Novel sulfonamides for the treatment of breast and cervical cancers

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Science (Research) in Chemistry

at

The University of Waikato

by

Nathaniel Buckley

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2022

Abstract

A series of novel sulfonamides were studied computationally to identify a potential mechanism of action for the treatment of breast and cervical cancers. Previous *in vivo* testing demonstrated varying degrees of effectiveness, with seven molecules showing the highest cytotoxicity. Computational calculations were chosen as the means for determining the drug target. Previous studies indicated β-Tubulin and Carbonic Anhydrase IX as the most likely proteins of interest, so these were selected for study.

The computational methodology selected for this study follows that of Krzywik et al.[1] For this, conventional docking studies were undertaken using Autodock Vina, followed by more accurate and computationally demanding molecular dynamics simulations using the MM/PBSA procedures in AMBER. Docking studies for the seven active compounds were performed using Vina, and showed promising results. The three most effective compounds (5AP, 6AP and 7BP) were then utilised in a series of MM/PBSA calculations to determine strength of interaction between the sulfonamides and drug targets.

The results of the AMBER calculations determined β-Tubulin to be the more likely drug target, although this result did show 6AP (the most effective *in vivo* molecule) to have the weakest binding, which was an unexpected result. A LIGPLOT study was done to determine a potential cause for this, showing a heavy emphasis on non-polar interactions between the drug molecules and proteins. Drug-like comparisons and pKa calculations were also performed, showing all three molecules to have very similar acidity and hydrogen bond capabilities, but suggesting lipophilicity and molecule size to be possible determining factors in the overall effect of the compounds. Transport across the cell membrane and through the cytoplasm were also suggested as potential reasons for the differing results, and further studies into these were suggested for possible confirmation.

Preface

The research for this thesis was done in collaboration with Dr Aneela Maalik at COMSATS University Islamabad, and is a continuation of her research into a novel set of sulfonamides developed by her team. The *in vivo* experimental results reported in this work were provided by Dr Maalik in a private communication[2]. A draft manuscript that includes both these experimental results and the computational studies of this thesis is currently in the final stages of preparation and will be submitted shortly.

Acknowledgements

This thesis represents the culmination of two years of part-time study. As such, a number of people have helped in some regard, and I would like you acknowledge this assistance.

First, to my supervisor, Associate Professor Jo Lane. Your guidance and help with the various calculations and research involved in this thesis has been invaluable, and I could not have completed it without you. Thank you for all of the work you have helped with, and for putting up with my relative inexperience with calmness and compassion.

Next, to Dr Aneela Maalik and your team. Thank you for providing the basis for this research, and assistance with clarifying parts of your own study when required. I hope the information I was able to provide is of use for your own projects.

To NeSI's Application Support Specialist Ben Roberts and the wider NeSI team, thank you for allowing me use of your supercomputers to carry out my research. Also, specifically to Ben, thank you for your help troubleshooting the various issues that arose from attempting the calculations, and guiding me through the use of Linux and the multiple programs utilised for this study. The calculations would not have been able to be completed without this help.

To my previous employers Ahmad Zareh and Kylie Burne, and my current employers Nathan and Andrea Reilly, thank you for allowing me the time to attend classes and meetings with my supervisor. I would not have been able to complete my studies to the level achieved without this relieving of stress from taking time off work.

Finally, to my family, my beautiful wife Riddhi and my perfect daughter Samira. This has been hard, balancing study, work and raising a new-born. You have given me not only the support and strength I needed to get through these last two years of study, but also bringing joy and laughter. I hope I have made you proud. I love you both.

Table of Contents

Table of Figures

Table of Abbreviations

CA IX – Carbonic Anhydrase IX

WHO – World Health Organisation

ER+ – Endocrine Receptor Positive

HER2+ – Human Epidermal Growth Factor 2 Positive

BRCA1/BRCA2 – Breast Cancer Gene 1 & 2

LCIS – Lobular Carcinoma in Situ

DCIS – Ductal Carcinoma in Situ

TNM – Tumour size, Number of nearby lymph nodes affected, Metastasis

HPV – Human Papilloma Virus

HTS – High-Throughput Screening

MD – Molecular Dynamics

CPU – Central Processing Unit

GPU – Graphical Processing Unit

GTP – Guanosine Triphosphate

GDP – Guanosine Diphosphate

BCN – Bicyclononyne

DMS – Dimethyl Sulfide

MM/PBSA – Molecular Mechanics combined with Poisson-Boltzmann and Surface Area solvation

MM/GBSA – Molecular Mechanics combined with General-Born and Surface Area solvation

PDB – Protein Databank

1. Introduction

Breast and cervical cancers are some of the most common forms of the disease currently affecting women in the world[6,7], and as such, a large amount of study has been done to treat these deadly conditions. However, current treatments have proven to not always be effective, and can have potentially harmful and unwanted side-effects. As such, research into this subject is ever-increasing, with new treatment avenues being explored.

Advances in the world of computational chemistry have led to the creation of new methods for understanding ligand-protein interactions, and assisted in the creation and testing of new medicines[8,9,10]. These advances have been boosted by the increase in computational power brought on by the development of supercomputers containing vast quantities of powerful GPUs, which allow large amounts of calculations to be processed at a speed hitherto unheard of[11]. Researchers have utilised these supercomputers to perform calculations, which previously may have taken a number of days, in a matter of hours or even minutes. This has greatly improved our ability to study and understand how ligands interact with proteins, and which potential compounds may be more effective in treating various medical conditions, such as the development of raltegravir for HIV[12].

This study will utilise computational chemistry techniques to study a selection of novel sulfonamides created and tested *in vitro* by Dr Aneela Maalik from COMSATS University Islamabad. The aim of this thesis is to determine which specific ligand-protein interactions are occurring in cancer cells, to confirm a possible mechanism of action for these sulfonamides, and thus gain a better understanding of their potential use into the future.

1.1 Breast Cancer

As the world's most prevalent form of cancer[6], and second leading cause of cancer deaths in women[13], much has been studied in regards to breast cancer. One such area of importance in diagnosing and treating this disease is the classification of differing types, as different forms require specific treatments.

Breast cancer traditionally are classified as one of four distinct types: Endocrine Receptor Positive (ER+), Human Epidermal Growth Factor Receptor 2 Positive (HER2+), Triple Positive (HER2+, Progesterone- and Oestrogen-Receptor Positive) or Triple Negative (all receptors negative)[13]. These are determined by the presence or absence of the certain receptors on the cellular membranes of the cancerous cells. However, recent research has led to a more specified classification system. Based on characteristics of the malignant cells such as specific gene expression, clinical features of the condition, in addition to presence or absence of hormone receptors, breast cancer can now be defined as Luminal (A and B predominantly), HER2 or Basal[14].

Although breast cancer is the most common form of the disease, a number of risk factors increase a person's risk of developing malignancy. A person's age and gender play a significant role, with older females having the highest prevalence of breast cancer[6,15]. A personal or family history (particularly with immediate relatives) can be a sign of the presence of mutations in genes such as BCRA1 and BRCA2, which code for tumour suppression, and therefore can lead to higher incidence of breast

cancer[15,16]. Changes in levels of the hormones oestrogen and progesterone, due to age, menopause or supplementation, can also affect the chance of a person developing breast cancer[15].

Development of cancerous lesions in breast tissue can occur in either the lobules (LCIS, or lobular carcinoma in situ) or the ducts (DCIS, or ductal carcinoma in situ), with the latter being the most common occurrence[15]. Of these, the vast majority (>99%) take the form of carcinomas, with the remaining amount being sarcomas[17]. If left untreated, lesions can metastasize to surrounding lymph nodes, most notably in the armpits, and from there to other organs of the body[17]. Overall prognosis of breast cancer is relatively good, with a 100% 5-year survival rate if caught early (Stage 1 of the TNM Classification System[18]). The more metastasis that occurs, the worse the survival rate becomes, with 93% over 5 years for Stage 2 cancers, 72% for Stage 3, and extensively spread Stage 4 lesions having only a 22% 5-year survival rate[15].

Current treatments for breast cancer fall under 4 main categories: surgical, radiation, chemotherapy and endocrine therapy[19]. Selection of treatment largely depends on the current disease state, such as how large the lesion is and level of metastasis, as well as patient-centric factors such as age, co-morbidities, physiological factors such as liver and kidney function, and socioeconomic factors[19]. In the majority of cases where little to no metastasis has occurred, surgical removal of the lesion is recommended[15]. Depending on the size of the lesion however, surgical oncology may require removal of part or all of the affected breast, which can have an array of social and mental effects on the patient being treated. In these cases, tumours may first be shrunk using a combination of radiation and medicinal therapies[15].

Radiation can be used to treat larger tumours and locally affected lymph nodes simultaneously, and it has been shown to vastly decrease the chances of relapse if used early enough[15,20]. The use of radiotherapy post-mastectomy has also shown to increase survival rates by up to 10%[20]. Radiation can also be used in metastasized cancers to reduce symptoms in palliative care[6,15].

In terms of medicinal therapies for breast cancer, there are a number of factors that must be considered. Arguably the most important of these is what receptors are present on the cell membrane of cancer cells, specifically HER2, progesterone and oestrogen receptors[6]. A combination or absence of one or each of these can lead to vastly different therapies, as they are often the targets for medicinal treatment. Tumours containing the first of these, HER2, can be treated with a number of monoclonal antibodies, the most common of which being trastuzumab (Herceptin)[17]. By contrast, tumours containing an abundance of Oestrogen receptors can be treated using aromatase inhibitors such as anastrozole, letrozole or exemestane[21]. More advanced cancers may require the use of oestrogen-receptor antagonists such as fulvestrant[21]. In the majority of cases where surgery is possible and utilised, adjuvant treatment with these medicines is common to prevent possible metastases (and therefore disease relapse) from occurring.

1.2 Cervical Cancer

The other form of cancer that was studied by Maalik et al. (and therefore is included in this thesis) as targets for treatment was cervical cancer. Whilst not as prevalent as breast cancer, this disease is in no way any less deadly, with the World Health Organisation (WHO) listing over 570,000 new diagnoses, and 311,000 deaths from cervical in just 2018 alone[7]. As such, treatment for this condition is also of significant importance.

The vast majority of cervical cancer is caused by the Human Papillomavirus (HPV)[7,22]. This species of virus has as many as 130 types, however, only 20 have been known to cause cervical cancer[22]. Of these, types 16 and 18 are the most prevalent in terms of cervical cancer development[22,23]. As the main transmission of this virus occurs due to physical contact during intercourse[22], this condition is often labelled as a sexually transmitted disease. However, simple infection by HPV is often not the sole causative factor, as infections tend to be self-limiting, and only proliferate if external factors such as co-infection, immunosuppression and oral contraceptive use prevent this from happening[22]. HPV causes cervical cancer through suppression of the tumour-suppressing gene TP53, which in turn leads to hyperproliferation of the affected cells, and therefore tumour growth[24].

Diagnosis and classification of cervical cancer is often difficult[25]. The usual initial test for detection is a Pap smear, which involves a trained health professional using a brush or spatula to collect a sample of tissue from a patient's cervix[25]. This is then sent to a testing facility to check for the presence of HPV or any abnormal cells, which poses issues of false positives or negatives, as well as potentially being cost-inefficient due to the training and equipment required[24,25]. Visual checks and biopsies may also be done to check for or confirm the presence of a lesion[24].

Once an HPV infection is established, cervical cancer can form as one of two main types: squamous cell carcinoma and adenocarcinoma[26]. Of these, squamous cell carcinomas, affecting the flatter squamous cells on the lower portion of the cervix[26], is vastly the more common form[24,26]. Adenocarcinomas, occurring in the glandular cells present in the upper cervix, is significantly less common[24,26]. However, both have the propensity to cause invasive disease progression by infection of nearby lymph nodes and metastasis[22,24]. As such, treatments for this disease can be relatively wide-ranging.

The first course for the form of treatment of cervical cancer, particularly from a public health stand-point, is through prevention of serious HPV infection[7]. As HPV is the predominant cause of cervical cancer, this approach has proven significantly effective. To achieve this, a mass vaccination campaign for young girls was initiated at the recommendation of and by the WHO[7]. The vaccination consists of two to three doses of HPV vaccine, starting as low as 9 years old (though most commonly given around 12-13[27]) and separated by 6 to 12 months depending on age of initiation[27]. HPV vaccination has led to the reduction of cervical cancer caused by HPV16 and HPV18 by up to 90% in those fully vaccinated[22,28].

Treatment of cervical cancer is very similar to that of breast cancer; initially, surgery is used, followed by radiation and chemotherapy to either assist or replace this depending on lesion size and rate of metastasis[24]. Conisation of the cervix, or complete hysterectomy are the primary surgeries performed for this condition, depending on rate of spread and fertility of the patient[24]. Where surgery alone is unsuccessful or subtherapeutic, chemotherapy measures may also be employed[24,29]. The most commonly-used medicine is cisplatin, as platinum-based compounds can act as a catalyst for inducing cell death is cancer cells[29]. Paclitaxel may also be employed in severe cases[24].

