DISSERTATION

NOVEL *IN SILICO*-DESIGNED SMYD3 INHIBITORS ELIMINATE UNRESTRAINED PROLIFERATION OF BREAST CARCINOMA CELLS

Submitted by

Ilham Mohammed Alshiraihi

Graduate Degree Program in Cell and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2021

Doctoral Committee:

Advisor: Mark Brown
Co-Advisor: Takamitsu Kato

Melissa Reynolds
Christopher Snow

ABSTRACT

Novel *in silico*-designed SMYD3 inhibitors eliminate unrestrained proliferation of breast carcinoma cells

SMYD3 is a lysine methyltransferase that regulates the expression of over 80 genes and is required for the uncontrolled proliferation of most breast, colorectal, and hepatocellular carcinomas. Elimination of SMYD3 restores normal expression patterns of these genes and halts aberrant cell proliferation. In this study, we used *in silico* screening to identify potential small molecule inhibitors of SMYD3 and tested the ability of these inhibitors to reduce its methyltransferase activity *in vitro*. Using breast cancer cell lines that overexpress SMYD3 and normal breast epithelial cell lines, we have confirmed the ability of one of these inhibitors, Inhibitor-4, to reduce cell proliferation, arrest the cell cycle, and induce apoptosis in breast cancer cells without affecting normal cell behavior. Our results provide a proof of concept for the *in silico* design of small molecule enzyme inhibitors and for the use of such an inhibitor to target SMYD3 for the treatment of cancer.

ACKNOWLEDGEMENTS

I would like to thank the many people who have helped me along the way in uncountable ways. First and foremost, I would like to thank my incredibly supportive husband, Ali Alturki, my kids, Bassam, Hossam and Serena, for putting up with me, and always encouraging me and believing in me. Also, I would like to thank my family, dad, mom, brothers, and sisters for making me the person I am and pushing me to pursue my dreams.

I would also like to thank my coworkers including Dillon Jarrell, Melissa Edwards, Claudia Wiese, Taru Dutt, Hend Ibrahim, Nora Nealon and Zeyad Arhouma for their assistance.

Finally, I would like to sincerely thank my graduate committee, including Dr. Melissa Reynolds and Dr. Christopher Snow for feedback and support. Also, I would like to thank my co-advisor, Dr. Takamitsu Kato, whose support, and advice along the way has been irreplaceable. Additionally, I would especially like to thank my advisor, Dr. Mark Brown for his support and guidance as well as the trust and freedom he has granted me.

PREFACE

SMYD (SET and MYND domain-containing) protein family members constitute a class of methyltransferases that regulate a wide range of normal cellular processes [1-3] and are also involved in several tumorigenic pathways [3-5]. SMYD3, the third member of the SMYD family, transfers methyl groups to lysine 4 on histone H3 (H3K4) and lysine 5 on histone 4 (H4K5), a residue that was previously thought to only undergo acetylation [1-3]. Overexpression of SMYD3 results in increased cell proliferation and activates many genes associated with cancer cell transformation [6] and metastasis [7]. Several studies have revealed that lung, breast, pancreatic, colorectal, and hepatocellular carcinoma are highly associated with SMYD3 overexpression [3, 8, 9]. *In vitro* studies using NIH3T3 cells (transfected with SMYD3) have demonstrated that SMYD3 involvement in uncontrolled proliferation is one of the crucial stages in tumorigenesis [10]. Furthermore, the growth of breast, hepatocellular, and colorectal carcinoma cell lines have been impaired significantly through SMYD3 knockdown [3, 8]. These studies demonstrate that the oncogenic impact of SMYD3 is mediated, in part, by its histone methylation activity and the resulting impact on the expression of oncogenes. These downstream genes include NKX2.8 [3], WNT10B [8], TERT [11], cMET [12], and CDK2 [13]. SMYD3 is also known to regulate cancer cell proliferation and viability through its interaction with vascular endothelial growth factor receptor 1 (VEGFR1) [14] and estrogen receptor (ER) [15], both of which are non-histone proteins. The role of SMYD3 in ER-mediated transcription through its histone methyltransferase activity is not fully known. SMYD3 acts as a cofactor of ERα and promotes its efficacy in response to

bound ligands. In addition, SMYD3 interacts with ER in the ligand binding domain and activates the transcriptional machinery of downstream genes [15]. Collectively, these studies indicate SMYD3 as a potential therapeutic target for cancer treatment.

The small molecule drug development process is notoriously expensive, time consuming, and inefficient. After target identification, identifying hit compounds with significant activity requires enormous small-molecule libraries and hours of experimentation. Optimizing hit compounds to identify leads that have activity in cells and that meet initial *in vitro* toxicity criteria often requires several rounds of iteration and molecular synthesis. Because of the difficulty and cost of this process, scientific literature is saturated with studies that identify proteins that are “promising therapeutic targets” but that proceed no further towards actual therapeutic development. In this study, we sought to establish a simple and affordable hit-to-lead methodology that could be implemented by average research laboratories that have elucidated druggable proteins. Using *in silico* screening to identify initial hits and restricting the initial library to purchasable compounds, we have demonstrated the ability to identify lead small molecule inhibitors without the need for physical compound libraries or in-house chemical synthesis.

Using SMYD3 as a target protein, we implemented our screening methodology and report a novel small molecule SMYD3 inhibitor (Inhibitor-4) that impairs breast cancer cell proliferation without affecting normal cells, thereby illustrating the potential of SMYD3 inhibitors in the clinical management of breast cancer as well as a proof of concept for this drug development platform. We used two breast cancer cell lines (MCF7 and MDA-MB-231) that were previously shown to overexpress SMYD3 [8, 15-

18] compared with the wild type breast epithelial cell line MCF10A [16, 17]. After initial hit identification *in vitro*, we purchased five small molecules, tested their efficacy as novel SMYD3 inhibitors, and discovered that Inhibitor-4 significantly reduces breast cancer proliferation, arrests the cell cycle, and induces apoptosis without impacting normal cells. In all experiments, we used a previously-identified SMYD3 small molecule inhibitor, BCI-121, as a positive control [19].

DEDICATION

This document is dedicated to my parents, husband Ali and kids Bassam, Hossam and Serena who lit up my life in the darkest moments and to my advisor who always believed in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1– INTRODUCTION

1.1 Breast cancer:

1.1.1. Breast structure:

The breast comprises two major types of tissues, including glandular and stromal (supporting) tissues. Glandular tissues contain the ducts and the milk-producing glands (lobules), while the stromal tissues contain fibrous and connective tissues of the breast. The breast also includes lymphatic tissues that remove cellular waste and fluids (Figure 1.1) [20].

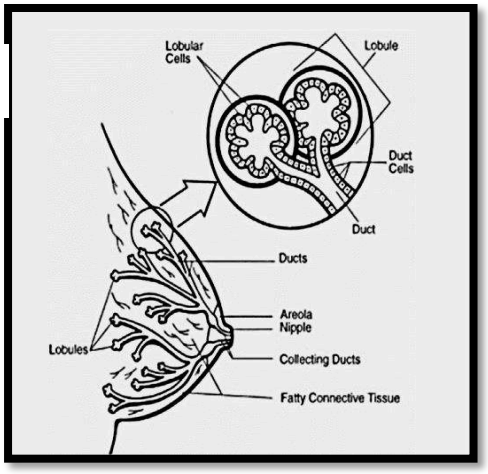


Figure 1.1: Human breast anatomy showing the breast ducts and lobules (milk-producing glands) [20].

Breast cells that line the ducts are the most common location for breast-related tumorigenesis (ductal cancers). Minor numbers of breast cancers also begin in the other breast tissues (Figure 1.2) [20].

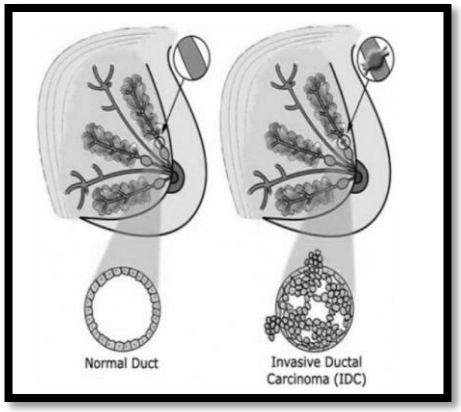


Figure 1.2: Human breast cancer showing cells that line the normal breast duct versus malignant invasive cells in ductal carcinoma [20].

1.1.2 Epidemiology of Breast cancer in humans:

Second to nonmelanoma skin cancer, breast cancer is the most common type of cancer in the world [21]. In 2017, more than 250,000 women were newly diagnosed with breast cancer and 12% of all women in the United States will be diagnosed with breast cancer during their lifetimes [22].

1.1.3. Breast cancer risk factors:

Many risk factors are associated with breast cancer, including age, genetics or family history, geographical variation, body weight, radiation exposure and hormone treatments (Table 1.1) [23].

Table 1.1: Breast cancer risk factors [23].

Factor	Relative risk	High risk group
Age	>10	Elderly
Geographical location	5	Developed country
Age at menarche	3	Menarche before age 11
Age at menopause	2	Menopause after age 54

Age at first full pregnancy	3	First child in early 40s
Family history	⩾2	Breast cancer in the first degree relative when young
Previous benign disease	4-5	Atypical hyperplasia
Cancer in another breast	>4	
Socioeconomic group	2	Groups I and II
Diet	1.5	High intake of saturated fat
Body weight:		
Premenopausal	0.7	Body mass index >35
Postmenopausal	2	Body mass index >35
Alcohol consumption	1.3	Excessive intake
Exposure to ionizing radiation	3	Abnormal exposure in young females after age 10
Taking exogenous hormones:		
Oral contraceptives	1.24	Current use
Hormone replacement therapy	1.35	Use for ⩾10 years
Diethylstilbestrol	2	Use during pregnancy

1.1.4. Breast cancer and family history:

In western countries, up to 10% of women's breast cancer cases are due to genetic susceptibility [24]. Breast cancer predisposition is usually inherited as an autosomal dominant inheritance pattern [25]. BRCA1 and BRCA2 are two well-documented genes highly associated with developing breast cancer in families with autosomal dominant inheritance [26]. BRCA1 is located on the long arm of chromosome 17, while BRCA2 is located on chromosome 13. Some mutations of these genes are represented with high

frequency in specific populations [23]. For instance, one in forty Ashkenazi Jewish women has a BRCA mutation. While BRCA1 is the dominant gene leading to breast cancer development in this population, BRCA2 is the dominant gene leading to breast cancer in Italian women [27]. Ashkenazi Jewish women carry either BRCA1 185 del AG (deletion of two base pairs occurs at position 185), BRCA1 5382 ins C (insertion of an extra-base pair at position 5382), or BRCA1 6174 del T (deletion of one base pair occurs at position 6174). However, about 50% of all familial breast cancer in Iceland involves BRCA2 999 DEL 5 (deletion of five base pairs in position 999). Also, breast cancer's high susceptibility occurs through rare inherited mutations in p53 and PTEN associated with familial syndromes (Li-Fraumeni and Cowden's, respectively). In addition, there are poorly understood gene associations that increase the likelihood of breast cancer by approximately three or four-fold above the overall population level. These less documented genes account for a significant portion of genetic disorders' overall contribution to breast cancer, though they are less likely to cause multi-case familial incidence [23].

1.1.5. Breast cancer types according to invasiveness:

Non-invasive breast cancer: Breast cancer cells in this type do not invade the fatty and connective tissues that are surrounding the primary tumor. The most common form of non-invasive breast cancer is ductal carcinoma in situ (DCIS) (90%). However, lobular carcinoma in situ is also considered a marker for increasing the risk of breast cancer, but it is less common [20].

Invasive Breast Cancer: Cancer cells in invasive breast cancer invade the surrounding connective and fatty tissues after breaking through the breast duct and lobular wall [20].

1.1.6. Frequency of Breast cancer:

- Lobular carcinoma *in situ* (Lobular neoplasia or LCIS) is a non-invasive accumulation of neoplastic cells in the breast lobules [20].
- Ductal carcinoma *in situ* (DCIS) is limited to the breast ducts and represents the most common type of non-invasive breast cancer [20].

1.1.7 Breast cancer and molecular subtypes according to (Mehrgou and Akouchekian, 2016) [28]:

Determining breast cancer's molecular subtypes is an effective tool to highlight and choose a suitable treatment plan and new therapies. Most research divides breast cancer into six main molecular subgroups (Figure 1.3) including:

Luminal A: In luminal A, luminal tumor cells initiate in the breast ducts' internal cellular lining. The tumor cells in this type tend to be positive for PR and ER (Progesterone and Estrogen receptors are activated by their hormones, and, thus, the tumor cells rapidly proliferate). On the other hand, this molecular subtype is negative for HER-2 (human epidermal growth factor receptor 2). HER-2 receptors manage normal breast cells' growth. However, aberrant expression of HER-2 leads to breast cancer. The tumor cells in this type are present in grades 1 and 2, resulting in a better best prognosis, including low frequency and high survival rate [28].

Luminal B: Tumor cells in this type are positive for PR and ER and positive for HER-2. The patients in this type have a poorer prognosis due to the larger tumor size, lower grade tumor, and lymph node involvement. This type can also be more readily detected at a young age relative to luminal A [28].

HER-2 type: This subgroup of tumors tends to be lower grade, negative for PR and ER, and includes progression into lymph nodes. ERH-2 type appears in 10% to 15% of breast cancer-positive women and includes poor prognosis, high-frequency rate, and common metastasis. HER-2 tumors appear in younger age groups than luminal A or B tumors [28].

Triple negative/basal-like: This type is referred to as triple-negative because the tumor cells are negative for the ER, PR, and HER-2 receptors. Basal-like is one of these categories, and the tumor cell tends to be like the outer cells (basal) that are lining the breast ducts. About 15% to 20% of breast cancer patients are triple-negative or basal-like. Young and African-American women are the most affected by this type of tumor. Triple-negative or Basal-like tumors have a poorer prognosis and more aggressive. As mentioned before, BRCA1 mutation causes a high susceptibility to breast cancer. Therefore, breast cancer women due to BRCA1 mutations are triple-negative and basal-like as well [28].

Claudin-low: This group of tumors is triple-negative. However, it differs from triple-negative by expressing cell-cell adhesion proteins (like E-cadherin that is reduced). Also, lymphocyte infiltration happens considerably. This type tends to have stem cells and mesenchymal cell features [28].

Normal-like: These tumors are usually small and have a high prognosis. Normal-like tumors appear in about 6% to 10% of all breast cancer cases [28].

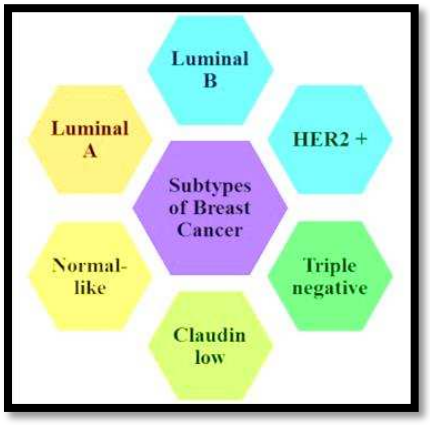


Figure 1.3: Schematic representation showing 6 molecular subtypes of breast cancer [28].

1.1.8 Stages of breast cancer [20].

Table 1.2: Staging of breast cancer.

Stage	Description
Stage0	*In situ* carcinoma (non-invasive): The tumor is limited and located in the milk-producing gland or ducts.
Stages I-IIIC (Invasive)	
I	The tumor has spread further to one to three lymph nodes near the breastbone or in the armpit, and the size is ¾ inches or less.
IIA	The tumor has spread further to one to three lymph nodes near the breastbone or in the armpit, and the size is ¾ inches or less.
IIB	The tumor has spread further to one to three lymph nodes near the breastbone or the armpit, and the size is larger than ¾ inches; however, in diameter, it is smaller than 2 inches.
IIIA	The tumor has spread further to up to nine lymph nodes in the armpit, and the size is 2 inches in diameter.
IIIB	The tumor has caused breast inflammation because it has spread to the chest wall or dermis.

IIIC	The tumor has grown to the lymph node's size and spread further to the lymph nodes in the armpit.
Metastatic cancer	
IV	Metastatic cancer: The tumor has spread to distant organs, tissues, or lymph nodes.

1.1.9. Management of breast cancer:

The following approaches are commonly considered in the clinical managed of breast cancer.

- **Surgery:** There are two main options available for breast cancer surgery [20].

  1. In breast-conserving surgery, the tumor is excised with a portion of the normal, surrounding tissue [20].

     - Lumpectomy: The breast tumor with a small portion of the surrounding tissue is excised [20].
     - Wide excision or partial mastectomy: A more significant amount of the normal surrounding tissue with the tumor is excised [20].
     - Quadrantectomy: In this type of surgery, approximately one-fourth of the breast is excised [20].

  2. Mastectomy: All the breast tissue is excised [20].

- **Radiation therapy:** Gamma rays or high-energy X-rays target a tumor or the tumor site post-surgery. Breast cancer treatment through electron beam radiotherapy to the breast scar can also be employed. In general, radiation therapy, either through

radiation rays or electron beam radiotherapy to the breast scar, is highly efficient in removing remaining cancer cells wither post-surgery or as a breast-conserving therapy [20].

- **Chemotherapy:** Cytotoxic drugs are used to control proliferation and spreading. It is one of the most effective therapeutic applications for breast cancer treatment and can be used in multiple rounds after surgery, before surgery, or as a breast-conserving therapy [20].

- **Nanotechnology:** Nanotechnology improves chemotherapeutic efficacy through applications for maintenance of anti-cancer drug activity and liposomal drug delivery. Liposomal doxorubicin, liposomal daunorubicin, and pegylated liposomal doxorubicin are efficient agents to treat breast cancer with less cardio toxicity and better safety profiles [20].

- **Gene therapy:** Accumulation of multiple molecular genetic defects result in cancer arises due to uncontrolled cell proliferation. Therefore, a diversity of gene therapy strategies has been used as a prospect for new cancer therapies. Tumor suppressor genes and proto-oncogene are responsible for the genesis of malignancy. Therefore, these genes have been used to guide gene therapy development toward restoring and ablating such genes [29].

- **Immunotherapy:** Immunotherapy is one of the more effective and less toxic therapies used for breast cancer and involves methods for inducing antibody binding to tumor-specific surface proteins on target cells [30]. Immune therapy is highly

effective in hematologic malignancy that requires bone marrow transplantations [30, 31].

- **Hormone therapy:** Tumors often express either estrogen receptors (ER) or progesterone receptors (PR), which can serve as targets for hormone therapy in cases where chemotherapy is not enough [32]. For breast cancer, there are two significant applications of hormone therapy that are typically employed for luminal A and/or B sub-types:

  - Selective estrogen receptor modulators: This treatment's strategy is to use drugs that bind to estrogen receptors to block estrogen-induced cell proliferation. One such drug is Tamoxifen, which can also be used as preventive therapy for women at high risk for developing breast cancer [33].

  - Aromatase inhibitors: In this type of hormone therapy, drugs are designed and used to inhibit the aromatase enzymes (the enzyme responsible for producing estrogen). Therefore, this enzyme is suppressed and blocks estrogen production, thus leading to reduced cell proliferation [33].

1.2. In silico screening:

In silico pharmacology is a rapidly expanding field of computational therapeutics or computational pharmacology. It also covers the worldwide development of various sources and techniques that use software to analyze and merge medical and biological data. *In silico* pharmacology uses input in simulation or in the building of computational models that can be applied to create prognosis, propose hypotheses, and eventually produce discoveries or developments in therapeutics and medicine [34].

1.2.1. Ligand-based virtual screening:

Ligand-based virtual screening is one of the various methods of *in silico* screening. The type of structural input used is essential to determine the compounds' complexity, degree, and final cost [35]. The central similarity-property is the main principle for these methods based on the similarity of molecules that should lead to the similarity of properties [36]. Therefore, chemical similarity calculations are the essence of ligand-based virtual screening [37]. Consistently, in a database all the compounds can be scored based on the similarity to one or more bioactive ligands. Accordingly, these methods offer significant productivity through the random selection of compounds in the database. The top-scoring compounds can be selected for experimental testing and rely on a cost-effective drug discovery method (Figure 1.4) [34].

