Targeting DHFR in Breast Cancer: Exploring Expression Patterns and Identifying Potential Inhibitors through Docking Studies

By

Faiaz Fahim Araf
20146057

A thesis submitted to the School of Pharmacy in partial fulfillment of the requirements for the degree of Bachelor of Pharmacy (Hons.)

School of Pharmacy
BRAC University
March, 2024

Declaration

It is hereby declared that

1. The thesis submitted is my own original work while completing degree at BRAC University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. I have acknowledged all main sources of help.

Student's Full Name & Signature:

Faiaz Fahim Araf
20146057

Approval

The project titled “Targeting DHFR in Breast Cancer: Exploring Expression Patterns and Identifying Potential Inhibitors through Docking Studies” submitted by Faiaz Fahim Araf (20146057) of Spring, 2020 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Pharmacy (Hons.) on March, 2024.

Supervised By:

Nashrah Mustafa
Lecturer
School of Pharmacy
BRAC University

Approved By:

Program Director:

Professor Dr. Hasina Yasmin
Assistant Dean
School of Pharmacy
BRAC University

Dean:

Professor Dr. Eva Rahman Kabir
Dean
School of Pharmacy
BRAC University

Ethics Statement

This study does not involve any animal or human trial.

Abstract

Dihydrofolate Reductase (DHFR) is a critical enzyme linked to tumor development in various cancers, including breast cancer. Breast cancer, a prevalent cancer diagnosed in approximately 2.3 million people each year, arises from complex and unknown mechanisms. This study comprises of a two-part approach. The initial analysis investigated DHFR expression across diverse cancers. The elevated expression of DHFR in breast cancer demonstrated its role in tumorigenesis. Current DHFR inhibitors have a tendency to develop resistance. To address this, the second part of the study utilized molecular docking study with various FDA-approved drugs of different therapeutic classes. With methotrexate as a reference, the study identified clemastine, desloratadine and pexidartinib hydrochloride as potential candidates for breast cancer treatment due to their strong binding affinities to DHFR, superimposition results, amino acid interactions and pharmacokinetic properties. These findings need further validation through molecular dynamic simulation, biological assays and *in vivo* studies.

Keywords: Dihydrofolate Reductase (DHFR); Breast Cancer; DHFR Expression; Molecular Docking

Dedication

Dedicated to those who have been suffering from Breast Cancer

Acknowledgement

I want to express my sincere thanks to my project supervisor, Nashrah Mustafa, Lecturer, School of Pharmacy for her assistance and support during my project. It was an honor and privilege to work under her guidance.

Table of Contents

List of Tables

List of Figures

List of Acronyms

BC	Breast Cancer
WHO	World Health Organization
HR	Hormone Receptors
HER2	Human Epidermal Growth Factor Receptor 2
BLBC	Basal Like Breast Cancer
NCI	National Cancer Institute
DHFR	Dihydrofolate Reductase
HRT	Hormone Replacement Therapy
CDC	Centers for Disease Control and Prevention
ADAR1	Adenosine Deaminase Acting on RNA 1
ER	Estrogen Receptor
HR+	Hormone Receptor-Positive
ERBB2	Human Epidermal Growth Factor Receptor 2
ORR	Objective Response Rate
CDK4/6	Cyclin-Dependent Kinase 4/6
BRCA1/2	Breast Cancer Gene 1/2
DHFR	Dihydrofolate reductase
NADPH	Nicotinamide-adenine dinucleotide phosphate
TMP	Thymidine Monophosphate
FA	Folic Acid
THF	Tetrahydrofolate
DHFS	Dihydrofolate Synthase
TS	Thymidylate Synthase

dUMP	Deoxyuridine Monophosphate
dTMP	Deoxythymidine Monophosphate
dAMP	Deoxyadenosine Monophosphate
SAM	S-adenosylmethionine
MTX	Methotrexate
miR-25-3p	MicroRNA 25-3p
miR-125a-3p	MicroRNA 125a-3p
AICART	Aminoimidazole Carboxamide Ribonucleotide Transformylase
FDA	Food and Drug Administration
NSAID	Non-steroidal Anti-Inflammatory Drugs
PDB	Protein Data Bank
SDF	Structured Data File
PDBQT	Protein Data Bank, Partial Charge (Q), & Atom Type (T)
UALCAN	University of ALabama at Birmingham CANcer data analysis Portal
GTEx	Genotype-Tissue Expression
KM plotter	Kaplan-Meier Plotter
CNS	Central Nervous System
GEPIA	Gene Expression Profiling Interactive Analysis
TCGA	The Cancer Genome Atlas
LAML	Acute Myeloid Leukemia
BRCA	Breast Invasive Carcinoma
ACC	Adrenocortical Carcinoma
BLCA	Bladder Urothelial Carcinoma
CESC	Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma
CHOL	Cholangio Carcinoma

COAD	Colon Adenocarcinoma
DLBC	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma
ESCA	Esophageal Carcinoma
GBM	Glioblastoma Multiforme
HNSC	HNSC: Head and Neck Squamous Cell Carcinoma
KICH	Kidney Chromophobe
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney Renal Papillary Cell Carcinoma
LAML	Acute Myeloid Leukemia
LGG	Brain Lower Grade Glioma
LIHC	Liver Hepatocellular Carcinoma
LUAD	Lung Adenocarcinoma
LUSC	Lung Squamous Cell Carcinoma
MESO	Mesothelioma
OV	Ovarian Serous Cystadenocarcinoma
PAAD	Pancreatic Adenocarcinoma
PCPG	Pheochromocytoma and Paraganglioma
PRAD	Prostate Adenocarcinoma
READ	Rectum Adenocarcinoma
SARC	Sarcoma
SKCM	Skin Cutaneous Melanoma
STAD	Stomach Adenocarcinoma
TGCT	Testicular Germ Cell Tumours
THCA	Thyroid Carcinoma
THYM	Thymoma

UCEC	Uterine Corpus Endometrial Carcinoma
UCS	Uterine Carcinosarcoma
UVM	Uveal Melanoma
TPM	Transcripts per Million
OS	Overall Survival
PFS	Progression-Free Survival
HR	Hazard Risk Ratio
RR	Relative Risk Ratio
OR	Odd Ratio
PARP	Poly (ADP-ribose) polymerase
TKI	Tyrosine Kinase Inhibitor
ASN	Asparagine
ARG	Arginine
GLN	Glutamine
THR	Threonine
SER	Serine
LEU	Leucine
ILE	Isoleucine
ALL	Acute Lymphoblastic Leukemia
GIST	Gastrointestinal Stromal Tumors
CLL	Chronic Lymphocytic Leukemia
UTR	Untranslated Region

Targeting DHFR in Breast Cancer Treatment

Chapter 1: Introduction

1.1 Introduction to Breast Cancer

Breast cancer (BC), a cancer characterized by abnormal and uncontrolled cell growth in the breast tissues. It is the most widespread cancer among women worldwide. With around 2.3 million newly diagnosed cases and 685,000 deaths in 2020, it remains the most prevalent cancer globally (Breast Cancer Statistics and Resources | Breast Cancer Research Foundation, 2023). The WHO has predicted a continued increase in global breast cancer cases worldwide highlighting the ongoing impact of this disease and a need for targeted treatment strategies. Breast cancer has been the main contributor to cancer-related mortalities in women, surpassing other female cancers in terms of the number of deaths (Breast Cancer, 2024). A major challenge associated with this cancer is its late diagnosis; it is frequently identified at an advanced stage, posing challenges for effective treatment and reducing survival probabilities (Breast Cancer Early Detection and Diagnosis | How to Detect Breast Cancer | American Cancer Society, 2024).

According to an American Cancer Society report, five-year relative survival rate for breast cancer is approximately 90%. This indicates that, on average, women with breast cancer have about a 90% likelihood of surviving for at least 5 years after diagnosis (Survival Rates for Breast Cancer | American Cancer Society, 2022.). Breast cancer is classified into four stages according to the degree of spreading: stage I is the initial phase; stage II signifies the beginning of cancer spread; stage III indicates a more advanced stage; and stage IV (metastasis) is the most advanced stage, denoting the spread of cancer to other body parts (Understanding Breast Cancer Survival Rates | Susan G. Komen®, 2024). If detected early (stage I), it has been found

to have an impressive 98%-100% overall survival rate. The diagnostic methods including screening, such as mammograms, ultrasounds, and biopsies, are crucial for early detection (Breast Cancer Early Detection and Diagnosis | How To Detect Breast Cancer | American Cancer Society, 2024). The subtypes of Breast Cancer include HER2-positive, Basal Like Breast Cancer (BLBC), Luminal A and Luminal B. Luminal A tumors are characterized by estrogen receptor (ER) expression, boast the most favorable prognosis and responsive to treatment. Luminal B tumors, expressing hormone receptors (HR) and HER2, demand more aggressive treatment due to shorter survival compared to Luminal A. HER2-expressed tumors are marked by HER2 overexpression and low ER expression. They benefit significantly from targeted therapies, transforming poor prognoses. Basal-like cancers are predominant among triple-negative tumors, with high-grade features and they require the incorporation of biomarker data like- Basal cytokeratins (CK5/6, CK14, and CK17), vimentin, fascin, nestin and moesin for informed clinical management (Johnson et al., 2021).

1.2 Risk Factors for Breast Cancer

Breast cancer is a condition linked to various risk factors, such as specific genetic mutations, particularly in the *BRCA1* and *BRCA2* genes which has been identified as significant contributors to its development(BRCA Gene Mutations: Cancer Risk and Genetic Testing Fact Sheet - NCI, 2020).

About 55-65% of women with a harmful *BRCA1* mutation and 45% with a harmful *BRCA2* mutation are expected to develop BC by age of 70 (The Role of BRCA Mutations in Breast Cancer, 2022). These mutations significantly influence the prevalence of BC cases in women with a familial history of the disease (BRCA1 and BRCA2 (PDQ®) - NCI, 2024). A recent study reported elevated levels of DHFR in *BRCA* mutated Breast Cancers (Li et al., 2024).

Hormonal aspects, including the onset of menstrual periods and age at first pregnancy, also influence this risk. Conditions like obesity, excessive alcohol use, and tobacco consumption have been identified as significant contributors to an increased risk of breast cancer according to WHO reports (World Health Organization: WHO & World Health Organization: WHO, 2023d). Additionally, lifestyle factors such as insufficient physical activity and postmenopausal hormone therapies such as estrogen-only HRT (hormone replacement therapy) and combined HRT (Estrogen + Progestin) also play a role in elevating the risk (CDC, 2023).

Studies have shown that Dihydrofolate Reductase (DHFR) expression levels increase in BC (Nakano et al., 2017a; Raimondi et al., 2019).DHFR has an important role in folate metabolism and is the primary target for the chemotherapy drug, methotrexate (Nakano et al., 2017a). Although DHFR mutations alone do not cause breast cancer, they may interact with other genetic factors leading to BC progression. Individuals with specific DHFR variants may have altered susceptibility to BC based on their overall genetic makeup (Dalivandan et al., 2021).

DHFR is an essential enzyme involved in the synthesis of tetrahydrofolate (Figure 1), a critical cofactor in DNA and RNA synthesis. Deregulation of folate metabolism caused by DHFR overexpression can impact its function, leading to altered cellular processes. RNA editing facilitated by ADAR1 modifies the binding sites for miR-25-3p and miR-125a-3p in the 3'-UTR of DHFR mRNA, resulting in elevated DHFR expression. This increase in expression enhances cell proliferation and makes cells more resistant to methotrexate, the most commonly used DHFR inhibitor. This implies that RNA editing by ADAR1 functions as a post-transcriptional mechanism that contributes to tumour resistance to methotrexate in BC by stabilizing DHFR mRNA and preventing its degradation (Nakano et al., 2017b).