1.3 Prior Research: Maalik et al.

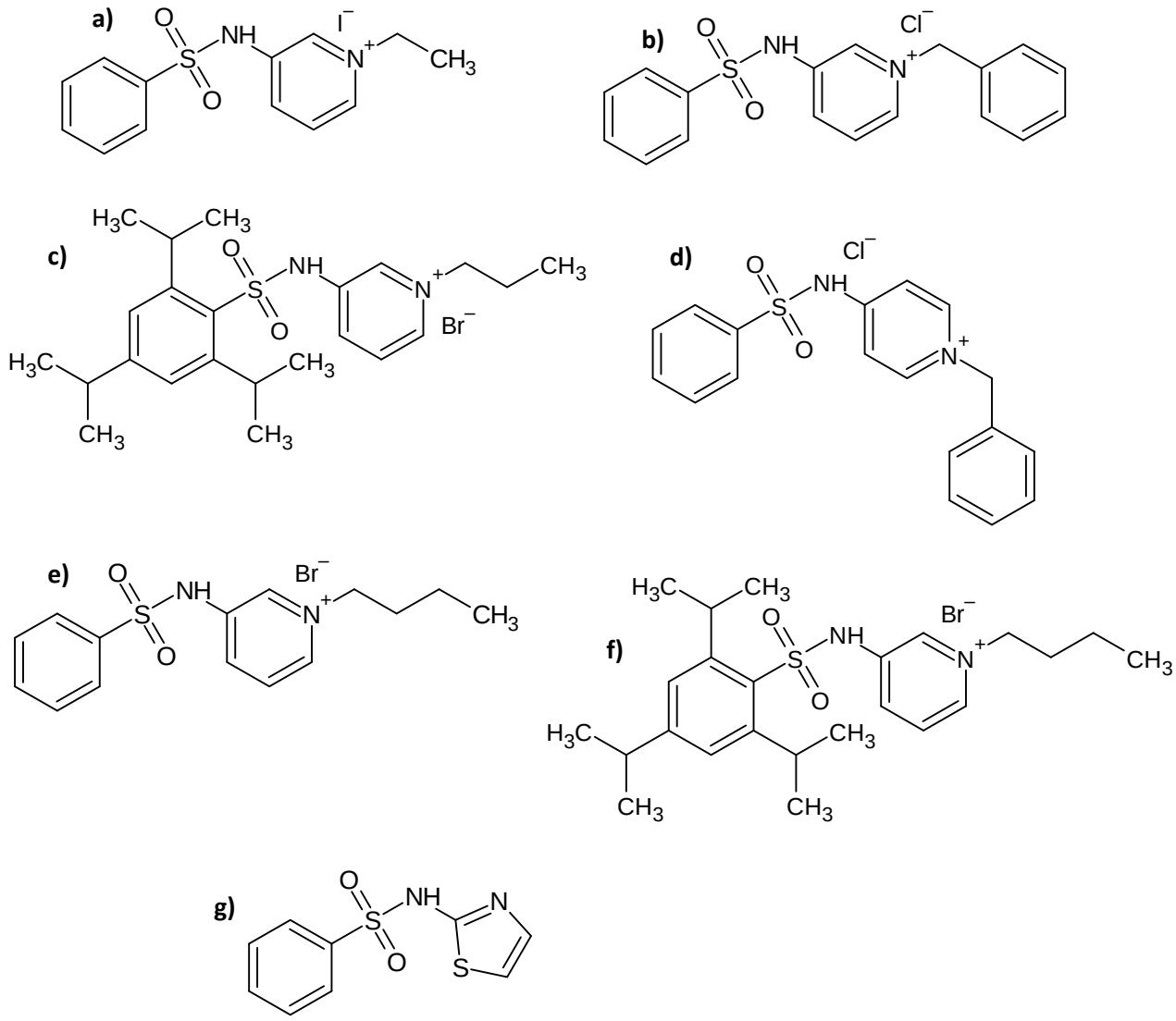


Fig. 1.1: Molecular structures of sulfonamides generated by Maalik et al. that were selected for study, designated as a) 6AP, b) 5AP, c) 7BP, d) 4AN2, e) 7AP, f) 6BP and g) 5A

As stated above, this study is in part a continuation of research by Dr Aneela Maalik from COMSATS University Islamabad. In her research, she has developed a number of sulfonamide-based alkyl-N-pyridyl derivatives, with the intention of using these for treating breast and cervical cancer.

Dr Maalik developed and synthesised 20 potential molecules for the study. These compounds were then tested against a HeLa cell line (an immortal line of cervical cancer cells[30]) and breast carcinoma cells. Effectiveness of the potential drugs was measured by studying apoptosis of the cells tested, and IC_{50} values for the compounds calculated. Apoptosis was studied via the use of the dye annexin V-FITC, which reacted with phosphatidylserine released by the cells upon cell death, and glow a strong green colour.

Of the molecules tested, seven showed relatively acceptable induction of apoptosis. These molecules were labelled as 6AP, 5AP, 7BP, 4AN2, 7AP, 6BP and 5A (in order of most effective to least effective, as seen in Figure 1.1 above). For the purposes of this study, only these compounds will be studied to determine the mode of action for these sulfonamides, as the remaining 13 molecules required significantly higher doses to achieve the same effect, thus making them much less suitable as medications. By contrast, the seven molecules chosen showed strong effects against the cell lines tested, with relatively low IC_{50} scores, indicating a lower dose is required to induce cell death. It is important to note that the majority of these molecules are in fact ionic in nature, having a net positive charge on the main structure (balanced by negatively-charged counter-ions), which may play a role in their overall effect.

1.4 Sulfonamides

The molecules Dr Maalik and her team synthesized are variations of sulfonamides, one class of potentially successful compounds that has been gaining traction over the past few years to treat cancer in humans[31]. These are a class of medicine encompassing molecules which contain a sulfonamide (SO_2NH) bond[32]. Other uses for these compounds range from antibiotics to diuretics and anti-inflammatories[32].

Potential medical uses for sulfonamides were first discovered in the late 1920's by Gerhard Domagk[33]. He found that prontosil would selectively reduce the amount of infectious streptococcus cells in a sample, reportedly when looking for a cure for his daughter's illness[34]. Further research into the 1930's revealed prontosil acted as a pro-drug for sulfanilamide, and the mode of action was determined to be disruption of the folic acid cycle via competitive inhibition of p-aminobenzoic acid (PABA)[34]. These discoveries led to a relative explosion of study into this class of medication throughout the 1940's and further[33].

Structurally, the key part of any sulfonamide is the sulfonamide bond[32,34]. However, commonly attached to this is a benzene ring, which when combined with the sulfonamide bond forms the basis of the PABA-inhibiting portion of the compound[34]. Side-chains do vary widely from this base compound, and occasionally the sulfonamide bond may be present as part of a ring itself[34]. In solution, sulfonamides are weakly acidic (with pKa values around 5.5-6.4), losing the hydrogen bound to the nitrogen in the sulfonamide bond in the process[35]. However, although the nitrogen easily loses its bound hydrogen, the strength of the sulfur-oxygen bonds prevents formation of a second nitrogen-sulfur bond, thus resulting in an unstable compound, and leading to the relatively minor pH change[35]. The presence of the amine bond does allow for the formation of hydrogen bonds, and the overall potentials of the sulfonamide region means this area is relatively polar[35]. In terms of physical properties, sulfonamides usually form a waxy solid or oily liquid[36]. The melting point can be relatively high however, being around 165°C for sulfanilamide[37].

The chemical properties of sulfonamides depend predominantly on the side chains, although they do tend to be at least somewhat polar[37,38]. This is due to the polarity of the sulfonamide bond itself, which is strong enough to even make non-polar compounds such as cryptopleurine, polar, as shown by Kwon et al.[38] This therefore means sulfonamides are relatively soluble in water and other polar solutions (such as cytoplasm)[38]. However, extensive non-polar side chains may create dipoles across the molecule, which can have an effect on the interactions the molecule may have within a cell[39].

As well as being potent antibacterials, diuretics and anti-inflammatories, recent research has focussed on the use of sulfonamides as anti-cancer therapies. Several *in vivo* studies have shown the effectiveness of these compounds in the slowing of cell proliferation, reduction in tumour growth, and even induction of cell apoptosis[1, 31,40,41,42,43]. In these recent previous studies, two key proteins have been identified as mechanisms by which sulfonamides induce these effects: β-Tubulin, and Carbonic Anhydrase IX[31,40,41,42,43].

1.4.1 β-Tubulin

Microtubules form an integral part of cells, both healthy and cancer cells alike. Their main role is to provide structure and strength to a cell, acting as mechanical-like rods to help stabilise the overall cell shape[44,45]. They also act as transport systems for many cell nutrients, and during mitosis, form a structure known as the mitotic spindle, which lines up and then separates chromosomes[1]. In highly proliferating systems such as cancerous tumours, this last role is of utmost importance. As such a key part of cell life and function, this has therefore made them ideal targets for a range of medications[42].

Microtubules consist of multiple dimers comprised by the proteins α-tubulin and β-tubulin[44,45]. These two proteins alternatively bind together in long, cylindrical chains to form the overall microtubule structure. The dimers are arranged in such a way that β-tubulin is always showing on the outer edge of the microtubule, with a tubulin dimer binding to it via the α-tubulin protein to extend the microtubule[44,45]. α-tubulin contains 426 amino acids, with β-tubulin having a larger primary chain of 437 amino acids[46]. A dimer will typically weigh around 97.6kDa[46], with the length being determined by the size of cell in question, though some can be as long as 20µm or even larger[47].

This extension of a microtubule is facilitated by the binding and hydrolysis of GTP, which allows conformational change of the tubulin proteins to occur. However, if left unbound the GTP is hydrolysed to GDP, which binds far less strongly to the microtubule structure and therefore leads to unravelling of the microtubule[1,45]. This unravelling then removes the structural strength and integrity of the microtubule, leading to cell-wide effects if left to occur frequently, in particular preventing the formation of the mitotic spindles[44].

Due to the breakdown of microtubules if GTP is allowed to be hydrolysed, this is an ideal target for potential medications. One particular medication that achieves this remarkably well is colchicine, a drug typically used to treat gout flare-ups[1]. Colchicine binds to a particular region of the β-tubulin protein (appropriately named the colchicine-binding site), located in the middle of the dimer, and prevents the conformational changes necessary for the β-tubulin to bind to an α-tubulin[1,42]. As such, the GTP is hydrolysed to GDP and the unravelling of the microtubule occurs.

Sulfonamides are theorised to bind to β-tubulin in the colchicine-binding site, and thus induce the same breakdown of the structures as colchicine[41,46]. A number of studies have been conducted which support this theory, notably by Banerjee et al.[42]. As such, for the molecules being studied in this thesis, the same or a similar mode of action is expected to be occurring.

1.4.2 Carbonic Anhydrase IX

Carbonic Anhydrase IX (CA IX) is an enzyme found on the extracellular membrane of human cells, and is responsible for the regulation of pH balance of these cells[31,43,48]. In tumour cells, it also regulates the effects of hypoxia on the cells, thus promoting the formation of an acidic environment that

encourages tumour growth[31,43]. Due to this important role in cancer cells, upregulation of this enzyme occurs, leading to a higher presence on the membrane and increased importance of the protein's role[31,48]. This therefore makes CA IX an ideal drug target for anti-cancer medications.

CA IX's primary enzymatic activity involves the conversion of carbon dioxide and water into bicarbonate and hydrogen ions in the extracellular medium[31]. These bicarbonate ions are then absorbed back into the cell, where they bind with free-floating protons to reform the carbon dioxide and water, thus removing the proton from the cytoplasm and reducing the acidity of the cell contents[31]. This reaction is achieved due to a zinc ion in the active site of the enzyme, which activates the water molecules to stimulate the reaction[31].

One potential issue with targeting CA IX however, is the relatively non-specific nature of the carbonic anhydrase family (which consists of at least 14 different enzymes)[43]. The result of this is that the binding of a medication to the protein may have low specificity, therefore leading to inefficient binding of the drug being used[31,43]. As a result, higher doses of drug may be required to achieve the required amount of inhibition, thus leading to potentially more adverse side effects caused by the medication.

The way sulfonamides inhibit CA IX is through direct competitive inhibition of the enzyme, by binding to the active site of the protein[31,43]. This inhibition prevents the enzyme from catalysing the bicarbonate reaction stated above. However, preliminary research indicates this has more of an effect on retarding cell growth, rather than inducing apoptosis[43]. This indicates that while sulfonamides may help in the reduction of tumour size and treatment of cancer by binding to CA IX, this might not be the only reaction occurring. Despite this, it is still very likely that at least some of the sulfonamides used will bind to this enzyme and have an overall effect, and as such, this is expected to also be a target of the sulfonamides being studied in this report.

1.5 Computational Analysis of Drug-Protein Interactions

The introduction of computers has vastly increased the ability for researchers to study chemical reactions and interactions. Experiments that have previously taken days or longer can now be done through the use of computer programs in much shorter time frames[8,10]. This has led to the development of many new medications, such as raltegravir for the treatment of HIV[12].

Computational analysis of potential new medications can take a number of forms. The first of these often utilised is high-throughput screening, or HTS[10]. This technique involves testing the binding of a large number of potential molecules against a target protein, and identifying those that bind strongly as candidates for further testing[8,10]. Although this method is successful at finding potential drugs (either directly or through modification of successful candidates), there is often large failure rates, due to the molecules being tested not being specific enough for the protein targeted[10]. As such, large libraries are often required for each HTS, limiting the facilities that are able to carry out these tests due to the relatively high computational load required[10].

By contrast, another, more specific way of discovering and testing potential new medications is through a series of calculations known as Molecular Dynamics (MD) [9,10]. This involves greater specificity with regards to the molecules and proteins being tested, and as such can give much more accurate results[9]. MD consists of calculations to determine probable orientations of molecule for docking with a target protein[8], as well as determining forces and energy involved in the interaction[9,49]. These simulations therefore allow for much more detailed understanding of drug-protein interactions,

and their accuracy and depth of study can be altered to fit with the hardware available to the researchers[49]. Examples of this type of calculation include docking studies by programs such as Autodock Vina, and MM/PBSA and MM/GBSA calculations[10], as utilised in this study.

One issue that does occur with the types of studies described above, is they tend to be very demanding on hardware[49]. To contend with this, supercomputers have been implemented to carry out the vast number of calculations necessary to perform an MD simulation[11]. These typically take the form of large banks of CPUs (often numbering in the thousands), with each CPU containing large numbers of cores capable of performing any number of different types of calculation[11]. The result of this massive arrangement of hardware is an ability to run many calculations simultaneously, either parallel to each other, or sequentially if this is required[11,49]. This greatly increases the speed at which the calculations can occur, as it reduces the lag between a calculation being initiated, and the memory becoming available for the computation to occur[11]. As such, memory- and core-demanding calculations such as MD simulations are able to be completed in much shorter time-frames, improving the rate at which new medications can be discovered, tested and understood[49].

In the most recent decade, a large shift has also begun to occur, with the advent of GPUs to supplement CPU usage for chemical calculations[50,51]. GPUs have become increasingly popular for these types of calculations due to the fact that they have lower energy requirements than CPUs, whilst maintaining a high efficiency for carrying out calculations[51]. Furthermore, GPUs are often significantly cheaper to purchase and replace when necessary, with even consumer-level units (significantly lower in price than more professional pieces) producing significant improvements to the efficiency of the supercomputers[50]. As such, it is now relatively common to find multiple GPUs per node on a supercomputers bank, vastly improving the efficiency of computational chemical calculations[51].