1.2.2. Target-based virtual screening methods:

Target-based virtual screening is a technique used to identify and generate novel bioactive compounds. The availability of structural information of the target that can be determined either experimentally or obtained computationally through homology modeling techniques is the main principle for target-based methods [38, 39]. Therefore, this approach is based upon searching for compounds that work on a known target biomolecule [40]. These methods provide a good prediction of the ligand-binding affinity to the protein (scoring) and the ligand's orientation and conformation in the binding pocket (docking). Consequently, novel inhibitors have been discovered for multiple enzyme targets - such as epidermal growth factor receptor kinase, vascular epidermal growth factor receptor two kinases, and cyclin-dependent kinase, using target-based virtual screening methods (Figure 1.4) [41].

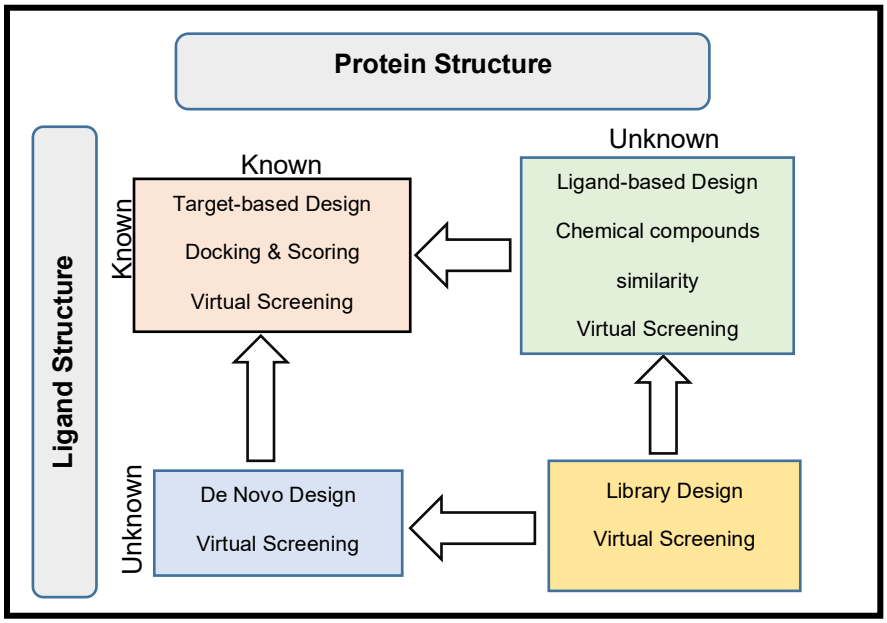


Figure 1.4: Schematic representation showing drug designing strategies. Determination of drug discovery strategies depend on either availability or unavailability of protein and/or ligand structure.

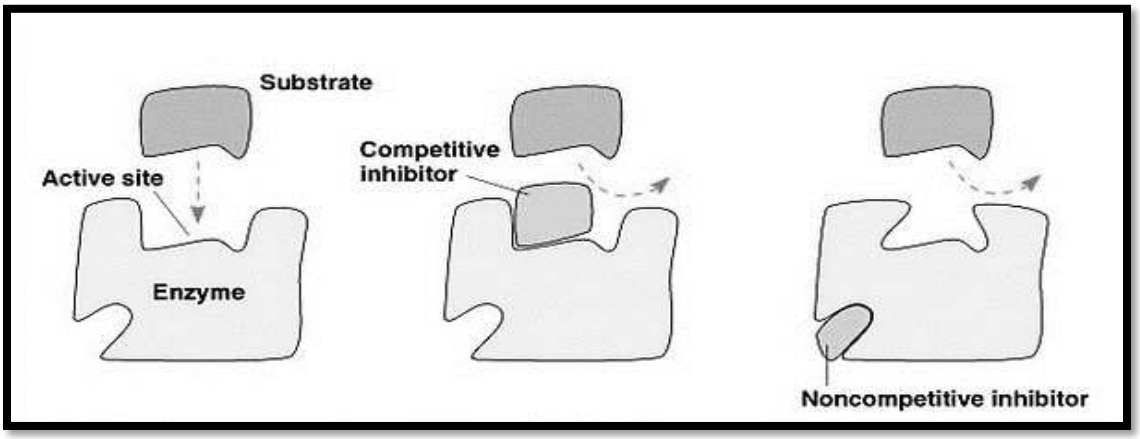
1.3. Enzyme inhibition:

In one form of enzyme inhibition, enzyme-substrate complex reactions are prevented through chemical compound binding covalently or noncovalently to the enzyme active site. Consequently, these chemical compounds inhibit or reduce the enzyme catalytic activity [42].

1.3.1. Types of enzymatic inhibition:

There are two types of enzyme inhibitors, reversible inhibitors and irreversible inhibitors. Reversible inhibitors are inhibitors that act reversibly with the enzyme. However, irreversible inhibitors are inhibitors that inactivate the enzyme irreversibly through covalent modification of the enzyme's essential residue [42].

In drug design, there are two subtypes of reversible inhibitors, competitive and non-competitive. The competitive inhibitors are designed to be the same as the substrate structure and bind reversibly to the enzyme's active site. Therefore, the competitive inhibitor competes with the substrate to bind to the active site of the enzyme. On the other hand, a non-competitive inhibitor is not related to the structural similarity of the substrate. The non-competitive inhibitor binds to the free enzyme or enzyme-substrate complex. It causes conformation change in the active site, resulting in the enzyme catalysis's inactivation (Figure 1.5) [42, 43].



a) Substrate binds to active site of enzyme normally.

b) Competitive inhibitor competes substrate for active

c) Noncompetitive inhibitor changes enzyme conformation and disrupt active site.

Figure 1.5: Mechanisms of reversable enzymatic inhibition. a) Normal binding. b) Competitive inhibition. c) Noncompetitive inhibition [43].

1.4. Chromatin structure and gene regulation:

In organisms, chromatin works as a flexible mechanism for regulating an organism's genome storage, starting from regulated gene expression processes to mitotic processes [44]. A whole chromosome is compacted sequentially through a sequence of highly ordered coiling in this mechanism, while regions of DNA are selectively made

more accessible to transcriptional machinery [45, 46]. A dynamic architecture in chromatin is exceptionally efficient by which approximately 2 meters of DNA in the nucleus are compressed while conserving functionality [47, 48]. In this regard, a repeated sequence of nucleosomes is the basic structural architecture of chromatin (Figure 1.6) [49]. The nucleosome core particle (NCP) comprises a DNA segment containing 146 base pairs and wrapped 1.7 left-handed super helical turns around an octamer of histones. The histone octamer consists of paired copies of each histone H2A, H2B, H3, and H4 (Figure 1.7). Histone core particles are connected through a linker DNA region. In addition to nucleosome structure, DNA and histone proteins undergo covalent modifications, one type of epigenetic inheritance system that does not rely upon DNA sequence changes [50]. In this regard, each histone core of the nucleosome has a histone N-terminal tail consisting of a sequence of amino acids and is considered a critical element in the regulatory processes that control both nucleosome structure and function [51]. Histone tails interact with the DNA core's poly-anionic backbone and contribute to nucleosome stability [52]. Most chromatin regulation, structure, and transcription are mediated by modifying histone tails (Figure 1.8). Histone modification refers to residue-specific covalent post-translational modifications (PTM) along with histone N-terminal tails or globular domains [53]. These modifications are considered a critical factor in regulating the accessibility of factors to DNA or stimulation of the recruitment of elements to transcriptional and chromatin assembly processes [54, 55].

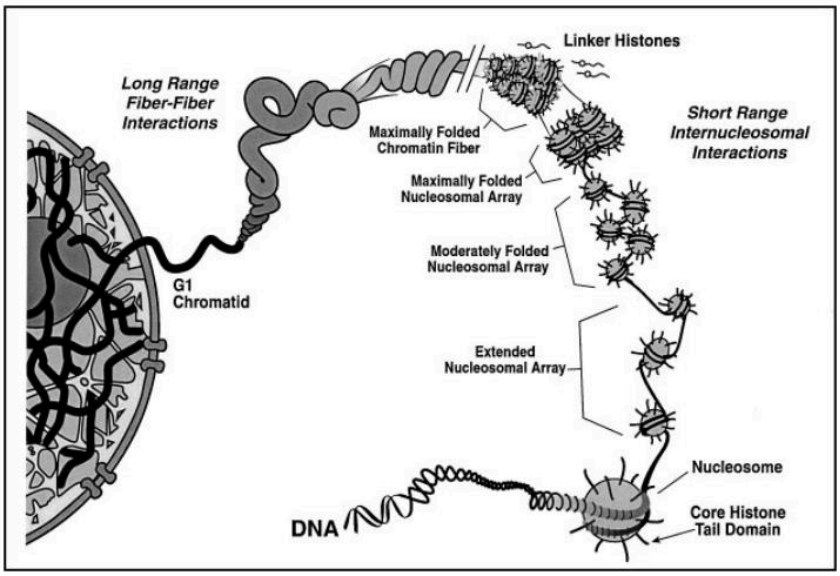


Figure 1.6: Chromatin organization. Schematic representation of chromatin fiber structure and condensation. Double helix DNA wrapped around chromatin core proteins to form nucleosome. The folding steps of extended nucleosome array into maximally folded chromatin fiber are shown in this schematic [56].

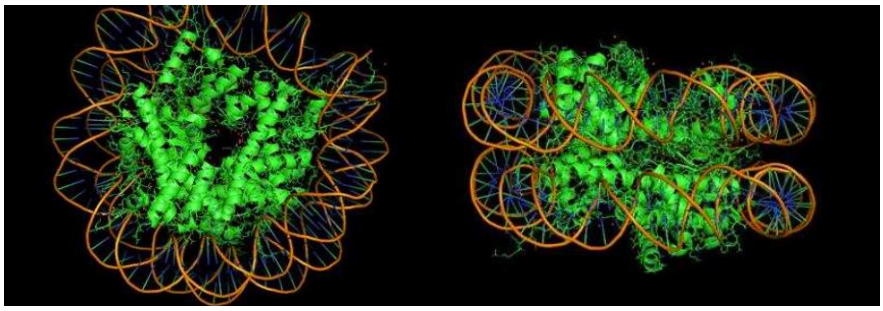


Figure 1.7: Human nucleosome structure. Human nucleosome consists of 164 base pairs DNA phosphodiester backbones wrapped 1.7 times around histone octamer protein containing two each of H2A, H2B, H3 and H4 (PDB 5AV5).

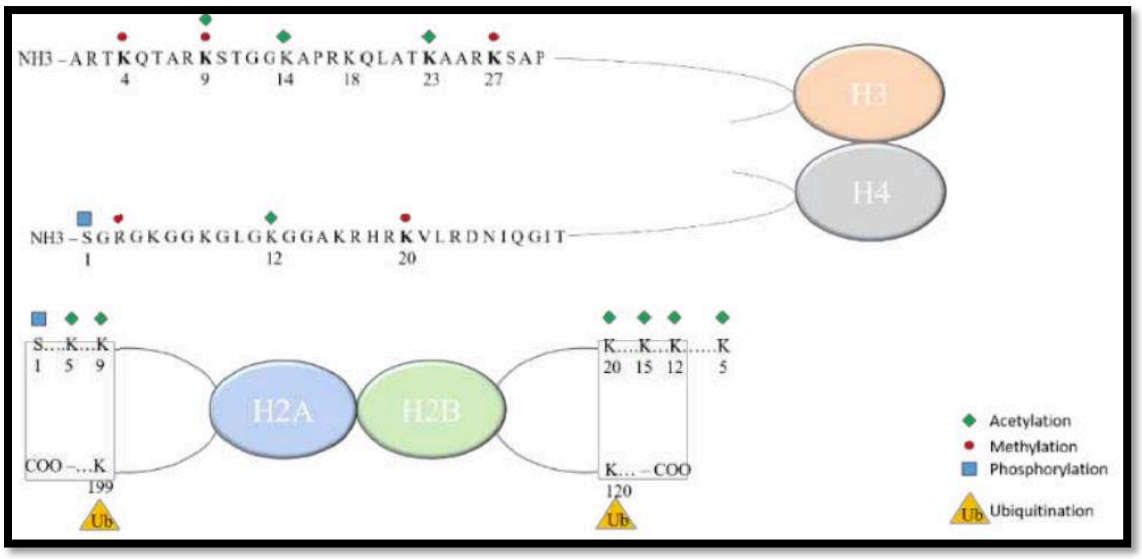


Figure 1.8: Post-translation modification sites on the histone tails that protrude from the histone core. The post-translational modifications shown are methylation (red), acetylation (green), ubiquitination (yellow) and phosphorylation (blue) [57].

1.5. SET domain and histone methylation:

Methylation, phosphorylation, acetylation, and ubiquitination are among the expected modifications to histones [58]. Methylation occurs on histone tails by adding one, two, or three methyl groups by histone methyltransferases that transport methyl groups of S-Adenosyl methionine to arginine or lysine side amino acids of H3 and H4 [59]. Methylation may cause gene activation or repression in a residue-specific manner [60]. Remarkably, the biology of lesions that lead to cancers and their clinical outcome is often connected with modifications to global levels of histone acetylation or methylation at specific sites [61]. Therefore, most histone lysine methyltransferases (HKMTs)' aberrant activity can induce many cancer types [3, 62]. Some lysine-specific methyltransferases have a SET domain, whereas other methyltransferases have a non-SET domain. SET-containing methyltransferase has SET domain, pre-SET and post-SET domains. These three domains are involved in the activity of methyltransferase. Pre-SET (or N-SET) and post-SET (or C-SET) domains flank the SET domain, and they

are cysteine-rich domains. Pre-SET domain has three zinc ions that form triangular zinc clusters, and this structure is critical for the stability of the protein, while the post-SET domain structure is required to complete the active site's activity. SET domains have a catalytic core rich in β-strands, a small number of α-helices and extended loops to form the folded structure [63]. SET domain stimulates the transport activity of the methyl group of S-Adenosyl methionine (SAM) toward protein targets, forming a methylated lysine residue and changing S-Adenosyl methionine (SAM) cofactor to S-adenosyl-L-homocysteine (AdoHcy) or (SAH) [64]. In this regard, the SET domain is highly evolutionarily conserved and found in many 'chromatin regulators.' Consequently, the SET domain was named for its appearance in three genes [The Suppressor of Variegation Su (var), Enhancer of Zeste E(z), and Trithorax], encode regulators in Drosophila [65]. This domain consists of approximately 130-amino acids, and its function was detected through the homology of the SET domain sequence to a set of plant methyltransferases (Rubisco LSMT). Jenuwein et al. (1998) [65] discovered the first SET- containing methyltransferase and also reported that mammalian SUV39H1 is SET domain-dependent and acts as a histone three specific methyltransferase (HMT), transferring a methyl group of S-Adenosyl methionine (SAM) to lysine 9 of the histone H3 N terminus [65].

Much evidence indicates that most SET domain proteins have possessed histone methyltransferase (HMT) activity which targets specific lysine residues within histone tails, resulting in gene expression regulation [66-71]. To date, the DOT1 family is the only non-SET lysine residue methyltransferase that has been reported. DOT1 methylates a lysine residue in the histone's globular core, unlike SET domain

methyltransferase that targets histone tail residues [72]. On the other hand, all known SET domain-containing histone methyltransferase stimulate methyl transfer to histone and non-histone proteins by the SET domain [1, 2, 9, 51, 73]. Therefore, these proteins regulate various pathways, including those critical for developing and appropriate cell cycle progression [74-76]. Subsequently, distinct gene expression states are regulated by transferring a methyl group to specific residues on histone tails. Generally, methylation of H3K4, H3K36, or H3K79 is associated with the activation of gene transcription, while lysine methylation on H3K9, H3K27, and H3K20 correspond to gene silencing [60].

1.5.1. SMYD family:

SMYD proteins are a subfamily named due to possessing both SET and MYND domain-containing proteins. SMYD family is unique because it is distinguished through a SET domain divided through a Myeloid-Nervy-DEAF1 (MYND) domain into two parts [1-3]. The MYND domain consists of a zinc finger motif involved in protein-protein interactions and is the main feature that distinguishes SMYDs from all other SET domain-containing proteins [1]. Even though SMYD family members regulate normal cellular processes, they are also involved in a broad range of tumorigenic pathways [3-5]. Aberrant upregulation of SMYD proteins correlates with altered expression of over eighty genes, including highly regulated homeobox genes, cell cycle regulators, and oncogenes (Figure 1.9) [3-5].

1.5.1.1. The role of SMYD1 in regulation of muscle and heart development:

In 2002, Gottlieb et al. identified SMYD1 and detected that SMYD1 plays a vital role during embryonic development and acts as a cardiac and skeletal muscle-specific regulator of proper cell differentiation and cardiac morphogenesis [2]. SMYD1 was shown to function as a histone deacetylase–dependent transcriptional repressor. Malformation of the right ventricle was observed by targeted deletion of SMYD1 in mice due to disruption of cardiomyocytes and morphogenesis [2]. It was later revealed that SMYD1 interacts with skNAC, acting as a specific transcription activator for heart and skeletal muscle [77].

1.5.1.2. The role of SMYD2 as a lysine methyltransferase and tumorigenesis:

SMYD2 is the second methyltransferase member of the SMYD family characterized [1] and reported to play an essential role in cell differentiation and embryogenesis [78]. SMYD2 transfers a methyl group to lysine residues located on histone or non-histone proteins. SMYD2 is associated with many cancers, including esophageal squamous, bladder carcinoma, gastric cancer, and acute lymphoblastic leukemia when it is aberrantly overexpressed, resulting in poor survival rates [5, 79-82]. Knockdown of SMYD2 in esophageal squamous carcinoma, bladder, and gastric cancer in cell line models eliminated cell proliferation [5, 80]. SMYD2, as a methyltransferase, was identified to methylate H3 lysine 36 and lysine 4 when it interacts with Hsp90 [83]. Therefore, methylation of H3 via SMYD2 has been linked to triggering the transcriptional machinery of many genes related to transcription, cell cycle regulation, and chromatin remolding [83].

Additionally, SMYD2 methylates many non-histone proteins. Until now, the best-described example of SMYD2 methylation of non-histone protein is p53, which acts as a tumor suppressor transcription factor [73, 84-86]. It has been reported that methylation of p53 at lysine 370 by SMYD2 reduces DNA and p53 binding, inhibiting p53 from binding to its specific gene promoters [73]. Knockdown SMYD2 in the same study with doxorubicin treatment as a chemotherapy drug led to arrest of the cell cycle and apoptosis that were mediated by p53. Therefore, inhibition of SMYD2 has the potential to elicit a better response to chemotherapy [87]. In addition to p53, SMYD2 methylates other non-histone proteins, such as the estrogen receptor (ER) [88, 89], PARP [90], retinoblastoma protein (RB) [79], and the chaperone protein Hsp90 [91, 92]. Studies have detected that SMYD2 methylates Hsp90, and this methylation pathway has been linked to the normal function of SMYD2 in normal muscle biology [93, 94] and cancer development [92].

