# Computer-Aided Drug Design and the Biological Evaluation of Anti-Cancer Drugs

**By**

**Razia Moorad**



**Thesis Presented for the Degree of**

**DOCTOR OF PHILOSOPHY**

**In the Department of Medical Biochemistry**

**Faculty of Health Sciences**

**University of Cape Town**

**South Africa**

**Supervisor: Dr Luiz F. Zerbini**

**Co-Supervisor: Professor Kelly Chibale**

**August 2015**

## Declaration

I, **Razia Moorad**, hereby declare that the work on which this dissertation/thesis is based on my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: ……

Date: …….

# Abstract

Computer-aided drug design has become a promising alternative to high-throughput screening by identifying potential hits *in silico* for *in vitro* evaluation. In this study a combination of ligand-based and structure-based virtual screening was performed to identify *in silico* hits. This was based on finding similar inhibitors to 6-amino-4-(4-phenoxyphenylethylamino) quinazoline, a potent inhibitor of the Nuclear Factor kappa B (NF-κB), a transcription factor that has a pivotal role in cancer survival and Pentamidine, an anti-parasitic drug that has recently been demonstrated to possess tumour-killing activity.

A hierarchical methodology consisting of a similarity search followed by structure-based virtual screening of the ZINC database was performed. In order to perform the docking studies, binding sites for 6-amino-4-(4-phenoxyphenylethylamino) quinazoline on the NF-κB/IκBα complex were identified through blind docking. In addition, the National Cancer Institute database was screened, utilising existing structure-activity relationship data from literature. A pharmacophore search was designed to test the hypothesis of the structural features necessary for activity as seen with quinazoline inhibitors of NF-κB. No virtual hits from the ZINC database were confirmed with *in vitro* activity. On the other hand, three compounds identified from the pharmacophore search were confirmed to inhibit cancer cell proliferation *in vitro*, with compound NSC727152 demonstrating the most potent activity. In order to determine if NSC727152 acted similarly to 6-amino-4-(4-phenoxyphenylethylamino) quinazoline by inhibiting NF-κB, the effects of NSC727152 on the expression of NF-κB targeted genes, including the Growth Arrest and DNA Damage 45 (GADD45) α and γ and the Interleukin 6 (IL-6) genes were evaluated. GADD45 α and γ have been shown to be regulated by NF-κB during cancer progression and aberrant IL-6 gene expression has been implicated in cancer progression and mortality and its expression is at least partially mediated via constitutive activation of NF-κB. In this study, it has been demonstrated that GADD45 α and γ are upregulated after treatment with NSC727152. A down-regulation of the IL-6 promoter activity and mRNA expression in cancer cells treated with NSC727152 has also been demonstrated in this study. However, no hits similar to Pentamidine were confirmed with *in vitro* activity.

In conclusion, the compound NSC727152 has been shown to inhibit NF-κB and further analysis is necessary to determine its full potential as an NF-κB inhibitor.

# Acknowledgements

It is a great pleasure to acknowledge the following people who have made this thesis possible:

Without the love, support and prayers of my parents and family this thesis would not exist. I am especially grateful to my dear parents, Banu and Nazeer, for their prayers, kind words of encouragement, advice and guidance. Your unconditional love, support and faith in me makes this thesis as much yours as it is mine. My love, appreciation, respect and admiration for you both is indescribable.

I would like to express my gratitude to my supervisors, Dr. Luiz. F. Zerbini and Professor Kelly Chibale, for the opportunity to work with them in their research laboratories. I am particularly grateful to Professor Kelly Chibale at the Department of Chemistry for his continued support, advice and work-ethic and for providing me the financial assistance when it was most needed.

I thank my fellow researchers and staff members, at the International Centre of Genetic Engineering and Biotechnology (ICGEB) and the Department of Chemistry, for their words of encouragement, assistance and for the fun times shared. I am very grateful to Dr Grace Mugumbate at the Department of Chemistry, for her scientific advice and insightful discussions. Your words of encouragement and remembrance of God was always gratifying. I would also like to thank Dr Juliano Paccez at the ICGEB, for his assistance in teaching me the new techniques that were needed to complete this thesis.