2. Methods

In addition to the methods listed below, example input files for the calculations can be found in Appendix A.

2.1 Ligand and Protein Preparation

The sulfonamides studied were selected from a series of molecules generated by Dr Aneela Maalik at COMSATS University Islamabad. Following review of *in vivo* efficacy, the seven more active molecules seen in Figure 1.1 were selected for initial docking studies. Following these, and in the interest of time and resources available, only 5AP, 6AP and 7BP were chosen for further MD analysis

Multiple conformations of the molecules were manually generated and minimised using molecular mechanics (UFF) in IQmol 2.14.1. For each molecule, a single conformer was found to be dominant, which was then optimized at the pc-1/B3LYP level of theory using QChem 5.2.1. Due to a lack of stereogenic centres in the chosen molecules, stereoisomers for the compounds would not exist, therefore no study on these was required.

Files for β-Tubulin and Carbonic Anhydrase IX were found on the RCSB Protein Data Bank, with PDB IDs of 6E7B and 4Q06 respectively. These particular files were selected after primary solvation attempts with other files proved unsuccessful. External molecules of GTP, G2P and Mg^{2+} for β-Tubulin, and V90, BCN, DMS and Zinc for CA IX were removed from the proteins via deletion of the entries in the PDB files.

2.2 Autodock Vina Docking

The ligand and protein files were uploaded to Autodock Tools 1.5.6 in preparation for docking. Grid boxes centred on the theorised binding site for the ligands were generated at sizes of 40x40x40 Å for both. Autodock Vina 1.1.2 was then used to generate primary docking scores using the standard application of the program, with exhaustiveness of 20 cycles for all combinations. The resulting ligand files were then loaded into PyMOL2 for conversion into *.pdb* files, as well as generation of ligand-protein compound files for further calculations. Snapshots of polar interactions were also done in PyMOL. LigPLOT v2.2 was used to generate images describing non-polar interactions between the ligands and the target proteins, using the Autodock Vina output files.

For the Autodock Vina simulations, the following equation is used to generate the docking score (c)[3]:

$$c=\sum_{i<j} f_{t_i t_j}(r_{ij})$$

Eq. 2.1: Calculation of Docking Scores produced by Autodock Vina, as described by Trott et al.[2]

This equation describes the summation of the interactions between atoms i (type t_i) and j (type t_j) over distance r_{ij}. It is important to note with this equation, that the target docking site is considered to be rigid, and does not include provisions for interactions with solvent molecules[3]. As such, Autodock Vina only provides a result depicting the free energy of docking between the ligand tested and the target protein.

2.3 MM/PBSA Calculations

In order to determine strength of binding for each molecule with the target proteins, MM/PBSA calculations were undertaken. MM/PBSA, or Molecular Mechanics combined with Poisson-Boltzmann and Surface Area solvation, is a method for estimating the change of free energy that occurs when small molecules interact with or bind to larger macromolecules such as proteins in solution[4,52]. In its simplest form, this is calculated by subtracting the free energies of the ligand and protein from the overall energy of the complex formed, as shown in Equation 2.1:

$$\Delta G_{\text{bind}} = \langle G_{\text{PL}} \rangle - \langle G_{\text{P}} \rangle - \langle G_{\text{L}} \rangle$$

Eq. 2.2: Base form of MM/PBSA calculation, as described by Genheden et al.[3]

In the above equation, G_T represents the total energy change of docking, G_{PL} the total energy of the docked ligand combined with the protein, and G_P and G_L representing the free energies of the protein and ligand respectively. For this calculation, the free energy present in a system is determined by the summation of all bond (*bnd*), electrostatic (*el*) and van der Waals forces (*vdW*), as well as polar (*pol*) and non-polar (*np*) contributions to solvent free energies for a set temperature (*T*) and entropy (*S*)[4]. These free energy calculations are performed using the following equation:

$$G = E_{\text{bnd}} + E_{\text{el}} + E_{\text{vdW}} + G_{\text{pol}} + G_{\text{np}} - TS$$

Eq. 2.3: Free energy equation used in MM/PBSA calculations to determine total energy of a system before and after binding, as described by Genheden et al.[3]

A number of steps were completed to carry out the above MM/PBSA calculations. Initial optimisation and proof-of-concept calculations were performed using Krzywik et al's paper *Synthesis, Antiproliferative Activity and Molecular Docking Studies of Novel Doubly Modified Colchicine Amides and Sulfonamides as Anticancer Agents*[1]. These experiments were run as closely to those in the reference as possible, however some keywords had to be estimated as attempts to obtain example input files for their calculations proved unsuccessful.

For all molecular dynamics calculations, AMBER 16 was used. The program pdb4amber was used to fix all required files for use with Amber. Antechamber was used to generate *.prepi* files, and parmchk2 was used to generate *.frcmod* files. Following altering of the files to ensure standardised naming of atoms throughout the molecules, these were then used in combination with the *.pdb* files for the ligand, proteins and compounds in the program tLeAP to generate the required topography and co-ordinate files. The force fields chosen were ff14SB for the proteins and GAFF for the ligands, and TIP3P water molecules were used for solvation.

The topography and co-ordinate files were then used in a series of calculations involving minimisation, heat, density and equilibration. The method employed was that of the Amber MM/PBSA Python Script tutorial by McGee et al[52], and utilised their standardised input files. Slight adjustments were made to cycle limits for these inputs however, to avoid unnecessary usage of computing time. For instance, the number of heat cycles was reduced from 25,000 to 15,000 as further calculations past this point proved unnecessary.

A production run was then performed on the equilibrated systems, producing a series of 4 'snapshot' files to complete the final energy calculations. These proved to be the most computationally intensive, with some taking up to 28 hours to complete. Standardised input files from McGee et al[52] were used in this step also.

The Python script version of the MM/PBSA and MM/GBSA calculations was used to determine the final results, as this was most applicable and easiest to access with regards to the experiment files and software available. The AMBER 16 framework was again used for these calculations, and results for both MM/PBSA and MM/GBSA were determined to allow for corroboration of data between two different calculation methods. Once again, standard input files were used for these calculations.

For these MM/PBSA calculations, the New Zealand eScience Infrastructure's Mahuika supercomputer was utilised.

3. Results

3.1: Pilot Experiments

We had initially hoped to replicate the results of Krzywik et al[1] in order to validate our implementation of their docking studies. However, we ran into several problems that made an exact comparison impossible. Most significantly, Krzywik et al[1] generated a unique structure for tubulin using the MOE software package and a bovine tubulin as the base, modifying this to fit the human form using the TUBA1A and TUBB genome sequences. However, the resulting PDB file of this new protein was not able to be found, and although numerous attempts were made to contact multiple members of the paper's research team for this structure, no replies were received. Instead, a different tubulin structure was found, and docking simulations were carried out to compare with those of Krzywik et al., to ensure the structure selected would be suitable for use with the sulfonamides from Dr Maalik et al.

It was therefore decided to carry out some simple Autodock Vina docking simulations to determine if any similar trends could be observed between Krzywik et al's results and those of another tubulin structure. If these were similar, then it is likely that the structure and method selected would be appropriate for use in our main sulfonamide experiments. A summary of these simulations is shown in Table 3.1.

Colchicine Derivative	Reference Paper Binding Energy (kJ/mol)	Autodock Vina Binding Affinity Score (kJ/mol)
6	-155.6	-39.7
11	-145.2	-41.8
12	-147.3	-31.0
14	-143.5	-30.1
15	-183.7	-32.6
16	-160.2	-38.9
18	-228.4	-28.0
19	-97.9	-28.5

Table 3.1: Comparison of binding energies from reference paper by Julia Krzywik et al[1] and Autodock Vina binding affinity score, in kJ/mol

As can be seen in the above table, it is unclear if the initial docking studies match up with the final binding energies calculated by Krzywik et al. A comparison of the study results with those obtained for this experiment was then carried out via the use of a pair of graphs, in order to identify any key similarities, as can be seen in Figure 3.1.

From this graph, it is clear that the Vina binding affinity scores do not always equate to binding energy calculations, though there are some promising correlations. For instance, six results (derivatives 6, 12, 14, 15, 16 and 19) do appear to follow a similar trend, with derivatives 6, 15 and 16 showing higher values than 12, 14 and 19. The latter three by contrast, have some of the lowest results of all derivatives studied, with 19 being significantly lower than most other results in both studies.

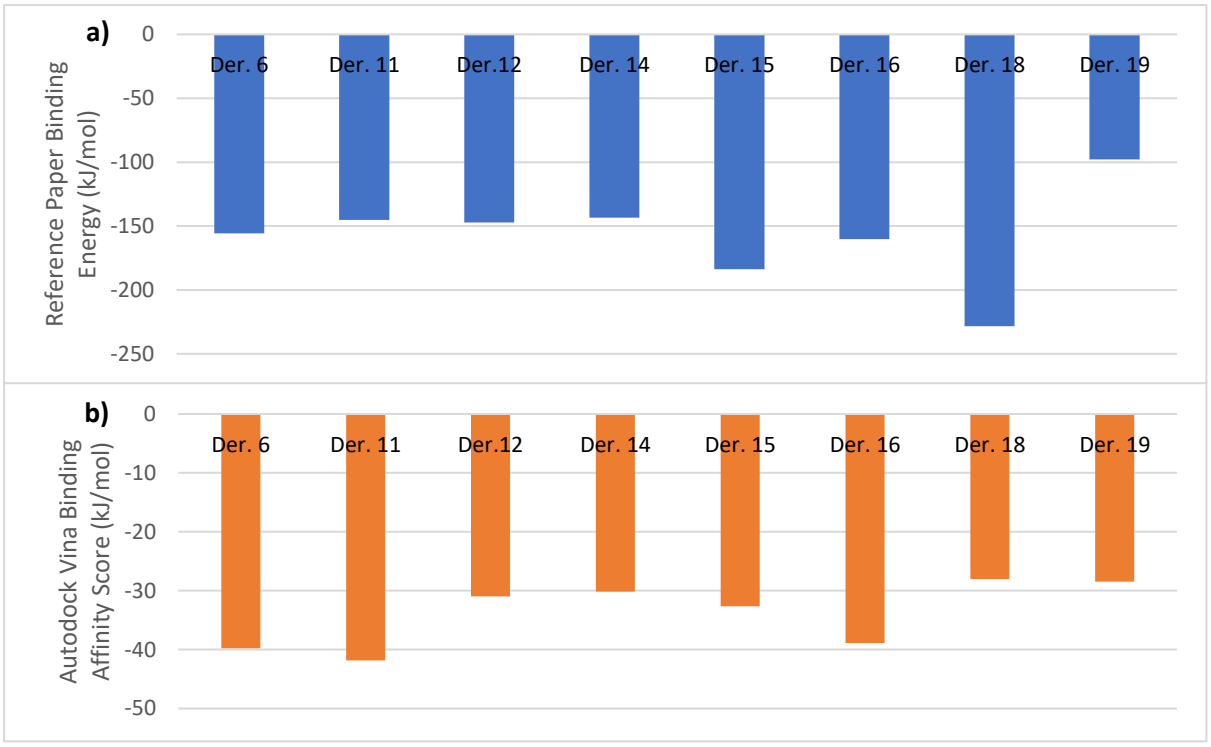


Fig. 3.1: Graphical comparison of a) binding energies from reference paper by Julia Krzywik et al[1] with b) Autodock Vina binding affinity score, in kJ/mol.

However, there is a rather large discrepancy with regards to the other results, namely 11 and 18. In the reference paper, derivative 18 showed the highest binding energy of the molecules studied, whereas in the Autodock Vina studies it was the smallest. Derivative 11 by contrast, had one of the smaller binding energies, but the highest Autodock Vina docking score.

As a final comparison to corroborate the methods, the exact locations and orientation of the docked molecules were compared between the reference paper and the results obtained for this study, such as those seen in Figure 3.2. This shows derivatives 6 and 12 (the most accurate results of the pilot study) bound in approximately similar sites and orientations in both examples, which supports the idea that the methodology and tubulin model chosen for the binding studies would be sufficient. The small variations detected were considered acceptable due to the random nature of Autodock Vina, with each seed having its own unique set of results, meaning no two experiments will produce identical results. The complete set of images comparing the results can be found in Appendix B.

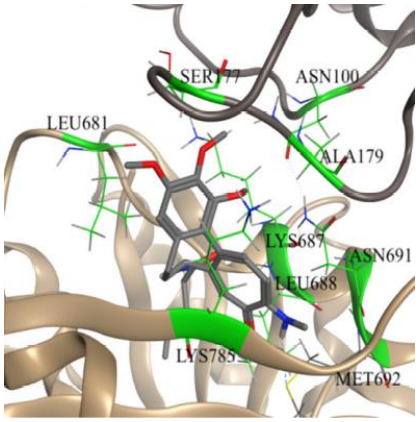
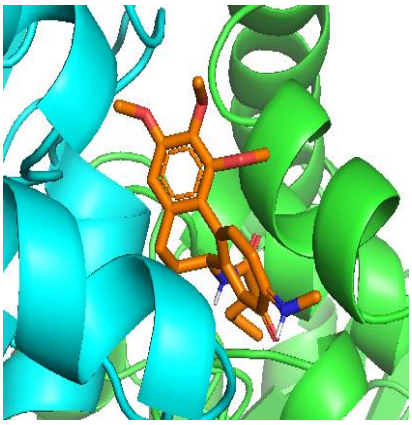
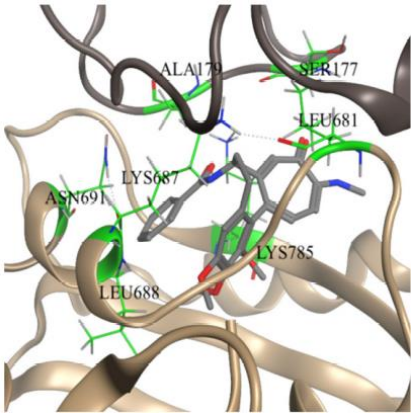
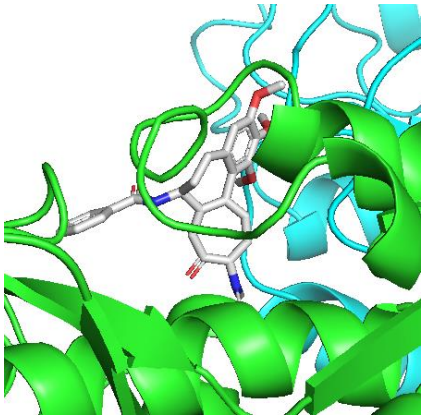
Derivative	Reference Paper Results	Autodock Vina Results
6		
12		

Figure 3.2: Comparison of 3-Dimensional outputs of derivatives 6 and 12 from Autodock Vina of docking studies between reference paper by Julia Krzywik et al and attempted replication of results

3.2: Autodock Vina Docking Scores

For the test molecules provided by Dr Maalik et al, the initial experiment completed was docking studies using Autodock Vina. The results of this experiment can be seen in Table 3.2.