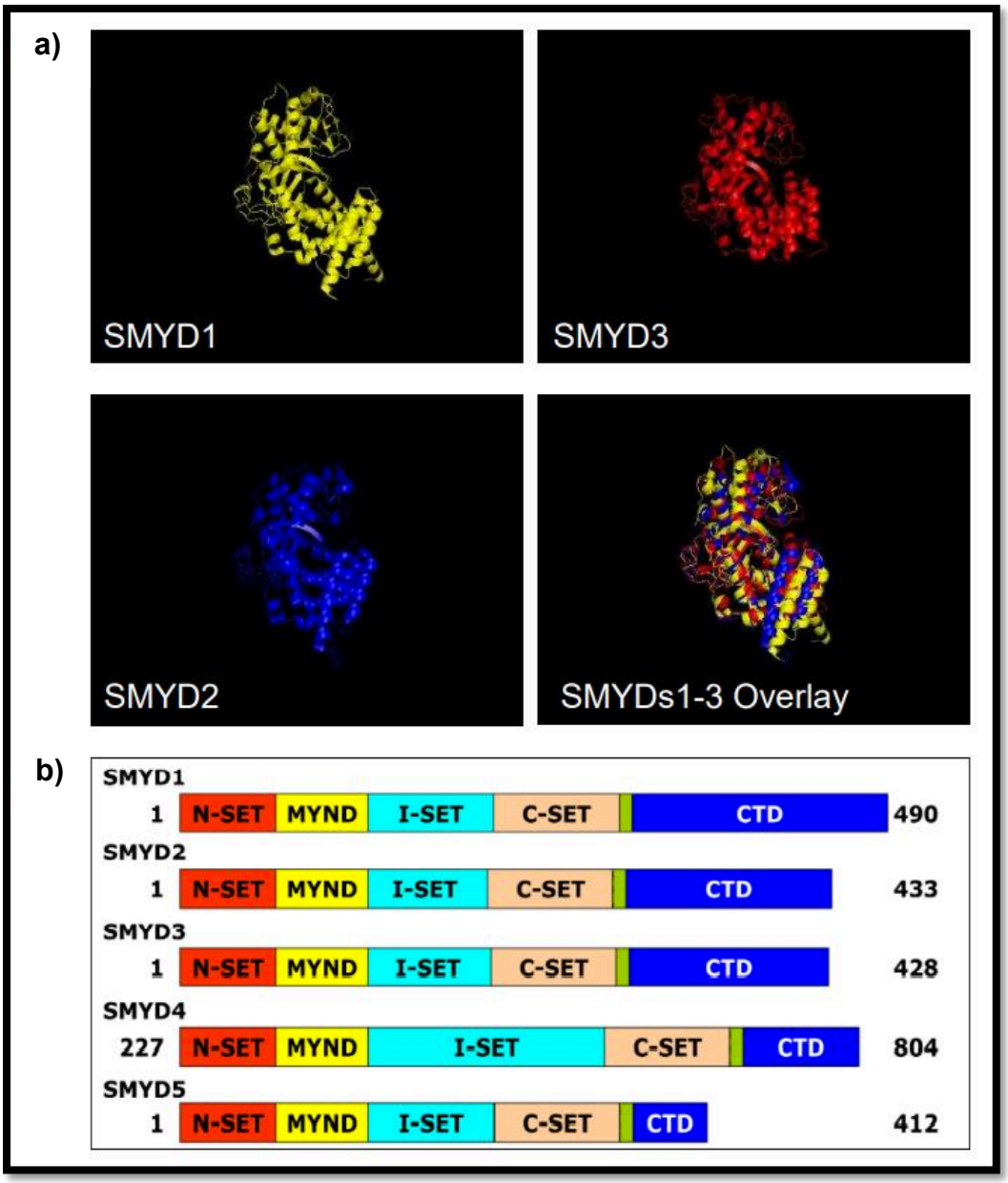


Figure 1.9: The homology of SMYD Family. a) Ribbon structure overlay. b) Domain alignment of five members of SMYD family showing the sequence length and structural identity of SMYDs proteins (SMYD1; PDB codes 3n71, SMYD2; 3tg4, and SMYD3; 3pdn).

1.5.1.3. SMYD3 as a transcription factor and tumorigenesis:

SMYD3 is the third member of the SMYD family of proteins and it has a role in normal cell development, proliferation, and viability. Therefore, SMYD3 overexpression is enough to increase cell growth and activate many genes associated with cancer cells'

transformation [6] and metastasis [7]. Multiple studies have shown that lung, breast, pancreatic, colorectal, and hepatocellular carcinoma are highly associated with SMYD3 overexpression [3, 8, 9].

Aberrant expression of SMYD3 is involved in uncontrolled cell proliferation. Therefore, breast cancer, hepatocellular, and colorectal carcinoma have been impaired significantly through SMYD3 knockdown [3, 8]. In the beginning, SMYD3 methylation of lysine four residue on histone H3 results in uncontrolled proliferation of hepatocellular, colorectal carcinoma, and breast cancer [3, 8]. In addition, Van Aller et al., 2012 [18] reported the role of SMYD3 to methylate lysine five residue on histone 4. SMYD3 activity is associated with the activation and transcription of many downstream genes including NKX2.8 [3], WNT10B [8], TERT [11], cMET [12], and CDK2 [13].

SMYD3 in cellular cytoplasm regulates cancer cell proliferation and viability through its interaction with VEGFR1 and estrogen receptor ER. In cancer cells, SMYD3 transfers methyl groups to VEGFR1, resulting in activation of the kinase signaling pathway [14]. In breast cancer cells, SMYD3 also methylates ER and works as a coactivator [15].

Moreover, Mazur et al., 2014 [9] and Gaedcke et al., 2010 [95] have reported that in colorectal cancer SMYD3 overexpression is associated with the KRAS mutation. Therefore, when SMYD3 methylates MAP3K2, the primary cytoplasmic substrate for SMYD3, it stimulates the MAP kinase signaling pathway and blocks dephosphorylation-mediated by PP2A, resulting in acceleration of the progression of pancreatic and lung cancer [9, 95]. SMYD3 also has been shown to interact with proteins such as RNA

polymerase II, Hsp90, which is a chaperone protein, and RNA helicase HELZ. These interactions are associated with uncontrolled proliferation [3].

1.5.1.5. Upregulation of SMYD3 and ovarian cancer:

SMYD3-silencing using siRNA-SMYD3 or SMYD3-small molecule inhibitor BCI-121 significantly inhibits ovarian cancer cell proliferation, arrests cell cycle at S phase, and promotes apoptosis [96]. PCR results detected that SMYD3 knockdown leads to overexpression of the cyclin-dependent kinase inhibitors (CDKN), that caused S phase arrest, as CDKN2B (p15INK4B), CDKN2A (p16INK4), CDC25A, CDKN3, and CDC25A. After silencing SMYD3, the cell apoptosis rate increased in concert with the downregulation of BIRC3 and upregulation of CD40LG. SMYD3 was also shown to induce triple-methylation to H4K20 after binding to the promoter region of CDKN2A, resulting in down-regulation of its expression. Conversely, SMYD3 leads to upregulation of BIRC3 after binding to its promoter and triple-methylating H3K4. Knocking down SMYD3 in nude mice was also shown to inhibit ovarian cancer growth [96].

Jiang et al., 2019 [96] also has investigated the link between SMYD3 local expression and the methylation-downstream impacts mediated by SMYD3. SMYD3 overexpression is highly associated with ovarian cancer growth. SMYD3 leads to oncogenic pathway progression by methylating lysine residues and causing integration in kinase-signaling cascade in the cellular cytoplasm. However, SMYD3 in the nucleus can methylate histone lysine residues and promote gene transcription through oncogenic pathway stimulation. SMYD3 is overexpressed in the cytoplasm and nucleus of the human ovarian epithelial cancer cell line (HEY). Briefly, SMYD3 silencing did not cause any protein expression changes of H3K4me3 or H4K20me3 in A2780 (human ovarian

cancer cell line). There were no relationships between SMYD3, H3K4me3, or H4K20me3 increases on CDKN2A and BIRC3 promoters in A2780 cells using CHIP assay. Therefore, overexpression of SMYD3 promotes cancer progression through only regulating lysine methylation of critical single molecules in A2780 cell cytoplasm without altering histone methylation. Conversely, SMYD3 silencing in HEY cells detected alteration in SMYD3-mediated histone and non-histone proteins [96].

SMYD3 overexpression is highly associated with ovarian cancer growth. Therefore, SMYD3-silencing using siRNA-SMYD3 or SMYD3-small molecule inhibitor BCI-121 significantly inhibits ovarian cancer cell proliferation, arrests cell cycle, and induces apoptosis [96].

1.5.1.6. SMYD3 small-molecule inhibitor decreases colorectal cancer growth:

Preserico et al., 2015 [19] has proved that SMYD3 is required for colorectal cancer growth through RNAi-mediated SMYD3 silencing. Therefore, these primary data drove this study to design and test SMYD3 inhibitors. One of these designed compounds, BCI-121, reduced SMYD3 activity *in vitro* and colorectal cancer cell lines (CRC) as proved by analyzing global H3K4me2/3 levels. The cellular growth inhibition by BCI-121 was comparable to the result that was observed upon SMYD3 genetic ablation. This study also confirmed that SMYD3 inhibitors are effective in CRC and other tumor cell lines from different origins, such as lung, pancreatic, ovarian, and prostate. Therefore, Preserico et al., 2015 study has confirmed the same results that we obtained from our study about the proof of principle that SMYD3 is an excellent target to inhibit using small

compounds to reduce SMYD3 activity valuable a novel therapeutic agent in cancer treatment [19].

1.7. Hypothesis:

Based on the critical role of SMYD3 during oncogenesis and its essential role in cancer proliferation, we hypothesize that SMYD3 catalytic inhibition will halt proliferation, arrest the cell cycle, and induce apoptosis of breast carcinoma cell lines.

Goal: Design a small molecule that can inhibit SMYD3 function through competitive binding with the enzyme ligand. We want to minimize the affinity of SMYD3 and its substrate binding. We will test these compounds' ability to halt cell proliferation, arrest the cell cycle, and induce apoptosis in breast cancer cell lines without harming normal cells.

1.8. Investigational aims:

Aim1: Design small molecules that can inhibit SMYD3 activity and test these inhibitors' ability to reduce SMYD3 activity using methyltransferase assays.

Aim2: Assess the efficiency of SMYD3 inhibitors to halt cell cycle progression and induce apoptosis in breast carcinoma cell lines.

CHAPTER 2- MATERIALS AND METHODS

2.1. In Silico Screening process

We implemented the Small Molecule Drug Discovery Suite (Schrodinger, Inc., NY, USA) to predict the binding affinity of a library of 137,990 molecules [97-99]. This library of molecules was downloaded from the ZINC15 database, and included all “purchasable” molecules with reported or predicted activity *in vitro* [100]. The 3D structure of SMYD3 used for *in silico* docking was uploaded from the Protein Data Bank (PDB) under PDB identification code 5EX3 [101]. After an initial simulation which docked each molecule into SMYD3′s protein-target binding pocket (not its s-adenosylmethionine binding pocket), the top ten hits (most-negative binding energy) were entered into the ZINC15 molecular similarity search engine, and the 50 most-similar compounds to each of the ten leading candidates were again scored using the Schrodinger software (500 total compounds). From this iteration, the top five compounds were purchased and assessed *in vitro* using SMYD3 methyltransferase assays. After initial experiments, Inhibitor-4 was found to be the most promising and, consequently, it advanced to the cell line experiments described below.

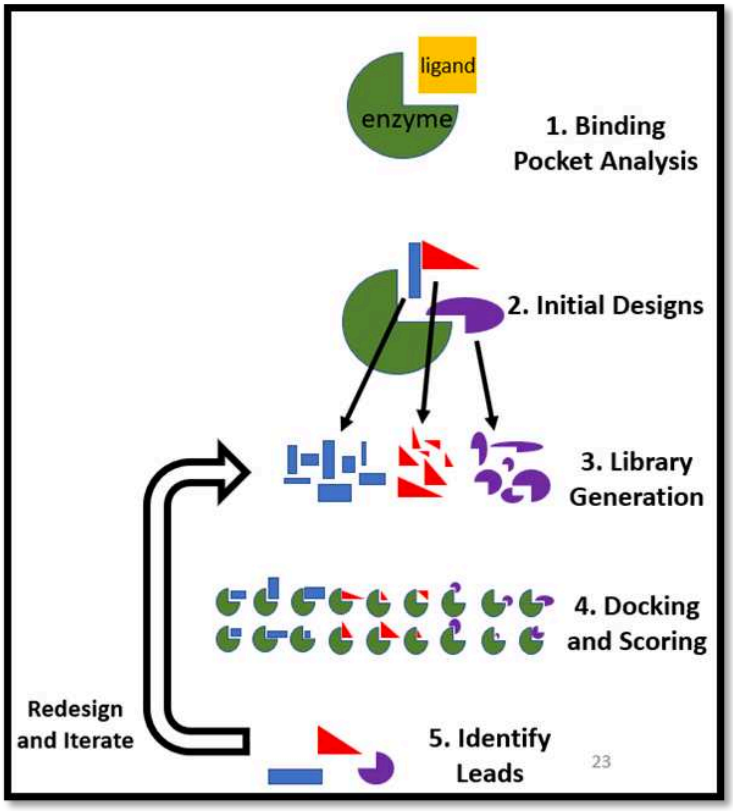


Figure 2.1: *In silico* screening process.

1. Identify and observe the enzyme pocket(s) and natural ligand(s).
2. Using the natural ligand, known inhibitors, key amino acids, and common drug structures, design 3-4 “first-iteration” small molecule inhibitors.
3. Create large library of similar compounds.
4. Dock and score the compounds in the library.
5. Select top candidates (highest binding affinities), redesign, and repeat steps 3-5.

2.2. Chemical Compounds

All screened compounds were dissolved in dimethyl sulfoxide (DMSO) as 5, 10 or 100 mM stock solutions. The positive control, BCI-121, is a previously-reported SMYD3 inhibitor shown to reduce the cellular proliferation of colorectal and ovarian cancer [19,

96]. It was purchased from Millipore Sigma (1817, Burlington, MA, USA) and dissolved in DMSO at 10 and 100 mM. BCI-121 was used in all experiments to investigate its impacts against breast cancer cell lines and as a positive control inhibitor. All compounds were stored at −20 °C until used for the experiments. Some d_6-dimethyl sulfoxide (D, 99.9%) containing 0.05% *v*/*v* TMS were used for the ^{1}H NMR stability study and the d_6-dimethyl sulfoxide was purchased from Cambridge Isotope Laboratories, Inc. and used as is. The stock solutions of 10.0 mM of BCI-121 and Inhibitor-4 were prepared immediately before use in d_6-DMSO.

2.3. In Vitro Methylation Assay

In vitro methylation was investigated using a colorimetric assay (BioVision, K986-100, Milpitas, CA, USA). SMYD3 inhibitors (160 nM) were incubated with H3 recombinant protein (1.6 µM; Sigma-Aldrich, SRP0177, St. Louis, MO, USA) for 10 min at room temperature. Next, SMYD3 recombinant protein (100 nM; Sigma-Aldrich, SRP0153, St. Louis, MO, USA) and s-adenosylmethionine (SAM) cofactor (500 µM, methyl donor ligand) were added to the SMYD3 inhibitor and H3 solution in the methylation buffer that was provided with the kit. The absorbance was read using a microplate reader (BioTek, Cytation 5, Winooski, VT, USA) at 570 nm in kinetic mode every 30 s at 37 °C for 45 min. The optical density (OD) of the inhibitors was normalized to the optical density (OD) of the control [3].

2.4. NMR Spectroscopy Analysis

The 10.0 mM stock solutions of Inhibitor-4 and BCI-121 were prepared freshly in d_6-DMSO containing 0.05% *v*/*v* tetramethylsilane (TMS) and diluted to the final

concentration of 5.0 mM in d_6-DMSO. The stabilities of Inhibitor-4 and BCI-121 were determined by 1D ^{1}H Nuclear Magnetic Resonance (NMR) spectroscopy on a Bruker 400 MHz NMR spectrometer at 25 °C using routine parameters [102]. 2D-NMR experiment will run to confirm the assignments (data not included) [102]. Chemical shifts were measured against TMS (0 ppm) as an internal reference. The spectra of Inhibitor-4 and BCI-121 were recorded at 0 and 24 h. The spectra were worked up and integrated using Mnova V.14 (MestreLab Research SL). The signals in the aromatic region were used to measure the ratio of starting material and hydrolysis product.

2.5. Cell Culture

All cells were purchased from ATCC and cultured according to manufacturer recommendations. The human breast epithelial cell line MCF10A (ATCC CRL-10317, Manassas, VA, USA) was used as a normal cell line and grown in DMEM/F12 (Invitrogen, 11330-032, Carlsbad, CA, USA) supplemented with 5% horse serum (Invitrogen, 16050-122, Carlsbad, CA, USA), 1% antibiotic/antimycotic mixture (Millipore Sigma, A5955, Burlington, MA, USA), insulin (10 µg/mL), EGF (20 ng/mL), cholera toxin (100 ng/mL), and hydrocortisone (500 ng/mL) [103]. The mammary gland breast cancer lines MCF7 and MDA-MB-231 (ATCC HTB-22 and 26, Manassas, VA, USA) were grown in DMEM media (Corning, 29818003, Corning, NY, USA) with 10% FBS (Atlas Biologicals, F-0500-A, Fort Collins, CO, USA) and 1% antibiotic/antimycotic mixture (Millipore Sigma, A5955, Burlington, MA, USA) [104]. All cell lines were grown in a humidified incubator at 5% CO_2 and 37 °C with regular passaging to avoid confluence.

2.6. Protein Extraction and Immunoblotting

Total protein was extracted from frozen cells using RIPA buffer (150 mM NaCl, 5 mM EDTA, 50 Tris-HCl pH 8.0, 1% NP-40, 0.5% Na-catecholate and 0.1% SDS) supplemented with protease inhibitor (PI 87785, Life tech, Carlsbad, CA, USA). Protein concentration was determined according to Bradford (*1976)* using bovine serum albumin as a standard [105]. Fifty micrograms of total protein were separated by SDS-PAGE, transferred to nitrocellulose membrane by electroblotting as described by Chiacchiera (2009) and probed with the antibodies specific for the indicated proteins [106]. Actin was used as an internal control for normalization. Antibodies for immunoblot detection of SMYD3 (Rabbit monoclonal antibody to SMYD3, ab183498, Abcam, Cambridge, MA, USA) and β-Actin (A5316-100 UL, Sigma-Aldrich, St. Louis, MO, USA) have been used as the primary antibodies. Bound antibodies on blots were detected by HRP--conjugated secondary antibodies (ab205718, Abcam, Cambridge, MA, USA). Detection was done using Clarity Western ECL Substrate (Bio-Rad, Hercules, MA, USA) and visualized using Image Lab Software (Bio-Rad). Densitometric evaluation was performed by ImageJ software (Version 2.0.0).

2.7. Immunocytochemistry

Exponentially growing cells were fixed with 4% paraformaldehyde for 15 min at room temperature. After PBS wash, cells were permeabilized with 0.1% SDS, 0.5% Triton X-100 in PBS for 10 min at room temperature. 10% goat serum in PBS was used for blocking for 30 min at room temperature. Primary antibody was diluted in 1:300 and treated for 1 h at 37 °C. Secondary antibody (Alexa488 conjugated anti-rabbit IgG) was diluted for 1:500 and treated for 30 min at 37 °C. DAPI in Vectashield antifade solution

was used for mounting. The images were capture by Zeiss Axiophot microscope with Qimaging Exi Aqua camera with Qcapture pro software. Blue signal was obtained with 50 millisecond exposure. Green signal was obtained with 200 millisecond exposure. Signals were quantified by densitometry using ImageJ2 software (Version 2.0.0).

2.8. In Vitro Cell Growth Inhibition Assay

Normal and breast cell lines were plated at a density of 20,000 cells/well onto a 6-well plate with different concentrations of Inhibitor-4 and BCI-121. After trypsinization, cell numbers were counted and scored as the number of proliferating cells after treatments at different time points (24, 48, 72 and 96 h) using a Coulter Counter Z2 (Beckman-Coulter Z2 Coulter Particle Count Counter and Size Analyzer, Brea, CA, USA). Data were analyzed and cell doubling time was calculated using GraphPad Prism 6 software (Graph Pad Software, La Jolla, CA, USA) through the exponential growing equation using the exponential growing stage [107].

2.9. Clonogenic Cell Survival

A colony formation assay was used to determine cell sensitivity to SMYD3 inhibitors. The self-renewal and proliferative capacities of cells were measured. To form colonies, cells were seeded onto 6 well plates and were treated with varying concentrations of BCI-121 and Inhibitor-4. The plated cells were incubated in a humidified incubator at 5% CO_2 and 37 °C for two weeks. Then, colonies were fixed with 100% ethanol and allowed to dry for 20 min at room temperature before staining. Colonies were stained using 0.1% crystal violet and allowed to dry before counting. Reproductively viable surviving cells were counted based on the microscopic colonies containing more than 50 cells.

From the cell survival fraction, survival curves were drawn using Graph Pad Prism 6 software (Graph Pad Software, La Jolla, CA, USA). At least three independent experiments for each cell line were conducted [108].

