Structure-based subsite specificity mapping of human cathepsin D using statine-based inhibitors. Protein science. 1997, 6, 1458-1466.

Malhotra G.K., Zhao X., Band H., Band V. Histological, molecular and functional subtypes of breast cancers. Cancer biology & therapy. 2010, 10, 955-960.

Malik M., Fenko M.D., Sheikh A.M., Wen G., Li X. A novel approach for characterization of cathepsin D protease and its effect on tau and β-amyloid proteins. Neurochemical research. 2011, 36, 754-760.

Malik M., Sheikh A.M., Wen G., Spivack W., Brown W.T., Li X. Expression of inflammatory cytokines, Bcl2 and cathepsin D are altered in lymphoblasts of autistic subjects. Immunobiology. 2011, 216, 80-85.

Matarrese P., Ascione B., Ciarlo L., Vona R., Leonetti C., Scarsella M., Mileo A.M., Catricala C., Paggi M.G., Malorni W. Cathepsin B inhibition interferes with metastatic potential of human melanoma: an in vitro and in vivo study. Molecular cancer. 2010, 9, 1-14.

Markicevic M., Kanjer K., Mandusic V., Buta M., Neskovic-Konstantinovic Z., Nikolic-Vukosavljevic D. Cathepsin D as an indicator of clinical outcome in early breast carcinoma during the first 3 years of follow-up. Biomarkers in medicine. 2013, 7, 747-58.

Maynadier M., Farnoud R., Lamy P.J., Laurent-Matha V., Garcia M., Rochefort H. Cathepsin D stimulates the activities of secreted plasminogen activators in the breast cancer acidic environment. International journal of oncology. 2013, 43, 1683-1690.

McArthur H., Dickler M. Biomarkers in breast cancer. Cancer biology & therapy. 2008, 1, 21-32.

McPherson K., Steel C., Dixon J.M. Breast cancer—epidemiology, risk factors, and genetics. Bmj. 2000, 7261, 624-628.

Metcalf P., Fusek M. Two crystal structures for cathepsin D: the lysosomal targeting signal and active site. The EMBO journal. 1993, 12, 1293.

Misek D.E., Kim E.H. Protein biomarkers for the early detection of breast cancer. International journal of proteomics. 2011, 1-9.

Mosmann T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. Journal of immunological methods. 1983, 65, 55-63.

Myc A., Pizzolo J.G., Dygulski K., Melamed M.R. Increase in acridine orange (ao) fluorescence intensity of monocytes cultured in plastic tissue culture plates as measured by flow cytometry. Cytometry. 1992, 13, 103-107.

Nair M.K., Varghese C.H., Swaminathan R. Cancer: Current scenario, intervention strategies and projections for 2015. Burden of Disease in India. 2005, 219.

Nomura T., Katunuma N. Involvement of cathepsins in the invasion, metastasis and proliferation of cancer cells. The Journal of Medical Investigation. 2005, 52, 1-9.

O'gorman M.R., Corrochano V. Rapid whole-blood flow cytometry assay for diagnosis of chronic granulomatous disease. Clinical and diagnostic laboratory immunology. 1995, 2, 227-232.

Plumb J.A. Cell sensitivity assays: clonogenic assay. Cancer Cell Culture: Methods and Protocols. 2004, 159-64.

Polyak K. Heterogeneity in breast cancer. The Journal of clinical investigation. 2011, 121, 3786-3788.

Ponce M.L. Tube formation: an in vitro matrigel angiogenesis assay. Angiogenesis Protocols: Second Edition. 2009, 183-188.

Radisky E.S. Cathepsin D: Regulation in mammary gland remodeling, misregulation in breast cancer. Cancer Biology & therapy. 2010, 5, 467-470.

Roszkowska-Jakimiec W., Wołczynski S., Chlabicz M. Effects of cathepsin D inhibitor from Vicia sativa L. seed hulls on human skin fibroblasts and breast cancer cells (in vitro studies). Rocz Akad Med Bialymst. 2004, 49, 234-235.

Riccardi C., Nicoletti I. Analysis of apoptosis by propidium iodide staining and flow cytometry. Nature protocols. 2006, 1, 1458-1461.

Sadeghi-Aliabadi H., Minaiyan M., Dabestan A. Cytotoxic evaluation of doxorubicin in combination with simvastatin against human cancer cells. Research in pharmaceutical sciences. 2011, 5, 127-133.

Saraswat-Ohri S., Vetvicka V. New insights into procathepsin D in pathological and physiological conditions. North American journal of medical sciences. 2011, 3, 222.

Sastry G.M., Adzhigirey M., Day T., Annabhimoju R., Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. Journal of computer-aided molecular design. 2013, 27, 221-234.

Sheikh A.M., Li X., Wen G., Tauqeer Z., Brown W.T., Malik M. Cathepsin D and apoptosis related proteins are elevated in the brain of autistic subjects. Neuroscience. 2010, 165, 363-370.

Siegel R.L., Miller K.D., Jemal A. Cancer statistics, 2015. CA: a cancer journal for clinicians. 2015, 1, 5-29.

Smith C.J., Osborn A.M. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. FEMS microbiology ecology. 2009, 67, 6-20.

Soderholm J., Heald R. Scratch n'screen for inhibitors of cell migration. Chemistry & biology. 2005, 12, 263-265.

Spyratos F., Hacene K., Rouesse J., Brunet M., Andrieu C., Desplaces A., Brouillet J.P., Defrenne A., Maudelonde T., Rochefort H. Cathepsin D: an independent prognostic factor for metastasis of breast cancer. The Lancet. 1989, 334, 1115-1118.

Steinfeld R., Reinhardt K., Schreiber K., Hillebrand M., Kraetzner R., Bruck W., Saftig P., Gartner J. Cathepsin D deficiency is associated with a human neurodegenerative disorder. The American Journal of Human Genetics. 2006, 78, 988-998.

Stingl J., Caldas C. Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. Nature Reviews Cancer. 2007, 7, 791-799.

Tan G.J., Peng Z.K., Lu J.P., Tang F.Q. Cathepsins mediate tumor metastasis. World J Biol Chem. 2013, 4, 91-101.

Tedone T., Correale M., Barbarossa G., Casavola V., Paradiso A., Reshkin S.J. Release of the aspartyl protease cathepsin D is associated with and facilitates human breast cancer cell invasion. The FASEB journal. 1997, 11, 785-792.

Terman A., Kurz T., Gustafsson B., Brunk U.T. Lysosomal labilization. IUBMB life. 2006, 58, 531-539.

Thompson A., Brennan K., Cox A., Gee J., Harcourt D., Harris A., Harvie M., Holen I., Howell A., Nicholson R., Steel M. Evaluation of the current knowledge limitations in breast cancer research: a gap analysis. Breast Cancer Res. 2008, 2, R26.

Tobia C., De Sena G., Presta M. Zebrafish embryo, a tool to study tumor angiogenesis. International Journal of Developmental Biology. 2011, 55, 505.