From this table, it is clear that there is at least some correlation between the *in vivo* efficacy studies and the Autodock Vina results. The more effective molecules in the previous experiment (6AP and 5AP, based on IC_{50} values) did show higher binding affinities for both molecules than at least half of all other sulfonamides tested, with 5AP scoring the highest score for both proteins. This shows that in at least some aspects of the ligands, part of their mode of action could include either β-Tubulin or CA IX, as the relatively high Autodock Vina docking scores suggest a link. However, like with the pilot investigations, further study is required to confirm these results.

Of note from the table, is that molecule 5AP has the highest overall docking score for both target proteins. This is despite the fact that during the *in vivo* experiments, 6AP was the most efficacious, with 5AP being in second rank. What this suggests is that docking of the sulfonamides may not be the deciding factor in the mode of action of these ligands, but instead a number of other factors could be working in unison.

Ligand	IC_{50} against HeLa cells after 72hr (μM)	β-Tubulin	CA IX
6AP	11.5	-31.0	-31.4
5AP	12.1	-38.5	-36.4
7BP	16.3	-27.2	-31.4
4AN2	17.3	-33.5	-33.9
7AP	20.1	-31.8	-28.9
6BP	24.3	-32.2	-27.6
5A	30.0	-28.0	-24.7

Table 3.2: Comparison of Autodock Vina binding affinity scores for the novel sulfonamides produced by Maalik et al with Tubulin and Carbonic Anhydrase IX, in kJ/mol. Ligands were ranked in order of in vivo *efficacy, from most to least effective based on IC_{50} values against HeLa cells after 72 hours.*

This table also shows two very clear trends: firstly, that on average the docking scores for tubulin were higher than those for CA IX, and secondly that the trend of the results for CA IX more closely matches the *in vivo* experimental results. These are interesting observations, as they show that while the results for tubulin seem to indicate a preference to bind to this protein, the results for the CA IX binding appear to corroborate the previous experiment's results. Furthermore, the relative closeness in these two sets of docking scores indicates that either protein may be a potential target for the sulfonamides. To confirm this, energy calculations in the form of MM/PBSA need to be done, to show the relationship between sulfonamide and protein more clearly.

3.3: MM/PBSA Results

Following the Autodock Vina docking scores, the MM/PBSA calculations were carried out. After further consultation with (and on the recommendation of) Dr Maalik, as well as in the interest of time and resources available, it was decided that only the results for the three most active compounds *in vivo* (6AP, 5AP and 7BP) would be calculated. Table 3.3 shows the results of these calculations.

Protein	Molecule	Vina Docking Score (kJ/mol)	MM/GBSA result (kJ/mol)	MM/PBSA result (kJ/mol)
β-Tubulin	6AP	-31.0	-1161.2	-1085.0
	5AP	-38.5	-1361.6	-1222.8
	7BP	-27.2	-1331.7	-1210.9
CA IX	6AP	-31.4	-363.6	-300.0
	5AP	-36.4	-415.8	-295.4
	7BP	-31.4	-402.6	-273.2

Table 3.3: Comparison of Autodock Vina binding affinity scores, MM/GBSA and MM/PBSA results for the selected novel sulfonamides produced by Maalik et al with tubulin and Carbonic Anhydrase IX, in kJ/mol.

The data presented in Table 3.3 makes a relatively clear case for the preferred target. The results of the MM/PBSA calculations produced much larger energy changes for β-tubulin than for CA IX, with some results (such as for 5AP) showing a near-1,000kJ/mol difference in energy calculated. These results, at least for the sulfonamides tested, seem to suggest that β-tubulin is indeed the protein these ligands act upon in the majority of cases. However, the energy change for CA IX do indicate at least partial binding to this protein may also occur.

A few interesting points can also be made regarding these results. Firstly, although the exact results of both the MM/GBSA and MM/PBSA results were not expected to directly match the Autodock Vina docking scores, due to differences in what they measure[4,3], similar trends would be expected based on the results of the pilot study. However, these sets of results not only differed from the Autodock Vina scores, but also the experimental results. This is particularly apparent with 6AP, which had the lowest change in energy calculated with β-tubulin, despite being the ligand with the highest *in vivo* efficacy. Indeed, the second key point is that the only result that appears to be in line with the *in vivo* experiments is the MM/PBSA calculation for CA IX, having the highest change for 6AP, followed by 5AP and finally 7BP. Based on this data, it could also be concluded that CA IX is the more likely target. However, other factors such as solubility and polarity of the compounds could also have resulted in the *in vivo* results differing from these calculations.

Another key difference that can be seen in the table above, is the fact that the MM/GBSA results show consistently more energy change than the MM/PBSA results. A difference between the two sets of results is to be expected, based on the way the two calculations are completed. MM/GBSA uses the Generalised Born model for calculating polar solvation free energy through approximating interactions between an atom of interest and that of the solvent[53]. However, MM/PBSA uses the Poisson-Boltzmann equation to calculate this energy, which more accurately defines the radii of the atoms being studied[54]. For the results in Table 3.2, this difference can be seen by the larger MM/GBSA results, which indicate the approximation used for the GB model overestimated the interactions and the radii of the atoms in the sulfonamides, and therefore the MM/PBSA results are likely to be the more accurate of the two sets.

A visualisation of the results of these calculations can be seen in Figure 3.3. In this graph, it can clearly be seen that there is a vast difference between the energy changes of the two proteins when the sulfonamides are docked. Furthermore, the standard deviations for these calculations (demonstrated by the error bars of each column) provide significant information about the results. Specifically, for β-tubulin, the relatively low energy change for 6AP compared to the other two sulfonamides is much more apparent, and the lack of any cross-over of the error bars for this protein indicates that this difference is likely to be real. This means that at least for this particular compound and protein, the MM/PBSA and *in vivo* results do not corroborate each other. This is not seen with the results of CA IX, which shows almost total overlap of all three sets of standard deviations, meaning there is no significant difference between all three molecules docking with this protein. As such, no true conclusions between the ligands can be drawn for this protein.

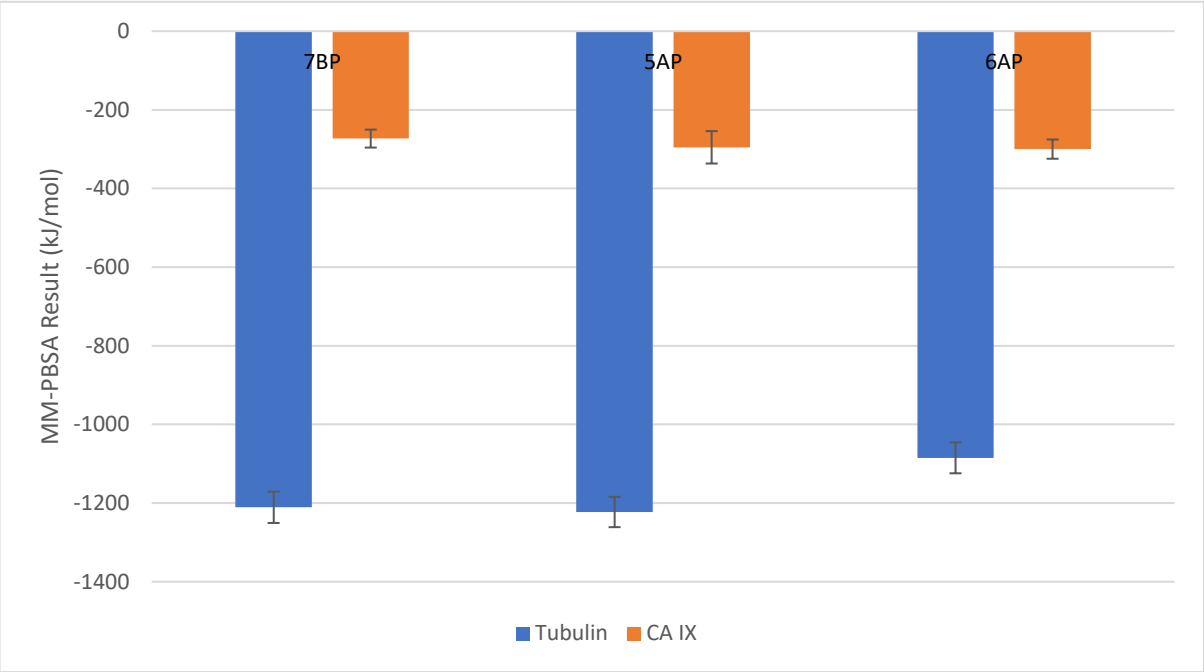


Fig. 3.3: Graphical illustration of MM/PBSA results for selected Sulfonamides with Tubulin and CA IX. Sulfonamides are listed in order of increasing in vivo *effectiveness, and error bars represent the standard deviation of each calculation performed.*

3.4: LIGPLOT Results and 3-Dimensional Imaging

To determine the exact nature of the interactions between the sulfonamides and the target proteins, LIGPLOT figures were generated and analysed. This program has found wide use for simplifying and visualising not only polar interactions and hydrogen bonds between ligands and proteins, but also non-polar interactions that are not always obvious from analysis of 3-dimensional images. These results can be seen in Figure 3.4.

The first of the results shown are for the most active compound *in vivo*, 6AP. In β-tubulin, 6AP can be seen to have a hydrogen bond interaction with Serine-131 in the protein sequence. This bond is the shortest of those shown in the figure, at only 2.97 Å between the two heavy atoms of oxygen and nitrogen, with the hydrogen in the bond being approximately 2 Å from the oxygen of the tubulin. This suggests there is a relatively strong bond between the protein and ligand. Furthermore, 11 other amino acids are shown to have hydrophobic interactions with the sulfonamide, indicating a relatively large network of contacts between the two. By contrast, there are no hydrogen bonds between 6AP and CA IX (the only combination not to have this kind or interaction). Instead, twelve residues are shown to have hydrophobic interactions with the molecule. Comparison of this result to those of the MM/PBSA calculations seems to indicate that this increase in contact between non-polar regions has a stronger role in the action of 6AP, as there was a greater change of energy calculated for CA IX than β-tubulin, compared with the results of the other two molecules.

For 5AP, the interaction with β-tubulin showed a hydrogen bond formed between Glycine-137 and one of the sulfur-bound oxygens in the ligand. There are also shown to be 16 hydrophobic interactions between the two, the most of any of the combinations shown. This may in some part explain the fact that this ligand-protein complex showed the most energy change of any tested, as more residue interactions naturally has the potential for more energy to be released by the interaction. With regards to the docking with CA IX, another hydrogen bond of 3.24 Å (the longest shown) can be seen, this time with Proline-200. Twelve hydrophobic interactions are also shown, which is the same as for 6AP.

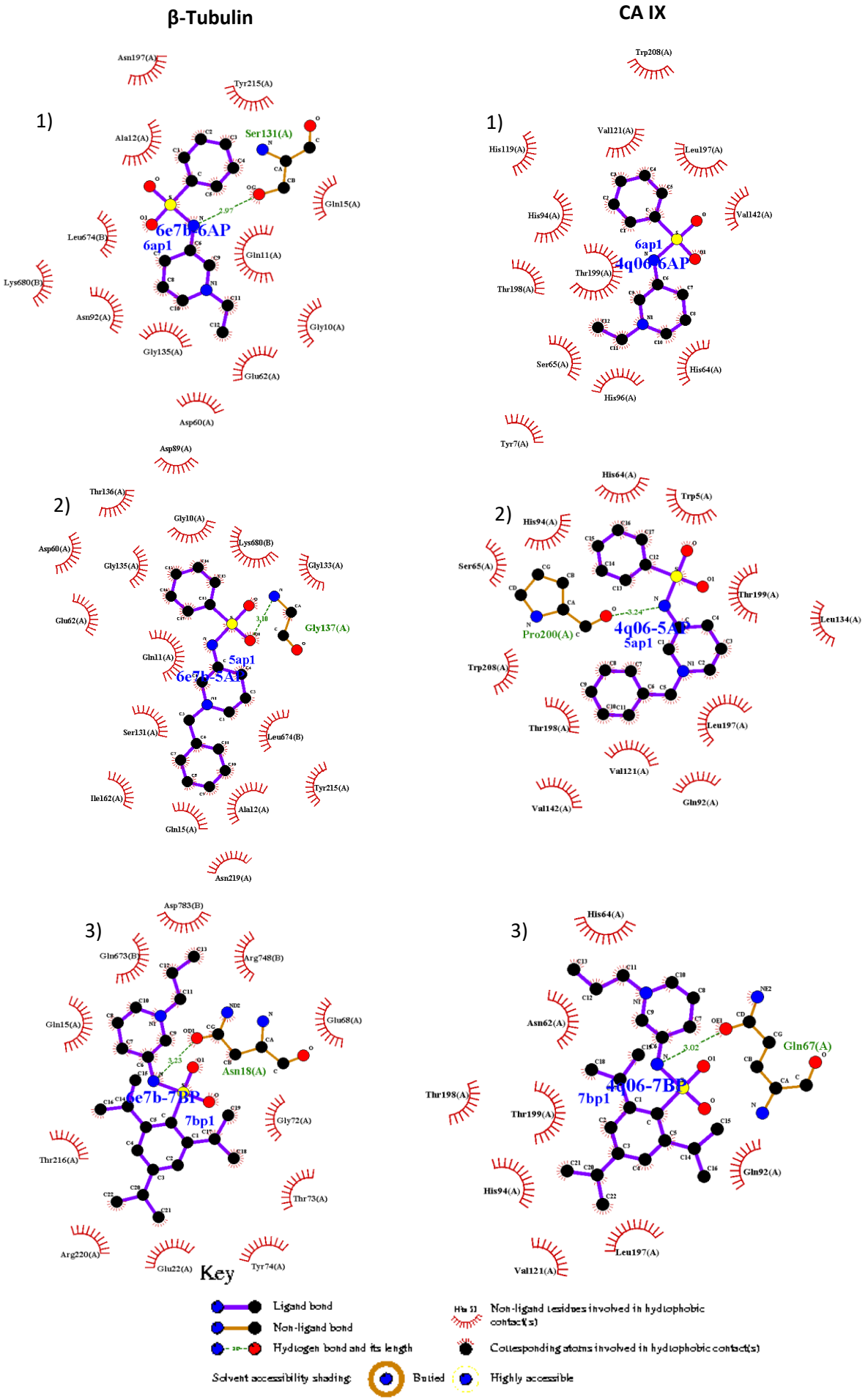


Fig. 3.4: LIGPLOT results for selected Sulfonamides with Tubulin and CA IX. Sulfonamides are in order of in vivo *effectiveness, from 1) 6AP, to 2) 5AP then 3) 7BP. Key reproduced from LIGPLOT Operating Manual*[4]

Interestingly however, the addition of the hydrogen bond seems to have little to no effect, or even possible a negative effect, on the energy released by this docking, in comparison to that of 6AP. This can be seen by the fact that the MM/PBSA calculations for the two complexes has 6AP showing a slightly larger change in energy.