1.3 Treatment Options for Breast Cancer

The primary treatment for women diagnosed with advanced BC typically involves a combined approach, incorporating surgery to remove as much of the cancer as possible, followed by chemotherapy (Treatment of Stage IV (Metastatic) Breast Cancer | American Cancer Society, 2024).

Currently, depending on the BC subtypes, the following line of therapies are generally used. For Hormone Receptor-Positive (HR+) and HER2-negative (ERBB2-), Aromatase Inhibitor plus CDK4/6 Inhibitor achieves an Objective Response Rate (ORR) of 53-59% and is used as first-line therapy. For later-line therapy, hormonal and/or targeted therapies are considered. If patients get resistant to several hormonal treatments, switching to a single chemotherapy drug is generally recommended.

For patients with HER2-positive (ERBB2+) BC, the first-line therapy typically involves a combination of taxane, trastuzumab, and pertuzumab, achieving an overall response rate (ORR) of 80%. Some patients with hormone receptor-positive (HR+)/ERBB2+ disease may be given endocrine therapy combined with ERBB2-targeted treatment as their first-line of therapy. In subsequent treatment stages, ERBB2-targeted therapy along with chemotherapy or endocrine therapy may be utilized. For HR+ cases, trastuzumab combined with chemotherapy is a viable option. Additionally, trastuzumab with endocrine therapy or lapatinib with capecitabine could be considered for later-line therapies.

For the triple-negative subtype, single-agent chemotherapy such as taxane (ORR = 36%), platinum (ORR = 31%), and anthracycline can be used as first-line treatment. In later lines of therapy for this subtype, single-agent chemotherapy including capecitabine, eribulin, vinorelbine, gemcitabine, olaparib, or talazoparib (in the presence of a germline *BRCA1/2* mutation) can be considered (Burguin et al., 2021).

Numerous clinical trials are in progress to assess the efficacy of different treatments for advanced BC (Research and Clinical Trials for Breast Cancer, 2024). For example, in a phase III clinical trial (NCT02470585), researchers are examining the effectiveness of veliparib when combined with carboplatin and paclitaxel for individuals newly diagnosed with advanced BC. However, BC treatment is rapidly transforming marked by notable progress in targeted therapies and personalized medicine. The identification of genetic mutations linked to breast cancer has not only improved our understanding of the disease but has also facilitated the development of more potent and precisely targeted treatments.

1.4 Synthetic Pathway of Folate Metabolism

DHFR facilitates the transfer of a hydride from the cofactor NADPH, acting as an electron donor, to DHF, inducing protonation to yield THF. Specifically, DHFR initiates the reduction of 7,8-DHF to 5,6,7,8-THF, employing reduced NADPH as a cofactor. Hence, DHFR associates with thymidylate synthase, which facilitates the reductive methylation of dUMP (deoxyuridine monophosphate) to dTMP (deoxythymidine monophosphate), utilizing 5,10-Methylene THF as a cofactor (Raimondi et al., 2019).

dTMP originates from dUMP through the action of the enzyme TS. TS transfers a methyl group from 5,10-Methylene THF to the C5 position of dUMP. This process is vital for thymine production, as thymine is not synthesized *de novo*, unlike the other three bases (Leclair, 2021).

The formation of dTMP is a crucial step in DNA replication and a target for drugs like methotrexate and 5-fluorouracil, which hinder cell growth. These drugs disrupt the folate cycle and thymidylate synthase, leading to reduced dTMP and increased dUMP levels. This disruption causes DNA damage and cell death, especially in rapidly dividing cells (Zheng & Cantley, 2019).

The formation of dTMP is a crucial step in DNA replication and a target for drugs like methotrexate and 5-fluorouracil, which hinder cell growth. These drugs disrupt the folate cycle and thymidylate synthase, leading to reduced dTMP and increased dUMP levels. This disruption causes DNA damage and cell death, especially in rapidly dividing cells (Biochemistry Free & Easy (Ahern and Rajagopal) - Biology LibreTexts, 2023).

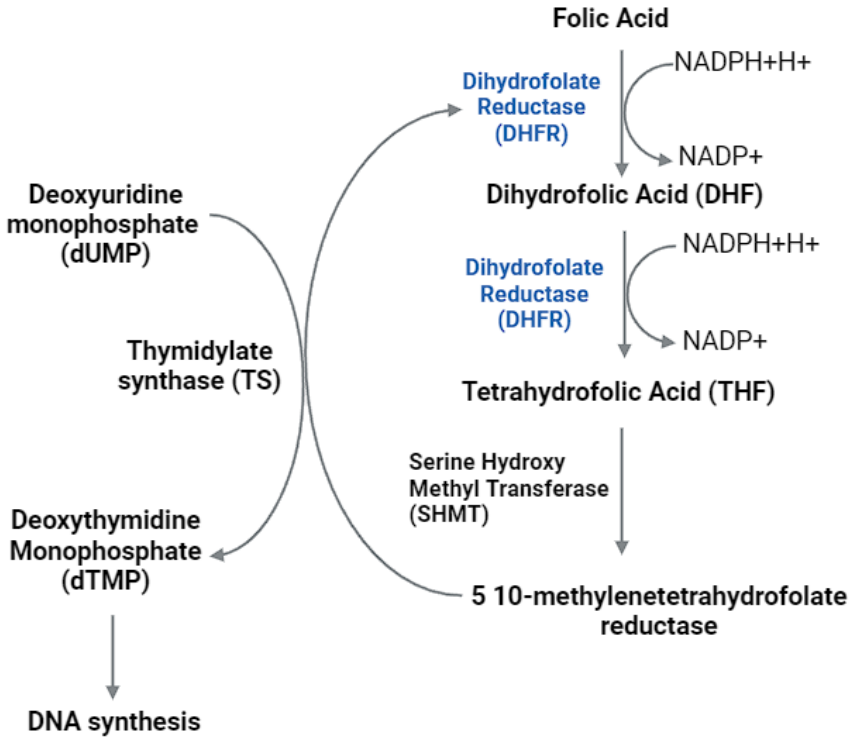


Figure 1: Normal function of DHFR protein

1.5 Structure of DHFR

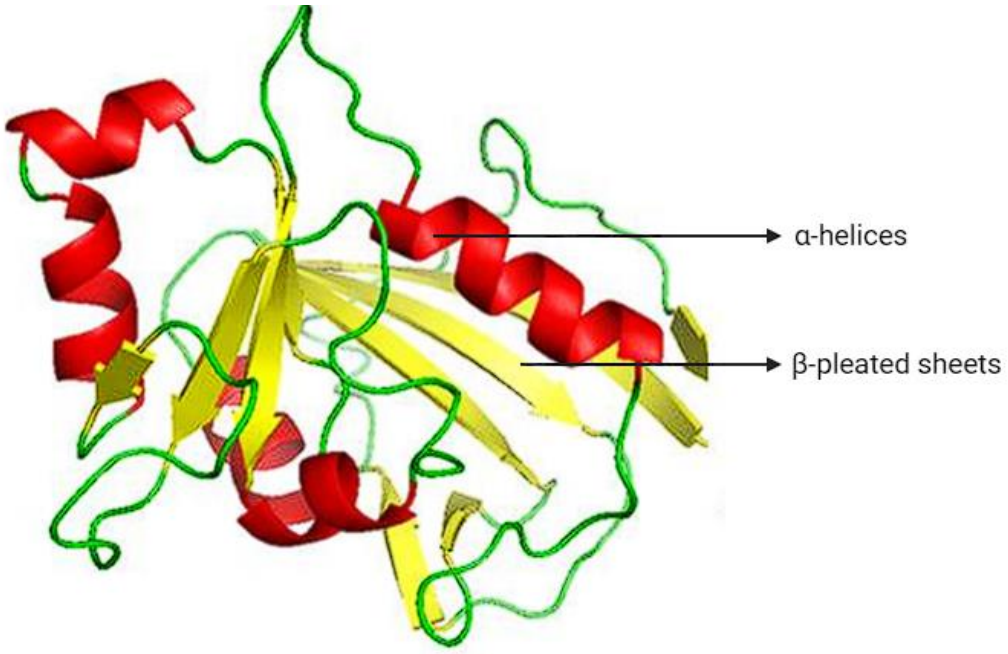


Figure 2: 3D structure of DHFR protein

DHFR is a small, water-soluble protein with a molecular weight between 18,000 and 25,000 Da. Over the years, extensive study and numerous efforts have been made to understand the structure of the DHFR isoforms. The primary structure of DHFR's polypeptide backbone folding is characterized by a central eight-stranded beta-pleated sheet. Among these strands, seven are parallel, while the eighth runs antiparallel. Connecting these beta strands are four alpha helices. The region spanning residues 9 to 24, referred to as "Met20" or "loop 1," along with other loops, constitutes a major subdomain surrounding the active site and this is where methotrexate binds to with DHFR (Osborne et al., 2001). Currently, the Protein Data Bank (PDB) has collected over one hundred structures derived from both eukaryotic and prokaryotic organisms either independently or in conjunction with various ligands. DHFR comprises eight sheets that create a stable framework, with seven sheets aligned in a parallel orientation and one in an antiparallel arrangement. All enzyme isoforms share a common structural feature of containing a minimum of four α-helices that intersect within the long loops of the sheets.

Moreover, one loop establishes the substrate binding site, while an additional two loops constitute the binding site for the coenzyme NADPH. DHFR lacks disulfide bridges and does not rely on metal ions to carry out its biochemical functions. It helps to stabilize the nicotinamide ring of NADPH, aiding in the transfer of hydride from NADPH to dihydrofolate. Additionally, it can open, close, or block the enzyme's active site. The amino acid Asp27 is essential as it helps protonate the substrate, keeping it in a favorable conformation for hydride transfer (Askari & Krajinovic, 2010).

DHFR plays a vital role in normal physiological processes. THF is required for the function of folate-dependent enzymes, thereby being important for the synthesis of purines and pyrimidines, the fundamental components of DNA and RNA. THF is essential for the intracellular transformation of synthetic folic acid, found in supplements and fortified foods, into its active forms that can engage in folate/homocysteine metabolism. THF plays a role in the remethylation process of homocysteine to methionine, a crucial step in generating S-adenosylmethionine (SAM), the primary methyl donor for the majority of biological methylation reactions. Methylation is critical for the modulation of gene expression, maintenance of DNA stability, and regulation of cellular differentiation (Askari & Krajinovic, 2010).

DHFR protein can be classified into two types: wild-Type DHFR and mutated DHFR. The unaltered form of dihydrofolate reductase (DHFR), known as the wild-type, is the standard, unmodified variant of the enzyme that is present in the normal physiological state. It plays a crucial role in various metabolic activities. On the other hand, mutated DHFR pertains to modified versions of the enzyme resulting from genetic alterations or mutations. These genetic changes can induce variations in the structure, function, or activity of the enzyme (Difference Between Wild Type and Mutant | Definition, Characteristics, Examples, Similarities, 2017).

Wild-Type DHFR is sensitive to drugs like methotrexate. On the other hand, mutations can cause resistance to drugs that target DHFR, such as trimethoprim (Bertino et al., 1996).