Toth M., Fridman R. Assessment of gelatinases (MMP-2 and MMP-9 by gelatin zymography. Metastasis Research Protocols: Volume I: Analysis of Cells and Tissues. 2001, 163-174.

van Meerloo J., Kaspers G.J., Cloos J. Cell sensitivity assays: the MTT assay. Cancer cell culture: methods and protocols. 2011, 237-245.

Vetvicka V., Fusek M. Procathepsin D as a tumor marker, anti-cancer drug or screening agent. Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Cancer Agents). 2012, 12, 172-175.

Weigel M.T., Dowsett M. Current and emerging biomarkers in breast cancer: prognosis and prediction. Endocrine-related cancer. 2010, 4, 245-262.

Weigelt B., Peterse J.L., Van't Veer L.J. Breast cancer metastasis: markers and models. Nature reviews cancer. 2005, 8, 591-602.

Westley B.R., May F.E. Prognostic value of cathepsin D in breast cancer. British journal of cancer. 1999, 79, 189.

Wu D., Yotnda P. Induction and testing of hypoxia in cell culture. JoVE (Journal of Visualized Experiments). 2011, 12, e2899.

Yelamanchili S.V., Chaudhuri A.D., Flynn C.T., Fox H.S. Upregulation of cathepsin D in the caudate nucleus of primates with experimental parkinsonism. Mol Neurodegener. 201, 6, 52.

ANNEXURE

Synthetic procedure and characterization for synthesis of substituted 1-(2-bromoethyl) indoline-2, 3-dione (**2a-c**).

To a solution corresponding substituted indoline-2,3-dione(1a-e) (1 eq) in dry N,N' dimethyl formamide (10 mL) was added K_2CO_3 (2.5 eq) and followed by 1,2- dibromoethane (0.913 g, 2 eq) and heated in sealed tube at 80 ^{0}C for 3 h (monitored by TLC & LCMS for completion). The reaction mixture was quenched with ice. The reaction mixture was extracted with ethyl acetate (3 × 40 mL). The combined organic extracts were washed with brine (2 × 40 mL) and water (3 × 50 mL), dried ($MgSO_4$) and concentrated under reduced pressure. The crude product was purified by silica gel column chromatography using hexane: ethyl acetate as eluent to give the corresponding substituted. 1-(2-bromoethyl) indoline-2,3-dione (**2a-e**)

1-(2-bromoethyl) indoline-2,3-dione (**2a**).

The compound was synthesized according to the general procedure using indoline-2,3-dione (1a) (1 g, 6.79 mmol) , 1,2 dibromoethane (2.55g, 13.6 mmol) and K_2CO_3 (2.2g, 16.9 mmol) to afford 2a (0.95 g, 55 %) as orange red solid.. ^{1}H NMR [DMSO-d6]: δH 3.64-3.92 (m, 4H), 7.15-7.46 (m, 4H). ^{13}C NMR [DMSO-d6] δc: 29.5, 55.3, 115.8, 116.5, 126.5, 130.2, 135.6, 147.3, 159.5, 180.2. ESI-MS m/z: 256 (M+2H) +. AnalCalcdfor $C_{10}H_8BrNO_2$: C, 47.27; H, 3.17; N, 5.51. Found: C, 47.26; H, 3.19; N, 5.50.

1-(2-bromoethyl)-5-methylindoline-2,3-dione **(2b)**.

The compound was synthesized according to the general procedure using 5-methylindoline-2,3-dione (1b) (1 g, 6.20 mmol) , 1,2 dibromoethane (2.33g, 12.4 mmol) and K_2CO_3 (2.1g, 15.5 mmol) to afford 2b (0.98 g, 59 %) as orange red solid.. ^{1}H NMR [DMSO-d6]: δH 2.52 (s, 3H), 3.71-4.02 (m, 4H), 7.23-7.81 (m, 3H). ^{13}C NMR [DMSO-d6] δc: 22.5, 26.5, 55.8, 115.5, 116.8,

125.2, 125.8, 138.2,146.3,159.5,180.2.ESI-MS m/z: 270 (M+2H) +. AnalCalcdfor $C_{11}H_{10}BrNO_2$: C, 49.28; H, 3.76; N, 5.22. Found: C, 49.29; H, 3.77; N, 5.20

1-(2-bromoethyl)-5-chloroindoline-2,3-dione **(2c).**

The compound was synthesized according to the general procedure using 5-chloroindoline-2,3-dione (1c) (1 g, 5.50 mmol) , 1,2 dibromoethane (2.06g, 11 mmol) and K_2CO_3 (2.1g, 15.5 mmol) to afford 2b (0.98 g, 59 %) as orange red solid.. ^{1}H NMR [DMSO-d6]: δH 3.82-4.12 (m, 4H), 7.86-8.13 (m, 3H). ^{13}C NMR [DMSO-d6] δc: 29.2, 55.6, 120.5, 123.8, 125.2, 135.6, 138.3, 147.5, 159.8, 180.3 .ESI-MS m/z: 290 (M+2H) +. AnalCalcdfor $C_{10}H_7BrClNO_2$: C, 41.63; H, 2.45; N, 4.85. Found: C, 41.64; H, 2.44; N, 4.83.

General procedure for synthesis of tert-butyl (1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)carbamate (**3a-c)**.

A solution of corresponding substituted *1-(2-bromoethyl)indoline-2,3-dione* (2a-c) (1 eq) in N,N dimethyl formamide (10 mL) and 4-N-Boc-aminopiperidine (1.1 eq)was taken in a seal tube. K_2CO_3 (3 eq) was added to the reaction mixture and heated to 100 ^{0}C for 4 h (monitored by TLC & LCMS), quenched with ice and concentrated under reduced pressure. The residue was partitioned between ethyl acetate (20 mL) and water (20 mL). Aqueous layer was basified with sodium carbonate solution and extracted with ethyl acetate (3 × 20 mL) and washed with water (3 × 10 mL) and brine (3 × 10 mL). Organic layer was dried over anhydrous sodium sulphate, filtered, and concentrated under reduced pressure to give the corresponding substituted tert-butyl (1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)carbamate (3a-c).

tert-butyl (1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)carbamate (**3a**).

The compound was synthesized according to the general procedure using 1-(2-bromoethyl)indoline-2,3-dione (2a) (1 g, 3.93 mmol) , 4-N-Boc-aminopiperidine (0.86g, 4.33mmol) and K2CO3 (1.35g, 9.84 mmol) to afford 3a (0.98 g, 69 %) as orange yellow solid.. ^{1}H NMR [DMSO-d6]: δH 1.42-3.85 (m, 22H), 7.03-7.75 (m,4H), 8.17 (b, 1H). ^{13}C NMR [DMSO-d6] δc: 29.5 (3C), 30.5 (2C), 49.5, 50.2, 52.5, (2C), 54.1, 80.3, 116.9, 117.8, 124.5, 130.2, 135.6, 149.2, 156.5, 160.2, 180.3. ESI-MS m/z: 374 (M+H) +. AnalCalcdfor $C_{20}H_{27}N_3O_4$: C, 64.32; H, 7.29; N, 11.25. Found: C, 64.34; H, 7.30; N, 11.24.

tert-butyl (1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)carbamate (**3b**).