By comparison with the other two compounds, 7BP has the fewest interactions with both proteins. For β-tubulin, hydrophobic contact with eleven residues can be seen, as well as a hydrogen bond with Asparagine-18, measuring 3.23 Å. In the case of CA IX, a relatively close hydrogen bond of 3.02Å exists with Glutamic Acid-67, as well as 8 non-polar interactions. The presence of fewer hydrophobic interactions with both proteins does corroborate the earlier results (both *in vivo* and MM/PBSA), lending more weight to the argument that it is these contacts that have a larger role in the overall docking of the ligands. One interesting note that can also be made, is that the hydrogen bond for both proteins was formed with the nitrogen atom of the sulfonamide bond, with 7BP being the only ligand to show this specific set of interactions.

To better illustrate the arrangements of the ligands docked with the target proteins, 3-dimensional images were generated. These can be seen in Figure 3.5 below:

Sulfonamide	β-Tubulin	CA IX

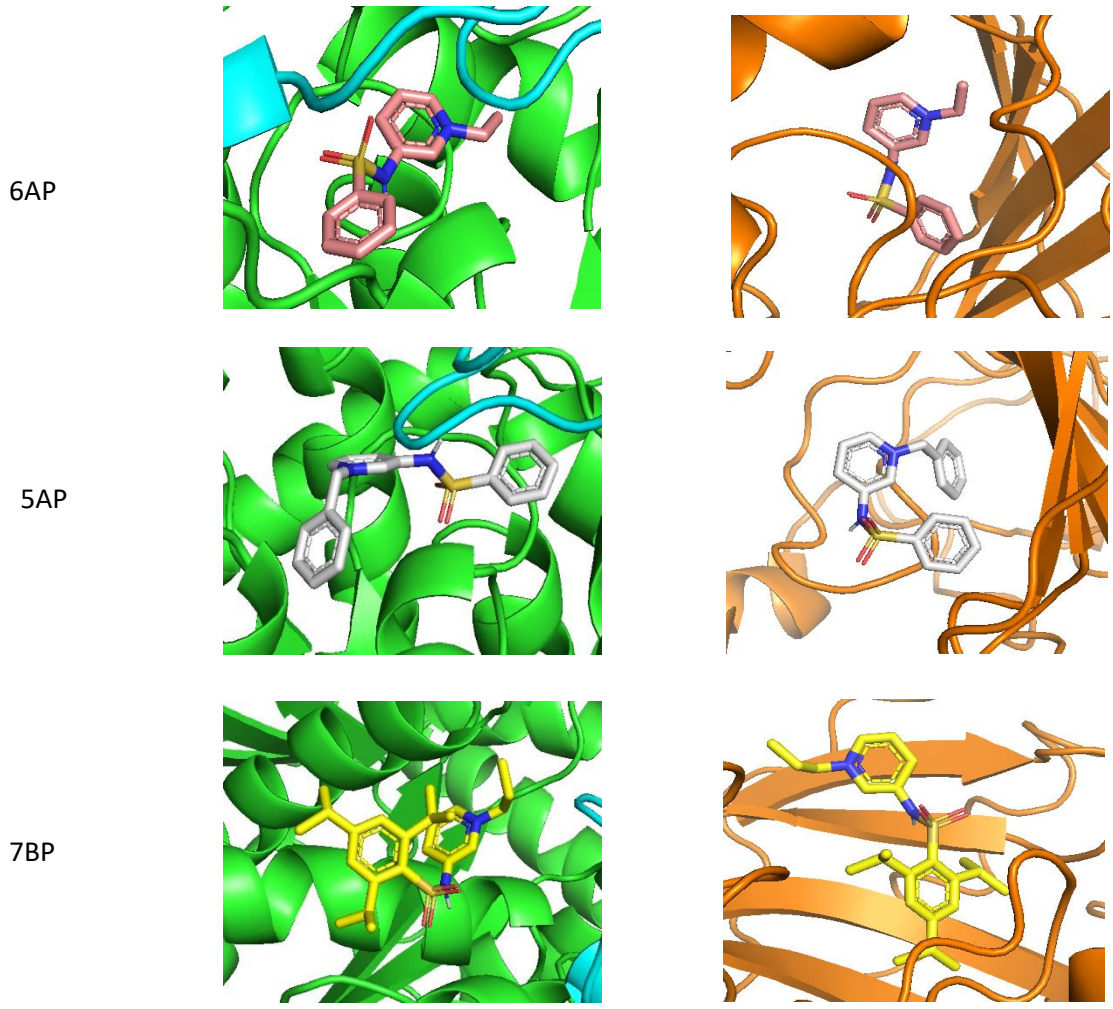


Figure 3.5: Comparison of 3-Dimensional outputs of docking of Sulfonamides 6AP, 5AP and 7BP with β-Tubulin and CA IX

From the above figure, a couple of key observations can be made. Firstly, for the tubulin results, all three molecules showed docking interactions relatively far into the protein with relation to the outer surface, whereas the CA IX dockings were all relatively external. This may explain the difference in results from the MM/PBSA calculations, as there would be less solvation energy change for the CA IX reactions, as a much larger part of each ligand is still interacting with the solvent and therefore less energy loss would occur.

The other major observation that may be made is that, at least for the CA IX images, the hydrophobic sections of each ligand (namely the sulfur-adjacent benzene rings) are predominantly docked further into the protein, and the hydrophilic sections such as the oxygens and hydrogens are external of the protein. This supports the data shown in the LIGPLOT results, as the majority of the interactions shown in Figure 3.4 are non-polar with at most only one polar interaction or hydrogen bond. It can be assumed therefore that at least in the case of CA IX, the more polar regions of the sulfonamides prefer interacting with the solvent than the protein itself, lend more weight to the difference in the MM/PBSA results seen between the two proteins.

4. Discussion and Conclusion

The research for this thesis produced a number of interesting results. In this chapter, a possible explanation for the results will be presented and explained. Limitations of the study will also be highlighted, and the potential impact of the results, as well as any possible continuing research into these sulfonamides will be explored.

4.1: Limitations of Study

As with the vast majority of scientific studies such as those performed for this thesis, a number of issues presented themselves over the course of the research. Where possible, these were overcome, however in a few instances this was not possible.

The first, and potentially most impactful limitation of the study came during the pilot experiments. For this, a tubulin model was created by Krzywik et al. using bovine tubulin as a basis, and modified to more closely match the human genome[1]. The study itself did not provide the exact PDB file or sequence used for the experiments, and attempts to contact the authors of the paper proved unsuccessful, meaning replication using this exact structure was impossible. Therefore, a human tubulin file was sourced through the Protein Databank, and as such, variations in the structure were likely to be present to those of the reference paper. In light of this, it was decided that instead of exactly replicating the MM/PBSA results of Krzywik et al., docking studies using Autodock Vina would be performed, and any trends found would be compared between the two sets of results. This therefore facilitated the identification of a suitable tubulin model for use in the sulfonamide calculations, whilst still providing a set of results to compare with, despite the inability to reproduce the exact structure used in Krzywik et al.'s studies.

Issues arose when trying to select the proteins used for the docking experiments. This was particularly prevalent with the tubulin model used, as a number of proteins were tested. Though able to create docking scores with Autodock Vina, the solvation step necessary for the MM/PBSA calculations in AMBER failed to produce workable files. A few reasons were identified to have caused these issues, such as missing residues and non-standard labelling of specific atoms, however eventually the file 6E7B was able to successfully solvate in both docked and free protein form, and as such this file was selected for use with the subsequent calculations.

The first run through with the calculations produced results which appeared to be egregious outliers in terms of expected values. Specifically, results were produced showing binding energies in excess of 20,000 kcal/mol, and energies of the proteins above 500,000 kcal/mol, far higher than expected from prior reading into the MM/PBSA calculation process. Through re-examination of the calculations, two key areas were identified in terms of potential fixes. Firstly, all files used were run through PyMOL to standardise naming and arrangements of atoms, and secondly the bound ligands were assigned specific molecule names in every file used. These two solutions corrected the calculations, and the results produced were more appropriate in scale, and so reported in this thesis.

In terms of the scope of the study itself, one potential limitation on the experiments arises from the specific proteins and ligands selected. Namely, though comprehensive study into the files tested did occur, and the files were selected as being the most likely target/ligand pairs, there was still only a relatively small sample size researched in the form of the three most active compounds. Furthermore, in terms of the proteins, only 2 were selected for the calculations: β-Tubulin and CA IX. However,

extensive study of the literature supported the selection of tubulin and CA IX as target proteins, so the impact of this limitation on the final study is likely to be minimal.

4.2: Explanation of Results

4.2.1: Pilot Study

The purpose of the pilot study was to identify a potential methodology for the future calculations, and identify a tubulin model that could be implemented in these simulations. However, exact replication of the results was not possible due to the reasons already stated in section 4.1 above. As such, it was decided that a comparison of the trends seen in the results of Krzywik et al. and the Vina docking scores produced for this thesis would be done, as exact matching of the values from the two sets of calculations were not expected to match. For the majority of the results, these trends did hold true, however the scores calculated for derivatives 11 and 18 did not follow those of the reference study.

One potential explanation for these differing results is the tubulin model selected. Though the protein used was chosen due to its apparent similarity to that used by Krzywik et al., without access to the full PDB file used, there is a relatively high possibility that the exact sequences did not match. As such, some key binding sites for the derivatives may be different between the two proteins, leading to a variance in the docking scores calculated.

In terms of the choice of calculation type, this could also potentially explain variances in two sets of results produced. The reason for this is the way the two different results are calculated. As seen in equation 2.1, Autodock Vina calculates a docking score with a rigid target protein, and no provisions are made for the effects of a solvent on the energies. However, the MM/PBSA calculations explicitly employ contributions (both polar and non-polar) to these solvation free energies, as seen in equation 2.3. It can therefore be suggested that a major reason for the difference between the two sets of results is due to the impact of the solvent on the docking of the ligands, specifically with regards to derivatives 11 and 18.

The images of the docked ligands in Figure 3.2 and Appendix B do help to support the results of the pilot study. Although it is difficult to determine the exact interactions present from these images alone, the ligands themselves appear to be situated in similar positions and orientations to that of Krzywik et al. Due to the previously stated difficulties with carrying out full MM/PBSA calculations, these images display the best evidence for the success of the pilot experiments. The closeness of the docking between the two sets of images indicates that though the docking scores may vary from the MM/PBSA calculation results, the simulated docking of the ligand to the target protein was replicated relatively successfully. It can therefore be concluded that the method followed to obtain these ligand-protein complexes was reasonably sound, and as such could be employed for the experimental method of the thesis.

4.2.2: Thesis Experiment Results

Autodock Vina

Based on the success of the pilot experiments, it was decided that the method used in the experimental tests would follow a similar structure to those of the prior study. Therefore, the first

calculation done was to determine the Autodock Vina docking scores for the ligands with each of the proteins identified.

The results of this experiment showed a number of interesting quirks, as illustrated by Table 3.1. Firstly, the docking scores for the entire set of test ligands were all relatively similar, despite the wide range of *in vivo* efficacies demonstrated by Maalik et al. Furthermore, of these docking scores, 6AP (the most effective ligand) had relatively middling scores for both target proteins, whereas 5AP had the strongest docking compared with all others tested. Based on the *in vivo* efficacy of 5AP, it can be argued that a higher docking score would be expected, however the sudden drop back down to that of 6AP does not follow this same trend.

Another notable piece of information generated by the Autodock Vina experiments is the closeness of the results for both sets of proteins, and the fact that on average, docking scores increase in tandem with *in vivo* efficacy. These support the theory that the selected proteins may be targets for the molecules, due to this corroboration of *in vivo* results. However, the relative closeness of the overall results mean that identification of a clear target protein cannot be made with these results alone. It is worth noting that for the majority of the molecules, the docking scores were higher for tubulin than CA IX, however the significant outlier of 7BP's Tubulin result does reduce the average for that protein's data set, and so may not be entirely accurate to real-world docking of the ligands.

The results of these calculations suggest that factors other than specific ligand-protein interactions may be impacting the overall efficacy of the molecules. However, further calculations using MM/PBSA should be analysed before any conclusions may be drawn.

MM/PBSA

Following the MM/PBSA calculations, there is a much clearer divide between the two proteins in terms of change in energy caused by the docking of the ligands. As covered earlier, the results indicate a much larger change in energy for tubulin than for CA IX, suggesting that this protein is the more likely target. However, the computational results do not completely match those of the *in vivo* experiments, as for tubulin, 6AP is shown to have the lowest change in energy, whereas 5AP has the highest. For CA IX, the results follow much more closely with those of Maalik et al., with 6AP showing the greatest change in energy.