Cancer cells differ from normal cells in several aspects. Normal cells become cancerous due to a series of mutations that cause uncontrolled growth and division (Witt & Tollefsbol, 2023). Unlike normal cells, which usually remain in their original location, cancer cells can invade nearby tissues and spread to distant parts of the body. In contrast to normal cells, which typically stay localized, cancer cells possess the capability to both invade neighboring tissues and spread to distant regions of the body (Witt & Tollefsbol, 2023). Mutation at different position of DHFR may cause resistance to different drugs *e.g.,* N51I mutation is correlated with resistance to pyrimethamine. C59R mutation is similarly associated with pyrimethamine resistance. S108N mutation is another mutation that contributes to resistance against pyrimethamine. I164L mutation is connected to resistance against high-grade sulfadoxine-pyrimethamine, although absent in the isolates examined in the study (Jiang et al., 2019).

Cancer cells go through genetic mutations that transform the cell from its normal state into a cancerous one. The genetic mutations can either be inherited or developed gradually with aging as genes undergo wear and tear, or arise due to exposure to factors that damage genes, such as ultraviolet (UV) radiation from the sun, alcohol or cigarette smoke (How Cancer Starts, Grows and Spreads | Canadian Cancer Society, 2021).

Cancer cells need a significant amount of energy for their growth and division as their metabolism is often different from that of normal cells. They also have the ability to escape the immune system, which typically identifies and eliminates abnormal cells (Cancer Cells: Types, How They Form, and Characteristics, 2023-b). Just like any other cells, cancer cells also require DHFR enzyme for functioning. As a result, anticancer drugs use DHFR enzyme as a target. DHFR is also target for antifolate medications, like methotrexate (MTX), that are used

in the treatment of cancer and certain inflammatory conditions. Through the inhibition of DHFR, antifolate drugs reduce the cellular reservoir of THF, leading to slowing down of DNA synthesis and cell proliferation (Askari & Krajinovic, 2010) which is a hallmark of cancer cells.

1.6 Causes and Effects of DHFR Mutation

DHFR protein can be mutated in several ways:

i. Increased expression: In certain cancer cells, such as BC there is more Dihydrofolate Reductase (DHFR) to make sure there is enough nucleotides for the cells to grow and divide rapidly. This helps the cells meet the high demands for building blocks needed during their rapid multiplication (DHFR Gene Mutation - MTHFR Gene HealthTM, 2023).

ii. Antifolate Resistance: Antifolate drugs (e.g., trimethoprim, methotrexate) target DHFR. However, mutations in DHFR can reduce drug binding affinity, leading to resistance. Cancer cells with mutated DHFR may evade the effects of antifolate therapies (DHFR Gene Mutation - MTHFR Gene HealthTM, 2023).

iii. Structural Alterations: Specific mutations can disrupt the active site of DHFR, affecting its catalytic function. For example, the L80F mutation destabilizes the DHFR protein or disrupts NADPH binding (Nakano et al., 2017). Other mutations may alter the binding affinity for diaminoheterocyclic molecules (Nakano et al., 2017b).

This study targets the overexpression of DHFR in BC.

1.6.1 Mutations Lead to Breast Cancer

DHFR is a crucial enzyme for the metabolism of folate, playing a notable role in progression of BC (Expression of DHFR in Breast Cancer - The Human Protein Atlas, 2021). The expression of DHFR in BC has been examined and observed through antibody staining (Expression of DHFR in Breast Cancer - The Human Protein Atlas, 2021).

Research reported in The Journal of Biological Chemistry revealed that adenosine-to-inosine RNA editing, facilitated by ADAR enzymes, may influence DHFR expression. The study demonstrated that ADAR1 has a positive regulatory role in DHFR expression through the editing of binding sites for miR-25-3p and miR-125a-3p in the 3' UTR of DHFR. This editing process enhances cellular growth and methotrexate resistance. An increased cellular DHFR expression is a factor contributing to tumour resistance to methotrexate (Nakano et al., 2017).

1.6.2 Mechanism of Action of Methotrexate- A Known DHFR Inhibitor

Methotrexate (MTX) demonstrates potent anti-cancer activity through a distinct mechanism. Operating as an anti-metabolite of folic acid (Vitamin B9), MTX interrupts cell division by inhibiting folate-related enzymes, particularly Dihydrofolate Reductase (DHFR). DHFR which converts dihydrofolate to tetrahydrofolate (THF) is a crucial coenzyme in nucleotide synthesis pathways essential for DNA and RNA production. Through DHFR inhibition, MTX hinders the crucial precursors required for cell proliferation, particularly in actively dividing cells and rapidly proliferating cancer cells. The drug's cytotoxic impact is most notable during the S phase of the cell cycle. Additionally, MTX, in its polyglutamated form, functions as an effective anti-folate agent, not only inhibiting DHFR but also influencing other enzymes like thymidylate synthase and AICART (Aminoimidazole Carboxamide Ribonucleotide Transformylase) involved in nucleotide synthesis. This disruption in DNA synthesis pathways ultimately induces cell apoptosis. Overall, MTX's anti-cancer effectiveness lies in its ability to selectively target rapidly dividing cells, rendering it a valuable chemotherapy agent for various malignancies (Koźmiński et al., 2020).

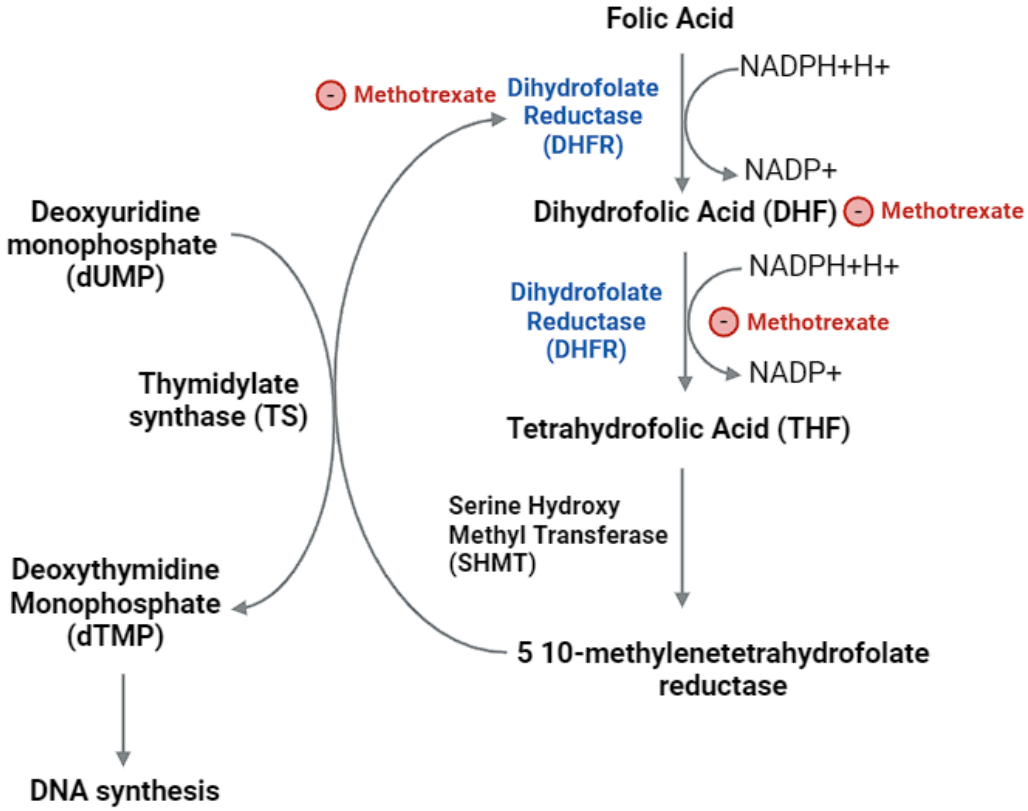


Figure 3: Mechanism of action of methotrexate

1.7 Drug Repurposing using *In Silico* Methods:

Drug repurposing (also known as drug repositioning or reprofiling), involves finding new therapeutic uses of existing medications. This approach has several advantages like reduced costs, time, and risk of failure compared to traditional drug discovery. It potentially enhances safety and efficacy when coupled to biological assays and *in vivo* studies. Despite these benefits, drug repurposing encounters challenges, including intellectual property issues, regulatory obstacles, and market incentives.

In silico methods and computational techniques using mathematical models and databases, play a vital role in drug repurposing.

These methods come in various types:

i. Target-based approaches identify new targets for existing drugs or *vice versa* using techniques like molecular docking, pharmacophore modelling, network analysis, or machine learning.

ii. Phenotype-based approaches find new uses for existing drugs or identify new drugs for known uses using gene expression analysis, text mining, clinical data mining, or drug repositioning networks.

iii. Integrative approaches combine diverse data types and methods, such as genomics, proteomics, metabolomics, and phenotypic data, to generate hypotheses or validate predictions for drug repurposing.

In silico methods enable high-throughput screening of multiple approved drugs and targets within a short timeframe, and their adaptability allows for easy integration of new data. Moreover, these methods are cost-effective and accessible, as they do not demand expensive laboratory equipment, making them available to researchers with diverse backgrounds. Importantly, *in silico* methods offer interpretability and transparency by providing mechanistic insights and evidence for drug repurposing. These findings can be subsequently verified through experimental testing or clinical trials, ensuring robust and reliable outcomes in the drug development process (Hodos et al., 2016).

Computational *in silico* models have the capacity to merge information from diverse sources, examine hypotheses that may be challenging to validate through experiments, and offer insights into the mechanisms and dynamics of cancer. These *in silico* models find utility in various areas, including the identification of biomarkers, drug targets, and optimal treatment strategies (Edelman et al., 2010).

1.7.1 Success of *In Silico* Methods in Cancer

The development of *in silico* models stands to gain from progress in data gathering and analysis, model integration and standardization, and collaborative efforts. Future enhancements may involve incorporating greater biological realism and complexity, encompassing elements like epigenetics, metabolism, and immunology (Edelman et al., 2010).

In silico methods were used on hydroxychloroquine (HCQ), an antimalarial medication. Researchers have utilized computational models to look into two critical aspects regarding HCQ's pharmacodynamics. Firstly, by employing a computational model of human atrial cardiomyocytes, they have analyzed HCQ's impact on atrial electrophysiology. Their findings hint at a promising antiarrhythmic role for low-dose HCQ in mitigating atrial fibrillation (AF). Secondly, *in silico* docking studies have been used to see HCQ's interactions with proteins of the SARS-CoV-2 virus. These investigations have revealed HCQ's increased binding affinity with essential viral proteins such as the spike glycoprotein and PLPRO protein, suggesting its potential as a candidate for antiviral drug development. Through these computational approaches, researchers continue to unravel the multifaceted pharmacological implications of HCQ, offering insights that could inform clinical decisions and therapeutic strategies. (Rao & Subash, 2020).

Target-Based *In Silico* Methods offer speed, efficiency, and predictive insights by leveraging computational simulations. These approaches reduce costs and aid rational drug design. On the other hand, Drug Repurposing capitalizes on existing drugs, minimizing risks, shortening development timelines, and maintaining cost-effectiveness. Both strategies contribute significantly to efficient drug discovery and development (Rao & Subash, 2020).

In silico methods can help researchers to identify unique signatures for target-specific amplification and detection of pathogens. It can also optimize the design and performance of

assays by reducing the number of iterations, empirical optimization, and validation steps. Furthermore, it reduces the time and cost of assay development and validation by utilizing available databases of whole genome sequences and bioinformatics tools (Lucia et al., 2020).