2.10. MTT Assay

Cells were plated at a density of 5000 cells/well in 96 well plates. After seeding, cells were treated with the vehicle (DMSO 0.1%, 0.15% and 0.2% *v*/*v*) or various concentrations of the screening inhibitors. The plated cells were incubated in CO_2 and treated for 24, 48, 72 and 96 h. Then, 10 µL MTT solutions (5 mg/mL) were added to each well followed by a 4 h incubation in CO_2 in the dark. Formazan crystals formed were dissolved in 100 µL of SDS followed by a second 4 h incubation in CO_2. The absorbance was read using a microplate reader (BioTek Instrument, Cytation 5, Winooski, VT, USA). The optical density (OD) of each sample was subtracted from the optical density (OD) of the background and the Formazan standard curve was determined. Cellular viability of all samples was calculated using the ratio of the inhibitor treated-groups versus vehicle-treated group. Graph bars were obtained using GraphPad Prism 6 software (Graph Pad Software, La Jolla, CA, USA) [109].

2.11. Cell Cycle Assay

Cell cycle distributions were analyzed using PI flow cytometry. Cells were plated at density of 5×10^5 cells per well onto 6-well plate. Cells were treated with 200 µM of Inhibitor-4 or BCI-121 and incubated in a humidified incubator at 5% CO_2 and 37 °C for 24 h. Following incubation, detached cells were collected, washed two times with phosphate buffered saline (PBS). Then, cells were fixed with 70% ethanol in PBS

overnight at 4 °C. The fixed cells were washed with PBS twice to remove ethanol thoroughly. The cells were resuspended in propidium iodide staining solution consisting of 20 µg/mL propidium iodide and 200 µg/mL RNase in 0.1% Triton X-100. The stained cells were incubated for 15 min in an incubator at 37 °C. DNA contents were measured subsequently using CyAn ADP analyzer flow cytometry (Beckman Coulter, Fort Collins, CO, USA). Each cell line was gated at 10,000 events and the cell cycle distributions were determined using FLOWJO 10.6 software (FlowJo LLC, Ashland, OR, USA) [110].

2.12. Apoptosis Assays

Cell apoptosis was detected using Annexin V, which binds to translocated phosphatidylserine (PS) in the plasma membrane as previously described [111]. Necrosis and late apoptosis were detected using PI to test loss of cell membrane integrity. Briefly, cells were plated and treated with 200 µM of either BCI-121 or Inhibitor-4 for 48 h. Then, the cells were washed with PBS, trypsinized, pelleted, and resuspended in Annexin binding buffer. The cells were stained first with APC Annexin V for 15 min and then with 2.5 µL of PI. The cell mixture was analyzed using a Cytek 4-laser Aurora instrument (Cytek, Fremont, CA, USA). from each sample, a minimum of 3 × 10^4 events was collected. SpectroFlo software (Cytek, Fremont, CA, USA) was used to analyze the multivariate data. APC Annexin V+/PI+, APC Annexin V-/PI-, APC Annexin V+/PI- or APC Annexin V-/PI+ represented late apoptotic cells, viable cells (intact), early apoptotic cells or necrosis, respectively [112].

In addition to, apoptosis induction by SMYD3 treatments was also assessed using Caspase 3/7 activation. Exponentially growing cells were treated with 200 µL of BC1-121 and Inhibitor-4. After 48 h of incubation, the early apoptosis was measured with the

activation of Caspase 3/7 by Caspase-Glo 3/7 kit (Promega, Madison, WI, USA). Glow luminesce of 15,000 cells was measured by Lumat LB9507 (Berthold technologies, Oak Ridge, TN, USA).

2.13. Statistical analysis:

The statistical significance of the results in this study was analyzed using GraphPad Prism 6 software (Graph Pad Software, La Jolla, CA, USA) for two-way ANOVA analysis. *p* value of less than 0.05 were considered statistically significant for all analyses.

CHAPTER 3- RESULTS

Using SMYD3 as a target protein, we implemented our screening methodology and reported a novel small molecule SMYD3 inhibitor (Inhibitor-4) that impairs breast cancer cell proliferation without affecting normal cells, illustrating the potential of SMYD3 inhibitors in the clinical management of breast cancer as well as a proof of concept for this drug development platform. We used two breast cancer cell lines (MCF7 and MDA-MB-231) that were previously shown to overexpress SMYD3 compared with the normal breast epithelial cell line MCF10A (Table 1). After initial hit identification *in vitro*, we purchased and tested five novel small molecule SMYD3 inhibitors and discovered that Inhibitor-4 significantly reduces breast cancer proliferation, arrests the cell cycle, and induces apoptosis without impacting normal cells. In all experiments, we used a previously-identified SMYD3 small molecule inhibitor, BCI-121, as a positive control.

Table 3.1: Summary of SMYD3 availability and activity in the cell lines used in this study.

Cancer Cell lines	Origin	SMYD3 expression	Assay	Methylation activity	References
MCF7/MDA-MB-231	Human epithelial breast cancer cells	High	Western blot, RT-qPCR	H4K5, H3K4	[8, 15-18]
MCF10A	Human epithelial breast cells	Very low	Western blot, RT-qPCR	H4K5, H3K4	[16, 17]

3.1. Inhibitor-4 Decreases SMYD3-Mediated H3 Methylation

After our iterative *in silico* screening using Schrodinger software (Glide®, Maestro®, LigPrep®, and Epik®), we purchased the top five hit compounds for testing. Hits were defined as the drug-like small molecules with the lowest free binding energy when docked in the protein-target binding pocket of SMYD3. The predicted free binding energies of the five lead compounds ranged from −7.2 kJ/mol to −9.1 kJ/mol, compared to the natural protein ligand's predicted free binding energy of only around −1 kJ/mol (fragment of VEGFR1). We used an *in vitro* methylation assay using purified Histone 3 (H3) to assess the ability of the five lead *in silico*-designed SMYD3 inhibitor candidates to decrease SMYD3 enzymatic activity. We demonstrated that Compound 4 (Inhibitor-4) significantly reduces SMYD3-mediated Histone 3 methylation (70% reduction), while the other novel compounds did not show significant differences. H3 was chosen because of previous studies that demonstrated SMYD3 methylates H3 preferentially (Figure 3. 1) [3].

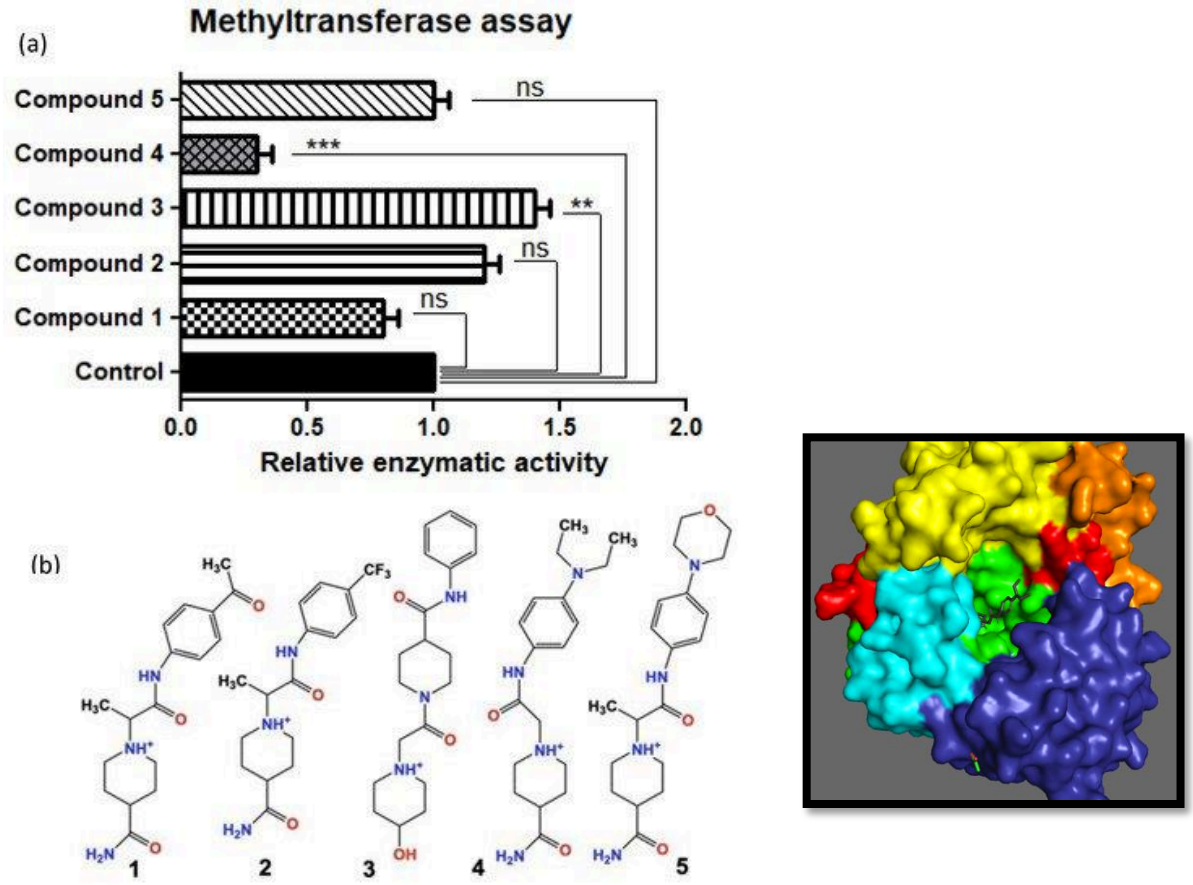


Figure 3.1: (a) Relative SMYD3 activity (top) with the top 5 candidates from *in silico* testing (bottom) using an *in vitro* methyltransferase assay (Colorimetric assay). (b) Compounds 1–5 are illustrated using ChemDraw and predicted binding orientation of competitive SMYD3 inhibitor. Compound 4 is here after referred to as Inhibitor-4. Error bars display standard error of means. Statistically significant differences from control are indicated by ** $\boldsymbol{p} < 0.01$, *** $\boldsymbol{p} < 0.001$ or ns $\boldsymbol{p} > 0.05$.

3.2. Inhibitor-4 and BCI-121 are Stable in d_6-DMSO Solution.

Because of the limited solubility of selected molecules in aqueous solution and in media, we dissolved BCI-121 and Inhibitor-4 in d_6-DMSO solution to record and analyze the 1D ^{1}H NMR spectra of both compounds. The major species attributed to Inhibitor-4 and BCI-121 were observed at time 0 and 24 h, as shown in Figure 3. 2a (BCI-121) and Figure 3. 2b (Inhibitor-4). The ^{1}H NMR peaks of the fresh and aged samples for

Inhibitor-4 showed no observable difference in the presence of the major component (67%) and minor component (33%) peaks as a function of time, suggesting that no hydrolysis is taking place during the experiment for Inhibitor-4. For BCI-121, 70% of the major species was present at time 0, however after 24 h this decreased slightly to 68%, suggesting that the positive control may be slightly less stable than Inhibitor-4.

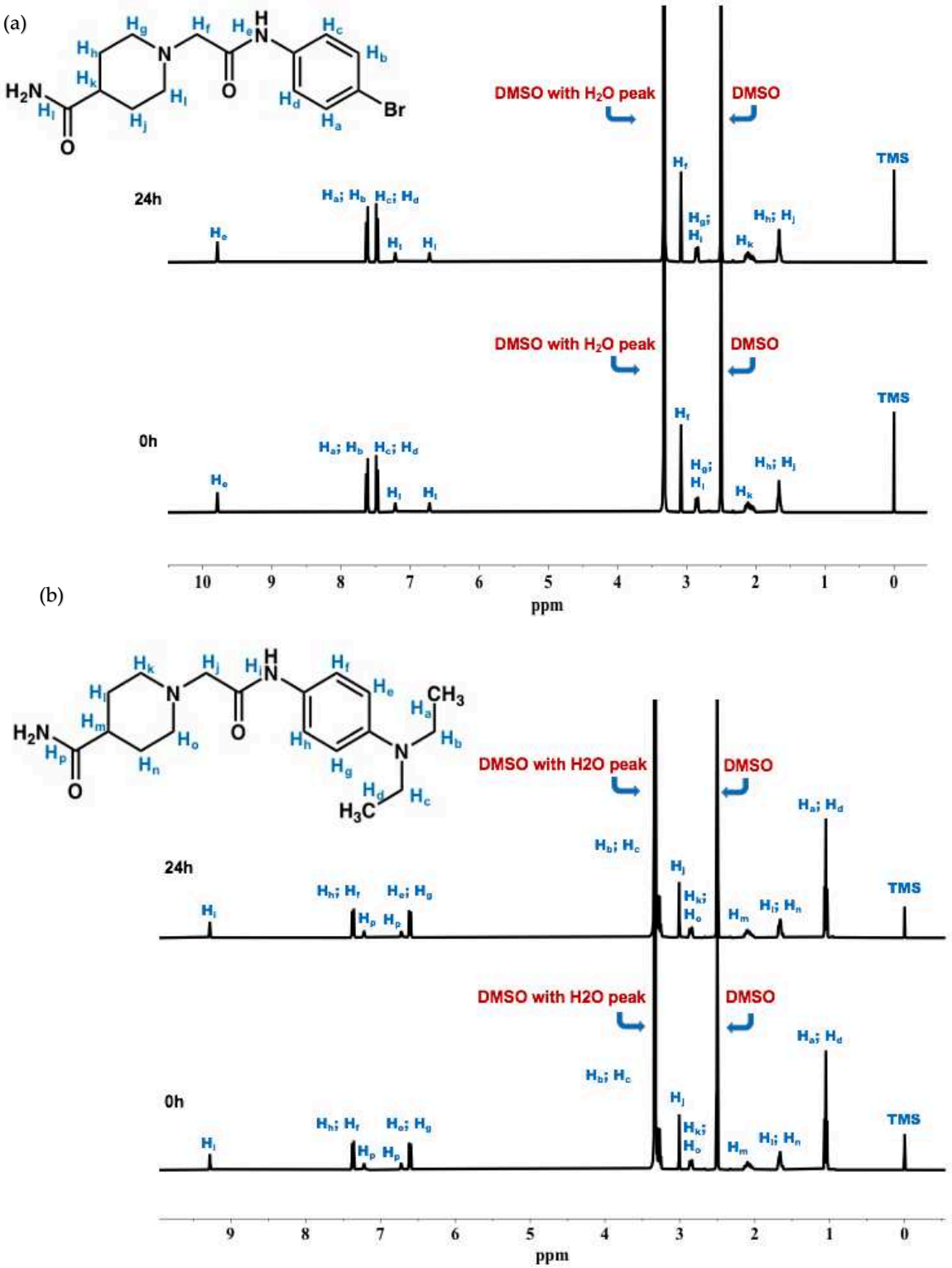


Figure 3.2: ^{1}H NMR spectra of fresh (0h) and aged (24h) of (a) BCI-121 and (b) inhibitor-4 in d_6-DMSO.

3.3. SMYD3 Is Overexpressed in Breast Cancer Cells

Western blot and immunocytochemistry were carried out to test the expression levels of SMYD3 using anti-SMYD3 antibody in normal and breast cancer cell lines. Western blot data have indicated that SMYD3 was highly expressed in breast cancer cell lines (1.8-fold in MCF7 and 2.6-fold in MDA-MB-231) compared to normal cell line (Figure 3.3 a, b, c).

Additionally, immunocytochemistry data have shown elevated levels of SMYD3 expression in breast cancer cell lines comparing to normal cell line (Figure 3.3 d, e). Therefore, increased SMYD3 expression could be correlated with breast carcinogenesis.

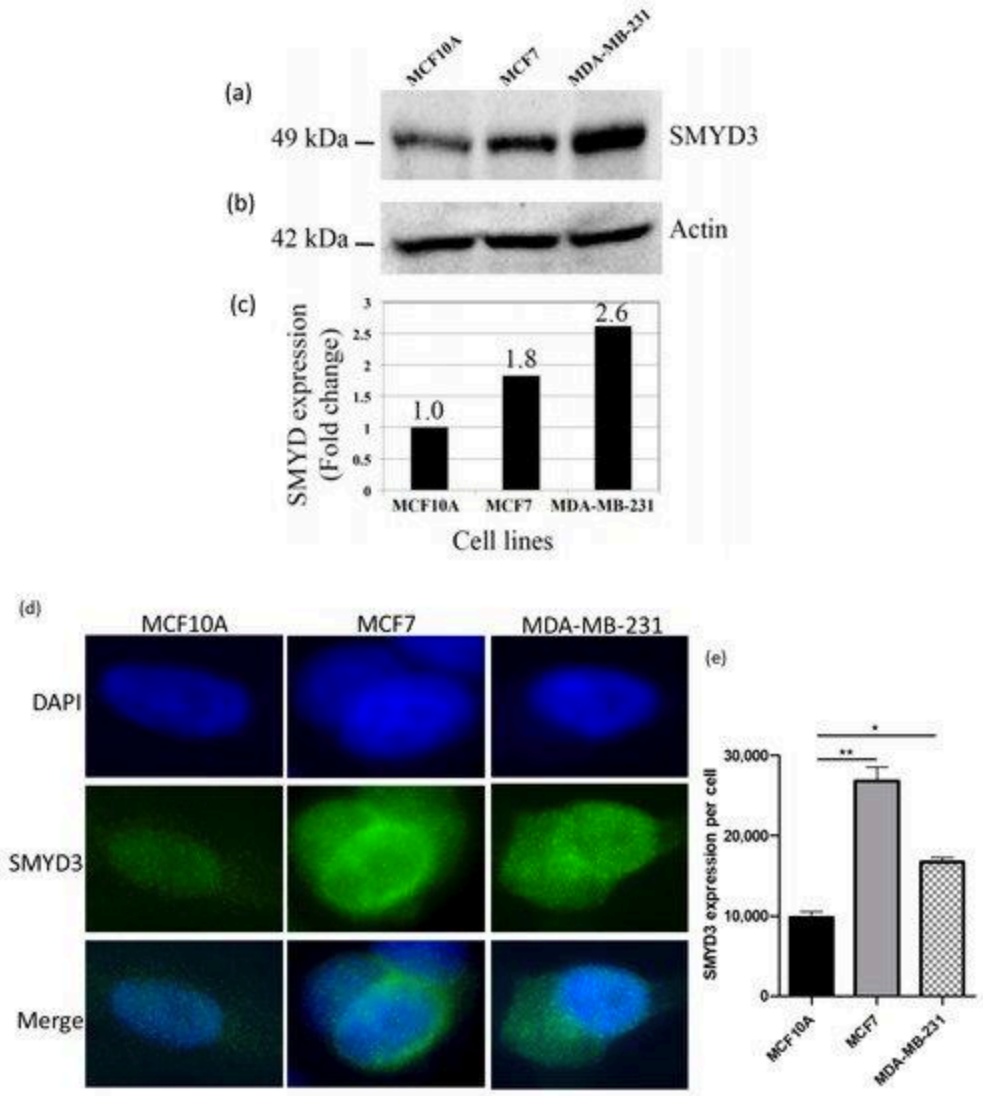


Figure 3.3: SMYD3 expression using western blot and immunocytochemistry: (**a**) Expression of SMYD3 protein in human cell lines using Western blot. (**b**) Expression of actin served as a quantitative control. (**c**) Western blot analysis shows fold change in SMYD3 expression in the cell lines. (**d**) Expression of SMYD3 protein using immunocytochemistry. (**e**) immunocytochemistry analysis shows SMYD3 intensity in the cell lines. Values are mean ± standard error of the means. Statistically significant differences from control are indicated by * $p < 0.05$, ** $p < 0.01$.

3.4. Inhibitor-4 Inhibits Growth of Breast Cancer Cells

The impact of SMYD3 inhibitors on growth of breast cancer cells was tested by adding 50, 100 and 200 μM of Inhibitor-4 or BCI-121 to breast cancer cell lines (MCF7 and MDA-MB-231) and normal breast epithelial cell line (MCF10A). The number of cells

was determined daily and the population doubling times were quantified (Figure 3.4). For MCF7 (breast cancer) cells, the basal doubling time for MCF7 was 38 h, while 40 h for MDA-MB-231. Using the positive control inhibitor, a concentration of 200 μM caused approximately 2-fold suppression of MCF7 cellular growth (Figure 3.4a). Using Inhibitor-4, however, a clear dose-dependent suppression in growth was observed with the first significant reduction observed at a concentration of 50 μM (Figure 3.4b). In the MDA-MB-231 cell line, a significant delay in the cellular growth was observed with 200 μM BCI-121 and only 50 μM Inhibitor-4 (Figure 3.4c, d).