The compound was synthesized according to the general procedure using 1-(2-bromoethyl)-5-methylindoline-2,3-dione (2b) (1 g, 3.73 mmol) , 4-N-Boc-aminopiperidine (0.82g, 4.10mmol) and K2CO3 (1.13g8.21 mmol) to afford 3a (1.1 g, 76 %) as orange solid.. ^{1}H NMR [DMSO-d6]: δH 1.42-3.93 (m, 25H), 7.39-7.95 (m,3H), 8.11 (b, 1H). ^{13}C NMR [DMSO-d6] δc:22.5, 28.6 (3C), 29.5 (2C), 49.2, 50.5, 52.8, (2C), 54.5, 80.2, 117.6, 118.3, 124.2, 124.5, 139.5, 143.8, 156.2, 159.6, 180.1. ESI-MS m/z: 388 (M+H) +. AnalCalcdfor $C_{21}H_{29}N_3O_4$: C, 65.09; H, 7.54; N, 10.84. Found: C, 65.11; H, 7.55; N, 10.84.

tert-butyl (1-(2-(5-chloro-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)carbamate (**3c**).

The compound was synthesized according to the general procedure using 1-(2-bromoethyl)-5-chloroindoline-2,3-dione (2c) (1 g, 3.46 mmol) , 4-N-Boc-aminopiperidine (0.76g, 3.81 mmol) and K2CO3 (1.19g, 8.68 mmol) to afford 3a (0.89 g, 64 %) as yellow solid.. ^{1}H NMR [DMSO-d6]: δH 1.42-4.01 (m, 22H), 7.65-8.06 (m,3H), 8.21 (b, 1H). ^{13}C NMR [DMSO-d6] δc:29.5 (3C), 30.2 (2C), 49.1, 49.5, 50.9 (2C), 52.5, 80.5, 120.2, 124.1, 127.5, 135.3, 140.2, 143.5, 156.5, 160.2, 180.3. ESI-MS m/z: 407 (M) +. AnalCalcdfor $C_{20}H_{26}ClN_3O_4$: C, 58.89; H, 6.42; N, 10.30. Found: C, 58.88; H, 6.40; N, 10.32.

General procedure for synthesis of substituted 1-(2-(4-aminopiperidin-1-yl)ethyl)indoline-2,3-dione *(4a-c).*

A solution of corresponding substituted tert-butyl (1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)carbamate (3a-c) (1 eq) in dichloromethane (10 mL) was taken in round bottom flask.The reaction mixture was cooled to 0 ºC followed by the addition of trifluro acetic acid (2 ml) dropwise. After addition the reaction mixture was stirred at rt for 3 h (monitored by TLC & LCMS), concentrated under reduced pressure.to give the corresponding substituted 1-(2-(4-aminopiperidin-1-yl)ethyl)indoline-2,3-dione (4a-c). Crude was taken for final coupling without purification.

General procedure for synthesis of substituted 1-(1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylthiourea *(5-7).*

A solution of 1-(2-(4-aminopiperidin-1-yl)ethyl)indoline-2,3-dione (4a) (1 eq) in dichloromethane (10 mL), was added triethyl amine (3 eq) at 0 ºC. Then corresponding substituted phenyl isothiocyanate was added to the reaction mixture and stirred at rt for 6 h (monitored by TLC & LCMS), quenched with ice and concentrated under reduced pressure. The residue was partitioned between ethyl acetate (20 mL) and water (20 mL). Aqueous layer was basified with sodium carbonate solution and extracted with ethyl acetate (3 × 20 mL) and washed with water (3 × 10 mL) and brine (3 × 10 mL). Organic layer was dried over anhydrous sodium sulphate, filtered, and concentrated under reduced pressure to give the corresponding substituted 1-(1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylthiourea (5-7).

1-(1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylthiourea (**5**).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)indoline-2,3-dione (**4a**) (0.1 g, 0.365 mmol) and phenyl isothiocyanate (0.059 g, 0.439 mmol) to afford **5** (0.075 g, 50 %) as yellow solid. M.p: 115-

117°C. ^{1}H NMR [DMSO-d_6]: δ_H 1.56 – 3.87 (m, 13H), 4.53 (b,1H), 6.63 – 7.88 (m, 9H), 10.56 (b, 1H).^{13}C NMR [DMSO-d_6] δ_c: 30.2 (2C), 50.5, 51.5, 54.2, 56.5, 116.9, 118.1, 126.2, 126.7 (2C), 128.7, 129.5 (2C), 130.2, 135.6, 137.5, 140.2, 149.5, 161.1, 178.6, 180.3.ESI-MS *m/z*: 409 $(M+H)^+$. AnalCalcdfor $C_{22}H_{24}N_4O_2S$: C, 64.68; H, 5.92; N, 13.71; Found: C, 64.69; H, 5.92; N, 13.70.

*1-(1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-(p-tolyl)thiourea (***6**).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)indoline-2,3-dione (4a) (0.1 g, 0.365 mmol) and para tolyl isothiocyanate (0.065 g, 0.439 mmol) to afford 6 (0.083 g, 53 %) as orange solid. M.p: 124-126°C. ^{1}H NMR [DMSO-d_6]: δ_H 1.62 – 3.95 (m, 16H), 4.21 (b,1H), 6.52 – 7.43 (m, 8H), 10.25 (b, 1H).^{13}C NMR [DMSO-d_6] δ_c: 22.5, 30.6 (2C), 50.5, 51.5 (2C), 52.5, 56.5, 117.1, 118.2, 126.7, 128.5 (2C),130.1(2C), 130.2, 135.3, 136.2, 138.5, 147.5, 160.2, 179.3, 180.5. ESI-MS m/z: 423 (M+H)+. AnalCalcdfor $C_{23}H_{26}N_4O_2S$: C, 65.38; H, 6.20; N, 13.26; Found: C, 65.39; H, 6.21; N, 13.25.

1-(4-chlorophenyl)-3-(1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)thiourea (7).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)indoline-2,3-dione (4a) (0.1 g, 0.365 mmol) and 4 chlorophenyl isothiocyanate (0.075 g, 0.439 mmol) to afford 7 (0.063 g, 38 %) as orange red solid. M.p: 125-127°C. . ^{1}H NMR [DMSO-d_6]: δ_H 1.63 – 4.02 (m, 13H), 4.65 (b,1H), 6.65 – 7.45 (m, 8H), 10.58 (b, 1H).^{13}C NMR [DMSO-d6] δc: 30.5 (2C) 50.2, 51.5, 54.2, 56.5, 117.1, 118.1, 126.5, 129.5 (2C), 129.8, 130.2 (2C), 134.2, 134.8, 135.3, 148.2, 150.5, 160.5, 178.2, 180.2.ESI-MS m/z: 442 (M)+. AnalCalcdfor $C_{22}H_{23}ClN_4O_2S$: C, 59.65; H, 5.23; N, 12.65; Found: C, 59.64; H, 5.22; N, 12.65

General procedure for the synthesis of substituted 1-(1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylurea (**8-10**).