1.8 Rationale of the Study

Breast cancer, the second leading cancer globally presents a significant challenge in the field of oncology due to its widespread prevalence, aggressive characteristics, and tendency to develop resistance to existing drugs over time. DHFR is an essential enzyme in the folate metabolism pathway, and thus in the production of nucleotides. Rapidly dividing cancer cells need a constant supply of nucleotides for DNA synthesis and cell growth. A few studies have shown that breast cancer tissues demonstrate elevated levels of DHFR compared to healthy tissues. The first part of the study aims to explore the expression patterns of DHFR in breast cancer tissues and normal tissues. It was hypothesised in the study that this increased expression makes the cancer cells more dependent on DHFR for survival and inhibiting the enzyme would help manage the disease by hindering the growth of cancer cells.

Despite the availability of several drugs that target DHFR proteins for breast cancer treatment, the challenge of drug resistance remains a major challenge to achieving successful outcomes. The second part of the study explored around 200 existing FDA approved drugs from various therapeutic classes such as anti diabetic drug, anticancer drugs, antihistamines, NSAIDs, antihypertensive drugs, statins and some natural compounds to screen and propose potential options that can effectively target DHFR protein and thus help in breast cancer treatment. Having a greater and comprehensive drug pool would help to overcome the challenge of drug resistance in patients with breast cancer.

2. Methodology

This section outlines the steps involved in the study.

2.1 Expression of DHFR in Breast Cancer

The following online tools were used to analyse the expression of DHFR in Breast Cancer.

i. UALCAN: UALCAN works as a comprehensive online tool for the analysis of cancer OMICS data. It has a portal facilitating the exploration, analysis, and visualization of genomic, transcriptomic, and proteomic data related to cancer (Chandrashekar et al., 2022a). UALCAN enables to investigate, analyze, and visually represent genomic, transcriptomic, and proteomic data sourced from The Cancer Genome Atlas (TCGA). With UALCAN, users have the capability to assess how gene expression influences patient survival across 33 different cancer types (Chandrashekar et al., 2022b). UALCAN was used in this study to get easy access to cancer related data and to perform pancancer gene expression analysis.

ii. GTEx: The Genotype-Tissue Expression (GTEx) Project shows the role of genetic variations in common human diseases. GTEx offers a comprehensive map of gene expression in diverse human tissues, enabling researchers to investigate how inherited genetic changes impact health (Genotype-Tissue Expression Project (GTEx), 2020). Evaluation of DHFR expression in normal breast tissues was conducted using data from the public GTEx in this study.

iii. Kaplan Meier Plotter: The Kaplan Meier plotter is a powerful tool used to study the link between gene expression and survival rates across various cancer types, examining over 35,000 samples. With the capability of conducting 18,000 analyses daily, it is a widely respected platform for discovering and validating survival markers. Unlike other tools, it performs calculations in real-time without relying on pre-calculated data. The database integrates gene expression and clinical information to assess the prognostic significance of a gene. Patient

samples are divided into groups based on different expression levels, and their survival rates are compared using Kaplan-Meier plots, providing hazard ratios and confidence intervals. Regular updates ensure the database's accuracy and comprehensiveness (Kaplan-Meier Plotter, 2024). The Kaplan-Meier plot is a statistical tool used in cancer research to estimate and visualize patient survival rates over time. Kaplan Meier plotter was used in this research to compare the OS and PFS of BC cells with high DHFR expression and BC cells with lower DHFR expression.

2.2 Preparation of Protein and Ligands

2.2.1 Protein Preparation

The protein structure (PDB ID: 1U72) was obtained from the protein data bank in the PDB format. Using PyMOL, the water molecules and the co crystallized structure, methotrexate (MTX) was removed from the protein (Yuan et al., 2017). Hydrogen atoms were added; then atom types and Kollman charges were assigned; energy minimization followed by optimization of the structure was carried out using AutoDock Tools (Tang et al., 2022). The curated protein was then saved as a PDBQT format file.

2.2.2 Ligand Molecules Preparation

The 3D structures of the ligand molecules were downloaded as SDF files (https://pubchem.ncbi.nlm.nih.gov/) and PDB files (https://go.drugbank.com/). The SDF format ligands were converted to PDB format using PyMOL. All the structures were converted to PDBQT format using AutoDock Tools and assigned charges.

2.3 Molecular Docking with AutoDock Vina

The potential binding site was identified and defined using PyMOL. Based on the identification, a grid box was generated around the binding site by setting the center and size

parameters in the AutoDock Vina configuration file. The grid file and config.txt file based on the values obtained from the grid.txt file was saved. Each ligand folder then had 2 PDB files of the protein and ligand, 2 PDBQT files of the protein and ligand, the config.txt file, and the grid.txt file.

To initiate the docking, the Windows "run" command was opened by typing 'cmd' in the Start menu and clicking 'OK.' This action opened a command prompt 'cd' followed by the space and the location of the files for docking were typed. Upon pressing 'Enter,' the desired directory was reached. Subsequently, a docking command was used. The command used for docking is given below:

"C:\Program Files (x86)\The Scripps Research Institute\Vina\vina.exe" --receptor protein.pdbqt --ligand ligand.pdbqt --config config.txt --log log.txt --out output.pdbqt.

The respective binding affinities were generated after the command was entered for each ligand.

2.3.1 Selection of Docking Results

At first, the co crytallised structure methotrexate was docked with our selected DHFR protein (PDB ID-1U72) and the binding affinity obtained was designated as the cut off value or the reference value since methotrexate is a known inhibitor of DHFR. The binding affinity of methotrexate to DHFR was found to be -9.1 kcal/mol. All the other values were compared to that of methotrexate. Better binding affinity values were chosen for the subsequent steps.

2.3.2. Superimposition of Selected Drugs with Methotrexate in the Binding Pocket

Superimposition of the selected drugs with the reference drug methotrexate was done in PyMOL to assess whether the selected drugs effectively bind into the targeted binding site of the enzyme.

2.3.3. Checking the Interactions using Discovery Studio

Discovery Studio (v 16.1.0.15350) was then used to check the amino acids involved in the binding interaction with DHFR (Adewole et al., 2022). Drugs that had higher numbers of common amino acids to methotrexate were chosen as these represented drugs binding to the same binding pocket of DHFR as methotrexate. The initial step involved opening the curated protein and the specific drug's output file in Discovery Studio (v 16.1.0.15350). Subsequently, the entire ligand group from the output file was chosen and transferred on the curated protein file. Hydrogen atoms were added. The receptor and ligand were defined. Next, the ligand interactions option was chosen. The ligand interactions with DHFR were found. Following these steps, a comprehensive display of all interactions was displayed.

The work that was done for molecular docking can be summarized using the following flowchart (Figure 4).

2.3.4 Pharmacokinetic Profile Analysis

The predicted pharmacokinetic properties of the selected drugs were then calculated using QikProp, Schrodinger (Hasan et al., 2022).

3. Results

3.1 DHFR Expression Analysis

3.1.1 Pancancer Mapping of DHFR in Breast Cancer

Initial analyses of DHFR expression in both normal and tumor tissues was conducted across 31 cancers using GEPIA (add acronym) with TCGA datasets (Figure 4). The findings showed elevated DHFR expression in most cancers compared to normal tissues, except for Acute Myeloid Leukemia (LAML). There was a notable increase in DHFR expression in Breast invasive carcinoma (BRCA) (*P≤0.05), indicating DHFR overexpression in Breast Cancer (BC).

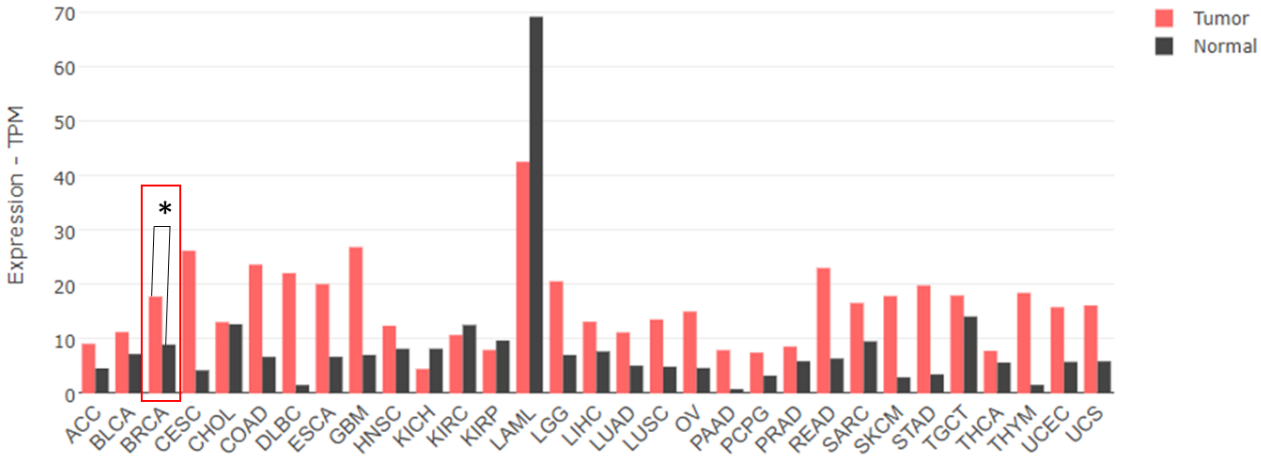


Figure 4: DHFR gene expression in normal and tumour tissues based on GEPIA data; Red: Tumour; Black: Normal tissues. Abbreviations: ACC: Adrenocortical carcinoma; BLCA: Bladder Urothelial Carcinoma; BRCA: Breast invasive carcinoma; CESC: Cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL: Cholangio carcinoma; COAD: Colon adenocarcinoma; DLBC: Lymphoid Neoplasm Diffuse Large B-cell Lymphoma; ESCA: Esophageal carcinoma; GBM: Glioblastoma multiforme; HNSC: Head and Neck squamous cell carcinoma; KICH: Kidney Chromophobe; KIRC: Kidney renal clear cell carcinoma; KIRP: Kidney renal papillary cell carcinoma; LAML: Acute Myeloid Leukemia; LGG: Brain Lower Grade Glioma; LIHC: Liver hepatocellular carcinoma; LUAD: Lung adenocarcinoma; LUSC: Lung squamous cell carcinoma; MESO: Mesothelioma; OV: Ovarian serous cystadenocarcinoma; PAAD: Pancreatic adenocarcinoma; PCPG: Pheochromocytoma and Paraganglioma; PRAD: Prostate adenocarcinoma; READ: Rectum adenocarcinoma; SARC: Sarcoma; SKCM: Skin Cutaneous Melanoma; STAD: Stomach adenocarcinoma; TGCT: Testicular Germ Cell Tumours; THCA: Thyroid carcinoma; THYM: Thymoma; UCEC: Uterine Corpus Endometrial Carcinoma; UCS: Uterine Carcinosarcoma; UVM: Uveal Melanoma. TPM: transcripts per million.

3.1.2 DHFR Expression in Normal Gynecological Tissues using The Genotype-Tissue Expression (GTEx)

After that, an examination of DHFR expression in normal breast tissues was conducted using data from the public GTEx (Figure 5). The data revealed the presence of DHFR mRNA in normal breast (mammary) tissues. From a sample size of 459, the median expression was found to be 5.190 transcripts per million. Among them, 168 were female with the median expression 5.530 TPM and 291 were male with the median expression found to be 5.032 TPM. This suggests that the expression is slightly higher in females than males.