For MCF10A (normal) cells, the effect of the SMYD3 inhibitors was limited. The basal doubling time for MCF10A was 28 h. Interestingly, no delay was noticed with 50, or 100 μM concentrations of either inhibitor. Treatment of the normal cells with 200 μM of Inhibitor-4 resulted in a minor, not significant, growth delay (approximately 5%), while treatment with 200 μM BCI-121 resulted in a major growth delay (Figure 3.4e, f). These results suggest that Inhibitor-4 shows more growth inhibition than BCI-121 and causes significant inhibition in cancer cell growth while only modestly impacting healthy cells.

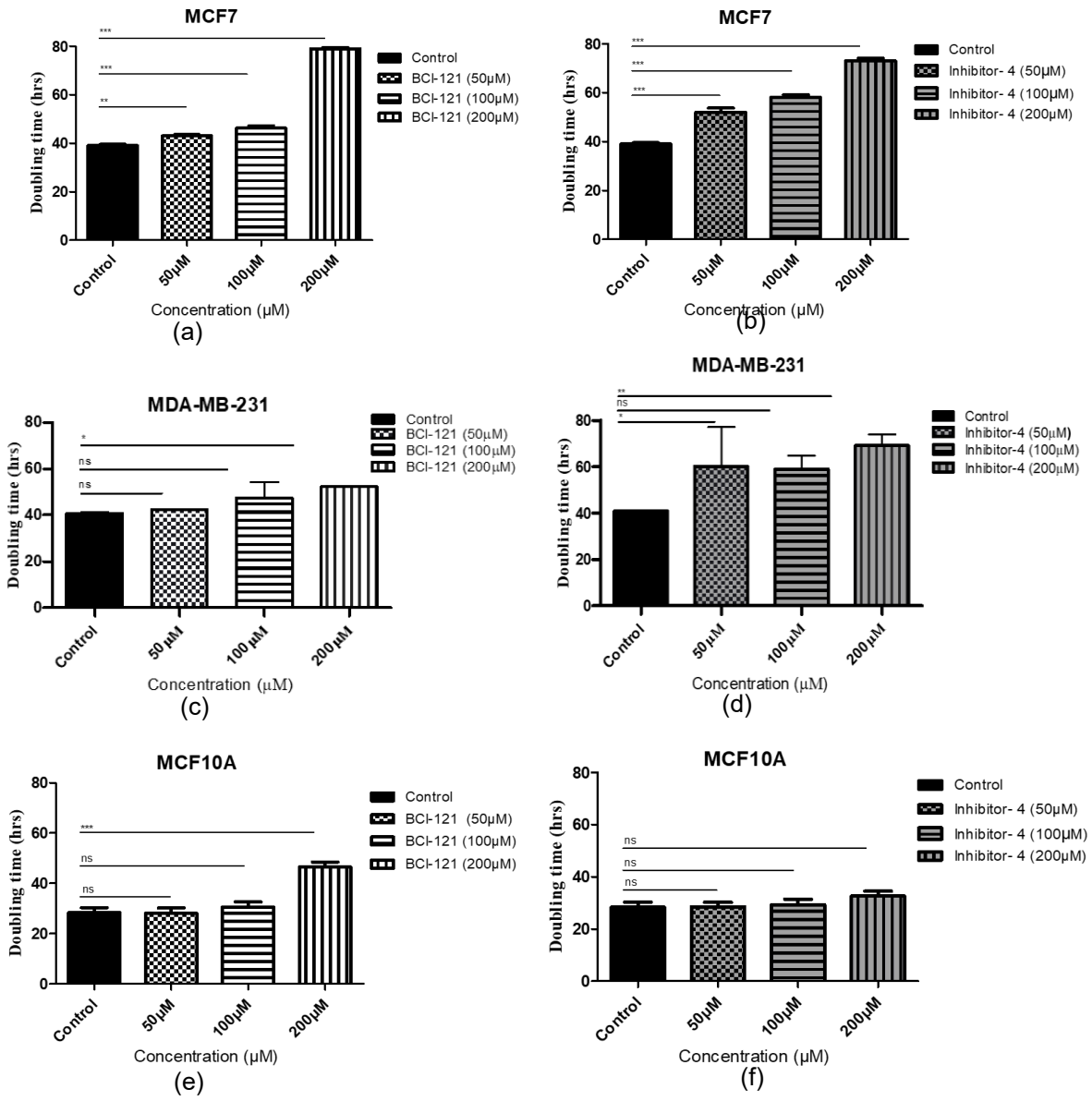


Figure 3.4: Cell population doubling time with SMYD3 inhibitor treatment. (**a**,**c**,**e**) Cells with BCI-121 as a positive control inhibitor. (**b**,**d**,**f**) Cells with SMYD3 Inhibitor-4. Values are mean ± standard error of the means. Statistically significant differences from control are indicated by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ or ns $p > 0.05$.

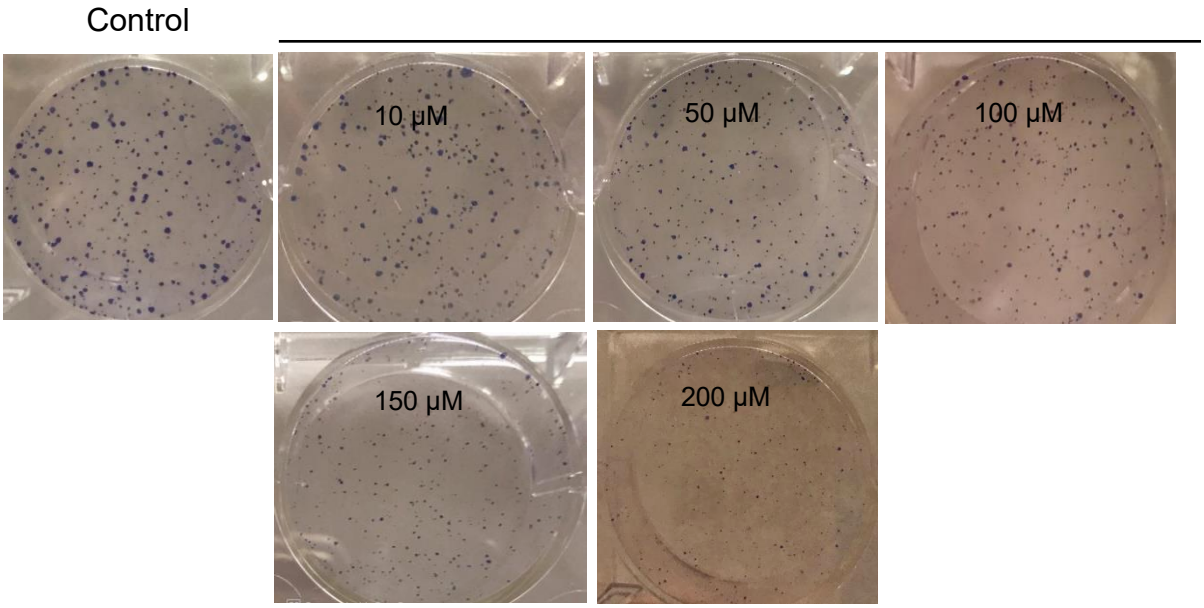
3.5. Inhibitor-4 Suppresses Breast Cancer Cell Colony Formation

To determine the effects of Inhibitor-4 on the colony formation of breast cancer cells and normal cell lines, the cells were treated with various concentrations of Inhibitor-4 and BCI-121 (10, 50, 100, 150 and 200 μM) and incubated for 2 weeks.

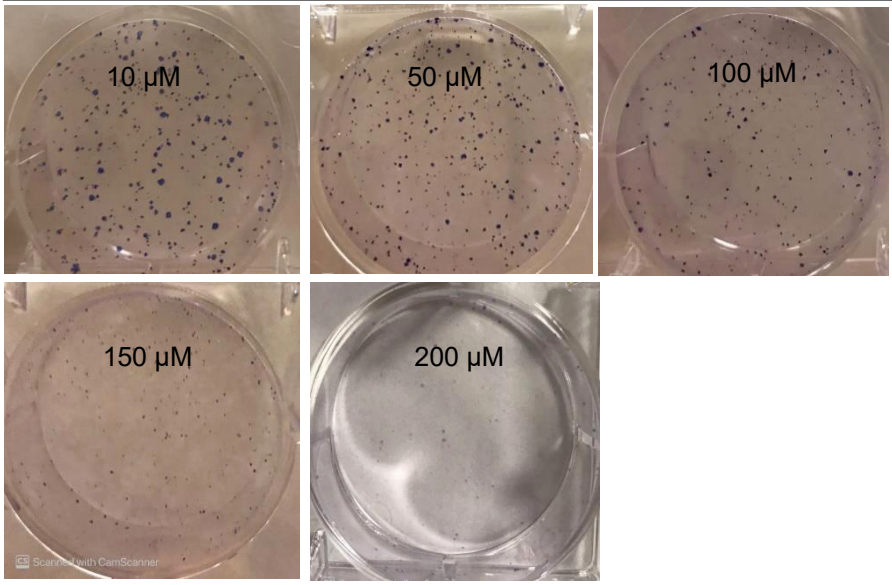
Treatment with Inhibitor-4 significantly suppressed clonogenic activity in MCF7 and MDAMB-231 cells at concentrations of 50, 100, 150 and 200 μM (Figure 3.5b, d, h, j) compared to normal MCF10A cell line (Figure 3.5f,l). Similarly, BCI-121 suppressed colony formation on MCF7 at nearly all concentrations (Figure 3.5a,g) and MDA-MB-231 cells at 150 and 200 μM concentrations (Figure 3.5c,i). Surprisingly, a significant decrease in colony formation of MCF10A (normal) cells was also observed at 200 μM concentration of BCI-121 (Figure 3.5e,k) compared to Inhibitor 4, which did not affect MCF10A survival (Figure 3.5f,l). This result again suggests the improved inhibition effect of Inhibitor-4 compared to BCI-121.

MCF-7 (Breast cancer cell line)

a) BCI-121

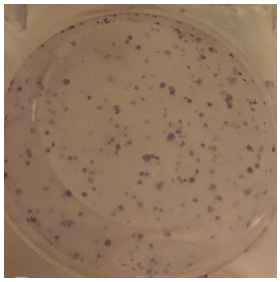


b) Inhibitor-4



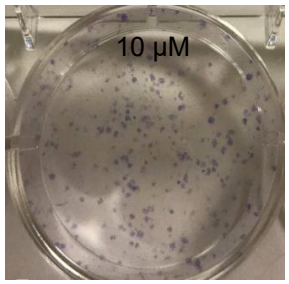
MDA-MB-231 (Breast cancer cell line)

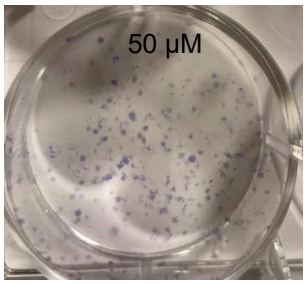
Control

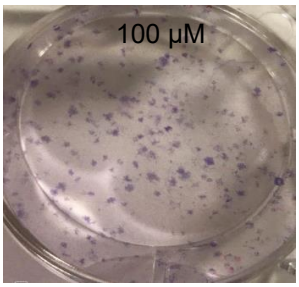


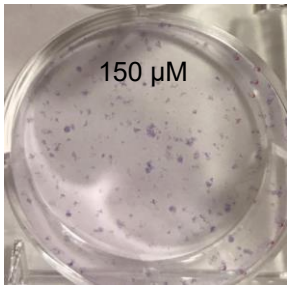
c)

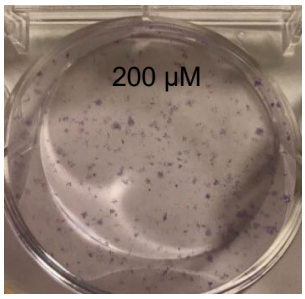
BCI-121





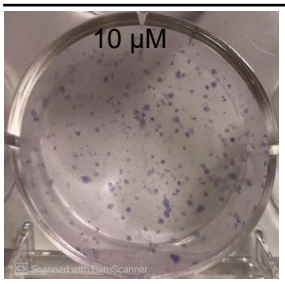


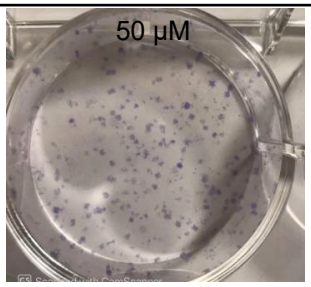


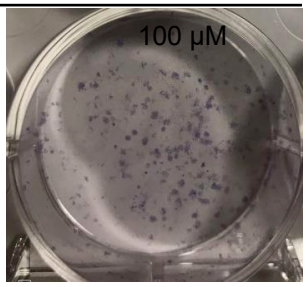


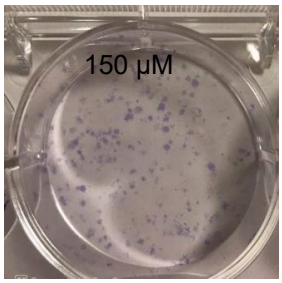
d)

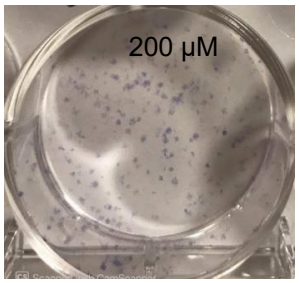
Inhibitor-4





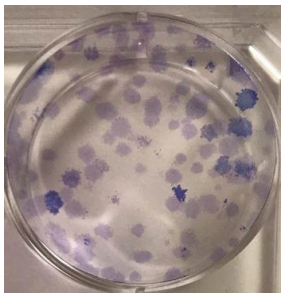




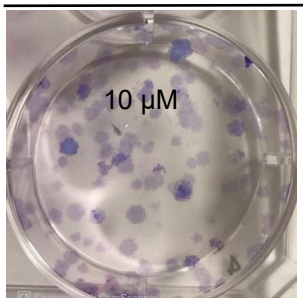


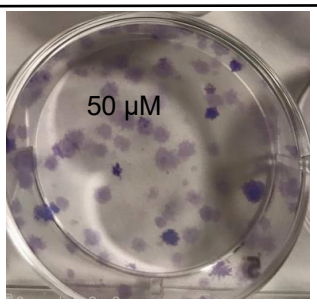
MCF10A (Normal breast cell line)

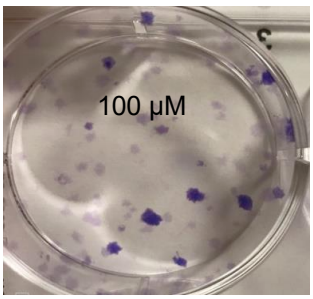
Control

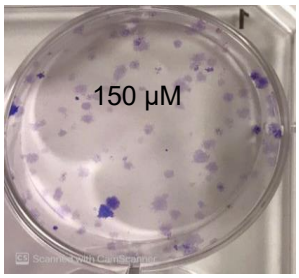


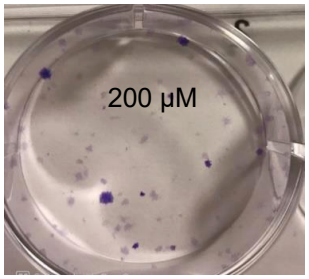
e) BCI-121



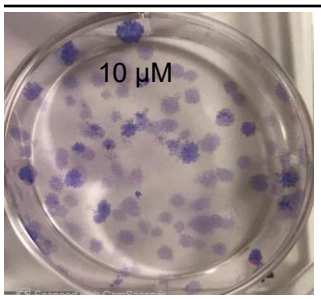


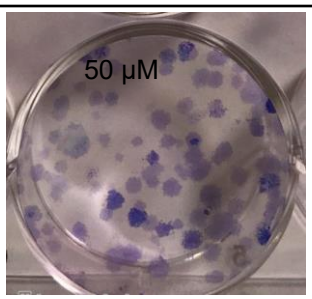


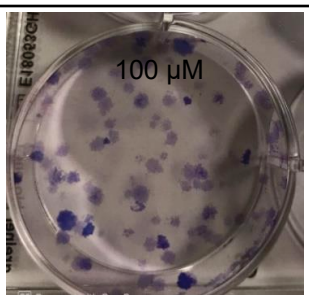


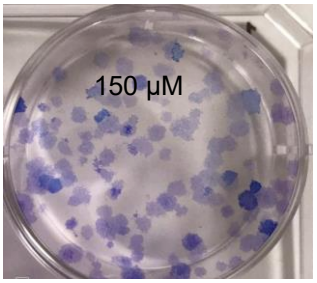


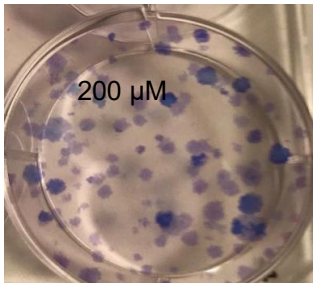
f) Inhibitor-4











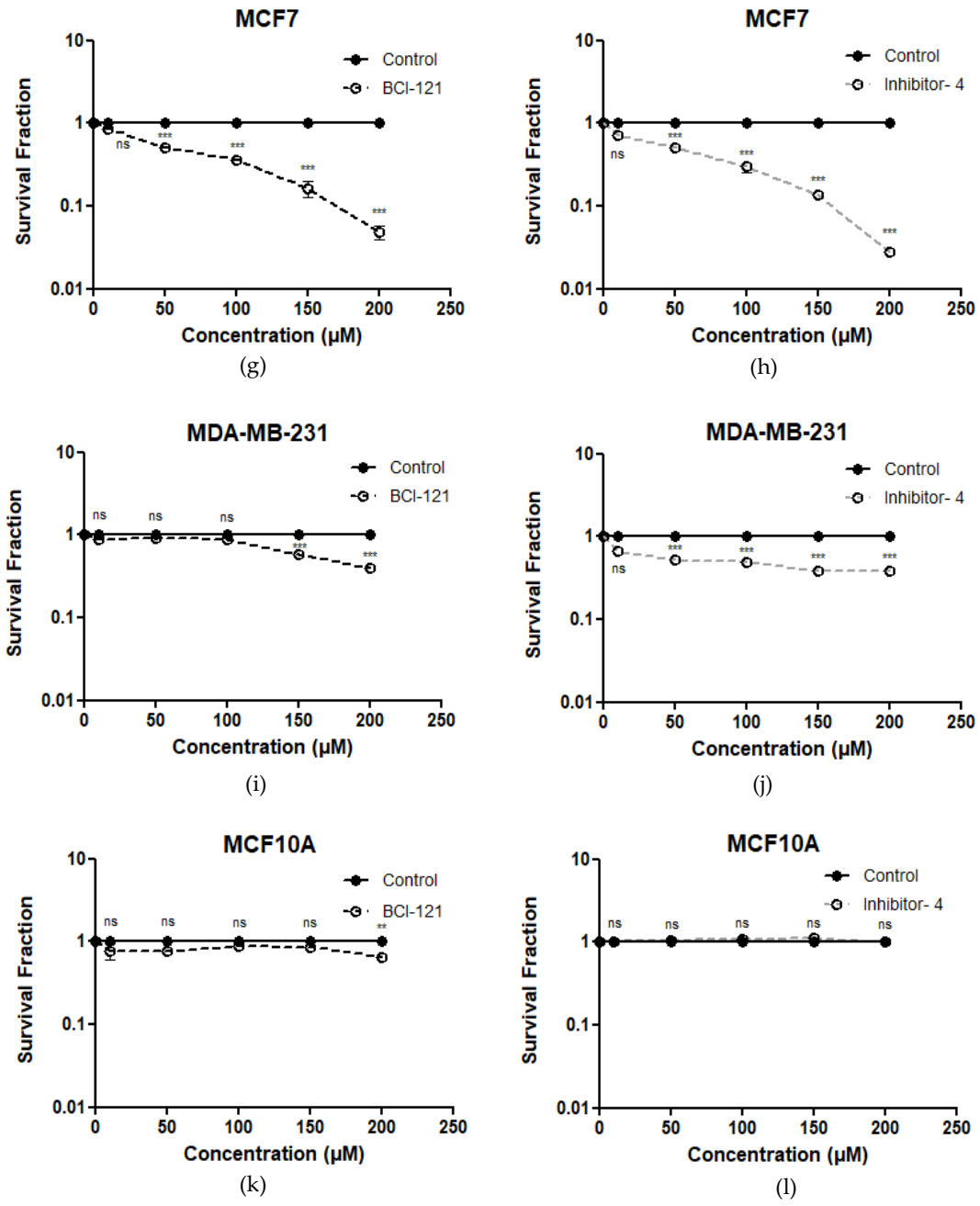


Figure 3.5: Clonogenic cell survival curve against BCI-121 (positive control inhibitor) and Inhibitor-4. (a, c, e, g, i, k) BCI-121 impact on breast cancer cell lines (a, c, g and i) and normal breast epithelial cell line (e and k). (b, d, f, h, j, l) Inhibitor-4 impact on breast cancer cell lines (b, d, h and j) and normal breast epithelial cell line (f and l). Error bars display standard error of means. Statistically significant differences from control are indicated by ** *P < 0.01,* *** *P < 0.001 or* ns *P > 0.05*. At least three independent experiments were carried out.