A solution of 1-(2-(4-aminopiperidin-1-yl)ethyl)indoline-2,3-dione (4a) (1 eq) in dichloromethane (10 mL), was added triethyl amine (3 eq) at 0 ºC. Then corresponding substituted phenyl isocyanate was added to the reaction mixture and stirred at rt for 6 h (monitored by TLC & LCMS), quenched with ice and concentrated under reduced pressure. The residue was partitioned between ethyl acetate (20 mL) and water (20 mL). Aqueous layer was basified with sodium carbonate solution and extracted with ethyl acetate (3 × 20 mL) and washed with water (3 × 10 mL) and brine (3 × 10 mL). Organic layer was dried over anhydrous sodium sulphate, filtered, and concentrated under reduced pressure to give the corresponding substituted 1-(1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylurea (8-10).

1-(1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylurea (8).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)indoline-2,3-dione (4a) (0.1 g, 0.365 mmol) and phenyl isocyanate (0.052 g, 0.439 mmol) to afford 8 (0.083 g, 56%) as dark yellow solid. M.p: 117-119°C.^{1}H NMR [DMSO-d6]: δH 1.57 – 3.85 (m, 13H), 7.27 – 7.78 (m, 9H), 10.55 (b,1H) 10.68 (b, 1H).^{13}C NMR [DMSO-d6] δc: 30.5 (2C), 49.5, 50.2, 50.5 (2C), 52.5, 116.8, 117.9, 120.2 (2C), 124.3, 127.9, 128.1 (2C), 128.5, 136.2, 140.5, 150.2, 155.6, 159.8, 180.5. ESI-MS m/z: 393 (M+H)+. AnalCalcdfor $C_{22}H_{24}N_4O_3$: C, 67.33; H, 6.16; N, 14.28; Found: C, 67.35; H, 6.15; N, 14.29.

1-(1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-(p-tolyl)urea (9).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)indoline-2,3-dione (4a) (0.1 g, 0.365 mmol) and para tolyl

isocyanate (0.05 8g, 0.439 mmol) to afford 9 (0.078 g, 53 %) as brown solid. M.p: 153-155°C. ^{1}H NMR [DMSO-d6]: δH 1.63 – 3.95 (m, 16H), 7.25 – 7.65 (m, 8H), 10.29 (b,1H) 10.38 (b, 1H).^{13}C NMR [DMSO-d6] δc: 22.5, 30.2 (2C), 49.5, 50.6, 51.9 (2C), 52.5, 116.5, 117.9, 122.5 (2C), 124.7, 129.5 (2C), 130.2, 135.6, 136.2, 138.3, 150.2, 155.6, 178.8, 180.2. ESI-MS m/z: 407 (M+H)+. AnalCalcdfor $C_{23}H_{26}N_4O_3$: C, 67.96.; H, 6.45; N, 13.78; Found: C, 67.97.; H, 6.44; N, 13.77.

1-(4-chlorophenyl)-3-(1-(2-(2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)urea (10).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)indoline-2,3-dione (4a) (0.1 g, 0.365 mmol) and 4 chlorophenyl isocyanate (0.067 g, 0.439 mmol) to afford 10 (0.065 g, 39 %) as yellow solid. M.p: 120-122°C. 1H NMR [DMSO-d6]: δH 1.58 – 3.93 (m, 13H), 7.32 – 7.82 (m, 8H), 10.42 (b,1H) 10.51 (b, 1H).13C NMR [DMSO-d6] δc:30.2 (2C), 50.2, 50.5, 51.5 (2C), 52.1, 117.2, 118.5, 121.5 (2C), 123.2, 128.6 (2C), 130.2, 132.6, 135.6, 138.2, 147.2, 155.6, 159.5, 180.2, . ESI-MS m/z: 427 (M)+. AnalCalcdfor $C_{22}H_{23}ClN_4O_3$: C, 61.90; H, 5.43; N, 13.12; Found: C, 61.91; H, 5.42; N, 13.12.

General procedure for synthesis of substituted 1-(1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylthiourea (11-13).

A solution of 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-methylindoline-2,3-dione (4b) (1 eq) in dichloromethane (10 mL), was added triethyl amine (3 eq) at 0 ºC. Then corresponding substituted phenyl isothiocyanate was added to the reaction mixture and stirred at rt for 6 h (monitored by TLC & LCMS), quenched with ice and concentrated under reduced pressure. The residue was partitioned between ethyl acetate (20 mL) and water (20 mL). Aqueous layer was basified with sodium carbonate solution and extracted with ethyl acetate (3 × 20 mL) and washed with water (3 × 10 mL) and brine (3 × 10 mL). Organic layer was dried over anhydrous

sodium sulphate, filtered, and concentrated under reduced pressure to give the corresponding substituted 1-(1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylthiourea (11-13).

1-(1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylthiourea (**11**).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-methylindoline-2,3-dione (**4b**) (0.1 g, 0.348 mmol) and phenyl isothiocyanate (0.056 g, 0.418 mmol) to afford **11** (0.085 g, 57 %) as dark yellow solid. M.p: 127-129°C. ^{1}H NMR [DMSO-d6]: δH 1.63 – 3.88 (m, 16H), 4.65 (b,1H), 6.95 – 7.85 (m, 8H), 10.58 (b, 1H).^{13}C NMR [DMSO-d_6] δ_c: 21.5, 30.2 (2C), 50.5, 52.5 (2C), 54.2, 56.5, 116.8, 117.9, 124.9, 125.6, 127.8 (2C), 129.5, 130.2 (2C), 136.8, 139.5, 146.2, 159.5, 178.5, 180.2. ESI-MS *m/z*: 423 $(M+H)^+$. AnalCalcdfor $C_{23}H_{26}N_4O_2S$: C, 65.38; H, 6.20; N, 13.26; Found: C, 65.37; H, 6.21; N, 13.25.

*1-(1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-(p-tolyl)thiourea (***12**).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-methylindoline-2,3-dione (4b) (0.1 g, 0.348 mmol) and para tolyl isothiocyanate (0.062 g, 0.418 mmol) to afford 12 (0.091 g, 60 %) as yellow solid. M.p: 135-137°C. ^{1}H NMR [DMSO-d6]: δH 1.67 – 4.11 (m, 19H), 4.35 (b,1H), 6.52 – 7.75 (m, 7H), 10.28 (b, 1H).^{13}C NMR [DMSO-d6] δc: 21.5 (2C), 30.2 (2C), 50.5, 52.1 (2C), 53.7, 45.5, 116.3, 117.5, 125.2, 126.5, 127.2 (2C), 130.2 (2C), 137.2, 137.9, 138.2, 146.2, 159.5, 177.9, 180.2. ESI-MS m/z: 437 (M+H)+. AnalCalcdfor $C_{24}H_{28}N_4O_2S$: C, 66.03; H, 6.46; N, 12.83; Found: C, 66.02; H, 6.47; N, 12.82.