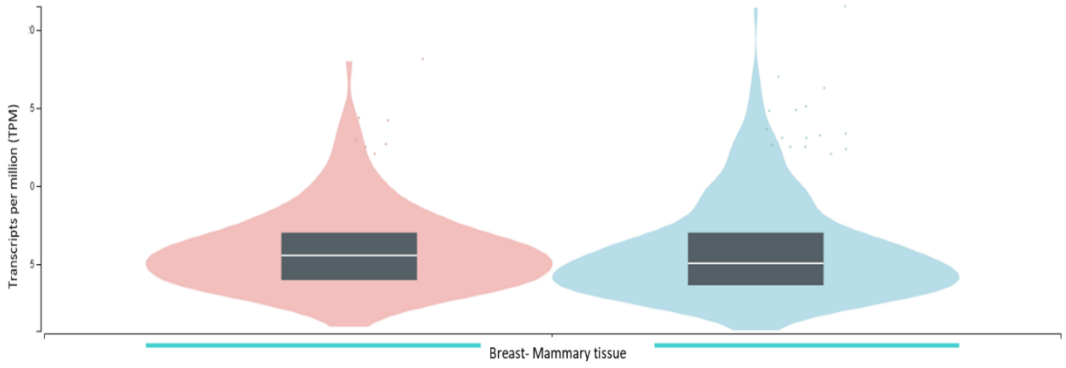


Figure 5: Expression of DHFR in normal tissues; Pink: female; Blue: male

3.1.3 Comparison of DHFR Expression in Normal and Breast Cancer tissues using GEPIA

The data suggests that the expression of DHFR increases significantly in the cancer tissues compared to the normal tissues.

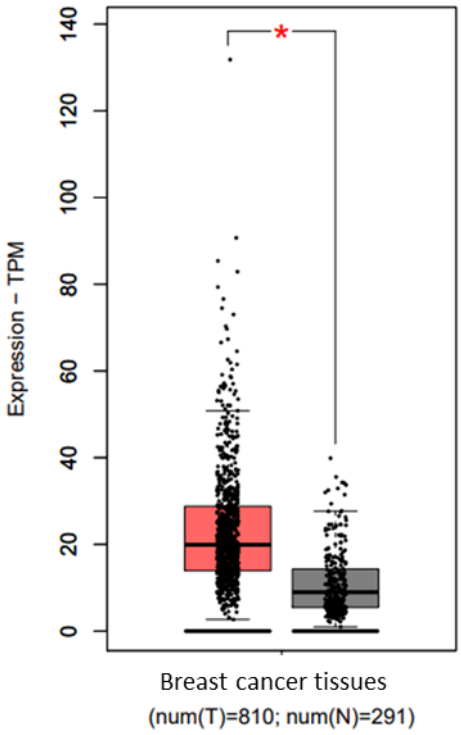


*Figure 6: Comparison of expression of DHFR using TGCA datasets in Breast Cancer; Cancer: red; Normal: ash. A significant increase in DHFR expression was observed for BC (*P≤0.001). TPM: Transcripts per million; n(T): sample size of tumours; n(N): sample size of normal tissues.*

3.1.4 Overall Survival and Progression Free Survival using Kaplan Meier (KM) plotter

Overall Survival (OS), a primary endpoint in cancer clinical trials shows how long patients survive after starting a particular treatment. Researchers look closely at OS to see how a treatment affects the overall life expectancy (Royle et al., 2023).

Progression-Free Survival (PFS) refers to the duration a group of people with cancer remain disease-free after a certain treatment. PFS is commonly used as a goal in controlled trials for cancer treatments. Unlike OS, which looks at deaths from any cause, PFS specifically focuses on the progression of the disease. Measuring PFS helps researchers understand the impact of diseases and treatments that do not necessarily lead to death, like pain, and disruptions in daily

life. It is useful for getting quicker trial results (What Progression-Free Survival Means After Cancer, 2023).

The predicted OS and PFS as a function of time has been illustrated in Figure 4 using the KM plotter. The patients were grouped as “BC patients with high DHFR expression” and “BC patients with low DHFR expression”. The probability for OS is less in patients having DHFR overexpression.

The data suggest that the OS for patients with BC with low DHFR expression was 120 months whereas the OS for patients having high DHFR expression was 69.6 months. Similarly, the PFS for patients with BC with low DHFR expression was 39.25 months whereas the OS for patients having high DHFR expression was 24.96 months.

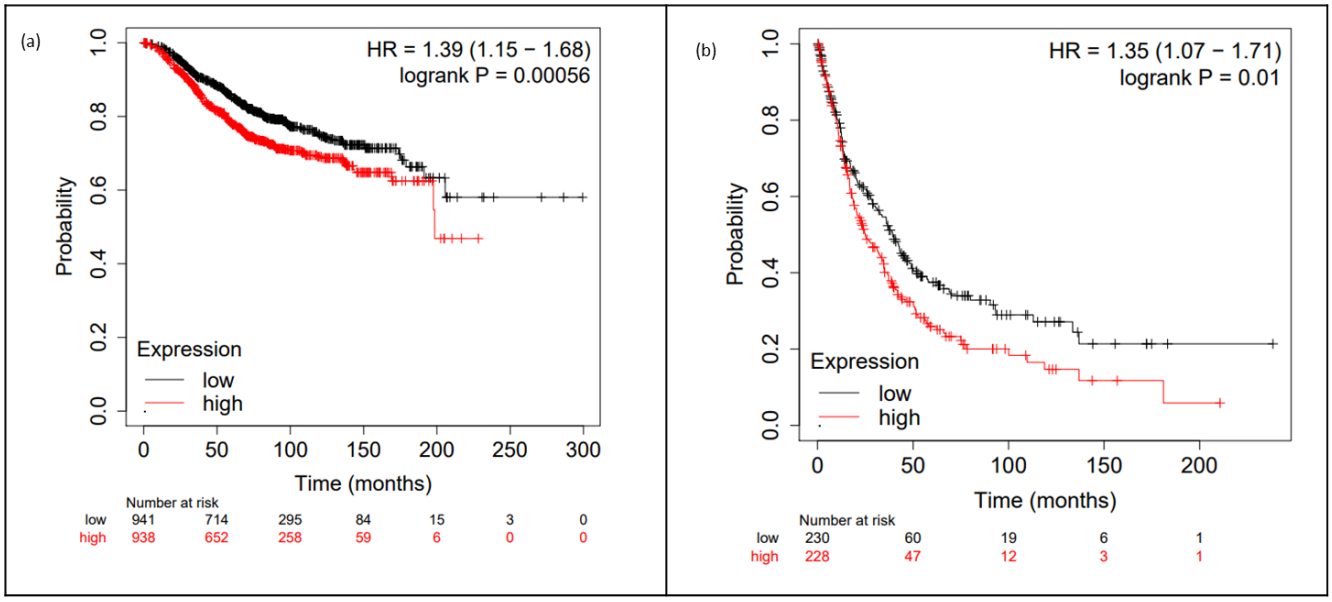


Figure 7: Kaplan-Meier plots showing the effect of DHFR expression in Breast Cancer (a) Overall survival (OS) in BC, (b) PFS in BC; HR represents Hazard Risk Ratio

If the Hazard Ratio is more than 1, it means there's a higher risk in the treatment group than in the control group. Whereas, if the Hazard Ratio is less than 1, it means there's a lower risk in the treatment group compared to the control group. The results show that the hazard ratio is

higher with higher DHFR expression (Hazard Ratio: Interpretation & Definition - Statistics By Jim, 2023).

3.2 Docking Analysis

3.2.1 Molecular Docking Results

Around 200 ligands were docked with DHFR and the results were compared to the binding affinity of methotrexate. Binding affinity values better than the reference drug were chosen. Methotrexate was chosen as the reference drug due to its effectiveness against DHFR protein (Teachey et al., 2008). The binding affinity of methotrexate was found to be 9.1 kcal/mol. There were 68 drugs which showed better binding affinity to DHFR compared to methotrexate (Table 1).

Table 1: Binding affinity values of reference drugs and selected drugs with DHFR (AutoDock 4)

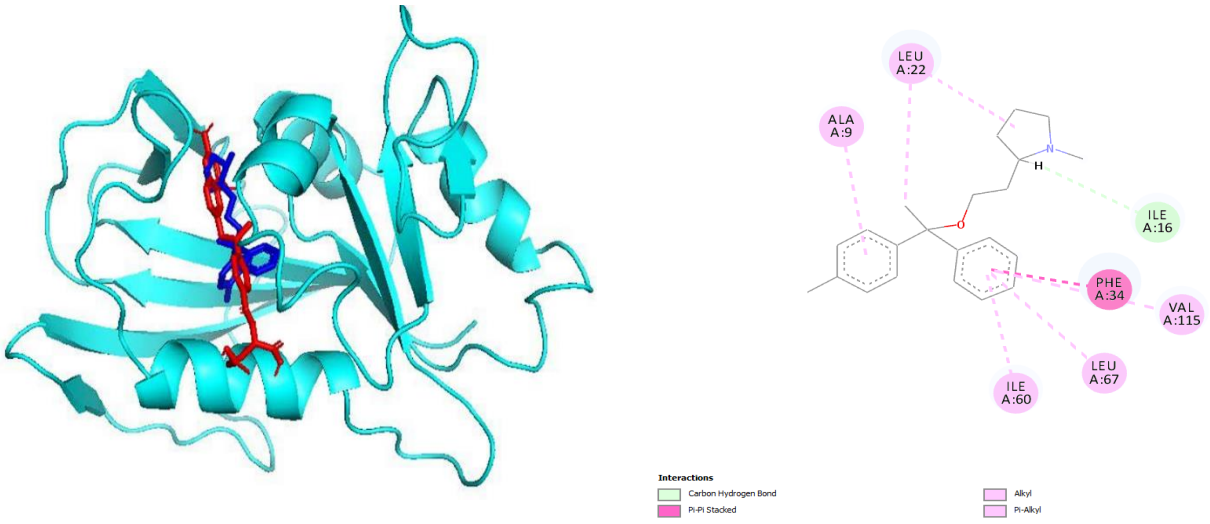
SI No	Drugs	Binding Affinities (kcal/mol)
1	Methotrexate (Reference drug)	-9.1
2	Imatinib mesylate (TKI)	-10.2
3	Ripretinib (TKI)	-9.6
4	Tucatinib (TKI)	-11.7
5	Tivozanib hydrochloride (TKI)	-10.1
6	Bosutinib (TKI)	-9.7
7	Axitinib (TKI)	-10.1
8	Asciminib hydrochloride (TKI)	-9.6
9	Pexidartinib hydrochloride (TKI)	-10.0

10	Infigratinib phosphate (TKI)	-9.8
11	Dasatinib (TKI)	-9.8
12	Regorafenib (TKI)	-10.0
13	Abemaciclib (TKI)	-10.5
14	Azelastine (Antihistamine)	-10.6
15	Desloratadine (Antihistamine)	-10.2
16	Clemastine (Antihistamine)	-9.4
17	Hydroxyzine (Antihistamine)	-9.7
18	Niraparib (PARP inhibitor)	-9.5
19	Rucaparib (PARP-inhibitor)	-9.7
20	Meloxicam (NSAID)	-9.3
21	Rofecoxib (NSAID)	-9.2
22	Topotecan (Anti-cancer drug)	-9.3
23	Doxorubicin Hydrochloride (Anti-cancer drug)	-9.7
24	Belzutifan (Anti-cancer drug)	-9.4
25	Duvelisib (Anti-cancer drug)	-9.3
26	Enasidenib Mesylate (Anti-cancer drug)	-9.6
27	Tamoxifen Citrate (Antiestrogen)	-9.5