As stated in Equation 2.3 above, the MM/PBSA calculations factor in a large number of binding energies, forces and interactions with not only the protein and ligand, but also solvent to determine the overall energy change of a protein-ligand interaction. Therefore, it can be inferred that the larger the MM/PBSA result is, the more energy is present in the interaction between the ligand and protein, and therefore the greater energy released by the docking or binding of the two. As such, from the results calculated for this experiment, the difference in magnitude between the tubulin and CA IX results demonstrates that tubulin is the more preferred target for the molecules. However, it should be noted that there is still a significant change in energy for CA IX, despite this not being the preferred target. This suggests that the binding of the molecules is not selective for one particular protein, and that not only could CA IX also play a part in the cytotoxicity of the sulfonamides, but there is potential for other proteins to be involved as well.

Possibly the most intriguing information that can be derived from the MM/PBSA calculations is the energy change of 6AP with tubulin. Despite being the most effective during the *in vivo* experiments, and the results for all other drug-protein interactions following the experimental results, this does stand out as having the lowest energy change during docking. Even when the standard deviation of all results is factored in, this result would still be the lowest of the three combinations tested with this

protein. A possible explanation for this is that 6AP is the smallest of the molecules tested, containing only 13 carbon atoms, compared to the 18 of 5AP and 23 of 7BP. As per Equations 2.2 and 2.3, a key part of the calculation is the energy of the ligand itself, as well as any interactions between atoms in the protein and the solvent the docking occurs in. Having fewer atoms as part of a molecule would naturally displace less aqueous solvent, therefore resulting in a lower solvation energy of the ligand, and thus result in a lower overall docking energy being generated. However, if this were to be the explanation for this lowered result, it would be expected that 7BP would have the largest change in energy seen in these calculations, with it being the largest of the three molecules. As this is not the case, other factors must also be influencing the 6AP-Tubulin result. To that end, the LIGPLOT results were analysed to discern any further explanation.

LIGPLOT

To better understand the results of the MM/PBSA calculations, a series of LIGPLOT images were generated to show specific interactions occurring between the ligands and proteins. The key information demonstrated by these images is the vast amount of non-polar interactions with the ligands, vastly outnumbering those few polar interactions and hydrogen bonds.

While it can be expected that a number of non-polar interactions were to occur between the drug and the protein (due to the drugs only containing a small number of polar regions, specifically around the amide bonds), the sheer number of interactions should duly be noted. Indeed, between 8 and 16 hydrophobic residues were shown to connect with the sulfonamides, compared with at most only one polar/hydrogen bond forming. Of these, 5AP had the most of these, likely due to the stretched nature of the molecule. This allowed more of the carbons in the ring structures to be available for reaction, thus leading to the higher interaction count. By contrast, 7BP, although it contains more carbons (and by extension non-polar carbon-carbon and carbon-hydrogen bonds) than both other molecules, it has the least number of non-polar interactions with each protein. This is because the density of the hydrophobic parts of the molecule leads to steric hinderance of a large proportion of the molecule, particularly in regards to the trio of propane residues on the benzene ring. The presence of these residues blocks interaction between the carbons of the benzene ring and non-polar amino acids in the proteins, thus resulting in 7BP only having 11 and 8 hydrophobic connections with Tubulin and CA IX respectively.

The abundance of non-polar interactions between the ligands and proteins also shed light on an important aspect of the molecules: large parts of their structure are hydrophobic. This is important due to the fact that the calculations were all done using ionised TIP3P water molecules as a solvent, which is very polar. This is important to note as, though the sulfonamides are all positively charged (and therefore should prefer the ionised solvent balancing this out), the copiousness of non-polar interactions highlighted in the LIGPLOT results suggests these are stronger than the hydrophilicity of the ionised nitrogen. It can therefore be noted that the hydrophobicity of the ligands likely plays a part in the relatively strong docking of the molecules to the proteins. Furthermore, Tubulin was shown to have more non-polar interactions with the molecules than CA IX on the whole (12 each for 6AP, but 16 for 5AP against 12, and 11 for 7BP against 8). This is likely a contributing factor to the results generated by the MM/PBSA calculations, as non-polar interactions are factored into the equation for these (see Equation 3 above). As such, it can be concluded that hydrophobic interactions between the ligands and the protein play a role in the strength of the docking with the proteins, and therefore likely has an influence on their viability as drugs (due to increased effect as a result of this docking).

One key piece of information these plots also provided (or rather, did not show) was with regards to the positive charge of the molecules, specifically around the ring structure and nitrogen atoms responsible for this charge. The results showed that in all cases studied, these rings had an abundance

of non-polar interactions, yet none were shown to have a hydrogen bond form with the ring-bound nitrogen. While this could be due to steric hinderance caused by the side-chains attached to the nitrogen, the majority showed evidence of non-polar interactions, indicating the relative strength of these connections. These results further support the theory that non-polar interactions have a significant effect on the docking of the sulfonamides to the proteins.

4.2.3: Conclusions of the Results

The results generated by the calculations carried out in this study have provided some interesting conclusions. Firstly, although there was a vast difference in *in vivo* effectiveness displayed by Maalik et al., the Autodock Vina docking scores were all relatively similar. Furthermore, the MM/PBSA calculations showed that tubulin is the more likely target for these sulfonamides compared to CA IX. However, CA IX did also show some potential effect too, so the molecules may not be completely selective in their mode of action. The LIGPLOT diagrams highlighted the abundance of non-polar interactions between the drugs and proteins, compared with polar and hydrogen bonds, thus showing these to be a determining factor in the strength of docking of the molecules.

The most interesting result generated however, was with regards to the MM/PBSA result for 6AP and tubulin. This showed significantly less energy release compared to the other molecules with tubulin, despite 6AP having the highest *in vivo* efficacy. Furthermore, the relatively narrow standard deviation of the result indicates a fairly consistent result, thus eliminating random error/changes in energy as being a possible explanation. In order to fully comprehend and explain this unusual result, further studies were required; namely, a comparison of the various drug-like properties the molecules have, such as Lipinski's Rule of Five.

4.3: Further Experiments

The unexpected 6AP-Tubulin result highlighted a need for further experimentation and calculations, in order to fully comprehend the mode of action of these sulfonamides, as well as identifying potential hinderances to the effect of these and other similar drugs on breast and cervical cancer cells. Indeed, and number of other factors affects the effectiveness of a molecule to act as a medicine, such as transport of the molecules through the cell membrane, to their movement throughout the cytoplasm. In order to quantify the impact of these potential issues on the effectiveness of the sulfonamides studied, a drug-like score comparison was performed.

4.3.1: Drug-like Comparison

To identify a number of the factors a molecule may have that influence its therapeutic viability, a comparison was done using the DruLiTo program. The results of this comparison can be seen in Table 4.1 below.

In the table, a number of key results stand out in terms of the variance between the molecules. The most obvious of these is the molecular weight and number of atoms in each sulfonamide. As previously discussed, 7BP is the largest molecule in terms of atom number, and this is reflected in a higher overall molecular weight of 403.24 g/mol. By contrast, 6AP is the smallest of the molecules, with only 33 atoms and a molecular weight of 263.09 g/mol. These results are inversely proportional to the *in vivo* efficacy of the drugs, which hints that the size of the molecule is a factor in how well the molecules act as medicines.

Sulfonamide	7BP	5AP	6AP
Molecular Weight (g/mol)	403.24	325.1	263.09
No. Atoms	63	40	33
No. Aromatic Groups	2	3	2
No. Rotatable Bonds	8	5	4
No. Rigid Bonds	21	20	15
No. Hydrogen Bonds	5	5	5
H-Bond Donators	1	1	1
H-Bond Acceptors	4	4	4
logP	5.334	1.641	1.318

Table 4.1: Comparison of results from DruLiTo program for determination of drug-like properties of sulfonamides 6AP, 5AP and 7BP. Molecules are listed in order of increasing in vivo *efficacy*

Another set of data that show variance between the sulfonamides is with regards to the bonds found in the molecules, specifically the rotatable bonds. As is to be expected with having the most atoms, 7BP contains the highest number of both rotatable and rigid bonds, with 6AP having the lowest. However, as with the molecular weight, the number of rotatable bonds is inversely proportional to the *in vivo* efficacy. This suggests that flexibility of the compound is not an important contributor to the effectiveness of the drugs.

By contrast, there are also a few results that are the same throughout all three compounds; namely, those that pertain to hydrogen bonds. All three sulfonamides have the same number of H-bond donators and acceptors between them, and as a result are only able to form one singular hydrogen bond. This corroborates the results of the LIGPLOT test, which showed little to no polar or hydrogen bonds being formed in the docking of the molecules. As such, it can be assumed these bonds are not vital to the medicinal properties of the compounds.

The largest and potentially most important result however, comes with the logP (or partition coefficient) of the sulfonamides. This property describes the ratio between solubility of a molecule in polar and non-polar diluents (commonly water and octanol)[55], with a large positive value indicating high lipophilicity, and a large negative value demonstrating high hydrophilicity. All three values for the sulfonamides tested are positive, indicating their preference for non-polar solvents, however the result for 7BP is significantly larger than the other two (with 6AP showing the lowest logP). These results match the closest to the relative *in vivo* efficacy of the compounds (with 6AP and 5AP having similar effect, and 7BP being much lower), suggesting this is a major contributing factor to the effectiveness of the molecules.

It is worth noting that the variation between the logP values of 7BP and the other two molecules is more significant that it may first appear, as this result is utilising a logarithmic scale. This means that though the difference between the logP values seems small (at only around 4), the actual difference is at a magnitude in the order of around 10,000. This suggests that 7BP is significantly for hydrophobic than the other two compounds. Furthermore, for the 7BP result, the logP of this molecule is above 5, whist those of the other two are significantly below this. This has potentially large ramifications for 7BP to act as a medicine (at least orally), as it is the only value studied to not comply with Lipinski's rule of 5[56]. This rule denotes a number of factors that affect a molecule's suitability to act as a drug with regards to oral absorption, specifically with regards to molecular weight (less than 500 g/mol), number of hydrogen bond donors (less than 5) and acceptors (less than 10), and logP (less than 5)[56]. Of all results shown in Table 4.1, all three molecules have molecular weights below 500 g/mol, as well as less than five hydrogen bond donors and ten acceptors. As such, this high logP value for 7BP could

contribute to a much lower oral absorption for this sulfonamide than for the other two, significantly hindering its ability to act as an oral drug to treat breast and cervical cancers.

From the results of the DruLiTo study, a number of factors that contribute to the effectiveness of the sulfonamides as anti-cancer therapies were identified. Namely, the lower molecular weight and partition coefficient of the compounds are the most important contributing factors, with the flexibility of the molecules playing a minor role, if one exists at all. However, further experiments will need to be conducted to determine the true significance of these factors on the effectiveness of the medicines.

4.3.2 Conclusions from the results and Further Potential Explanation

The two experiments above have identified a number of potential explanations for the unexpected outcomes of the MM/PBSA calculations, as well as rule out others. The key results that were shown pertain to the DruLiTo comparison, namely the molecular weight and the logP values. These matched the *in vivo* experimental efficacies the closest of all results generated for this thesis, and as such are the most likely determining factors for the effectiveness of the molecules as drugs. By contrast, the flexibility of the molecules did not show any significant impact in terms of matching the experimental results from Maalik et al., and as such are not likely to have an impact in terms of the effectiveness of the molecules.

Looking past the results of this thesis, there are a few other factors that could contribute to the relative effectiveness of 6AP and 5AP. Namely, the ability of these compounds to cross the cell membrane and enter cancer cells, and any methods by which they can be transported through the cells to their target proteins, could be determining factors. However, further studies into these areas would need to be done before any conclusions can be drawn.

4.4: Potential Impact of Results

The results of this study have the potential to have ramifications in terms of the reference study done by Maalik et al.

In terms of the immediate impact, the results shown are able to give a better understanding of the way in which the sulfonamides act on breast and cervical cancer cells, as well as identify that β-Tubulin is the more likely target for the drugs of the two tested. This could potentially lead to further study into the compounds, such as modifications to improve binding to the molecules to their target proteins, or improve the hydrophilicity of the compounds. The results of this thesis were passed on to the team of the reference study, so they may be utilised in any potential further research.

On a wider scale, the results of this study and those of Maalik et al. have helped to identify and categorise a number of potential new anti-cancer therapies. Increasing evidence supporting the theory that Tubulin and CA IX are likely targets for these treatments could result in more drugs being developed to act specifically on these two proteins. Furthermore, the use of sulfonamides to successfully destroy cancer cells could lead to an increase in this class of medications being used in oncological therapies, with the potential for reduced side effects from these relatively less harmful compounds. Continuing research in this topic would be needed, however.

4.5: Future Research

As discussed already during this thesis, there are still a number of factors that could be studied to corroborate these results. Furthermore, the results of this study could lead to a wide array of future research into the use of sulfonamides as anti-cancer therapies.

In order to corroborate the results of the MM/PBSA calculations, the first experiment that could be carried out would be to perform these simulations with the remaining four molecules that were only studied using Autodock Vina in the present work. This could provide further evidence of not only the identification of β-Tubulin as the preferred target of the two proteins tested, but also support the theories that logP and hydrophobic interactions are the more significant deciding factors in the efficacy of the compounds. This could also help to provide further explanation for the unexpected 6AP-Tubulin result.

To continue the study into the specific compounds of this research, there are a few directions that could be taken. Firstly, to better understand and support the results of the MM/PBSA and *in vivo* experiments, and as identified in Section 4.3.3 above, further research into the methods by which the compounds are absorbed by the cancer cells, and moved through the cytoplasm to their target proteins could be carried out. For example, potential transported proteins imbedded in the cell membrane could be identified and studied.

Another set of experiments that could be carried out following these two studies (and the next logical step in drug development) would be to perform live animal studies. These should be done to not only help confirm the effectiveness of the compounds in treating breast and cervical cancers, but also to help identify potential side effects of the medications. For example, a study could be done involving mice affected by cancerous tissue, with the effects of the various compounds being studied to help select the most appropriate to proceed to human trials.

Following the results of this thesis, as well as or alongside the potential studies mentioned above, another experiment that could be conducted would be to identify any particular aspects of the molecules that could potentially be a hinderance to their use as medications, and modify these. This could potentially result in not only a new set of compounds that are more effective than those already studied, but also to reduce potential side effects, thus making them safer for their eventual human use.