1-(4-chlorophenyl)-3-(1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)thiourea (13).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-methylindoline-2,3-dione (4b) (0.1 g, 0.348 mmol) and 4 chloro phenyl isothiocyanate (0.070 g, 0.418 mmol) to afford 11 (0.058 g, 36 %) as orange solid. M.p: 141-1143 °C. ^{1}H NMR [DMSO-d6]: δH 1.58 – 3.95 (m, 16H), 4.62 (b,1H), 6.45 – 8.02 (m, 7H), 10.57 (b, 1H).13C NMR [DMSO-d6] δc: 21.3, 30.2 (2C), 50.5, 52.5 (2C), 53.2, 54.2, 116.9, 117.8, 125.2, 125.6, 130.5 (2C), 131.5 (2C), 132.5, 136.5, 138.2, 146.2, 159.5, 178.2, 180.2. ESI-MS m/z: 456 (M)+. AnalCalcdfor $C_{23}H_{25}ClN_4O_2S$: C, 60.45; H, 5.51; N, 12.26; Found: C, 60.44; H, 5.50; N, 12.27.

General procedure for synthesis of substituted 1-(1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylurea *(14-16).*

A solution of 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-methylindoline-2,3-dione (4b) (1 eq) in dichloromethane (10 mL), was added triethyl amine (3 eq) at 0 ºC. Then corresponding substituted phenyl isothiocyanate was added to the reaction mixture and stirred at rt for 6 h (monitored by TLC & LCMS), quenched with ice and concentrated under reduced pressure. The residue was partitioned between ethyl acetate (20 mL) and water (20 mL). Aqueous layer was basified with sodium carbonate solution and extracted with ethyl acetate (3 × 20 mL) and washed with water (3 × 10 mL) and brine (3 × 10 mL). Organic layer was dried over anhydrous sodium sulphate, filtered, and concentrated under reduced pressure to give the corresponding substituted 1-(1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylurea (14-16).

1-(1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylurea (**14**).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-methylindoline-2,3-dione (**4b**) (0.1 g, 0.348 mmol) and phenyl isocyanate (0.049g, 0.418 mmol) to afford **14** (0.088 g, 62 %) as yellow solid. M.p: 136-138

°C. ^{1}H NMR [DMSO-d6]: δH 1.61 – 4.05 (m, 16H), 7.51 – 8.21 (m, 8H), 10.51 (b,1H) 10.65 (b, 1H).^{13}C NMR [DMSO-d_6] δ_c: 22.5, 30.2 (2C), 48.5, 50.1, 51.5 (2C), 52.1, 117.3, 118.2, 122.5 (2C), 124.5, 125.3, 128.1, 129.5 (2C), 140.2, 138.5, 148.2, 155.6, 159.5, 180.5. ESI-MS *m/z*:407 $(M+H)^+$. AnalCalcdfor $C_{23}H_{26}N_4O_3$: C, 67.96; H, 6.45; N, 13.78; Found: C, 67.98; H, 6.44; N, 13.77.

*1-(1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-(p-tolyl)urea (***15**).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-methylindoline-2,3-dione (4b) (0.1 g, 0.348 mmol) and para tolyl isocyanate (0.055 g, 0.418 mmol) to afford 15 (0.078 g, 53 %) as brown solid. M.p: 139-141°C. ^{1}H NMR [DMSO-d6]: δH 1.59 – 3.85 (m, 19H), 7.35 – 8.15 (m, 7H), 10.56 (b,1H) 10.69 (b, 1H).^{13}C NMR [DMSO-d6] δc: 21.2 (2C), 30.2 (2C), 48.5, 49.8, 52.5 (2C), 54.5, 116.9, 117.5, 120.5 (2C), 124.9, 125.3, 130.2 (2C), 136.9, 137.5, 138.5, 146.2, 155.2, 160.5, 180.2. ESI-MS m/z: 421 (M+H)+. AnalCalcdfor $C_{24}H_{28}N_4O_3$: C, 68.55; H, 6.71; N, 18.57; Found: C, 68.54; H, 6.70; N, 18.56.

1-(4-chlorophenyl)-3-(1-(2-(5-methyl-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)urea (16).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-methylindoline-2,3-dione (4b) (0.1 g, 0.348 mmol) and 4 chloro phenyl isothiocyanate (0.063 g, 0.418 mmol) to afford 16 (0.051 g, 33 %) as yellow orange solid. M.p: 145-146 °C. ^{1}H NMR [DMSO-d6]: δH 1.65 – 4.15 (m, 16H), 7.52 – 8.32 (m, 7H), 10.62 (b,1H) 10.71 (b, 1H).^{13}C NMR [DMSO-d6] δc: 22.5 30.2 (2C), 48.1, 49.5, 51.5 (2C), 52.1, 116.8, 117.1, 121.5 (2C), 124.9, 125.6, 130.3 (2C), 132.6, 137.5, 138.2, 144.2, 155.6, 159.2, 180.5. ESI-MS m/z: 440 (M)+. AnalCalcdfor $C_{23}H_{25}ClN_4O_3$: C, 62.65; H, 5.71; N, 12.71; Found: C, 62.66; H, 5.70; N, 12.72.

***General procedure for synthesis of substituted 1-(1-(2-(5-chloro-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylthiourea** (17-19).*

A solution of 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-chloroindoline-2,3-dione (4c) (1 eq) in dichloromethane (10 mL), was added triethyl amine (3 eq) at 0 ºC. Then corresponding substituted phenyl isothiocyanate was added to the reaction mixture and stirred at rt for 6 h (monitored by TLC & LCMS), quenched with ice and concentrated under reduced pressure. The residue was partitioned between ethyl acetate (20 mL) and water (20 mL). Aqueous layer was basified with sodium carbonate solution and extracted with ethyl acetate (3 × 20 mL) and washed with water (3 × 10 mL) and brine (3 × 10 mL). Organic layer was dried over anhydrous sodium sulphate, filtered, and concentrated under reduced pressure to give the corresponding substituted 1-(1-(2-(5-chloro-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylthiourea (17-19).