3.2.2 Superimposition Results

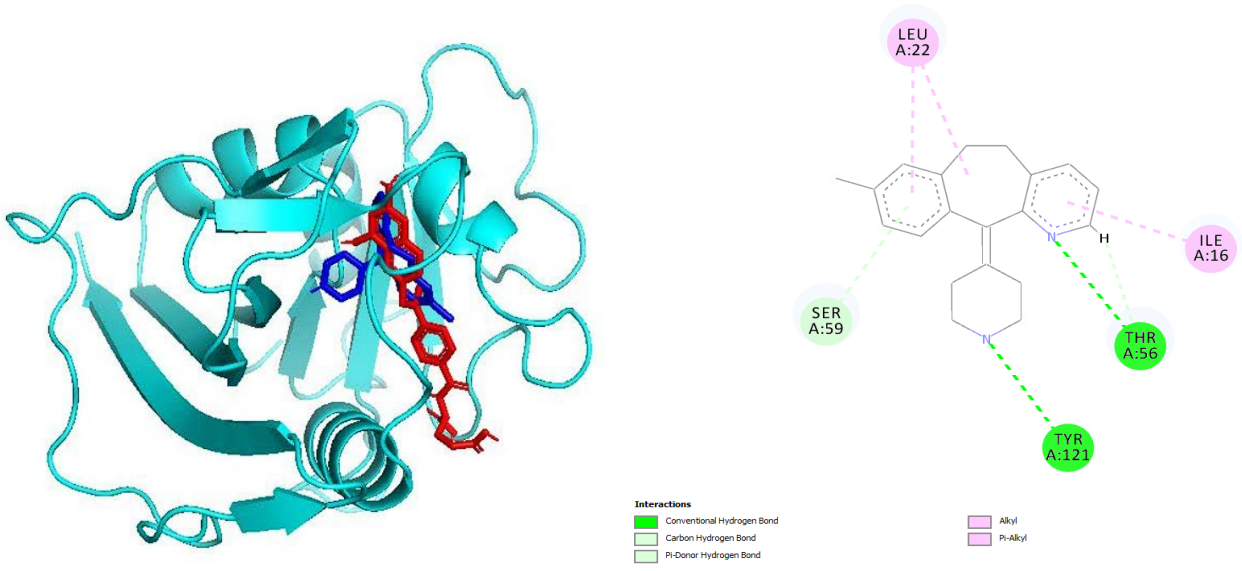
By superimposing molecules in the binding pocket with the reference drug, we can evaluate if a drug fits into the active or targeted binding site of the enzyme, aiding in the prioritization of molecules for further experimentation. The following figures show some of the superimposition results (Figure 8-10). Clemastine, desloratadine and pexidartinib

hydrochloride superimposed with methotrexate in the DHFR binding pocket. The interactions of the mentioned drugs obtained from Discovery Studio (v 16.1.0.15350) are also displayed.



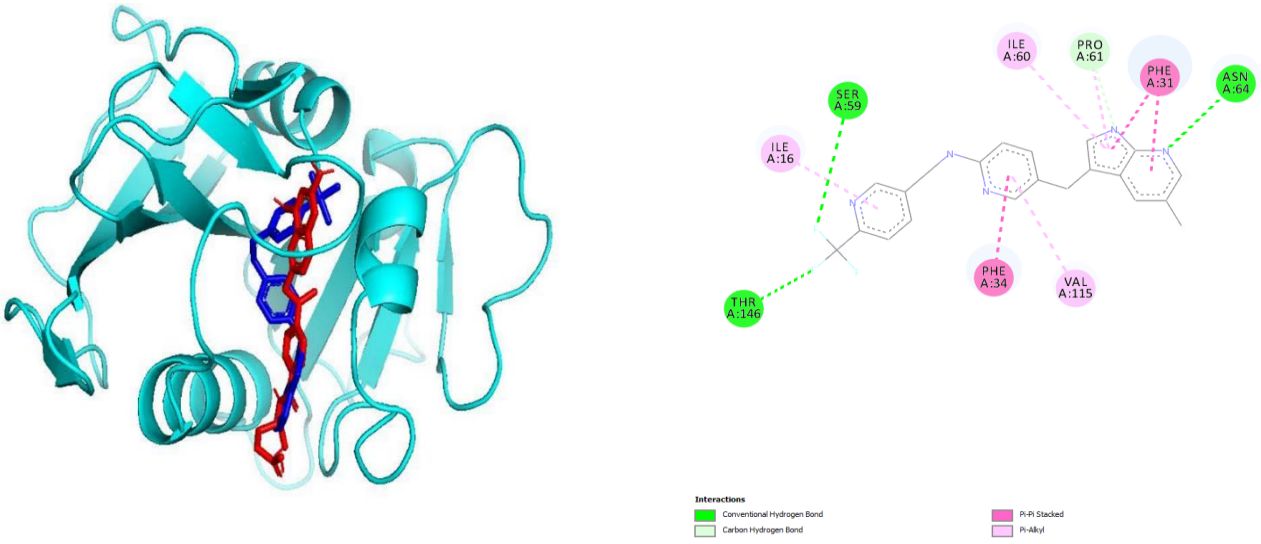
(a) ***(b)***

Figure 8: *Superimposed binding mode of Methotrexate (red) and Clemastine (blue) with DHFR (visualized in PyMOL) (b)2D diagram of DHFR-Clemastine interaction (visualized in Discovery Studio version 16.1.0.15350)*



(a) ***(b)***

Figure 9: *Superimposed binding mode of Methotrexate (red) and Desloratadine (blue) with DHFR (visualized in PyMOL) (b)2D diagram of DHFR-Desloratadine interaction (visualized in Discovery Studio version 16.1.0.15350)*



(a) ***(b)***

Figure 10: *Superimposed binding model of Methotrexate (red) and Pexidartinib Hydrochloride (blue) with DHFR (visualized in PyMOL) (b) 2D diagram of DHFR-Pexidartinib Hydrochloride interaction (visualized in Discovery Studio version 16.1.0.15350)*

3.2.3 Discovery Studio Results

Discovery Studio (v 16.1.0.15350) was used to check the protein-ligand interactions in order to determine whether the chosen drugs bind to the same pocket as the reference drug. In addition, the types off bonds were also determined. The common amino acids involved in the DHFR protein-ligand interaction are given in the table below (Table 2).

Table 2: Common amino acids involved in binding with the reference drug (Discovery Studio v 16.1.0.15350)

Drugs	Amino Acids
Methotrexate (Reference drug)	ASN64, ARG70, GLN35, THR146, SER59, LEU22, ILE60, ILE16
Imatinib mesylate (TKI)	SER59, LEU22, ILE60
Ripretinib (TKI)	SER59, THR146, ILE16, LEU22

Tucatinib (TKI)	SER59, ILE60, ILE16
Tivozanib hydrochloride (TKI)	SER59, ASN64, ILE60
Bosutinib (TKI)	GLN35, SER59, ASN64, LEU22, ILE60
Axitinib (TKI)	ILE60, ILE16, THR146
Asciminib hydrochloride (TKI)	ASN64, SER59, GLN35
Pexidartinib hydrochloride (TKI)	SER59, ASN64, THR146, ILE60, ILE16
Infigratinib phosphate (TKI)	THR146, ILE60, LEU22, ILE16
Dasatinib (tyrosine kinase inhibitor (TKI))	THR146, ILE16, LEU22
Regorafenib (TKI)	SER59, ILE16, ILE60, LEU22
Abemaciclib (TKI)	THR146, LEU22, ILE60, ILE16
Azelastine (Antihistamine)	LEU22, ILE60
Desloratadine (Antihistamine)	SER59, LEU22, ILE 16
Clemastine (Antihistamine)	ILE16, LEU22, ILE60
Hydroxyzine (Antihistamine)	THR146, ILE16, LEU22, ILE 60
Niraparib (PARP inhibitor)	SER59, LEU22, ILE60
Rucaparib (PARP-inhibitor)	THR146, SER59, LEU22
Meloxicam (NSAID)	SER59, THR146, LEU22
Rofecoxib (NSAID)	SER59, ILE16, LEU22, THR146
Topotecan (Anti-cancer drug)	SER59, ILE16, LEU22, ILE60
Doxorubicin hydrochloride (Anti-cancer drug)	THR146, SER59, LEU22, ILE60
Belzutifan (Anti-cancer drug)	THR146, SER59, ILE16, LEU22
Duvelisib (Anti-cancer drug)	GLN35, ASN64, ILE60, LEU22
Enasidenib mesylate (Anti-cancer drug)	SER59, THR146, ILE60, LEU22

Tamoxifen citrate (Antiestrogen)	SER59, LEU22, ILE60, ILE16

The protein-ligand interaction was visualized using Discovery Studio (v 16.1.0.15350). The amino acids involved in the binding of methotrexate with DHFR are as follows: ASN64, ARG70, GLN35, THR146, SER59, LEU22, ILE60 and ILE16. The drugs that showed common amino acids include ripretinib (TKI), bosutinib (TKI), pexidartinib hydrochloride (TKI), infigratinib phosphate (TKI), regorafenib (TKI), abemaciclib (TKI), hydroxyzine (Antihistamine), rofecoxib (NSAID), topotecan (Anti-cancer drug), doxorubicin hydrochloride (Anti-cancer drug), belzutifan (Anti-cancer drug), duvelisib (Anti-cancer drug), enasidenib mesylate (Anti-cancer drug), tamoxifen citrate (Antiestrogen). This shows that the mentioned drugs bind to the same pocket of DHFR as methotrexate.

Common side-effects were taken into consideration as the more serious side-effects are not that common and so are left out of consideration. Rofecoxib was screened out as this drug is no longer available due to its association with an increased risk of heart attacks and strokes. imatinib mesylate was screened out as it may cause tumor lysis syndrome and cause issues in urination. Ripretinib was also screened out as this drug might elevate the blood pressure. Axitinib was screened out as its common side effects include elevation of blood pressure. Asciminib hydrochloride was screened out as it may commonly lower the blood cell counts. Infigratinib phosphate was screened out as it commonly causes changes in the liver function. Regorafenib was left out as it commonly causes high blood pressure. Niraparib was not selected because it can cause sleep issues. Topotecan was screened out as it lowers blood cell counts. Doxorubicin hydrochloride causes hair loss and so was screened out. Belzutifan was screened out as it causes anemia and increases blood pressure. Duvelisib was left out as it lowers blood cell counts.

The drugs left after screening the other drugs out due to having serious side effects include azelastine, abemaciclib, bosutinib, clemastine, dasatinib, desloratadine, enasidenib, hydroxyzine, meloxicam, pexidartinib hydrochloride and rucaparib.

QikProp Results

The pharmacokinetic parameters were analyzed using QikProp, Schrodinger (Table 3)

Table 3: Analysis of pharmacokinetic parameters

Molecules	**Absorption**	**Distribution**			**CNS permeability**		
	%HOA	**QPPCaco**	**QPPMDCK**	**QPlogKhsa**	**CNS**	**QPlogBB**	**PSA**
Methotrexate (Reference-drug)	100	0.064	0.024	-0.903	-2	-4.703	232.803
Azelastine	100	1063.613	1390.976	0.933	2	0.647	39.805
Abemaciclib	77.439	123.567	173.578	0.795	1	0.31	68.001
Bosutinib	73.73	100.404	248.227	0.501	1	0.201	77.957
Clemastine	100	2194.519	3155.986	0.985	2	0.722	10.95
Dasatinib	83.991	152.464	259.089	0.162	-1	-0.726	108.579
Desloratadine	100	988.051	1331.585	0.77	2	0.768	24.788
Enasidenib	100	912.342	4220.167	0.373	-1	-0.548	92.723
Hydroxyzine	84.542	175.342	227.401	0.081	1	0.39	40.357
Meloxicam	86.447	285.547	220.429	-0.39	-2	-1.064	107.033

Pexidartinib hydrochloride	100	859.786	4837.662	0.872	0	-0.459	64.382
Rucaparib	84.597	204.26	176.398	0.395	0	-0.157	68.842

**Percentage of human oral absorption (%HOA), intestinal permeability (QPPCaco2) in nm/s, Coefficient of binding to human serum albumin (QPlogKhSa), Central Nervous System activity (CNS), Van der Waals surface area of polar nitrogen and oxygen atoms (PSA), IC50 on QPlogHERG, renal permeability (QPPMDCK) in nm/s, brain/blood partition coefficient (QPlogBB).*

Percentage of Human Oral Absorption, denoted as %HOA, represents the proportion of an orally administered drug that reaches the systemic circulation (Drug Absorption - Clinical Pharmacology - Merck Manuals Professional Edition, 2023). Projected human oral absorption on a scale from 0 to 100% is estimated using a quantitative multiple linear regression model. This characteristic typically aligns closely with Human Oral Absorption, as they both assess the same trait. Values over 80% is considered satisfactory (User Manual, 2012).