4.6: Conclusion

In conclusion, the research produced for this thesis has shown a number of interesting results. The novel sulfonamides provided by Maalik et al. were narrowed down to only the most effective following *in vivo* experiments, and β-Tubulin and Carbonic Anhydrase IX were selected as the most likely targets for their mode of action. Autodock Vina docking and MM/PBSA calculations were performed, and Tubulin was identified as the most likely target protein.

However, the results of the MM/PBSA calculations did not entirely match those of the experimental data. 6AP, the most effective *in vivo*, was shown to have the weakest docking to tubulin, indicating that there may be other factors affecting the success rate of the molecules. This was also potentially

explained by consideration of the solvent free energy contributions to the MM/PBSA calculations, the cause of which was the small number of atoms in the compound, potentially leading to less interactions with solvent molecules.

Other properties of the molecules were also studied to highlight any other factors that may have contributed to the results of the two studies. Molecular weight and the relative lipophilicity of the were identified as the most likely contributors, with the smaller, more hydrophilic compounds showing the highest efficacy. However, further research not only into these compounds, but other sulfonamides as potential anti-cancer therapies is encouraged, so that we can better treat these devastating diseases.

References

(1) Krzywik, J.; Mozga, W.; Aminpour, M.; Janczak, J.; Maj, E.; Wietrzyk, J.; Tuszyński, J. A.; Huczyński, A. Synthesis, Antiproliferative Activity and Molecular Docking Studies of Novel Doubly Modified Colchicine Amides and Sulfonamides as Anticancer Agents. *Mol. Basel Switz.* **2020**, *25* (8), E1789. https://doi.org/10.3390/molecules25081789.

(2) Maalik, A. Synthesis, Characterization, Cytotoxic and Early Apoptosis Induction of Sulfonamide-Based Alkyl-N-Pyridyl Derivatives, 2020.

(3) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461. https://doi.org/10.1002/jcc.21334.

(4) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discov.* **2015**, *10* (5), 449–461. https://doi.org/10.1517/17460441.2015.1032936.

(5) Wallace, A. C.; Laskowski, R. A. LIGPLOT v.4.0 - Operating manual http://www.csb.yale.edu/userguides/graphics/ligplot/manual/index.html (accessed 2022 -02 -07).

(6) Breast cancer https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed 2021 -09 -20).

(7) Cervical cancer https://www.who.int/westernpacific/health-topics/cervical-cancer (accessed 2021 -09 -20).

(8) Garofalo, M.; Link to external site, this link will open in a new window; Grazioso, G.; Link to external site, this link will open in a new window; Cavalli, A.; Sgrignani, J.; Link to external site, this link will open in a new window. How Computational Chemistry and Drug Delivery Techniques Can Support the Development of New Anticancer Drugs. *Molecules* **2020**, *25* (7), 1756. http://dx.doi.org/10.3390/molecules25071756.

(9) Bussi, G.; Branduardi, D. Free-Energy Calculations with Metadynamics: Theory and Practice. In *Reviews in Computational Chemistry Volume 28*; John Wiley & Sons, Ltd, 2015; pp 1–49. https://doi.org/10.1002/9781118889886.ch1.

(10) Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W. Computational Methods in Drug Discovery. *Pharmacol. Rev.* **2014**, *66* (1), 334–395. https://doi.org/10.1124/pr.112.007336.

(11) Minami, K. Supercomputers and Application Performance. In *The Art of High Performance Computing for Computational Science, Vol. 2: Advanced Techniques and Examples for Materials Science*; Geshi, M., Ed.; Springer: Singapore, 2019; pp 1–9. https://doi.org/10.1007/978-981-13-9802-5_1.

(12) Summa, V.; Petrocchi, A.; Bonelli, F.; Crescenzi, B.; Donghi, M.; Ferrara, M.; Fiore, F.; Gardelli, C.; Gonzalez Paz, O.; Hazuda, D. J.; Jones, P.; Kinzel, O.; Laufer, R.; Monteagudo, E.; Muraglia, E.; Nizi, E.; Orvieto, F.; Pace, P.; Pescatore, G.; Scarpelli, R.; Stillmock, K.; Witmer, M. V.; Rowley, M. Discovery of Raltegravir, a Potent, Selective Orally Bioavailable HIV-Integrase Inhibitor for the Treatment of HIV-AIDS Infection. *J. Med. Chem.* **2008**, *51* (18), 5843–5855. https://doi.org/10.1021/jm800245z.

(13) Banerjee, S.; Tian, T.; Wei, Z.; Shih, N.; Feldman, M. D.; Peck, K. N.; DeMichele, A. M.; Alwine, J. C.; Robertson, E. S. Distinct Microbial Signatures Associated With Different Breast Cancer Types. *Front. Microbiol.* **2018**, *9*, 951. https://doi.org/10.3389/fmicb.2018.00951.

(14) Classification and prognosis of invasive breast cancer: from morphology to molecular taxonomy | Modern Pathology https://www.nature.com/articles/modpathol201033 (accessed 2021 -09 -21).

(15) Alkabban, F. M.; Ferguson, T. Breast Cancer. In *StatPearls*; StatPearls Publishing: Treasure Island (FL), 2021.

(16) Gorodetska, I.; Kozeretska, I.; Dubrovska, A. BRCA Genes: The Role in Genome Stability, Cancer Stemness and Therapy Resistance. *J. Cancer* **2019**, *10* (9), 2109–2127. https://doi.org/10.7150/jca.30410.

(17) Feng, Y.; Spezia, M.; Huang, S.; Yuan, C.; Zeng, Z.; Zhang, L.; Ji, X.; Liu, W.; Huang, B.; Luo, W.; Liu, B.; Lei, Y.; Du, S.; Vuppalapati, A.; Luu, H. H.; Haydon, R. C.; He, T.-C.; Ren, G. Breast Cancer Development and Progression: Risk Factors, Cancer Stem Cells, Signaling Pathways, Genomics, and Molecular Pathogenesis. *Genes Dis.* **2018**, *5* (2), 77–106. https://doi.org/10.1016/j.gendis.2018.05.001.
(18) Rosen, R. D.; Sapra, A. TNM Classification. In *StatPearls*; StatPearls Publishing: Treasure Island (FL), 2022.
(19) Management of Early Breast Cancer - Evidence-Based Best Practice Guideline. 255.
(20) Joshi, S. C.; Khan, F. A.; Pant, I.; Shukla, A. Role of Radiotherapy in Early Breast Cancer: An Overview. *Int. J. Health Sci.* **2007**, *1* (2), 259–264.
(21) Sex hormones and hormone antagonists in malignant disease - New Zealand Formulary https://nzf.org.nz/nzf_4835 (accessed 2021 -10 -01).
(22) Fowler, J. R.; Maani, E. V.; Jack, B. W. Cervical Cancer. In *StatPearls*; StatPearls Publishing: Treasure Island (FL), 2021.
(23) Bosch, F. X.; Sanjosé, S. de. Human Papillomavirus in Cervical Cancer. *Curr. Oncol. Rep.* **2002**, *4* (2), 175–184. https://doi.org/10.1007/s11912-002-0079-y.
(24) Waggoner, S. E. Cervical Cancer. *Lancet Lond. Engl.* **2003**, *361* (9376), 2217–2225. https://doi.org/10.1016/S0140-6736(03)13778-6.
(25) Devi, M. A.; Ravi, S.; Vaishnavi, J.; Punitha, S. Classification of Cervical Cancer Using Artificial Neural Networks. *Procedia Comput. Sci.* **2016**, *89*, 465–472. https://doi.org/10.1016/j.procs.2016.06.105.
(26) Types of Cervical Cancer: Common, Rare and More https://www.cancercenter.com/cancer-types/cervical-cancer/types (accessed 2021 -10 -02).
(27) Human papillomavirus (HPV) https://www.immune.org.nz/diseases/human-papillomavirus-hpv (accessed 2021 -10 -02).
(28) Human Papillomavirus (HPV) Vaccines - National Cancer Institute https://www.cancer.gov/about-cancer/causes-prevention/risk/infectious-agents/hpv-vaccine-fact-sheet (accessed 2021 -10 -03).
(29) Šarenac, T.; Mikov, M. Cervical Cancer, Different Treatments and Importance of Bile Acids as Therapeutic Agents in This Disease. *Front. Pharmacol.* **2019**, *10*, 484. https://doi.org/10.3389/fphar.2019.00484.
(30) Vessels for Collective Progress: The Use of HeLa Cells in COVID-19 Research. *Science in the News*, 2020.
(31) Lopez, M.; Bornaghi, L. F.; Innocenti, A.; Vullo, D.; Charman, S. A.; Supuran, C. T.; Poulsen, S.-A. Sulfonamide Linked Neoglycoconjugates−A New Class of Inhibitors for Cancer-Associated Carbonic Anhydrases. *J. Med. Chem.* **2010**, *53* (7), 2913–2926. https://doi.org/10.1021/jm901888x.
(32) BPJ 49: Antibiotic series: Appropriate use of sulfonamide antibiotics https://bpac.org.nz/bpj/2012/december/sulfonamides.aspx (accessed 2021 -10 -03).
(33) SHAMBAUGH, G. E., JR. History of Sulfonamides. *Arch. Otolaryngol.* **1966**, *83* (1), 1–2. https://doi.org/10.1001/archotol.1966.00760020003001.
(34) Wang, J.-G.; Li, Z.-M.; Ma, N.; Wang, B.-L.; Jiang, L.; Pang, S. S.; Lee, Y.-T.; Guddat, L. W.; Duggleby, R. G. Structure-Activity Relationships for a New Family of Sulfonylurea Herbicides. *J. Comput. Aided Mol. Des.* **2005**, *19* (11), 801–820. https://doi.org/10.1007/s10822-005-9028-9.
(35) Cotton, F. Albert.; Stokely, P. F. Structural Basis for the Acidity of Sulfonamides. Crystal Structures of Dibenzenesulfonamide and Its Sodium Salt. *J. Am. Chem. Soc.* **1970**, *92* (2), 294–302. https://doi.org/10.1021/ja00705a012.
(36) Sulfonamide | chemical compound https://www.britannica.com/science/sulfonamide (accessed 2021 -10 -04).
(37) PubChem. Sulfanilamide https://pubchem.ncbi.nlm.nih.gov/compound/5333 (accessed 2021 -10 -04).

(38) Kwon, Y.; Song, J.; Lee, H.; Kim, E.-Y.; Lee, K.; Lee, S. K.; Kim, S. Design, Synthesis, and Biological Activity of Sulfonamide Analogues of Antofine and Cryptopleurine as Potent and Orally Active Antitumor Agents. *J. Med. Chem.* **2015**, *58* (19), 7749–7762. https://doi.org/10.1021/acs.jmedchem.5b00764.

(39) Souza, A.; Castro, H.; Brito, M.; Andrighetti-Fröhner, C.; Magalhães, U.; Oliveira Strege, K.; Gaspar-Silva, D.; Pacheco, L.; Joussef, A.; Steindel, M.; Simões, C.; Santos, D.; Albuquerque, M.; Rodrigues, C.; Nunes, R. Leishmania Amazonensis Growth Inhibitors: Biological and Theoretical Features of Sulfonamide 4-Methoxychalcone Derivatives. *Curr. Microbiol.* **2009**, *59*, 374–379. https://doi.org/10.1007/s00284-009-9447-2.

(40) Mohan, R.; Banerjee, M.; Ray, A.; Manna, T.; Wilson, L.; Owa, T.; Bhattacharyya, B.; Panda, D. Antimitotic Sulfonamides Inhibit Microtubule Assembly Dynamics and Cancer Cell Proliferation. *Biochemistry* **2006**, *45* (17), 5440–5449. https://doi.org/10.1021/bi0523409.

(41) Luo, Y.; Zhou, Y.; Song, Y.; Chen, G.; Wang, Y.-X.; Tian, Y.; Fan, W.-W.; Yang, Y.-S.; Cheng, T.; Zhu, H.-L. Optimization of Substituted Cinnamic Acyl Sulfonamide Derivatives as Tubulin Polymerization Inhibitors with Anticancer Activity. *Bioorg. Med. Chem. Lett.* **2018**, *28* (23–24), 3634–3638. https://doi.org/10.1016/j.bmcl.2018.10.037.

(42) Banerjee, M.; Poddar, A.; Mitra, G.; Surolia, A.; Owa, T.; Bhattacharyya, B. Sulfonamide Drugs Binding to the Colchicine Site of Tubulin: Thermodynamic Analysis of the Drug−Tubulin Interactions by Isothermal Titration Calorimetry. *J. Med. Chem.* **2005**, *48* (2), 547–555. https://doi.org/10.1021/jm0494974.

(43) Mboge, M. Y.; Chen, Z.; Wolff, A.; Mathias, J. V.; Tu, C.; Brown, K. D.; Bozdag, M.; Carta, F.; Supuran, C. T.; McKenna, R.; Frost, S. C. Selective Inhibition of Carbonic Anhydrase IX over Carbonic Anhydrase XII in Breast Cancer Cells Using Benzene Sulfonamides: Disconnect between Activity and Growth Inhibition. *PLOS ONE* **2018**, *13* (11), e0207417. https://doi.org/10.1371/journal.pone.0207417.

(44) Morgan, D. O. *The Cell Cycle: Principles of Control*; Primers in biology; Published by New Science Press in association with Oxford University Press ; Distributed inside North America by Sinauer Associates, Publishers: London : Sunderland, MA, 2007.

(45) Fojo, A. T.; Fojo, A. T. *The Role of Microtubules in Cell Biology, Neurobiology, and Oncology*; Humana Press: Totowa, NJ, UNITED STATES, 2008.

(46) Bank, R. P. D. RCSB PDB - 6E7B: 13-pf 3-start GMPCPP-human alpha1B/beta3 microtubules https://www.rcsb.org/structure/6E7B (accessed 2021 -10 -05).

(47) Lacroix, B.; Letort, G.; Pitayu, L.; Sallé, J.; Stefanutti, M.; Maton, G.; Ladouceur, A.-M.; Canman, J. C.; Maddox, P. S.; Maddox, A. S.; Minc, N.; Nédélec, F.; Dumont, J. Microtubule Dynamics Scale with Cell Size to Set Spindle Length and Assembly Timing. *Dev. Cell* **2018**, *45* (4), 496-511.e6. https://doi.org/10.1016/j.devcel.2018.04.022.