I am grateful for the opportunity to work with Dr Alexander Alex and for his assistance in performing the docking calculations. His scientific advice, discussions and suggestions contributed immensely to the completion of this thesis.

I would like to thank my financial funders, ICGEB, for awarding me a full-fellowship to complete this PhD at the University of Cape Town.

I will forever remain grateful to Almighty God, for the good health and well-being that was necessary to complete this thesis.

# List of Abbreviations

| | |
|---|---|
| NCI | National Cancer Institute |
| SAR | Structure-Activity Relationship |
| FDA | Food and Drug Administration |
| ADME | Absorption Distribution Metabolism and Excretion |
| NCE | New Chemical Entities |
| CADD | Computer-Aided Drug Design |
| vHTS | Virtual High-Throughput Screening |
| HTS | High Through-put Screening |
| cMAPs | Connectivity Maps |
| ccRCC | Clear Cell Renal Cell Carcinomas |
| QNZ | N4- [2-(4-phenoxyphenyl) ethyl]-4, 6-quinazoline diamine |
| NF-κB | Nuclear Factor-κB |
| NSAIDS | Non-Steroidal Anti Inflammatory Drugs |
| LBVS | Ligand-Based Virtual Screening |
| SBVS | Structure-Based Virtual Screening |
| QSAR | Quantitative Structure-Activity Relationships |
| 2D | 2-Dimensional |
| 3D | 3-Dimensional |
| ECFP | Extended Connectivity Finger Print |
| ROCS | Rapid Overlay of Chemical Structures |
| Glide SP | Glide Standard Precision |
| Glide XP | Glide Extra Precision |
| LGA | Lamarkian Genetic Algorithm |
| ELISA | Enzyme-Linked Immunoadsorbent Assay |
| FRET | Fluorescence Resonance Energy Transfer |
| AdoMet-DC | S-Adenosylmethionine Decarboxylase |
| DAO | Diamine Oxidase |

| | |
|---|---|
| IL-6 | InterLeukin-6 |
| $IC_{50}$ | Half Maximal Inhibitory Concentration |
| RMSD | Root Mean Square Deviation |
| vROCS | Virtual ROCS Interface |
| SDF | Structure Data File |
| PDB | Protein Data Bank |
| IκBα | Inhibitory κBα |
| IKK | Inhibitory Kappa Kinase |
| GA | Genetic Algorithm |
| NLS | Nuclear Localisation Signal |
| Ser | Serine |
| Gln | Glutamine |
| Lys | Lysine |
| Phe | Phenylalanine |
| Tyr | Tyrosine |
| PKA | Protein Kinase A |
| Asp | Aspartic Acid |
| Asn | Asparagine |
| TPQ | Trihydroxyphenylalanine Quinone |
| Trp | Tryptophan |
| ROC Curve | Receiver Operating Characteristic Curve |
| SLogP | Atomic Log P |
| AUC | Area Under the Curve |
| Sp | Specificity |
| Se | Sensitivity |
| TN | True Negative |
| FP | False Positive |
| TP | True Positive |
| FN | False Negative |
| DMAMA | 5'-deoxy-5'-(Dimethylamino)-8-methyladenosine |