3.6. Inhibitor-4 Reduces Cell Viability in MCF7 Cells

The effect of Inhibitor-4 on the viability of wild type and cancer cell lines was evaluated using an MTT assay at different time points (24, 48, 72, and 96 h). Cells were treated with the vehicle (DMSO 0.1%, 0.15% and 0.2%), BCI-121, or Inhibitor-4 (Figure 3.6). Treatment with BCI-121 caused significant decreases in cell viability in both breast cancer (MCF7 and MDA-MB-231) and wild type (MCF10A) cell lines at multiple time points, particularly at concentrations of 150 and 200 μM (Figure 3.6a, c, e). However, Inhibitor-4 caused significant decreases in cell viability only in the cancer cell lines (MCF7 at 150 and 200 μM, MDA-MB-231 at 200 μM; Figure 3.6b, d). No concentration of Inhibitor-4 impacted MCF10A cellular viability (Figure 3.6f). Collectively, these data suggest that Inhibitor-4 is a promising, cancer-specific inhibitor that reduces cancer cell line viability and growth without affecting normal cells.

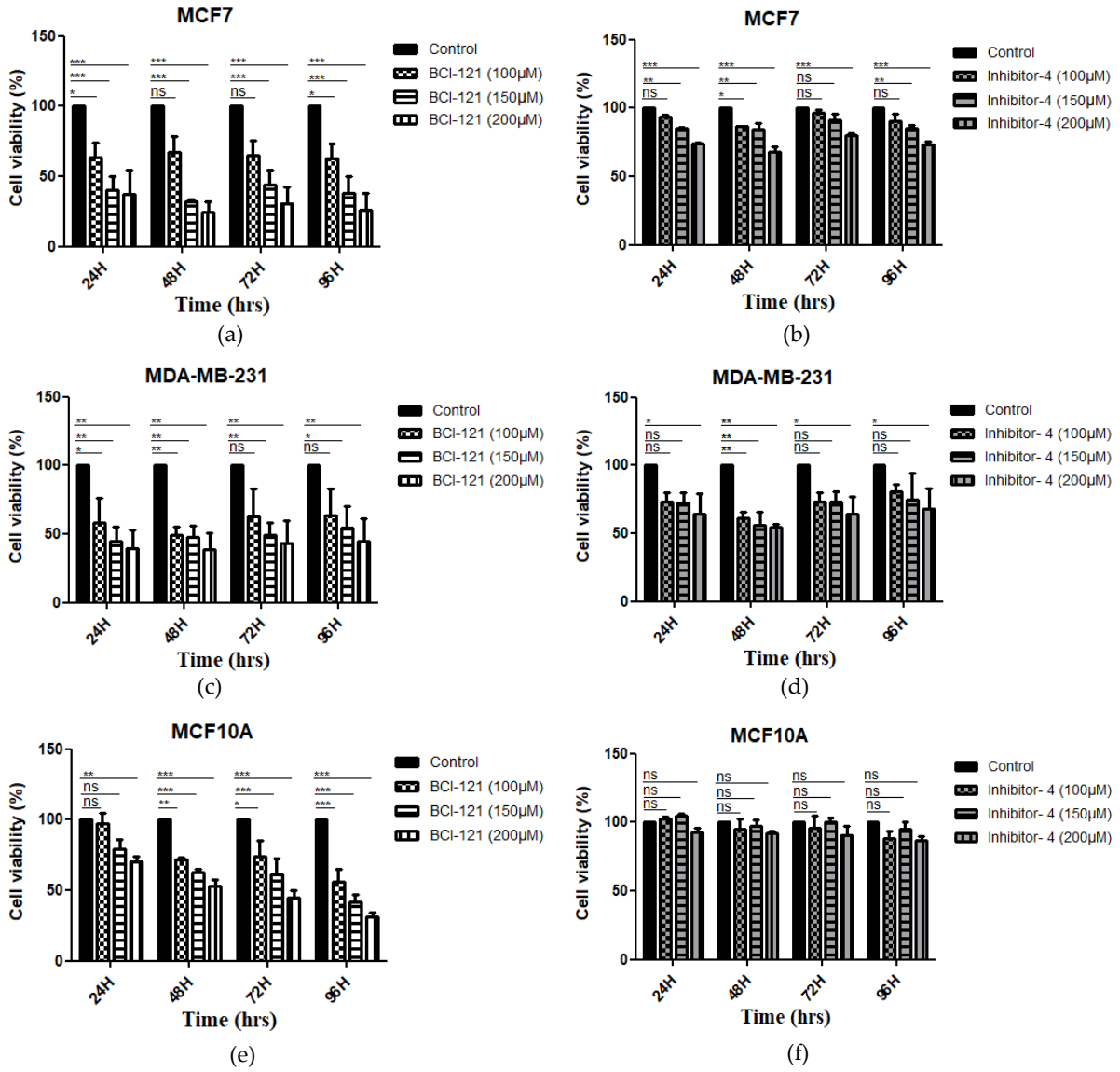
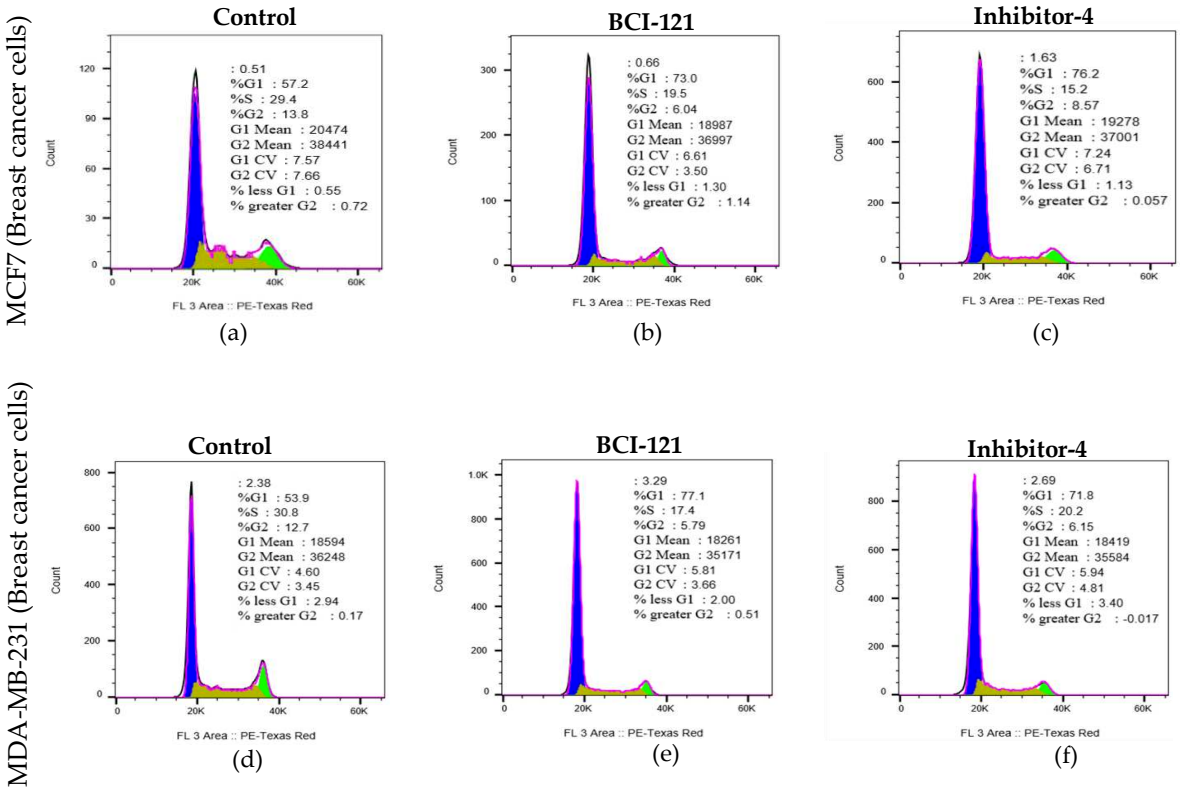


Figure 3.6: Cell viability using MTT assay. (a, c and e) BCI-121 effect on MCF7, MDA-MB-231 (breast cancer cell lines) and MCF10A (normal breast epithelial cell line). (b, d and f) Inhibitor-4 impact on MCF7, MDA-MB-231 (breast cancer cell lines) and MCF10A (normal breast epithelial cell line). Values are mean ± standard error of the means. Statistically significant differences from control are indicated by ** $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ or* ns $P > 0.05$.

3.7. Inhibitor-4 Induces Cell Cycle Arrest in Breast Cancer Cells

To investigate whether the growth inhibitory effect of Inhibitor-4 on breast cancer cells was due to cell cycle arrest, we conducted cell cycle analysis using Propidium Iodide (PI) staining. Cells were treated with 200 µM of BCI-121 or Inhibitor-4 for 24 h. As shown in Figure 3.7 treatments with both the positive control inhibitor and Inhibitor-4 induced G1 arrest and reduced S phase in MCF7 cells (Figure 3.7b, c, j, k) compared to MCF7 control (Figure 3.7a). Also, both treatments led to G1 arrest in MDA-MB-231 (Figure 3.7e, f, l, m) compared to MDA-MB-231 control (Figure 3.7d). Therefore, BCI-121 and Inhibitor-4 prompted an increase in G1 fractions. However, treatments with BCI-121 and Inhibitor-4 did not induce cell cycle arrest (Figure 3.7h, i) compared to control (Figure 3.7g) or show statistical differences in normal MCF10A cells (Figure 3.7n, o).



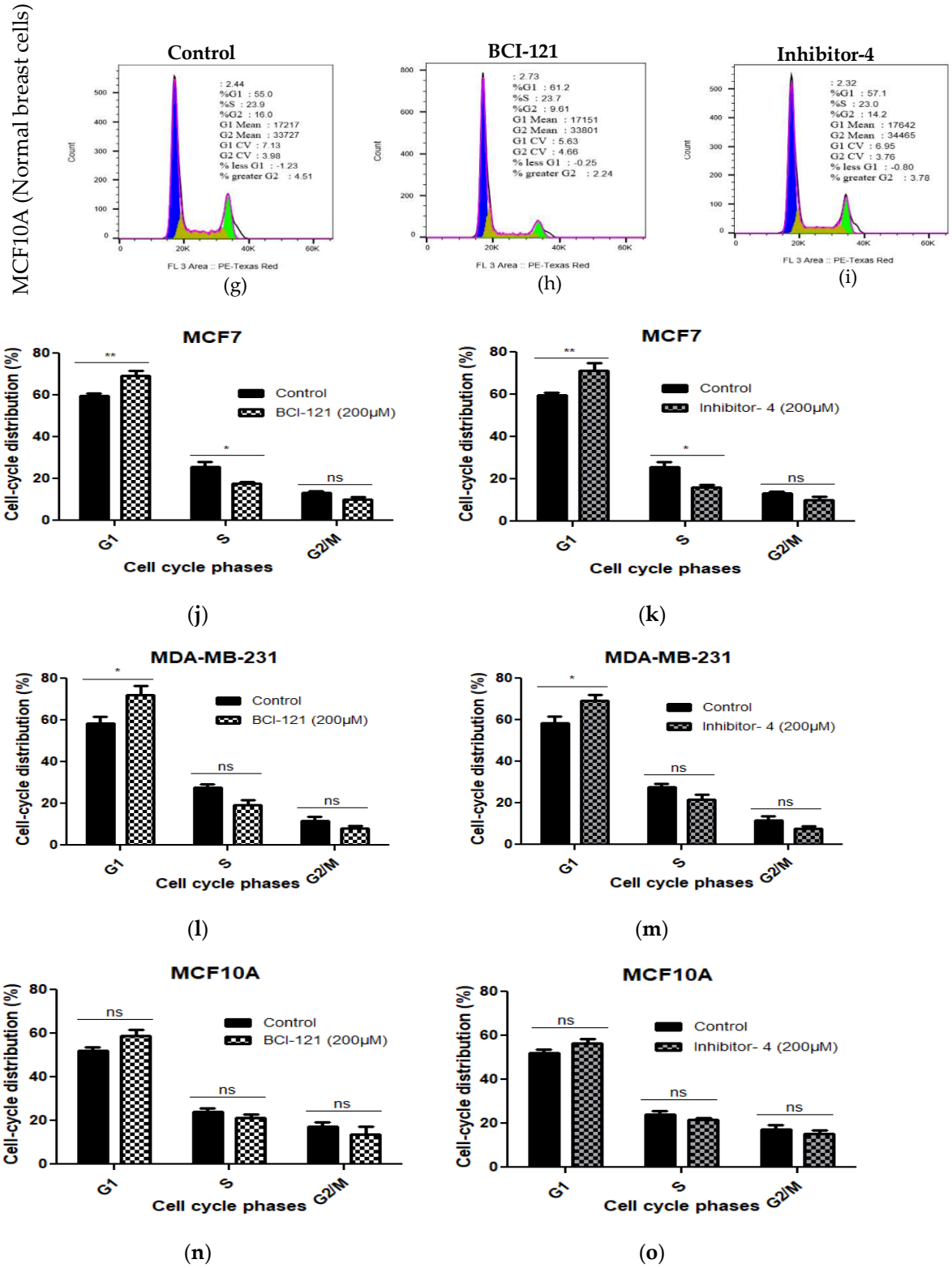


Figure 3.7: The cell cycle distribution was assessed using PI in MCF7, MDA-MB-231 (breast cancer cell lines) and MCF10A cells (normal breast epithelial cell line) with SMYD3 inhibitor treatments for 24 h and was investigated by flow cytometry. (a, d and

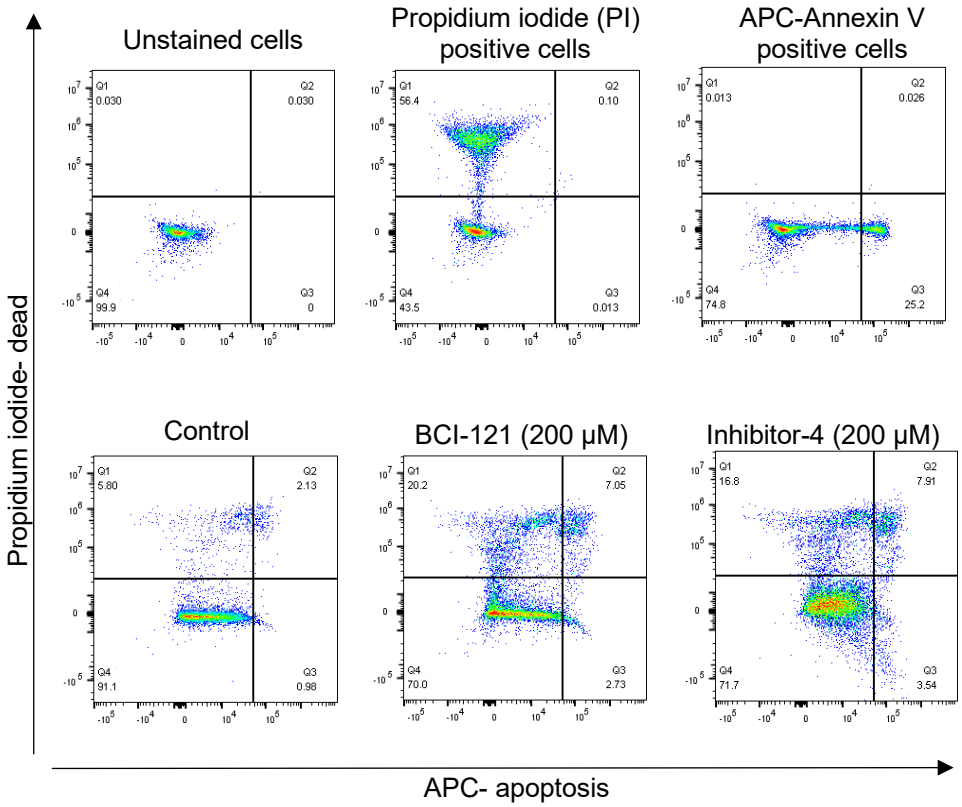
g) Untreated control distributions for MCF7, MDA-MB-231 and MCF10A. (b, e and h) Impact of BCI-121 on MCF7, MDA-MB-231 and MCF10A. (c, f and i) Effect of Inhibitor-4 on MCF7, MDA-MB-231 and MCF10A. (j, l, n) indicate statistical significant differences on cell cycle phases of the three cell lines treated with BCI-121. (k, m, o) show statistically significant differences on cell cycle phases of the three cell lines treated with Inhibitor-4. Values are mean ± standard error of the means. Statistically significant differences from control are indicated by * $p < 0.05$, ** $p < 0.01$ or ns $p > 0.05$.

3.8. Inhibitor-4 Promotes Apoptosis in Breast Cancer Cells

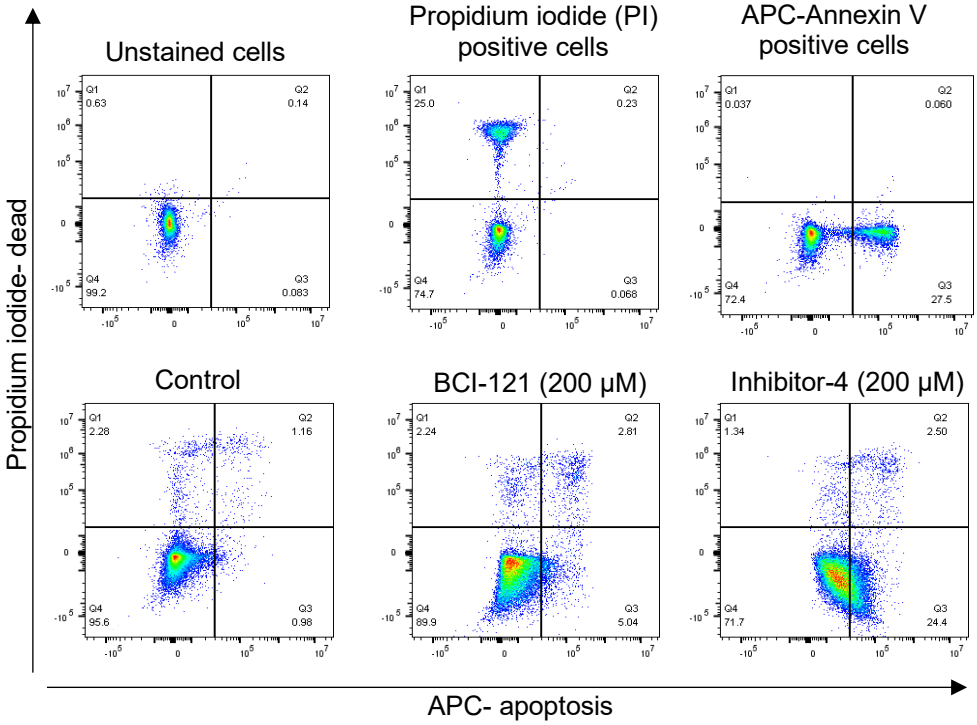
To reveal whether Inhibitor-4 induces apoptosis on breast cancer cell line or not, we performed apoptosis assay using APC Annexin V/PI followed by flow cytometry analysis. After 48 h of Inhibitor-4 treatment, the percentage of live cells decreased to 71% in both breast cancer cell lines (from 91% in MCF7 and 95% in MDA-MB-231) as demonstrated by flow cytometry (Figure 3.8 a, b, e, g). Also, treatment with Inhibitor-4 showed increase in late apoptosis and necrosis percentages in MCF7 (3.6a, e), while MDA-MB-231 showed early apoptosis with treatment of Inhibitor-4 (Figure 3.8b, g). BCI-121 caused late apoptosis in MCF7 and both early and late apoptosis in MDA-MB-231 cells, in addition to necrosis in MCF7 cells (Figure 3.8a, b, d, f). Neither treatment caused significant differences in apoptosis nor necrosis in MCF10A cells (Figure 3.8c, h, i).