1-(1-(2-(5-chloro-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylthiourea (**17**).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-chloroindoline-2,3-dione (4c) (0.1 g, 0.325 mmol) and phenyl isothiocyanate (0.052 g, 0.39 mmol) to afford **17** (0.08 g, 55 %) as yellow solid. M.p: 110-112°C. ^{1}H NMR [DMSO-d6]: δH 1.63 – 3.95 (m, 13H), 4.82 (b,1H), 7.05 – 8.32 (m, 8H), 10.56 (b, 1H).^{13}C NMR [DMSO-d_6] δ_c: 30.2 (2C), 50.5 , 51.5 (2C), 54.5, 56.5, 120.2, 123.2, 124.8 (2C), 125.7, 129.5, 130.2 (2C), 135.6, 138.2, 139.5, 145.6, 161.2, 178.5, 180.5. ESI-MS *m/z*: 443 $(M+H)^+$. AnalCalcdfor $C_{22}H_{23}ClN_4O_2S$: C, 59.65; H, 5.23; N, 18.57; Found: C, 59.63;

1-(1-(2-(5-chloro-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-(4-chlorophenyl)thiourea (19).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-chloroindoline-2,3-dione (4c) (0.1 g, 0.325 mmol) and 4 chloro phenyl isothiocyanate (0.066 g, 0.39 mmol) to afford 11 (0.063 g, 40 %) as yellow solid. M.p: 156-158°C. ^{1}H NMR [DMSO-d6]: δH 1.57 – 3.85 (m, 13H), 4.67 (b,1H), 6.65 – 8.01 (m, 7H), 10.56 (b, 1H).^{13}C NMR [DMSO-d6] δc: 30.2 (2C) 49.2, 50.5 (2C), 52.1, 55.2, 120.5, 124.5, 125.5, 129.8 (2C), 130.5 (2C), 134.2, 135.6, 137.5, 139.5, 145.3, 160.2, 177.8, 180.5. ESI-MS m/z: 477 (M)+. AnalCalcdfor $C_{22}H_{22}Cl_2N_4O_2S$: C, 55.35; H, 4.64; N, 11.74; Found: C, 55.36; H, 4.65; N, 11.72.

General procedure for synthesis of substituted 1-(1-(2-(5-chloro-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylurea *(20-22).*

A solution of 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-chloroindoline-2,3-dione (4c) (1 eq) in dichloromethane (10 mL), was added triethyl amine (3 eq) at 0 ºC. Then corresponding substituted phenyl isocyanate was added to the reaction mixture and stirred at rt for 6 h (monitored by TLC & LCMS), quenched with ice and concentrated under reduced pressure. The residue was partitioned between ethyl acetate (20 mL) and water (20 mL). Aqueous layer was basified with sodium carbonate solution and extracted with ethyl acetate (3 × 20 mL) and washed with water (3 × 10 mL) and brine (3 × 10 mL). Organic layer was dried over anhydrous sodium sulphate, filtered, and concentrated under reduced pressure to give the corresponding substituted 1-(1-(2-(5-chloro-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylurea (20-22).

1-(1-(2-(5-chloro-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-phenylurea (**20**).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-chloroindoline-2,3-dione (4c) (0.1 g, 0.325 mmol) and phenyl isocyanate (0.046 g, 0.39 mmol) to afford **20**(0.75 g, 54 %) as yellow solid. M.p: 133-135°C.

^{1H} NMR [DMSO-d6]: δH 1.58 – 3.95 (m, 13H), 7.56 – 8.25 (m, 8H), 10.59 (b,1H) 10.67 (b, 1H).^{13}C NMR [DMSO-d_6] δ_c: 30.2 (2C), 48.6, 49.5, 50.8 (2C), 52.5, 120.2, 122.5 (2C), 125.8, 127.5, 129.5, 130.2 (2C), 135.6, 139.5, 140.2, 145.2, 155.6, 160.1, 180.2. ESI-MS *m/z*: 426 $(M)^+$. AnalCalcdfor $C_{22}H_{23}ClN_4O_3$: C, 61.90; H, 5.43; N, 13.12; Found C, 61.91; H, 5.42; N, 13.13.

*1-(1-(2-(5-chloro-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-(p-tolyl)urea (***21**).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-chloroindoline-2,3-dione (4c) (0.1 g, 0.325 mmol) and para tolyl isocyanate (0.051 g, 0.39 mmol) to afford 21 (0.65 g, 45 %) as orange solid. M.p: 145-147°C. 1H NMR [DMSO-d6]: δH 1.62 – 4.07 (m, 16H), 7.18 – 8.12 (m, 7H), 10.49 (b,1H) 10.65 (b, 1H).13C NMR [DMSO-d6] δc: 21.5, 30.2 (2C), 49.5, 50.1, 50.5 (2C), 54.2, 120.5, 122.5 (2C), 125.6, 127.5, 130.2 (2C), 135.3, 136.5, 138.5, 140.2, 145.2, 155.6, 160.5, 180.5. ESI-MS m/z: 440 (M)+. AnalCalcdfor $C_{23}H_{25}ClN_4O_3$: C, 62.65; H, 5.71; N, 12.71; Found: C, 62.64; H, 5.72; N, 12.70.

1-(1-(2-(5-chloro-2,3-dioxoindolin-1-yl)ethyl)piperidin-4-yl)-3-(4-chlorophenyl)urea (22).

The compound was synthesized according to the general procedure using 1-(2-(4-aminopiperidin-1-yl)ethyl)-5-chloroindoline-2,3-dione (4c) (0.1 g, 0.325 mmol) and 4 chloro phenyl isocyanate (0.059 g, 0.39 mmol) to afford 22 (0.065 g, 39 %) as yellow solid. M.p: 156-158°C. 1H NMR [DMSO-d6]: δH 1.65 – 3.95(m, 13H), 7.62 – 8.29 (m, 7H), 10.56 (b,1H) 10.65 (b, 1H).13C NMR [DMSO-d6] δc: 30.2 (2C), 49.5, 50.5, 51.8 (2C), 52.1, 120.5, 121.6 (2C), 124.5, 125.5, 130.2 (2C), 134.5, 135.6, 138.6, 139.5, 145.7, 155.6, 160.2, 180.5. ESI-MS m/z: 461 (M)+. AnalCalcdfor $C_{22}H_{22}Cl_2N_4O_3$: C; 57.28; H, 4.81; N, 12.14; Found: C;57.29; H, 4.80; N, 12.15.

APPENDIX

LIST OF PATENTS

Yogeeswari, P., Sriram, D., **Hasitha, S.A.,** Brahmam, M., **Indian Patent Application No. 1763/DEL/2015 (12/06/2015)**: Cathepsin D inhibitors and compositions thereof for treating breast cancer.

LIST OF PUBLICATIONS

From Thesis Work:

1. **Anantaraju HS,** Battu MB, Viswanadha S, Sriram D, Yogeeswari P. Cathepsin D inhibitors as potential therapeutics for breast cancer treatment: Molecular docking and bioevaluation against triple-negative and triple-positive breast cancers. **Molecular diversity**, *Springer*. 2015 Nov 13:1-5.
2. **Anantaraju HS** [1], Bobesh k Andrews[2], Srikant Viswanadha [3], Shubham Dwivedi[1], Swapna Yellanki[4], Pushkar Kulkarni[4], Sonal saxena[6], Komu Naga Mohan[6], Dharmarajan Sriram[2, 5], Perumal Yogeeswari[1,"] Molecular docking, Synthesis and Bioevaluation of indoline-2, 3-diones **as** novel small molecule Cathepsin D inhibitors. **Chemical Biology and Drug Design** (2016)