| | |
|---|---|
| Glu | Glutamic Acid |
| His | Histidine |
| Cys | Cysteine |
| MTT | 3-(4, 5-dimethylthiazol-2-yl) 2, 5 diphenyl tetrazolium bromide |
| qPCR | Quantitative Polymerase Chain Reaction |
| ATTC | American Type Culture Collection |
| VHL | Von Hippel-Lindau |
| DMEM | Dulbecco's Modified Eagle Medium |
| MEM | Minimum Essential Medium |
| RPMI-1640 | Roswell Park Memorial Institute-1640 |
| FBS | Fetal Bovine Serum |
| PBS | Phosphate Buffered Saline |
| DMSO | Dimethylsulfoxide |
| NMR | Nuclear Magnetic Resonance |
| DMF | N, N,-Dimethylformamide |
| MD | Molecular Dynamics |
| GADD45 | Growth Arrest and DNA Damage 45 |
| β-Me | β-Mercaptoethanol |
| GAPDH | Glyceraldehyde-3-phosphate Dehydrogenase |
| DFO | Desferrioxamine |
| ROS | Reactive Oxygen Species |
| MEKK4 | Mitogen-activated Protein Kinase Kinase 4 |
| MAPK | Mitogen-Activated Protein Kinase |

# Table of Contents

# Chapter 1: Introduction

## 1.1 The Anticancer Drug Discovery Process

Drug discovery has been in practice since antiquity (E. Ravina, 2011). This is evident by the ancient Egyptians meticulously documenting their advances made in surgery and pharmacology in several compendia, which also provides evidence of the earliest efforts made in anticancer drug discovery (S. Nobili et al, 2009; E. Ravina, 2011). The Edwin Smith Papyrus, which dates as far back as 3000 BC, mainly included descriptions of the surgical discoveries made and contains descriptions of the first occurrences of cancer or as it was described 'ulcers' of the breast (E. Ravina, 2011; M. Stiefel et al, 2006). In addition to the surgical removal of these tumours, herbal remedies were used to alleviate pain and over the centuries these remedies increased in therapeutic value by the inclusion of animal toxins and minerals (E. Ravina, 2011; J. Drews, 2000; S. Nobili et al, 2009).

In the early 1900's Paul Ehrlich defined the term 'chemotherapy' as the use of chemicals to treat disease and demonstrated that infectious diseases such as syphilis could be treated with drugs (E. Ravina, 2011; V. T. DeVita et al, 2008). Paul Ehrlich's interests also extended to the treatment of cancer, where he identified alkylating agents and aniline dyes as potential anticancer therapies but apparently he was not optimistic about their chances of success (V. T. DeVita et al, 2008). In addition, Paul Ehrlich contributed immensely to the advancements made in animal studies, being the first to document their effectiveness in evaluating the medicinal properties of drugs which later lead to major contributions in anticancer therapy (V. T. DeVita et al, 2008). The discovery of radium in 1898, by Marie and Pierre Curie lead to radiotherapy and in addition to surgery, these two treatment options dominated the field of anticancer treatment up until the 1960's (B. C. Baguley, 2002; V. T. DeVita et al, 2008).

The focus was put on anticancer research during World War II, in 1943, when the toxic effects in soldiers exposed to the accidental mustard gas spills were identified (V. T. DeVita et al, 2008). Studies were performed in secrecy in the United States of America, alongside research to develop chemical warfare, that resulted in the identification of nitrogen mustard and furthermore its medicinal action in treating lymphoma (B. C. Baguley, 2002). Soon after,

Sidney Faber, who was working with folic acid, identified Aminopterin, a close analogue of folic acid which eventually led to the development of Methotrexate which achieved leukaemia remission in children (M. Arruebo et al, 2011; V. T. DeVita et al, 2008). In 1958, researchers at the National Cancer Institute (NCI) discovered that Methotrexate could be used to cure solid tumours in humans, providing the first available treatment for solid tumours. In the years that followed, several alkylating drugs were synthesised such as chlorambucil and cyclophosphamide (M. Arruebo et al, 2011). These discoveries sparked tremendous interest in anticancer drug discovery and as a result many pharmaceutical companies and research centres were established during the 1950s and 1960s (E. Ratti et al, 2001). The anticancer drug discovery process was in its early stages, where researchers still heavily relied on the serendipitous discovery of compounds and utilised *in vivo* animal models to determine their therapeutic effects (E. Ratti et al, 2001). It was soon realised that the effects of absorption, metabolism and the pharmacokinetic effects of the compounds often resulted in a few successful lead candidates. In addition to the inability to produce back-up compounds in the event of compounds failing to produce the desired cytotoxic effects, a more rational approach utilising animal *in vitro* screening cell lines was developed (M. Suggitt et al, 2005). The use of *in vitro* animal models allowed for Structure-Activity Relationships (SAR) to be established, increasing the number of compounds that were tested and the role of the medicinal chemist (E. Ratti et al, 2001; V. T. DeVita et al, 2008).