Apoptosis induction through SMYD3 inhibitors was also tested using Caspase-3/7 activity assay. The data have shown increases in Caspase-3/7 activity in MDA-MB-231, however, no significant differences in MCF7, which is Caspase-3/7 independent apoptosis pathway, and MCF10-A (Figure 3.8j, k, l).

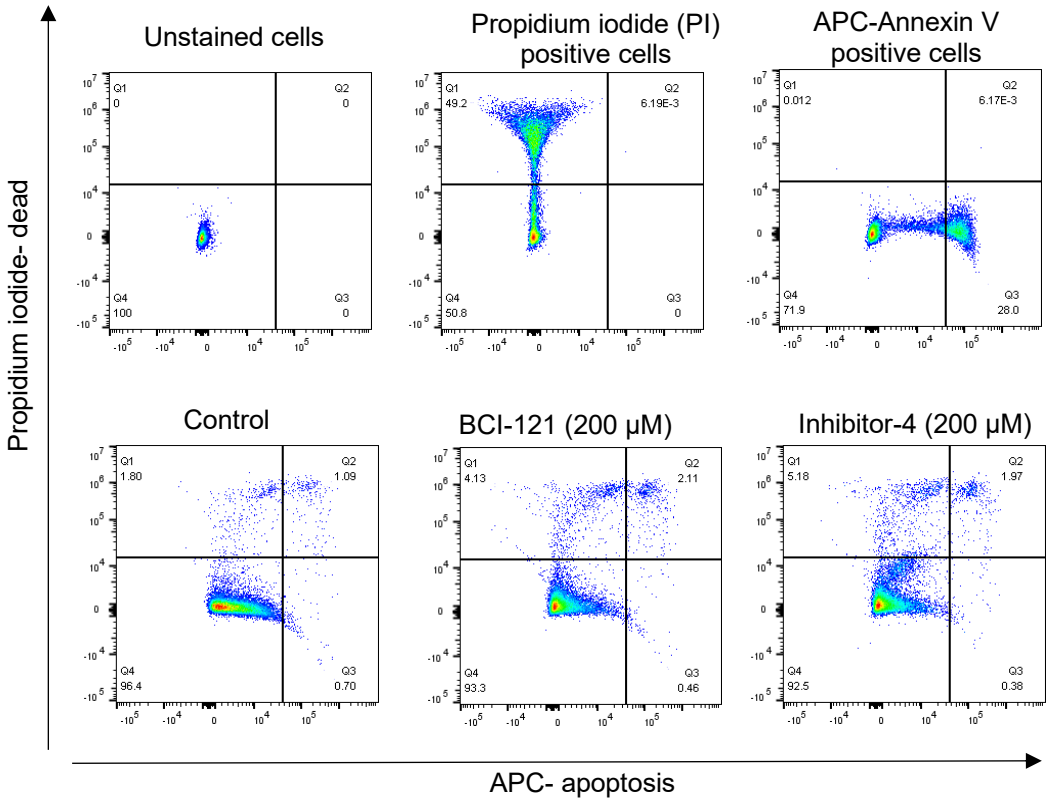
a)

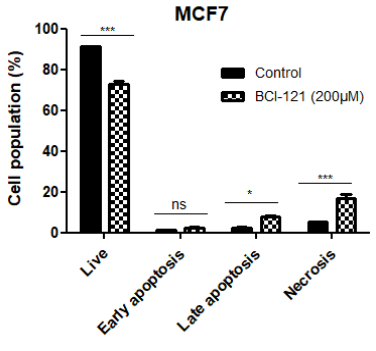

MCF7 (Breast cancer cells)
Unstained cells
Propidium iodide (PI) positive cells
APC-Annexin V positive cells
Control
BCI-121 (200 µM)
Inhibitor-4 (200 µM)
Propidium iodide- dead
APC- apoptosis

b)

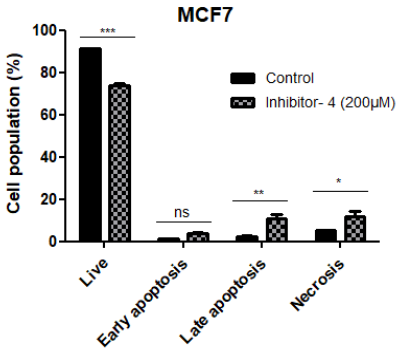

MDA-MB-231 (Breast cancer cells)
Unstained cells
Propidium iodide (PI) positive cells
APC-Annexin V positive cells
Control
BCI-121 (200 µM)
Inhibitor-4 (200 µM)
Propidium iodide- dead
APC- apoptosis

c)


MCF10A (Normal breast cells)
Unstained cells
Propidium iodide (PI) positive cells
APC-Annexin V positive cells
Control
BCI-121 (200 µM)
Inhibitor-4 (200 µM)
Propidium iodide- dead
APC- apoptosis
Q1 0
Q2 0
Q4 100
Q3 0
Q1 49.2
Q2 6.19E-3
Q4 50.8
Q3 0
Q1 0.012
Q2 6.17E-3
Q4 71.9
Q3 28.0
Q1 1.80
Q2 1.09
Q4 96.4
Q3 0.70
Q1 4.13
Q2 2.11
Q4 93.3
Q3 0.46
Q1 5.18
Q2 1.97
Q4 92.5
Q3 0.38


MCF7
Cell population (%)
Control
BCI-121 (200µM)
***
ns
*
***
Live
Early apoptosis
Late apoptosis
Necrosis

(**d**)


MCF7
Cell population (%)
Control
Inhibitor- 4 (200µM)
***
ns
**
*
Live
Early apoptosis
Late apoptosis
Necrosis

(**e**)

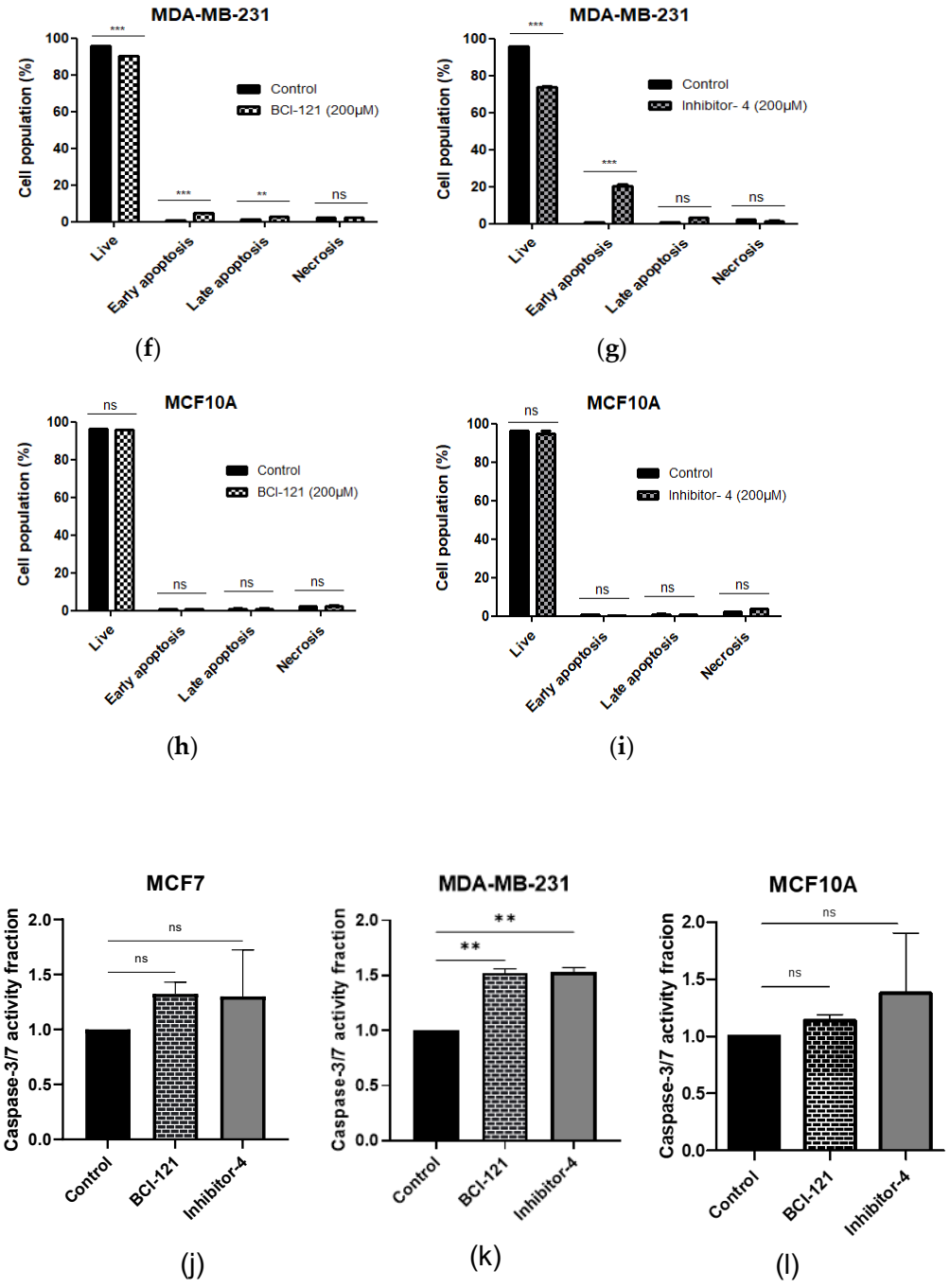


Figure 3.8: Cell apoptosis was assessed using APC Annexin V/PI and flow cytometry. (**a**, **d, e**) MCF7 (breast cancer), (**b, f, g**) MDA-MB-231 (breast cancer) and (**c, h, i**) MCF10A (normal breast) cell lines were treated with SMYD3 inhibitors for 48 h. (**d**, **f**, **h**) Apoptosis after 48 h of BCI-121 treatment on each cell line. (**e**, **g**, **i**) Apoptosis after 48 h of Inhibitor-4 treatment on each cell line. (**j**–**l**) Cell apoptosis was investigated with SMYD3 inhibitors using Caspase-3/7 activity assay. Values are mean ± standard error of the means. statistically significant differences from control are indicated by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ or ns $p > 0.05$.

CHAPTER 4– DISCUSSION AND CONCLUSION

4.1. Discussion:

Aberrant expression of SMYD3 has been shown to be oncogenic and is essential for the proliferation of most colorectal, hepatocellular, and breast carcinomas, as well as prostate cancer [3, 8]. Over 80 genes (including highly regulated homeobox genes, cell cycle regulators, and oncogenes) display altered expression because of aberrant upregulation of SMYD proteins [1-3]. Specifically, SMYD3 over-expression is highly associated with cancer development by regulating tumor proliferation, metastasis, invasion, and apoptosis [113]. Several studies have shown that SMYD3 regulates the oncogenic RAS signaling pathway by integrating a cytoplasmic-kinase signaling cascade, resulting in accelerated cell proliferation and differentiation [9]. Another study demonstrated that SMYD3 is essential for estrogen receptor-mediated transcription in breast cancer cells by down-regulating SMYD3 via RNA interference [15]. SMYD3 mediated-H2A.Z methylation has also been shown to trigger cyclin A1 gene expression, leading to cell cycle activation in breast cancer cells [114]. Knock down of SMYD3 in ovarian cancer tissues leads to upregulation of CDKN2B (p15INK4B), CDKN2A (p16INK4), CDC25A and CDKN3 as members of cyclin-dependent kinase inhibitors (CDK) [96]. Inducing apoptosis via silencing of SMYD3 has also been observed in ovarian cancer in vivo and has been accredited to the upregulation of CD40LG and downregulation of BIRC3 [96]. BIRC3 is a member of the inhibitors of apoptosis proteins (IAP) family and relates to many cancers in cases of aberrant overexpression because it can prevent apoptotic signals [115, 116]. Therefore, it is likely that SMYD3 inhibitors can trigger apoptosis by down-regulating BIRC3. In addition, another study demonstrated

that the MCF7 cell line lacks Caspase 3, which is essential for apoptosis, however in the absence of Caspase 3, Caspase 6 can be activated as an alternative mechanism to trigger apoptosis. As a result, under cellular stress, MCF7 cells undergo apoptosis in response to Caspase 6 and necrosis in response to TNF-α stimulation [117-119]. Despite the connection between SMYD3-overexprssion and several types of carcinogenesis, few studies have targeted SMYD3 inhibition in the context of breast cancer through the design of the inhibitors.

In this study, we sought to design small molecule inhibitors for the inhibition of SMYD3-mediated methylation (Figure 3.1), proliferation (Figure 3.4), colony formation (Figure 3.5), and viability (Figure 3.6) in breast cancer cells. Specifically, we demonstrated that *in silico* enzyme models can predict effective competitive enzyme inhibitors by screening vast molecular libraries and predicting binding energies. This approach to small molecule design significantly reduces the time, expense, and equipment that have been required for traditional benchtop small molecule screening until now.

Using Schrodinger® software and several *in vitro* assays, we demonstrated that one of the hit compounds identified *in silico* (Inhibitor-4) was able to reduce breast cancer cellular growth and viability without affecting normal breast epithelial cells. *In vitro*, Inhibitor-4 was shown to inhibit SMYD3-mediated histone methylation. In breast cancer cells, Inhibitor-4 extended cell doubling time (Figure 3.4). We also demonstrated that Inhibitor-4 arrests the cell cycle in breast cancer cells without affecting normal cells (Figure 3.7), which demonstrates an improvement over BCI-121, a previously-developed SMYD3 inhibitor. Finally, the novel SMYD3 inhibitor presented here caused apoptosis in breast cancer cell lines without affecting the normal breast cell line (Figure

3.8). However, testing the *in vivo* SMYD3 specificity of Inhibitor-4 needs to investigate the impacts of Inhibitor-4 on SMYD3-knockdown cells. This could be performed for future characterization and validation.

4.2. Conclusions:

Epigenetic control is tightly regulated in humans, and aberrant epigenetic marks are associated with several disease pathologies. Histone lysine methylation is a critical epigenetic regulator and is modulated chiefly by SET-domain-containing enzymes. Specifically, SMYD3 is a SET and MYND domain-containing enzyme that methylates both histone and non-histone targets. The overexpression of SMYD3 has been discovered in several cancer types, including breast, colorectal, and hepatocellular carcinomas. Therefore, the inhibition of SMYD3 is promising for the therapeutic treatment of these cancer types.

In conclusion, our study established SMYD3 as a potential target for the clinical management of breast cancer. As a result, we have successfully developed and used *in silico* compound screening inhibitors that target the catalytic domain of SMYD3. Consequently, Inhibitor-4 has shown a very critical role in silencing SMYD3 methyltransferase activity *in vitro*. In phenotypic assays, it has been shown to halt SMYD3-mediated cell proliferation, arrest the cell cycle and induce apoptosis in breast cancer cell lines compared with the standard breast cell line.

4.3 Future directions:

Here we have covered previous studies that proved SMYD3 expression increases during carcinogenesis. SMYD3 is overexpressed in several cancers, including breast,

colorectal and hepatocellular carcinomas. Consequently, we developed and confirmed a SMYD3 catalytic inhibitor's ability to halt cell cycle progression and induce apoptosis in breast carcinoma cells. Our future directions include evaluating the ability of the SMYD3 inhibitor to halt cell proliferation, arrest the cell cycle and induce apoptosis in colorectal and hepatocellular carcinomas.

In addition, we aim to target another member of the SMYD family (SMYD2), which is involved in the induction of leukemias. SMYD2 is typically involved in the regulation of transcription, cell proliferation, apoptosis. Its methylation targets include histones, p53, and RB. SMYD2-mediated methylation of p53K370 inhibits p53-mediated apoptosis. SMYD2 is involved in the regulation of lineage progression during hematopoiesis; its overexpression is essential for the transformation and survival of cells involved in chronic myeloid leukemia (CML) and Mixed-lineage leukemia (MLL). Consequently, we aim to design SMYD2 catalytic inhibitors and confirm these inhibitors' ability to halt cell cycle progression and induce apoptosis.

Regarding *in silico* screening, our research continues to optimize SMYD3 inhibitors *in vitro* and test them using various cancer cell lines. To do this, we will perform structure-activity relationship (SAR) studies to optimize Inhibitor-4. The core structures of Inhibitor-4 and Bcl-121 are identical and vary only in the small R-group attached to the benzyl ring. Bcl-121 has only a Bromine atom in this location, while Inhibitor-4 has nitrogen covalently linked to two ethyl groups. These two R-groups differ significantly in hydrophilicity. Therefore, modification of this region of the inhibitors is a promising location for altering specificity and efficacy.

We also aim to assess the cellular mechanisms of SMYD3 inhibitors to halt cell cycle progression and induce apoptosis in breast, colorectal and hepatocellular carcinoma cell lines. RNA sequencing will be applied to understand better the gene expression profiles responsible for cell proliferation and apoptosis in breast, colorectal and hepatocellular carcinomas. SMYD3 methylates histone and non-histone proteins. Therefore, we aim to study the effects of SMYD3 inhibitors in pathways associated with DNA damage repair (HR repair proteins), angiogenesis (VEGFR1), proliferation (MAP3K2), and metastasis (MMP9).

Ultimately, we will conduct our study in an animal model to establish proof-of-concept for the use of SMYD3 inhibitors in the clinical management of breast, colorectal, and hepatocellular carcinomas.

REFERENCES

1. Brown, M.A., et al., *Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex.* Molecular cancer, 2006. **5**(1): p. 26.
2. Gottlieb, P.D., et al., *Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis.* Nature genetics, 2002. **31**(1): p. 25-32.
3. Hamamoto, R., et al., *SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells.* Nature cell biology, 2004. **6**(8): p. 731-740.
4. Brown, M.A., et al., *C-terminal domain of SMYD3 serves as a unique HSP90-regulated motif in oncogenesis.* Oncotarget, 2015. **6**(6): p. 4005.
5. Komatsu, S., et al., *Overexpression of SMYD2 relates to tumor cell proliferation and malignant outcome of esophageal squamous cell carcinoma.* Carcinogenesis, 2009. **30**(7): p. 1139-1146.
6. Cock-Rada, A.M., et al., *SMYD3 promotes cancer invasion by epigenetic upregulation of the metalloproteinase MMP-9.* Cancer research, 2012. **72**(3): p. 810-820.
7. Luo, X.-G., et al., *Histone methyltransferase SMYD3 promotes MRTF-A-mediated transactivation of MYL9 and migration of MCF-7 breast cancer cells.* Cancer letters, 2014. **344**(1): p. 129-137.
8. Hamamoto, R., et al., *Enhanced SMYD3 expression is essential for the growth of breast cancer cells.* Cancer science, 2006. **97**(2): p. 113-118.
9. Mazur, P.K., et al., *SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer.* Nature, 2014. **510**(7504): p. 283-287.
10. Luo, X.G., et al., *Effects of SMYD3 overexpression on transformation, serum dependence, and apoptosis sensitivity in NIH3T3 cells.* IUBMB life, 2009. **61**(6): p. 679-684.
11. Liu, C., et al., *The telomerase reverse transcriptase (hTERT) gene is a direct target of the histone methyltransferase SMYD3.* Cancer research, 2007. **67**(6): p. 2626-2631.
12. Zou, J.-N., et al., *Knockdown of SMYD3 by RNA interference down-regulates c-Met expression and inhibits cells migration and invasion induced by HGF.* Cancer letters, 2009. **280**(1): p. 78-85.
13. Ren, T.-n., et al., *Effects of SMYD3 over-expression on cell cycle acceleration and cell proliferation in MDA-MB-231 human breast cancer cells.* Medical Oncology, 2011. **28**(1): p. 91-98.
14. Kunizaki, M., et al., *The lysine 831 of vascular endothelial growth factor receptor 1 is a novel target of methylation by SMYD3.* Cancer research, 2007. **67**(22): p. 10759-10765.
15. Kim, H., et al., *Requirement of histone methyltransferase SMYD3 for estrogen receptor-mediated transcription.* Journal of Biological Chemistry, 2009. **284**(30): p. 19867-19877.