Estimated Caco-2 cell permeability in nm/sec was predicted using QikProp, Schrodinger. Caco2 cells mimic the gut-blood barrier. QikProp predictions are for non-active transport. A value greater than 500 nm/sec was considered satisfactory (User Manual, 2012).

Predicted MDCK cell permeability in nanometers per second (nm/sec) was taken into account. MDCK cells are recognized as a reliable representation of the blood-brain barrier. QikProp predictions are specifically for non-active transport. A value greater than 500 nm/sec was considered within the range (User Manual, 2012).

Binding to human serum albumin is anticipated, with a specified range between -1.5 and 1.5 (User Manual, 2012).

Predicted brain/blood partition coefficient was calculated. The required range spans from -3.0 to 1.2. Predicted central nervous system activity was assessed on a scale from -2 (inactive) to +2 (active). The van der Waals surface area of polar nitrogen, oxygen, and carbonyl carbon

atoms was determined. The accepted range is between 7 and 200 square angstroms ($Å^2$) (User Manual, 2012).

The pharmacokinetic characteristics were then analyzed to propose drug candidates for Breast Cancer. For azelastine, %HOA was 100%. The QPPCaco was 1063.613 nm/sec which is better than the recommended range. QPPMDCK value was greater than 500 nm/sec which fell under the recommended range. The QPlogKhsa is between the range of –1.5 – 1.5 kcal/mol. The QPlogBB is between the range of –3.0 – 1.2 and the CNS value is 2. The PSA value is also within the range of 7-200 Å. So, azelastine fulfills all the pharmacokinetic parameters and was selected. For Abemaciclib, the %HOA was 77.439 which was below the 80%, so it was screened out. For bosutinib, the %HOA was below the required 80% and so it was screened out. For clemastine, the %HOA value was over the required 80%. The QPPCaco value was also over 500 nm/sec. The QPPMDCK value was over the required 500 nm/sec. The QPlogKhsa value was also between the recommended range –1.5 – 1.5. The QPlogBB is also between the desired value of –3.0 – 1.2. The CNS value is also 2 which is between the desired value. The PSA value range was within the recommended range (7-200 $Å^2$). So, clemastine could also be a potential candidate.

For dasatinib, the % HOA was above the required 80% but the QPPCaco value was less than 500 nm/sec and so it was screened out. For desloratadine, the %HOA value was over the required 80%. The QPPCaco value was over 500 nm/sec. The QPPMDCK value was over the required 500 nm/sec. The QPlogKhsa and QPlogBB were also within the range (–1.5 – 1.5 kcal/mol and–3.0 – 1.2 respectively). The CNS value was satisfactory. The PSA value range also matched the desired range which is between 7-200 $Å^2$. So, desloratadine was selected. For hydroxyzine, rucaparib and meloxicam, although the %HOA value is over the required 80% but the QPPCaco and QPPMDCK values were well below the required 500 nm/sec mark and they were discarded from consideration. For pexidartinib, the %HOA, QPPCaco, QPPMDCK,

QPlogKhsa, QPlogBB and CNS values were excellent. The PSA value range also matches the desired range which is between 7-200 $Å^2$. So, Pexidartinib hydrochloride was also considered.

3.3 Discussion

DHFR, under normal physiological state catalyzes the reduction of dihydrofolate to tetrahydrofolate, a critical step in DNA, RNA, purine and pyrimidine synthesis. Tetrahydrofolate is vital for one-carbon unit transfer in various biosynthetic processes, establishing DHFR as a key enzyme in cellular metabolism and growth. However, in certain cancers, the expression of DHFR increases markedly contributing to abnormal growth and proliferation in cancer cells (Expression of DHFR in Cancer - Summary - The Human Protein Atlas, 2021).

This study explores the expression of DHFR in BC tissues. Initially, an expression analysis was conducted to assess the significant upregulation of DHFR in BC patients, particularly in Breast Invasive Carcinoma (BRCA), emphasizing DHFR overexpression in BC. Using GEPIA it was found that the expression of DHFR significantly increases in BC tissues compared to normal tissues. Morales et al. have shown the increase of DHFR in BC (Morales et al., 2009).

Genotype-Tissue Expression (GTEx) results showed that DHFR expression in normal breast tissues is markedly higher in females compared to males, which explains why females are more prone to having BC compared to males. The Kaplan-Meier plot (Figure 7) results show that the OS for patients with BC with high DHFR expression is almost half as OS for patients with BC with low DHFR expression. The Kaplan-Meier plots also suggest that, PFS for patients with BC with low DHFR expression was significantly higher compared to OS for patients with high DHFR expression. These results highlight the importance of inhibiting DHFR in BC.

For this study, methotrexate, a known DHFR inhibitor, was selected as the reference drug. Molecular docking was performed to identify molecules with higher binding affinity to DHFR.

The results were compared to methotrexate. Drugs with good binding affinity values were chosen for subsequent study (Table 1). It is crucial that these drugs bind to the same DHFR binding pocket as methotrexate, so superimposition of the drugs to methotrexate in the binding pocket was carried out. Discovery Studio (v 16.1.0.15350) was used to assess the types of bonds, amino acids involved in the binding to DHFR.

From the data from Discovery Studio (v 16.1.0.15350), the following drugs were selected: ripretinib (TKI), bosutinib (TKI), pexidartinib hydrochloride (TKI), infigratinib phosphate (TKI), regorafenib (TKI), abemaciclib (TKI), hydroxyzine (Antihistamine), rofecoxib (NSAID), topotecan (Anti-cancer drug), doxorubicin hydrochloride (Anti-cancer drug), belzutifan (Anti-cancer drug), duvelisib (Anti-cancer drug), enasidenib mesylate (Anti-cancer drug).

Next, an assessment of pharmacokinetic parameters using QikProp, Schrodinger was conducted. Inadequate absorption in the body renders a drug ineffective. Compounds exhibiting less than 80% absorption were excluded. Evaluation of distribution parameters ensured proper drug dissemination throughout the body. Drugs crossing the Blood-Brain Barrier (BBB) were eliminated from consideration due to potential adverse effects. The drugs should also be properly excreted from the body and if they are not excreted and stay in the body for long period of time, it can be detrimental for the body.

Finally, after analysing all the data, clemastine, desloratadine and pexidartinib hydrochloride were proposed for BC based on their binding affinities, superimposition results, interactions and pharmacokinetic profiles. Clemastine formed a carbon hydrogen bond, two alkyl bonds and a pi-alkyl bond in the common amino acids of methotrexate with the protein. Desloratadine formed a pi-donor hydrogen bond, an alkyl bond and two pi-alkyl bonds in the common amino acids of methotrexate with the protein. Desloratadine formed a pi-donor hydrogen bond, an

alkyl bond and two pi-alkyl bonds in the common amino acids of methotrexate with the protein. Pexidartinib hydrochloride formed two conventional hydrogen bond, halogen (fluorine), a conventional hydrogen bond and two pi-alkyl bonds in the common amino acids of methotrexate with the protein. Hydrogen bonds are considered strong interaction (Przybysz et al., 2016).

This study has a few limitations. Only around 300 drugs were used for the study. Using a larger library could improve the findings of the study. Second, molecular dynamic simulation was not carried out, which could have helped confirm and strengthen our findings from the docking experiments. Consideration of these factors in future research could enhance the overall reliability of the study.

4. Conclusion:

The first part of the study demonstrated high expression of DHFR in breast cancer. It also demonstrated that elevated DHFR expression can lead to lower overall and progression free survival.

The study proposes clemastine (antihistamine), desloratadine (antihistamine) and pexidartinib hydrochloride (TKI) as suitable candidates for treatment of BC based on the molecular docking study results. The binding affinities of these drugs were -9.4 kcal/mol, -10.2 kcal/mol and -10.0 kcal/mol respectively which were better than the binding affinity of the reference, methotrexate (-9.1 kcal/mol). Methotrexate interacted with ASN64, ARG70, GLN35, THR146, SER59, LEU22, ILE60, ILE16 which were considered to be part of the binding pocket of DHFR. Clemastine and desloratadine were found to interact with three amino acids that were in common with methotrexate. Pexidaritinib hydrochloride had five amino acids in common with methotrexate. Clemastine formed a carbon-hydrogen bond, two alkyl bonds, and a pi-alkyl bond within the common amino acids of methotrexate with the protein. Desloratadine formed a pi-donor hydrogen bond, an alkyl bond, and two pi-alkyl bonds with the common amino acids of methotrexate with the protein. Pexidartinib hydrochloride formed two conventional hydrogen bonds; halogen (fluorine), a conventional hydrogen bond, and two pi-alkyl bonds with the common amino acids of methotrexate with the protein, indicating strong interactions. They all demonstrated satisfactory predicted pharmacokinetic properties as well. Further validation can be done using biological assays using cell lines followed by *in vivo* studies.

Reference:

Adewole, K., Ishola, A., & Olaoye, I. (2022). In silico profiling of histone deacetylase inhibitory activity of compounds isolated from Cajanus cajan. *Beni-Suef University Journal of Basic and Applied Sciences*, *11*(1). https://doi.org/10.1186/s43088-021-00191-y

Askari, B. S., & Krajinovic, M. (2010). Dihydrofolate Reductase Gene Variations in Susceptibility to Disease and Treatment Outcomes. In *Current Genomics* (Vol. 11).