Considerable advances were made in anticancer drug discovery and by the 1960's notable remission rates were achieved in treating childhood leukaemia's, although remission was often brief and disease reoccurrence was common with the current single agent treatment (M. Arruebo et al, 2011; V. T. DeVita et al, 2008). Despite this, confidence was instilled in researchers which led to the development of several combination therapy programmes, which achieved higher remission rates not only in treating childhood leukaemia but also in treating solid tumours such as Hodgkins disease (J. N. Latosinska et al, 2013; V. T. DeVita et al, 2008). After the 1970s, a considerable amount of funding was reserved for anticancer research and as a consequence the NCI had replaced the animal cell lines with a large screening panel of human tumours and the transplantable animal tumours were replaced with the human xenografts in nude mice (M. Suggitt et al, 2005; V. T. DeVita et al, 2008). These complex and expensive models dealt with the prevailing toxicity issues which were mainly due to the differences in animal and human pathologies. Although the increased expenses resulted in a fewer number of compounds being screened, the same number of leads were

being identified which included the taxanes such as Paclitaxel (V. T. DeVita et al, 2008; S. Nobili et al, 2009).

While success was achieved in finding new chemotherapeutic agents up until the 1990s, regulatory bodies such as the Food and Drug Administration (FDA) had put in place more stringent regulations regarding the safety and administration of drugs and as such many compounds failed to meet these regulations. In addition, in anticancer drug discovery the molecular dynamics of cancer was an added complexity, adding to the failure rate and high drug attrition rate from the drug discovery process (J. N. Latosinska et al, 2013). A further limitation encountered by researchers in anticancer drug discovery was to achieve selectivity for cancer cells, which revived Paul Ehrlichs 'magic bullet' philosophy, leading to the principle of selective cytotoxicity (J. N. Latosinska et al, 2013). This necessitated for advancements to be made in molecular biology and during the 1990s, a major focus was placed on finding drugs that inhibited the angiogenesis of cancer cells, that is, the formation of tumour blood vessels. The first drug of this class, Avastin, a humanised monoclonal antibody was discovered in 1997 and approved by the FDA in 2004 for its use in several types of metastatic cancers. Several other inhibitors of angiogenesis were discovered such as Sunitinib and Cetuximab (J. N. Latosinska et al, 2013). By this time, the drug discovery process had become an expensive, complex, time consuming and market-driven approach to find novel drugs candidates, urging for new technologies and a more rational approach to achieve success (J. N. Latosinska et al, 2013).

As a result, complex synthetic methods such as combinatorial chemistry were developed, with the premise that screening a larger number of compounds would result in a larger number of active compounds being identified (J. P. Kennedy et al, 2008). Although this resulted in technological advancements to screen the large number of compounds, it was soon realised that the expected outcomes from this approach was not the case and the focus was then shifted to identifying biological targets and their relationship to the structure and function of chemical molecules (J. Drews, 2000). The identification of key therapeutic targets was made possible by advancements in the fields of genomics and proteomics, which enabled an understanding of disease characteristics or phenotypes. This resulted in a variety of biological targets such as transcriptional factors, enzymes and receptors, which played a key role in disease progression to be identified (R. Goulding et al, 2009). The target-driven approach also resulted in a shift towards requiring a constant dialogue between chemists and biologists (E. Ratti et al, 2001; J. Drews, 2000). This adaptation was also reflected in the