16. Liu, L., et al., *Genetic alterations of histone lysine methyltransferases and their significance in breast cancer.* Oncotarget, 2015. **6**(4): p. 2466.
17. Ma, S.-J., et al., *Correlations of EZH2 and SMYD3 gene polymorphisms with breast cancer susceptibility and prognosis.* Bioscience reports, 2018. **38**(1).
18. Van Aller, G.S., et al., *Smyd3 regulates cancer cell phenotypes and catalyzes histone H4 lysine 5 methylation.* Epigenetics, 2012. **7**(4): p. 340-343.
19. Peserico, A., et al., *A SMYD3 small-molecule inhibitor impairing cancer cell growth.* Journal of cellular physiology, 2015. **230**(10): p. 2447-2460.
20. Sharma, G.N., et al., *Various types and management of breast cancer: an overview.* Journal of advanced pharmaceutical technology & research, 2010. **1**(2): p. 109.
21. Bray, F., et al., *Cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin, 2018. **68**(6): p. 394-424.
22. Waks, A.G. and E.P. Winer, *Breast cancer treatment: a review.* Jama, 2019. **321**(3): p. 288-300.
23. McPherson, K., C. Steel, and J. Dixon, *ABC of breast diseases: breast cancer—epidemiology, risk factors, and genetics.* BMJ: British Medical Journal, 2000. **321**(7261): p. 624.
24. Armaou, S., et al., *Contribution of BRCA1 germ-line mutations to breast cancer in Greece: a hospital-based study of 987 unselected breast cancer cases.* British journal of cancer, 2009. **101**(1): p. 32-37.
25. Martínez-Ferrandis, J., et al., *Mutational analysis of BRCA1 and BRCA2 in Mediterranean Spanish women with early-onset breast cancer: Identification of three novel pathogenic mutations.* Human mutation, 2003. **22**(5): p. 417-418.
26. Ayub, S.G., et al., *Mutational analysis of the BRCA2 gene in breast carcinoma patients of Kashmiri descent.* Molecular medicine reports, 2014. **9**(2): p. 749-753.
27. Calderón-Garcidueñas, A.L., et al., *Clinical follow up of Mexican women with early onset of breast cancer and mutations in the BRCA1 and BRCA2 genes.* Salud publica de Mexico, 2005. **47**: p. 110-115.
28. Mehrgou, A. and M. Akouchekian, *The importance of BRCA1 and BRCA2 genes mutations in breast cancer development.* Medical journal of the Islamic Republic of Iran, 2016. **30**: p. 369.
29. Osborne, C., P. Wilson, and D. Tripathy, *Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications.* The oncologist, 2004. **9**(4): p. 361-377.
30. Dougan, M. and G. Dranoff, *The immune response to tumors.* Current Protocols in Immunology, 2009. **85**(1): p. 20.11. 1-20.11. 4.
31. Dougan, M. and G. Dranoff, *Immunotherapy of cancer*, in *Innate immune regulation and cancer immunotherapy*. 2012, Springer. p. 391-414.
32. Clarke, M., et al., *Early Breast Cancer Trialists Collaborative Group (EBCTCG). Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials.* Lancet, 2005. **366**(9503): p. 2087-2106.
33. Lewis, J.S. and V.C. Jordan, *Selective estrogen receptor modulators (SERMs): mechanisms of anticarcinogenesis and drug resistance.* Mutation

Research/Fundamental and Molecular Mechanisms of Mutagenesis, 2005. **591**(1-2): p. 247-263.
34. Ekins, S., J. Mestres, and B. Testa, *In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling.* British journal of pharmacology, 2007. **152**(1): p. 9-20.
35. Lengauer, T., et al., *Novel technologies for virtual screening.* Drug discovery today, 2004. **9**(1): p. 27-34.
36. Johnson, M.A. and G.M. Maggiora, *Concepts and applications of molecular similarity*. 1990: Wiley.
37. Willett, P., *Similarity-based approaches to virtual screening*. 2003, Portland Press Ltd.
38. Klebe, G., *Virtual ligand screening: strategies, perspectives and limitations.* Drug discovery today, 2006. **11**(13-14): p. 580-594.
39. Shoichet, B.K., *Virtual screening of chemical libraries.* Nature, 2004. **432**(7019): p. 862-865.
40. Iwata, H., R. Kojima, and Y. Okuno, *An in silico Approach for Integrating Phenotypic and Target-based Approaches in Drug Discovery.* Molecular informatics, 2020. **39**(1-2): p. 1900096.
41. Muegge, I. and I.J. Enyedy, *Virtual screening for kinase targets.* Current medicinal chemistry, 2004. **11**(6): p. 693-707.
42. Sharma, R., *Enzyme inhibition and bioapplications*. 2012: BoD–Books on Demand.
43. Metzler, D.E., *Biochemistry (2 Volume Set): The Chemical Reactions of Living Cells*. 2003: Elsevier.
44. Turner, B.M., *Cellular memory and the histone code.* Cell, 2002. **111**(3): p. 285-291.
45. Bonasio, R., S. Tu, and D. Reinberg, *Molecular signals of epigenetic states.* science, 2010. **330**(6004): p. 612-616.
46. Lee, J.-S., E. Smith, and A. Shilatifard, *The language of histone crosstalk.* Cell, 2010. **142**(5): p. 682-685.
47. Felsenfeld, G. and M. Groudine, *Controlling the double helix.* Nature, 2003. **421**(6921): p. 448-453.
48. Richmond, T.J., *Predictable packaging.* Nature, 2006. **442**(7104): p. 750-751.
49. Kornberg, R.D. and Y. Lorch, *Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome.* Cell, 1999. **98**(3): p. 285-294.
50. Trerotola, M., et al., *Epigenetic inheritance and the missing heritability.* Human genomics, 2015. **9**(1): p. 17.
51. Fletcher, T.M. and J.C. Hansen, *The nucleosomal array: structure/function relationships.* Critical Reviews™ in Eukaryotic Gene Expression, 1996. **6**(2-3).
52. Hayes, J.J., D.J. Clark, and A.P. Wolffe, *Histone contributions to the structure of DNA in the nucleosome.* Proceedings of the National Academy of Sciences, 1991. **88**(15): p. 6829-6833.
53. Rister, J. and C. Desplan, *Deciphering the genome's regulatory code: the many languages of DNA.* Bioessays, 2010. **32**(5): p. 381-384.
54. Hager, G.L., J.G. McNally, and T. Misteli, *Transcription dynamics.* Molecular cell, 2009. **35**(6): p. 741-753.

55. Vitolo, J.M., C. Thiriet, and J.J. Hayes, *The H3-H4 N-terminal tail domains are the primary mediators of transcription factor IIIA access to 5S DNA within a nucleosome.* Molecular and Cellular Biology, 2000. **20**(6): p. 2167-2175.
56. Hansen, J.C., *Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions.* Annual review of biophysics and biomolecular structure, 2002. **31**(1): p. 361-392.
57. Zhang, Y. and D. Reinberg, *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.* Genes & development, 2001. **15**(18): p. 2343-2360.
58. Khorasanizadeh, S., *The nucleosome: from genomic organization to genomic regulation.* Cell, 2004. **116**(2): p. 259-272.
59. Rice, J.C., et al., *Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains.* Molecular cell, 2003. **12**(6): p. 1591-1598.
60. Sims III, R.J., K. Nishioka, and D. Reinberg, *Histone lysine methylation: a signature for chromatin function.* TRENDS in Genetics, 2003. **19**(11): p. 629-639.
61. Seligson, D.B., et al., *Global histone modification patterns predict risk of prostate cancer recurrence.* Nature, 2005. **435**(7046): p. 1262-1266.
62. Ayton, P.M. and M.L. Cleary, *Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins.* Oncogene, 2001. **20**(40): p. 5695-5707.
63. Xiao, B., J.R. Wilson, and S.J. Gamblin, *SET domains and histone methylation.* Current opinion in structural biology, 2003. **13**(6): p. 699-705.
64. Dillon, S.C., et al., *The SET-domain protein superfamily: protein lysine methyltransferases.* Genome biology, 2005. **6**(8): p. 227.
65. Jenuwein, T., et al., *SET domain proteins modulate chromatin domains in eu-and heterochromatin.* Cellular and Molecular Life Sciences CMLS, 1998. **54**(1): p. 80-93.
66. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications.* Cell research, 2011. **21**(3): p. 381-395.
67. Briggs, S.D., et al., *Trans-histone regulatory pathway in chromatin.* Nature, 2002. **418**(6897): p. 498-498.
68. Lachner, M., R.J. O'Sullivan, and T. Jenuwein, *An epigenetic road map for histone lysine methylation.* Journal of cell science, 2003. **116**(11): p. 2117-2124.
69. Rea, S., et al., *Regulation of chromatin structure by site-specific histone H3 methyltransferases.* Nature, 2000. **406**(6796): p. 593-599.
70. Rice, J.C. and C.D. Allis, *Code of silence.* Nature, 2001. **414**(6861): p. 259-261.
71. Varier, R.A. and H.M. Timmers, *Histone lysine methylation and demethylation pathways in cancer.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2011. **1815**(1): p. 75-89.
72. Min, J., et al., *Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase.* Cell, 2003. **112**(5): p. 711-723.
73. Huang, J., et al., *Repression of p53 activity by Smyd2-mediated methylation.* Nature, 2006. **444**(7119): p. 629-632.
74. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-1080.

75. Kouzarides, T., *Histone methylation in transcriptional control.* Current opinion in genetics & development, 2002. **12**(2): p. 198-209.
76. O'Carroll, D., et al., *The polycomb-group GeneEzh2 is required for early mouse development.* Molecular and cellular biology, 2001. **21**(13): p. 4330-4336.
77. Sims, R.J., et al., *m-Bop, a repressor protein essential for cardiogenesis, interacts with skNAC, a heart-and muscle-specific transcription factor.* Journal of Biological Chemistry, 2002. **277**(29): p. 26524-26529.
78. Sesé, B., et al., *SMYD2 is induced during cell differentiation and participates in early development.* International Journal of Developmental Biology, 2013. **57**(5): p. 357-364.
79. Cho, H.-S., et al., *RB1 methylation by SMYD2 enhances cell cycle progression through an increase of RB1 phosphorylation.* Neoplasia, 2012. **14**(6): p. 476-IN8.
80. Komatsu, S., et al., *Overexpression of SMYD2 contributes to malignant outcome in gastric cancer.* British journal of cancer, 2015. **112**(2): p. 357-364.
81. Pimkhaokham, A., et al., *Nonrandom chromosomal imbalances in esophageal squamous cell carcinoma cell lines: possible involvement of the ATF3 and CENPF genes in the 1q32 amplicon.* Japanese journal of cancer research, 2000. **91**(11): p. 1126-1133.
82. Sakamoto, L.H.T., et al., *SMYD2 is highly expressed in pediatric acute lymphoblastic leukemia and constitutes a bad prognostic factor.* Leukemia research, 2014. **38**(4): p. 496-502.
83. Abu-Farha, M., et al., *The tale of two domains: proteomics and genomics analysis of SMYD2, a new histone methyltransferase.* Molecular & cellular proteomics, 2008. **7**(3): p. 560-572.
84. Jiang, Y., et al., *Crystal structures of histone and p53 methyltransferase SmyD2 reveal a conformational flexibility of the autoinhibitory C-terminal domain.* PloS one, 2011. **6**(6): p. e21640.
85. Sajjad, A., et al., *Lysine methyltransferase Smyd2 suppresses p53-dependent cardiomyocyte apoptosis.* Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2014. **1843**(11): p. 2556-2562.
86. Wang, L., et al., *Structure of human SMYD2 protein reveals the basis of p53 tumor suppressor methylation.* Journal of Biological Chemistry, 2011. **286**(44): p. 38725-38737.
87. Katayama, M., et al., *Gene trio signatures as molecular markers to predict response to doxorubicin cyclophosphamide neoadjuvant chemotherapy in breast cancerpatients.* Brazilian Journal of Medical and Biological Research, 2010. **43**(12): p. 1225-1231.
88. Jiang, Y., et al., *Structural insights into estrogen receptor α methylation by histone methyltransferase SMYD2, a cellular event implicated in estrogen signaling regulation.* Journal of molecular biology, 2014. **426**(20): p. 3413-3425.
89. Zhang, X., et al., *Regulation of estrogen receptor α by histone methyltransferase SMYD2-mediated protein methylation.* Proceedings of the National Academy of Sciences, 2013. **110**(43): p. 17284-17289.
90. Piao, L., et al., *The histone methyltransferase SMYD2 methylates PARP1 and promotes poly (ADP-ribosyl) ation activity in cancer cells.* Neoplasia, 2014. **16**(3): p. 257-264. e2.

91. Abu-Farha, M., et al., *Proteomic analyses of the SMYD family interactomes identify HSP90 as a novel target for SMYD2.* Journal of molecular cell biology, 2011. **3**(5): p. 301-308.
92. Hamamoto, R., et al., *SMYD2-dependent HSP90 methylation promotes cancer cell proliferation by regulating the chaperone complex formation.* Cancer letters, 2014. **351**(1): p. 126-133.
93. Donlin, L.T., et al., *Smyd2 controls cytoplasmic lysine methylation of Hsp90 and myofilament organization.* Genes & development, 2012. **26**(2): p. 114-119.
94. Voelkel, T., et al., *Lysine methyltransferase Smyd2 regulates Hsp90-mediated protection of the sarcomeric titin springs and cardiac function.* Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2013. **1833**(4): p. 812-822.
95. Gaedcke, J., et al., *Mutated KRAS results in overexpression of DUSP4, a MAP-kinase phosphatase, and SMYD3, a histone methyltransferase, in rectal carcinomas.* Genes, chromosomes and cancer, 2010. **49**(11): p. 1024-1034.
96. Jiang, Y., et al., *Overexpression of SMYD3 in Ovarian Cancer is Associated with Ovarian Cancer Proliferation and Apoptosis via Methylating H3K4 and H4K20.* Journal of Cancer, 2019. **10**(17): p. 4072.
97. Friesner, R.A., et al., *Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein− ligand complexes.* Journal of medicinal chemistry, 2006. **49**(21): p. 6177-6196.
98. Greenwood, J.R., et al., *Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution.* Journal of computer-aided molecular design, 2010. **24**(6-7): p. 591-604.
99. Shelley, J.C., et al., *Epik: a software program for pK a prediction and protonation state generation for drug-like molecules.* Journal of computer-aided molecular design, 2007. **21**(12): p. 681-691.
100. Sterling, T. and J.J. Irwin, *ZINC 15–ligand discovery for everyone.* Journal of chemical information and modeling, 2015. **55**(11): p. 2324-2337.
101. Fu, W., et al., *Structural basis for substrate preference of SMYD3, a SET domain-containing protein lysine methyltransferase.* Journal of Biological Chemistry, 2016. **291**(17): p. 9173-9180.
102. Koehn, J.T., et al., *A synthetic isoprenoid lipoquinone, menaquinone-2, adopts a folded conformation in solution and at a model membrane interface.* The Journal of organic chemistry, 2017. **83**(1): p. 275-288.
103. Debnath, J., S.K. Muthuswamy, and J.S. Brugge, *Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures.* Methods, 2003. **30**(3): p. 256-268.
104. Roger, S., P. Besson, and J.-Y. Le Guennec, *Involvement of a novel fast inward sodium current in the invasion capacity of a breast cancer cell line.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 2003. **1616**(2): p. 107-111.
105. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Analytical biochemistry, 1976. **72**(1-2): p. 248-254.
106. Chiacchiera, F., et al., *p38 α blockade inhibits colorectal cancer growth* in vivo *by inducing a switch from HIF1 α-to FoxO-dependent transcription.* Cell Death & Differentiation, 2009. **16**(9): p. 1203-1214.

107. Maeda, J., et al., *Genomic instability and telomere fusion of canine osteosarcoma cells.* PloS one, 2012. **7**(8): p. e43355.
108. Kato, T.A., R. Okayasu, and J.S. Bedford, *Signatures of DNA double strand breaks produced in irradiated G1 and G2 cells persist into mitosis.* Journal of cellular physiology, 2009. **219**(3): p. 760-765.
109. Van Meerloo, J., G.J. Kaspers, and J. Cloos, *Cell sensitivity assays: the MTT assay*, in *Cancer cell culture*. 2011, Springer. p. 237-245.
110. Krishan, A., *Rapid flow cytofluorometric analysis of mammalian cell cycle by propidium iodide staining.* The Journal of cell biology, 1975. **66**(1): p. 188-193.
111. Vermes, I., et al., *A novel assay for apoptosis flow cytometric detection of phosphatidylserine expression on early apoptotic cells using fluorescein labelled annexin V.* Journal of immunological methods, 1995. **184**(1): p. 39-51.
112. Aubry, J.P., et al., *Annexin V used for measuring apoptosis in the early events of cellular cytotoxicity.* Cytometry: The Journal of the International Society for Analytical Cytology, 1999. **37**(3): p. 197-204.
113. Huang, L. and A.-M. Xu, *SET and MYND domain containing protein 3 in cancer.* American journal of translational research, 2017. **9**(1): p. 1.
114. Tsai, C.-H., et al., *SMYD3-mediated H2A. Z. 1 methylation promotes cell cycle and cancer proliferation.* Cancer research, 2016. **76**(20): p. 6043-6053.
115. Deveraux, Q.L. and J.C. Reed, *IAP family proteins—suppressors of apoptosis.* Genes & development, 1999. **13**(3): p. 239-252.
116. Kempkensteffen, C., et al., *Expression parameters of the inhibitors of apoptosis cIAP1 and cIAP2 in renal cell carcinomas and their prognostic relevance.* International journal of cancer, 2007. **120**(5): p. 1081-1086.
117. Jänicke, R.U., et al., *Caspase-3 is required for DNA fragmentation and morphological changes associated with apoptosis.* Journal of Biological Chemistry, 1998. **273**(16): p. 9357-9360.
118. Kagawa, S., et al., *Deficiency of caspase-3 in MCF7 cells blocks Bax-mediated nuclear fragmentation but not cell death.* Clinical cancer research, 2001. **7**(5): p. 1474-1480.
119. Tewari, M., et al., *Yama/CPP32β, a mammalian homolog of CED-3, is a CrmA-inhibitable protease that cleaves the death substrate poly (ADP-ribose) polymerase.* Cell, 1995. **81**(5): p. 801-809.

LIST OF ABBREVIATIONS

ABBREVIATION	MEANING	PAGE
SMYD	SET and MYND	iii
SET	The Suppressor of Variegation Su (var), Enhancer of Zeste E(z) and Trithorax	iii
MYND	Myeloid-Nervy-DEAF1	iii
H	Histone	iii
K	Lysine	iii
NKX2.8	NK2 Homeobox 8	iii
Wnt	Wnt family member 10	iii
TERT	Telomerese reverse transcriptase	iii
cMET	MET proto-oncogene, receptor tyrosine kinase	iii
CDK2	Cyclin-dependent kinase 2	iii
VEGFR-1	Vascular endothelial growth factor receptor-1	iii
ER	Estrogen receptors	iii
BRCA	Breast cancer	3
p53	Tumor suppressor gene	4
DCIS	ductal carcinoma in situ	4
LCIS	Lobular carcinoma in situ	4

PR	Progesterone	5
HER-2	human epidermal growth factor receptor 2	5
NCP	The nucleosome core particle	14
PTM	post-translational modifications	14
SAM	S-Adenosyl methionine	16
AdoHcy or SAH	S-adenosyl-L-homocysteine	16
Hsp90	Chaperone protein	20
RB	Retinoblastoma protein	20
MAP3K2	Mitogen-activated protein kinase 2	22
CDKN	cyclin-dependent kinase inhibitors	23
BIRC3	Baculoviral IAP repeat containing 3	23