Bertino, J. R., Göker, E., Gorlick, R., Li, W. W., & Banerjee, D. (1996). Resistance Mechanisms to Methotrexate in Tumors. *The Oncologist*, *1*(4), 223–226. https://doi.org/10.1634/theoncologist.1-4-223

Biochemistry Free & Easy (Ahern and Rajagopal) - Biology LibreTexts. (2023). Retrieved March 1, 2024, from https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Free_and_Easy_(Ahern_and_Rajagopal)

BRCA Gene Mutations: Cancer Risk and Genetic Testing Fact Sheet - NCI. (2020). Retrieved March 1, 2024, from https://www.cancer.gov/about-cancer/causes-prevention/genetics/brca-fact-sheet

BRCA1 and BRCA2 (PDQ®) - NCI. (2024). Retrieved March 1, 2024, from https://www.cancer.gov/about-cancer/causes-prevention/genetics/brca-genes-hp-pdq

Breast cancer. (2024). Retrieved March 1, 2024, from https://www.who.int/news-room/fact-sheets/detail/breast-cancer

Breast Cancer Early Detection and Diagnosis | How To Detect Breast Cancer | American Cancer Society. (2024). Retrieved March 1, 2024, from

https://www.cancer.org/cancer/types/breast-cancer/screening-tests-and-early-detection.html

Breast Cancer Early Detection and Diagnosis | How To Detect Breast Cancer | American Cancer Society. (2024). https://www.cancer.org/cancer/types/breast-cancer/screening-tests-and-early-detection.html

Breast Cancer Statistics And Resources | Breast Cancer Research Foundation. (2023). Retrieved March 1, 2024, from https://www.bcrf.org/breast-cancer-statistics-and-resources/

Burguin, A., Diorio, C., & Durocher, F. (2021). Breast cancer treatments: Updates and new challenges. *Journal of Personalized Medicine*, *11*(8). https://doi.org/10.3390/jpm11080808

Cancer Cells: Types, How They Form, and Characteristics. (2023). Retrieved March 1, 2024, from https://www.verywellhealth.com/what-are-cancer-cells-2248795

Chandrashekar, D. S., Karthikeyan, S. K., Korla, P. K., Patel, H., Shovon, A. R., Athar, M., Netto, G. J., Qin, Z. S., Kumar, S., Manne, U., Crieghton, C. J., & Varambally, S. (2022a). UALCAN: An update to the integrated cancer data analysis platform. *Neoplasia (United States)*, *25*, 18–27. https://doi.org/10.1016/j.neo.2022.01.001

Chandrashekar, D. S., Karthikeyan, S. K., Korla, P. K., Patel, H., Shovon, A. R., Athar, M., Netto, G. J., Qin, Z. S., Kumar, S., Manne, U., Crieghton, C. J., & Varambally, S. (2022b). UALCAN: An update to the integrated cancer data analysis platform. *Neoplasia (United States)*, *25*, 18–27. https://doi.org/10.1016/j.neo.2022.01.001

Dalivandan, S. T., Plummer, J., & Gayther, S. A. (2021). Risks and function of breast cancer susceptibility alleles. In *Cancers* (Vol. 13, Issue 16). MDPI AG. https://doi.org/10.3390/cancers13163953

DHFR gene mutation - MTHFR Gene health™. (2023). Retrieved March 1, 2024, from https://mthfrgenehealth.com/dhfr-gene-mutation/

Difference Between Wild Type and Mutant | Definition, Characteristics, Examples, Similarities. (2017). Retrieved March 1, 2024, from https://pediaa.com/difference-between-wild-type-and-mutant/

Drug Absorption - Clinical Pharmacology - Merck Manuals Professional Edition. (2023). Retrieved March 7, 2024, from https://www.merckmanuals.com/professional/clinical-pharmacology/pharmacokinetics/drug-absorption#maincontent

Edelman, L. B., Eddy, J. A., & Price, N. D. (2010). In silico models of cancer. In *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* (Vol. 2, Issue 4, pp. 438–459). https://doi.org/10.1002/wsbm.75

Expression of DHFR in breast cancer - The Human Protein Atlas. (2021). Retrieved March 1, 2024, from https://www.proteinatlas.org/ENSG00000228716-DHFR/pathology/breast+cancer

Expression of DHFR in cancer - Summary - The Human Protein Atlas. (2021). Retrieved March 7, 2024, from https://www.proteinatlas.org/ENSG00000228716-DHFR/pathology

Genotype-Tissue Expression Project (GTEx). (2020). Retrieved March 2, 2024, from https://www.genome.gov/Funded-Programs-Projects/Genotype-Tissue-Expression-Project

Hasan, M. M., Khan, Z., Chowdhury, M. S., Khan, M. A., Moni, M. A., & Rahman, M. H. (2022). In silico molecular docking and ADME/T analysis of Quercetin compound with its evaluation of broad-spectrum therapeutic potential against particular diseases. *Informatics in Medicine Unlocked*, *29*. https://doi.org/10.1016/j.imu.2022.100894

Hazard Ratio: Interpretation & Definition - Statistics By Jim. (2023). Retrieved March 5, 2024, from https://statisticsbyjim.com/probability/hazard-ratio/

Hodos, R. A., Kidd, B. A., Shameer, K., Readhead, B. P., & Dudley, J. T. (2016). In silico methods for drug repurposing and pharmacology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *8*(3), 186–210. https://doi.org/10.1002/wsbm.1337

How cancer starts, grows and spreads | Canadian Cancer Society. (2021). Retrieved March 1, 2024, from https://cancer.ca/en/cancer-information/what-is-cancer/how-cancer-starts-grows-and-spreads

Jiang, T., Chen, J., Fu, H., Wu, K., Yao, Y., Eyi, J. U. M., Matesa, R. A., Obono, M. M. O., Du, W., Tan, H., Lin, M., & Li, J. (2019). High prevalence of Pfdhfr-Pfdhps quadruple mutations associated with sulfadoxine-pyrimethamine resistance in Plasmodium falciparum isolates from Bioko Island, Equatorial Guinea. *Malaria Journal*, *18*(1). https://doi.org/10.1186/s12936-019-2734-x

Johnson, K. S., Conant, E. F., & Soo, M. S. (2021). Molecular Subtypes of Breast Cancer: A Review for Breast Radiologists. In *Journal of Breast Imaging* (Vol. 3, Issue 1, pp. 12–24). Oxford University Press. https://doi.org/10.1093/jbi/wbaa110

Kaplan-Meier plotter. (2024). Retrieved March 20, 2024, from https://kmplot.com/analysis/

Koźmiński, P., Halik, P. K., Chesori, R., & Gniazdowska, E. (2020). Overview of dual-acting drug methotrexate in different neurological diseases, autoimmune pathologies and

cancers. In *International Journal of Molecular Sciences* (Vol. 21, Issue 10). MDPI AG. https://doi.org/10.3390/ijms21103483

Leclair, R. J. (2021). *CELL BIOLOGY, GENETICS, AND BIOCHEMISTRY FOR PRE-CLINICAL STUDENTS*. https://LibreTexts.org

Li, P. C., Zhu, Y. F., Cao, W. M., & Li, B. (2024). ER-positive and BRCA2-mutated breast cancer: a literature review. In *European Journal of Medical Research* (Vol. 29, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s40001-023-01618-1

Lucia, J. S., Sozhamannan, S., Gans, J. D., Koehler, J. W., Soong, R., Lin, N. J., Xie, G., Olson, V., Roth, K., & Beck, L. (2020). Appendix Q: Recommendations for developing molecular assays for microbial pathogen detection using modern in silico approaches. *Journal of AOAC International*, *103*(4), 882–899. https://doi.org/10.1093/JAOACINT/QSAA045

Morales, C., García, M. J., Ribas, M., Miró, R., Muñoz, M., Caldas, C., & Peinado, M. A. (2009). Dihydrofolate reductase amplification and sensitization to methotrexate of methotrexate-resistant colon cancer cells. *Molecular Cancer Therapeutics*, *8*(2), 424–432. https://doi.org/10.1158/1535-7163.MCT-08-0759

Nakano, M., Fukami, T., Gotoh, S., & Nakajima, M. (2017a). A-to-I RNA editing up-regulates human dihydrofolate reductase in breast cancer. *Journal of Biological Chemistry*, *292*(12), 4873–4884. https://doi.org/10.1074/jbc.M117.775684

Nakano, M., Fukami, T., Gotoh, S., & Nakajima, M. (2017b). A-to-I RNA editing up-regulates human dihydrofolate reductase in breast cancer. *Journal of Biological Chemistry*, *292*(12), 4873–4884. https://doi.org/10.1074/jbc.M117.775684

Osborne, M. J., Schnell, J., Benkovic, S. J., Dyson, H. J., & Wright, P. E. (2001). Backbone dynamics in dihydrofolate reductase complexes: Role of loop flexibility in the catalytic mechanism. *Biochemistry*, *40*(33), 9846–9859. https://doi.org/10.1021/bi010621k

Przybysz, P., Dubowik, M., Kucner, M. A., Przybysz, K., & Buzała, K. P. (2016). Contribution of hydrogen bonds to paper strength properties. *PLoS ONE*, *11*(5). https://doi.org/10.1371/journal.pone.0155809

Raimondi, M. V., Randazzo, O., Franca, M. La, Barone, G., Vignoni, E., Rossi, D., & Collina, S. (2019). DHFR inhibitors: Reading the past for discovering novel anticancer agents. In *Molecules* (Vol. 24, Issue 6). MDPI AG. https://doi.org/10.3390/molecules24061140

Rao, K. S., & Subash, P. (2021). *In Silico Drug Repurposing: An Effective Tool to Accelerate the Drug Discovery Process*. www.intechopen.com

Royle, K. L., Meads, D., Visser-Rogers, J. K., White, I. R., & Cairns, D. A. (2023). How is overall survival assessed in randomised clinical trials in cancer and are subsequent treatment lines considered? A systematic review. In *Trials* (Vol. 24, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13063-023-07730-1

Survival Rates for Breast Cancer | American Cancer Society. (2022). Retrieved March 1, 2024, from https://www.cancer.org/cancer/types/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-survival-rates.html

Tang, S., Chen, R., Lin, M., Lin, Q., Zhu, Y., Ding, J., Hu, H., Ling, M., & Wu, J. (2022). Accelerating AutoDock Vina with GPUs. *Molecules*, *27*(9). https://doi.org/10.3390/molecules27093041

Teachey, D. T., Sheen, C., Hall, J., Ryan, T., Brown, V. I., Fish, J., Reid, G. S. D., Seif, A. E., Norris, R., Chang, Y. J., Carroll, M., & Grupp, S. A. (2008). mTOR inhibitors are

synergistic with methotrexate: An effective combination to treat acute lymphoblastic leukemia. *Blood*, *112*(5), 2020–2023. https://doi.org/10.1182/blood-2008-02-137141

The Role of BRCA Mutations in Breast Cancer. (2022). Retrieved March 1, 2024, from https://www.verywellhealth.com/brca-mutations-and-breast-cancer-4158206

Treatment of Stage IV (Metastatic) Breast Cancer | American Cancer Society. (2024). Retrieved March 1, 2024, from https://www.cancer.org/cancer/types/breast-cancer/treatment/treatment-of-breast-cancer-by-stage/treatment-of-stage-iv-advanced-breast-cancer.html

Understanding Breast Cancer Survival Rates | Susan G. Komen®. (2024). Retrieved March 1, 2024, from https://www.komen.org/breast-cancer/facts-statistics/breast-cancer-statistics/survival-rates/

User Manual, Q. (2012). *Schrödinger Press QikProp 3.5 User Manual QikProp User Manual*.

What Are the Risk Factors for Breast Cancer? | CDC. (2021). Retrieved March 1, 2024, from https://www.cdc.gov/cancer/breast/basic_info/risk_factors.htm

What Progression-Free Survival Means After Cancer. (2023). Retrieved March 5, 2024, from https://www.verywellhealth.com/progression-free-survival-pfs-2252163

Witt, B. L., & Tollefsbol, T. O. (2023). Molecular, Cellular, and Technical Aspects of Breast Cancer Cell Lines as a Foundational Tool in Cancer Research. *Life*, *13*(12), 2311. https://doi.org/10.3390/life13122311

Yuan, S., Chan, H. C. S., & Hu, Z. (2017). Using PyMOL as a platform for computational drug design. In *Wiley Interdisciplinary Reviews: Computational Molecular Science* (Vol. 7, Issue 2). Blackwell Publishing Inc. https://doi.org/10.1002/wcms.1298

Zheng, Y., & Cantley, L. C. (2019). Toward a better understanding of folate metabolism in health and disease. In *Journal of Experimental Medicine* (Vol. 216, Issue 2, pp. 253–266). Rockefeller University Press. https://doi.org/10.1084/jem.20181965