(48) *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications*; Frost, S. C., McKenna, R., Eds.; Subcellular Biochemistry; Springer Netherlands: Dordrecht, 2014; Vol. 75. https://doi.org/10.1007/978-94-007-7359-2.

(49) Hospital, A.; Goñi, J. R.; Orozco, M.; Gelpi, J. L. Molecular Dynamics Simulations: Advances and Applications. *Adv. Appl. Bioinforma. Chem.* **2015**, *8*, 37–47. https://doi.org/10.2147/AABC.S70333.

(50) Kutzner, C.; Páll, S.; Fechner, M.; Esztermann, A.; de Groot, B. L.; Grubmüller, H. More Bang for Your Buck: Improved Use of GPU Nodes for GROMACS 2018. *J. Comput. Chem.* **2019**, *40* (27), 2418–2431. https://doi.org/10.1002/jcc.26011.

(51) Kowalski, K.; Bair, R.; Bauman, N. P.; Boschen, J. S.; Bylaska, E. J.; Daily, J.; de Jong, W. A.; Dunning, T.; Govind, N.; Harrison, R. J.; Keçeli, M.; Keipert, K.; Krishnamoorthy, S.; Kumar, S.; Mutlu, E.; Palmer, B.; Panyala, A.; Peng, B.; Richard, R. M.; Straatsma, T. P.; Sushko, P.; Valeev, E. F.; Valiev, M.; van Dam, H. J. J.; Waldrop, J. M.; Williams-Young, D. B.; Yang, C.; Zalewski, M.; Windus, T. L. From NWChem to NWChemEx: Evolving with the Computational Chemistry Landscape. *Chem. Rev.* **2021**, *121* (8), 4962–4998. https://doi.org/10.1021/acs.chemrev.0c00998.

(52) Amber Advanced Tutorials - Tutorial 3 - MM-PBSA - Introduction http://ambermd.org/tutorials/advanced/tutorial3/ (accessed 2022 -02 -06).
(53) John Mongan, †; Carlos Simmerling, #; J. Andrew McCammon, †; David A. Case, §; Alexey Onufriev*, ◆. Generalized Born Model with a Simple, Robust Molecular Volume Correction http://pubs.acs.org/doi/full/10.1021/ct600085e (accessed 2022 -02 -22). https://doi.org/10.1021/ct600085e.
(54) Tuccinardi, T. What Is the Current Value of MM/PBSA and MM/GBSA Methods in Drug Discovery? *Expert Opin. Drug Discov.* **2021**, *16* (11), 1233–1237. https://doi.org/10.1080/17460441.2021.1942836.
(55) Avdeef, A.; Box, K. J.; Comer, J. E. A.; Hibbert, C.; Tam, K. Y. PH-Metric LogP 10. Determination of Liposomal Membrane-Water Partition Coefficients of Lonizable Drugs. *Pharm. Res.* **1998**, *15* (2), 209–215. https://doi.org/10.1023/A:1011954332221.
(56) Benet, L. Z.; Hosey, C. M.; Ursu, O.; Oprea, T. I. BDDCS, the Rule of 5 and Drugability. *Adv. Drug Deliv. Rev.* **2016**, *101*, 89–98. https://doi.org/10.1016/j.addr.2016.05.007.

Appendix A

md.in

```
Explicit solvent molecular dynamics constant pressure 50 ns MD
 &cntrl
   imin=0, irest=1, ntx=5,
   ntpr=500000, ntwx=500000, ntwr=500000, nstlim=25000000,
   dt=0.002, ntt=3, tempi=300,
   temp0=300, gamma_ln=1.0, ig=-1,
   ntp=1, ntc=2, ntf=2, cut=9,
   ntb=2, iwrap=1, ioutfm=1,
```

heat.in

```
Heat
 &cntrl
  imin=0,
  ntx=1,
  irest=0,
  nstlim=10000,
  dt=0.002,
  ntf=2,
  ntc=2,
  tempi=0.0,
  temp0=300.0,
  ntpr=100,
  ntwx=100,
  cut=8.0,
  ntb=1,
  ntp=0,
  ntt=3,
  gamma_ln=2.0,
  nmropt=1,
  ig=-1,
 /
&wt type='TEMP0', istep1=0, istep2=9000, value1=0.0, value2=300.0 /
&wt type='TEMP0', istep1=9001, istep2=10000, value1=300.0, value2=300.0 /
&wt type='END'
```

min.in

```
Minimize
 &cntrl
  imin=1,
  ntx=1,
  irest=0,
  maxcyc=200000,
  ncyc=1000,
  ntpr=100,
  ntwx=0,
  cut=8.0,
```

density.in

```
heat subs
 &cntrl
  imin=0,irest=1,ntx=5,
  nstlim=25000,dt=0.002,
  ntc=2,ntf=2,
  cut=8.0, ntb=2, ntp=1, taup=1.0,
  ntpr=500, ntwx=500,
  ntt=3, gamma_ln=2.0,
  temp0=300.0, ig=-1,
  ntr=1, restraintmask=':1-242',
  restraint_wt=2.0,
```

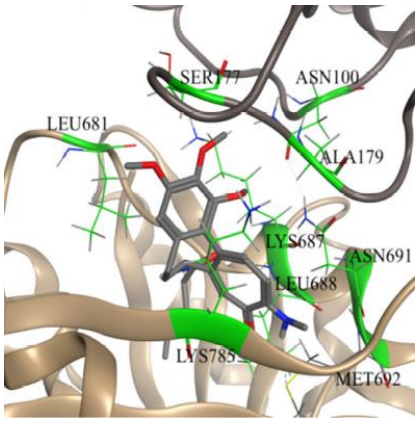
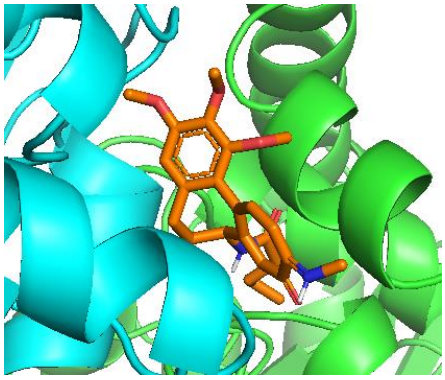
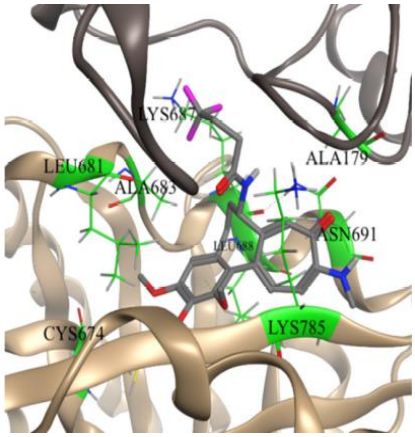
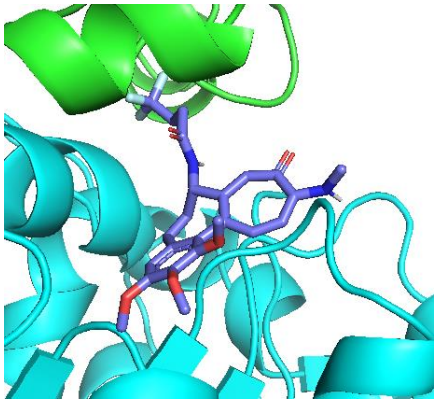
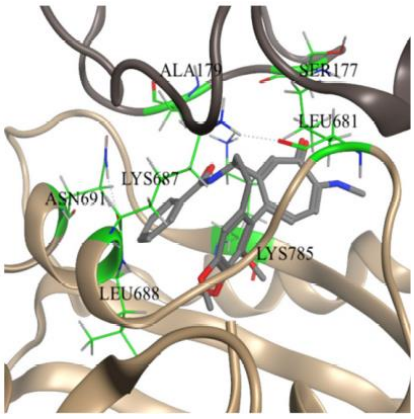
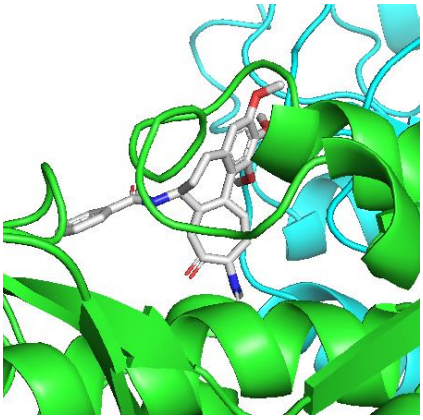
equilibrium.in

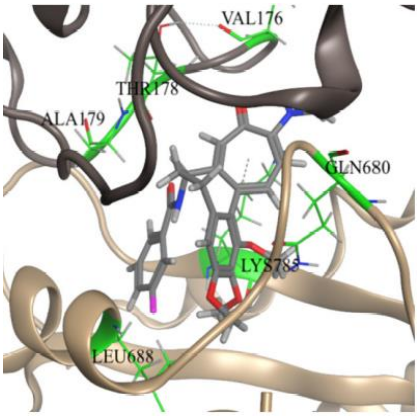
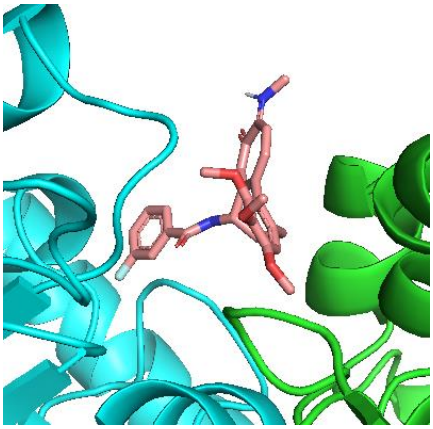
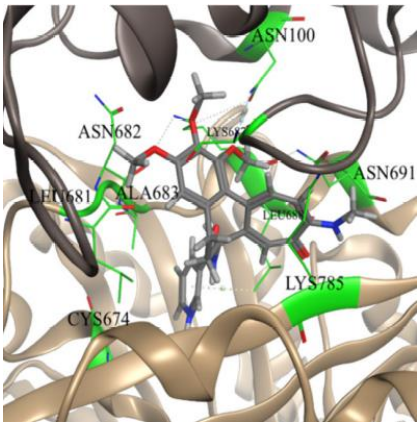
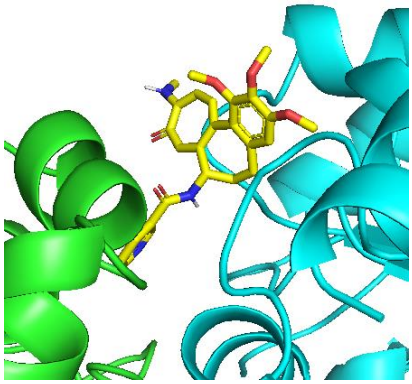
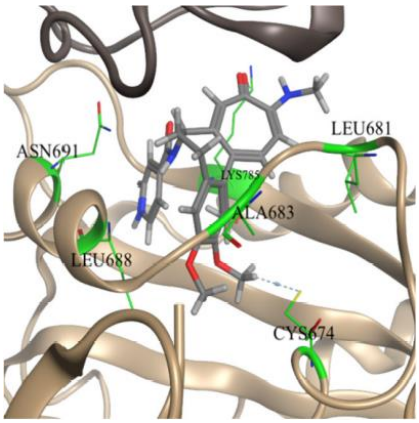
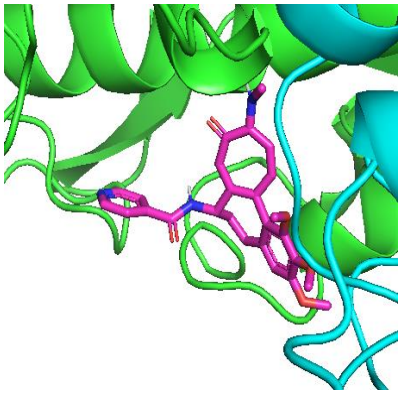
```
heat subs
 &cntrl
  imin=0,irest=1,ntx=5,
  nstlim=25000,dt=0.002,
  ntc=2,ntf=2,
  cut=8.0, ntb=2, ntp=1, taup=2.0,
  ntpr=1000, ntwx=1000,
  ntt=3, gamma_ln=2.0,
  temp0=300.0, ig=-1,
```

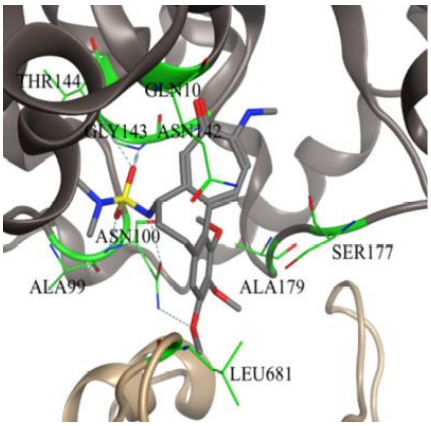
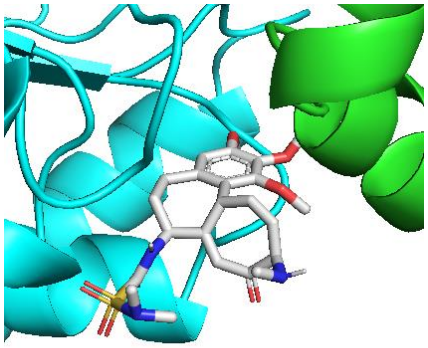
production.in

```
Production
 &cntrl
  imin=0,
  ntx=5,
  irest=1,
  nstlim=50000,
  dt=0.002,
  ntf=2,
  ntc=2,
  temp0=300.0,
  ntpr=5000,
  ntwx=5000,
  cut=8.0,
  ntb=2,
  taup=2.0,
  ntp=1,
  ntt=3,
  barostat=1,
  gamma_ln=2.0,
  ig=-1,
```

Appendix B

Derivative	Reference Paper Results	Autodock Vina Results
6		
11		
12		

Derivative	Reference Paper Results	Autodock Vina Results
14		
15		
16		

Derivative	Reference Paper Results	Autodock Vina Results
18		
19