Other Publications

1. Agarwal DS, **Anantaraju HS**, Sriram D, Yogeeswari P, Nanjegowda SH, Mallu P, Sakhuja R. Synthesis, characterization and biological evaluation of bile acid-aromatic/heteroaromatic amides linked via amino acids as anti-cancer agents. Steroids. 2015 Dec 31.
2. Reddy ER, Trivedi R, Sarma AV, Sridhar B, **Anantaraju HS**, Sriram D, Yogeeswari P, Nagesh N. Sugar-boronate ester scaffold tethered pyridyl-imine palladium (ii) complexes: synthesis and their in vitro anticancer evaluation. Dalton Transactions. 2015;44(40):17600-16.
3. Devi PB, Sridevi JP, Kakan SS, Saxena S, Jeankumar VU, Soni V, **Anantaraju HS**, Yogeeswari P, Sriram D. Discovery of novel lysine ε-aminotransferase inhibitors: An intriguing potential target for latent tuberculosis. Tuberculosis. 2015 Dec 31;95(6):786-94.
4. Sridevi JP, Suryadevara P, Janupally R, Sridhar J, Soni V, **Anantaraju HS**, Yogeeswari P, Sriram D. Identification of potential Mycobacterium tuberculosis topoisomerase I inhibitors: A study against active, dormant and resistant tuberculosis. European Journal of Pharmaceutical Sciences. 2015 May 25;72:81-92.
5. Polkam N, Rayam P, Anireddy JS, Yennam S, **Anantaraju HS**, Dharmarajan S, Perumal Y, Kotapalli SS, Ummanni R, Balasubramanian S. Synthesis, in vitro anticancer and antimycobacterial evaluation of new 5-(2, 5-dimethoxyphenyl)-1, 3, 4-thiadiazole-2-amino derivatives. Bioorganic & medicinal chemistry letters. 2015 Apr 1;25(7):1398-402.
6. Sridevi JP, **Anantaraju HS**, Kulkarni P, Yogeeswari P, Sriram D. Optimization and validation of Mycobacterium marinum-induced adult zebrafish model for evaluation of oral anti-tuberculosis drugs. International journal of mycobacteriology. 2014 Dec 31;3(4):259-67.

7. Mallika Alvala, Ravi Alvala, **Anantaraju HS**, Pulla Venkat Koushik, Madhu Babu Battu, Variam Ullas Jeankumar, Perumal Yogeeswari, Dharmarajan Sriram. Novel SIRT1 inhibitors-mediated apoptosis in prostate cancer: An in-vitro and in-vivo evaluation. Life Sciences. July 2015 (In Press).

BIOGRAPHY OF HASITHA SHILPA ANANTARAJU

Ms. Hasitha Shilpa Anantaraju completed her Bachelor of Pharmacy from Andhra University, Vishakapatnam, Andhra Pradesh, India and Master of Science (M.S. (Pharm.)) in Pharmacology and Toxicology from National Institute of Pharmaceutical Education and Research (NIPER), Hyderabad, India. She has been awarded INSPIRE fellowship by Department of Science and Technology (DST), India for a period of five years (2013-2018) under the supervision of Prof. P. Yogeeswari. She has published 8 scientific publications in well-renowned international journals and filed an Indian patent.

BIOGRAPHY OF PROF.P.YOGEESWARI

Prof. P. Yogeeswari is presently working in the capacity of Professor and Associate Dean (Sponsored Research and Consultancy Division), Department of Pharmacy, Birla Institute of Technology and Science, Pilani, Hyderabad Campus. She received her Ph.D. degree in the year 2001 from Banaras Hindu University; Varanasi. She has been involved in research for the last 14 years and in teaching for 13 years. APTI honoured her with YOUNG PHARMACY TEACHER AWARD for the year 2007. In 2010, ICMR honored her by awarding "Shakuntala Amir Chand Award" for her excellent biomedical research. She has been awarded for IASP 2014 "Excellence in Pain Research and Management in Developing Countries" under the basic science research category at Argentina in October 2014. She has collaborations with various national and international organizations that include National Institute of Health, Bethesda, USA, National Institute of Mental Health and Neurosciences, Bangalore, Karolinska Institute, Stockholm, Sweden, National Institute of Immunology, New Delhi, India, Pasteur Institute, University of Lille, France, Bogomoletz Institute of Physiology National Academy of Science, Ukraine, and Faculty of Medicine of Porto, Porto, Portugal,. She has to her credit more than 190 research publications and one Indian Patent, Application No: 1138/CHE/2009. She is an expert reviewer of many international journals like Journal of Medicinal Chemistry (ACS), Journal of Chemical Information & Modeling (ACS, USA), Bioorganic Medicinal Chemistry (Elsevier), Recent Patents on CNS Drug Discovery (Bentham), etc. She has also co-authored a textbook on organic medicinal chemistry with Dr. D Sriram titled "Medicinal Chemistry" published by Pearson Education and one book chapter in in Jan 2013 by IGI Global. She is a lifetime member of Association of Pharmacy Teachers of India and Indian Pharmacological Society. She has successfully completed many sponsored projects and currently handling projects sponsored by DST, DBT INDO-BRAZIL, ICMR-INSERM, and CSIR. She has guided sixteen Ph.D students and currently five students are pursuing their Ph.D. work.

BIOGRAPHY OF Dr. SRIKANT VISWANADHA

Dr. Srikant Viswanadha is presently working as Vice president, Drug Discovery, Incozen Therapeutics Pvt. Ltd. Hyderabad. He received his Ph.D. degree in nutritional biochemistry in the year 2003 from Virginia Polytechnic Institute and State University; Blacksburg, VA, USA. He has a post-doctoral training in National Institute of Diabetes, Digestive, and Kidney Diseases (NIDDK), National Institutes of Health (NIH), USA. He received a NIDDK travel award in 2006 for 'Research Excellence' and John Lee Pratt Fellowship in 2000 at Virginia Tech. He has been involved in research for the last 12 years. He has 10 research papers, 1 book chapter, 11 international poster presentations and 6 international patents to his credit.

Cathepsin D, an Aspartic Protease as a Potent Therapeutic Option for Breast Cancer: Design, Synthesis and Bioevaluation against Triple-Positive and Triple-Negative Breast Cancers

THESIS

Submitted in partial fulfillment

Of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

ANANTARAJU HASITHASHILPA

ID No 2013PHXF010H

Under the Supervision of

Prof. P. YOGEESWARI

BITS Pilani Pilani | Dubai | Goa | Hyderabad

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

2014

Chapter 6. Recapitulation and conclusion

The current research in investigating biomarkers for early detection, prognosis and the prediction of treatment responses in breast cancer is rapidly expanding. On the other hand, no validated biomarker currently exists for use in routine clinical practice, and breast cancer detection and management remains dependent on invasive procedures. The efforts to find a cure for breast cancer are ongoing. Researchers develop and test new drugs and other treatments all the time. The focus on monoclonal antibodies, anti-angiogenesis drugs and immunotherapy has always been a demand in the lab. The promising example is Trastuzumab (Herceptin), a FDA- approved monoclonal antibody. Targeted therapeutic strategies had a mixed response in terms of successes and failures.

Our current study is to identify non-peptidic small molecule inhibitors against Cathepsin D, an aspartic protease which was found to over express in metastatic breast tumors. CatD came into lime light as an important prognostic biomarker found in the cytosol of breast cancer patients during the clinical trials. Furthermore, CatD was found to play a vital role in ECM degradation of tumor cells resulting in breast cancer invasion and metastasis. We performed a thorough literature search and found there is an every need to discover CatD inhibitors to treat metastatic breast cancer.

In summary,

- Using molecular modeling methods, in a high throughput screening against human CatD (PDB: 1LYB) and from Asinex and BITS databases, we successfully identified a novel scaffold of inhibitors of CatD.
- CatD enzyme activity and anti-cancer activity of the lead molecules was evaluated, both against the enzyme and in cell based assays.

- Based on the substructure analysis of the lead identified, we synthesized and characterized 17 compounds as analogues of lead molecule M29 to improve the potency.
- Of the molecules synthesized, compound CIS-1 was found to be the most active in cellular assays. Percentage inhibition of 76% using CatD enzyme assay for compound CIS- while the GI50 in MDA-MB-231 cells was 1.02±0.05μM.
- In addition to the anti-proliferative activity, compound CIS-1 inhibited ROS and increased lysosomal degradation detected using Acridine Orange in cancer cells.
- Cell cycle analysis indicated that compound CIS-1 caused a G0/G1 arrest in cells when treated at its GI_{50} concentration of 1.02 μM.
- Angiogenesis assays like wound healing assay, boyden chamber assay and tube formation assay also indicated the role of CIS-1 in preventing angiogenesis of cancer cells.
- An *invivo* zebrafish angiogenesis assay was further carried out to validate the anti-angiogenic effect of the most potent lead CIS-1.
- Prolonged treatment with CatD inhibitors resulted in downregulation of endothelial growth factors, matrix metallo proteases which has in turn effected the levels of AKT and mTOR.
- Combinations of CIS-1 with PARP-1 inhibitor Olaparib were found to be beneficial compared to that of single agents alone in the treatment of metastatic TNBC.

In conclusion, findings from this study have a strong molecular modeling backbone in optimizing and identifying CatD inhibitors and a validated logic in structurally modifying the identified leads to achieve more promising potent molecules. Our findings also suggested the mechanism through which the lead molecules were effective in causing CatD inhibition and the downstream effects involved in causing the inhibition. Our study also has a relevant *invivo*

evaluation and a combination of PARP-1 inhibitiors and CatD inhibitors which could provide a rationale treatment strategy in metastatic breast cancer patients.

Further studies in using *invivo* rodent tumor model mimicking human mammary cancer is warranted for obtaining insights into additional closer effect of identified leads. Our study illustrated the need in developing CatD inhibitors, the role of CatD inhibitors and the relationship between CatD and invasion factors and its combination effect with PARP-1 inhibitors. Thus, the use of CatD inhibitors could be a potential anti-tumor TNBC therapy. In addition, the CatD inhibitors would be important in other cancers such as colon, lung and prostate cancers.

Chapter 7. Future perspectives

Recent studies demonstrated the role of CatD in various cancers along with inflammatory disorders, Alzheimer's disease, lipid adipogenesis and many other neurological disorders. So, there is every scope to test the inhibitors identified in this study on all other ailments for their role in CatD inhibition. The inhibitors we achieved should be further tested in an *invivo* rodent tumor model setup that mimics the human metastatic breast cancer. The tumor model can be achieved by injecting 4T1 mouse mammary tumor cells into BALB/C strain of mice or by injecting MDA-MB-231 human breast cancer cells into nude mice. We procured 4T1 cells and analyzed the level of gene expressions that the lead CIS-1 down regulated during our study. Also, the correlation between PARP-1 inhibitors and CatD inhibitors can be extended for their combined effect. Apart from docking studies prediction of physicochemical properties, the lead CIS-1 can further be evaluated for its plasma and metabolism studies during the *invivo* study.

This study is a challenge and achievement in discovering small molecule non-peptidic CatD inhibitors and also to deduce the mechanism through which the lead molecules might act. This could serve as an important hope in scientists to combat cancer which is the "sword of hope".

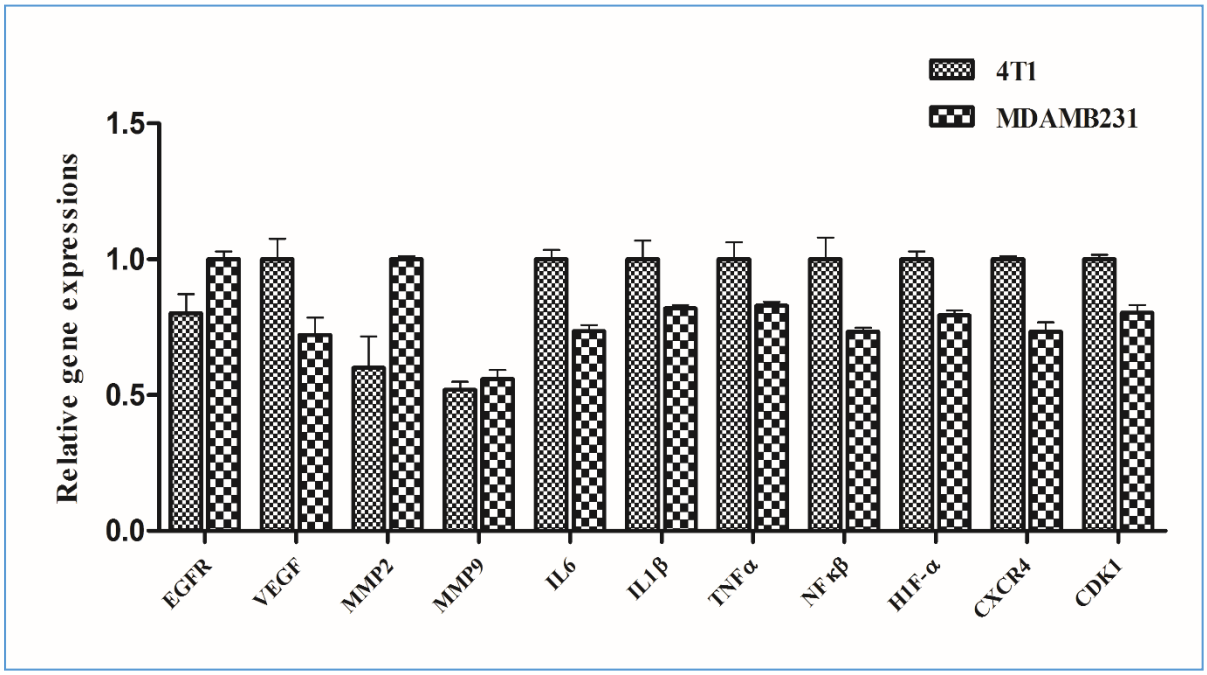


Figure 58: An RTPCR analysis which is a comparative study of levels of the listed gene expressions in 4T1 Vs MDA-MB-231 in order to develop *invivo* 4T1 mouse tumor model. 4T1: Mouse mammary tumor cells; MDA-MB-231: Human breast adenocarcinoma cells.