change of screening assays that were commonly used to test for active drugs (E. Ratti et al, 2001; J. Drews, 2000; R. Borchardt et al, 2007). Moving from the traditional whole-animal, diseased based screening assays, such as whole cell assays, to biochemical assays using purified proteins had its advantages and disadvantages which are discussed further in this chapter. The main consequence of this change put the focus on characterising the 'drug-like' properties of drug compounds which included the Absorption, Distribution, Metabolism and Excretion (ADME), along with the physicochemical properties (R. Borchardt et al, 2007). It was quickly realised that the early characterisation of these properties was important to increase the chances of success and reduce the rate of drug attrition from the process. Under the pressure of the high costs and resources spent and the considerable time associated in getting drugs to the clinic, alternative cost-effective methods were sought (A. S. Reddy et al, 2007). This resulted in the development of computational tools to assist medicinal chemists to rationally design new drug candidates, with their intended biological targets in mind. Protein crystallography was another field that made great advancement during this time, which complemented the computational tools that were being developed (A. S. Reddy et al, 2007). Furthermore, collaborations between industry and academia were propelled (A. S. Reddy et al, 2007), much like the research laboratory this project was conducted in.

The value of computational data in predicting the expected outcomes of drug candidates was encouraging and contributed immensely to attaining the goal of drug discovery, which included finding safer medicines to treat disease (S. A. Patil, 2012). The target-driven approach meant that researchers had identified protein targets and validated their roles in the progression of the disease, which were often found to be over-expressed in disease tissues and had low expression levels in normal tissues (R. Goulding et al, 2009). The role of the medicinal chemists became increasingly important in the drug discovery process, selecting compounds from a virtual screening study, synthesising them and identifying active hits with non-promiscuous binding from *in vitro* assays and improving on their activity by conducting SAR studies to design appropriate analogues for biological testing (A. S. Reddy et al, 2007). Data regarding the ADME and physicochemical properties of active compounds that were extracted from *in vitro* and *in vivo* studies were utilised to optimise lead compounds during the lead optimisation stage with the goal of maximising efficiency and reducing side effects in animal models (J. G. Lombardino et al, 2004; S.A. Patil, 2012). Vigorous safety and toxicology testing are performed on promising drug candidates, comprising of the preclinical

and clinical development of New Chemical Entities (NCEs) before they are made available to the market (E. Ratti et al, 2001) (Figure 1).

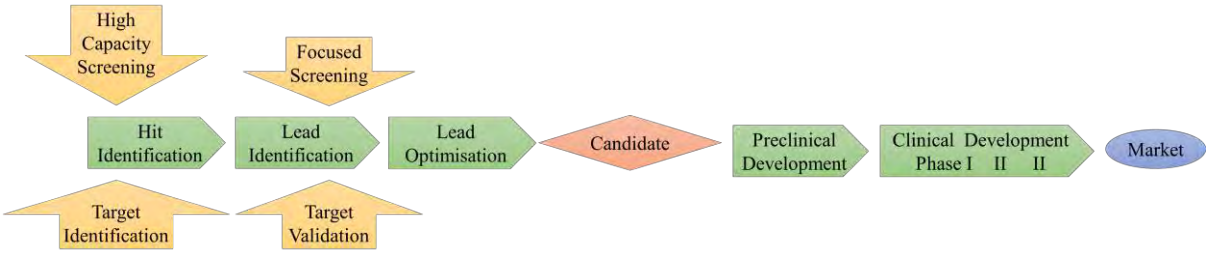



**Figure 1: The modern drug discovery process. The process was refined to include the target-driven approach adopted and includes the high throughput screening of compounds to facilitated hit identification. The application of computer-aided drug discovery is commonly applied from the hit discovery phase to the pre-clinical phase (E. Ratti et al, 2001).**

The design of Computer Aided Drug Design (CADD) technologies involved many fields such as biology, biophysics, structural biology and computational scientists (A. Wadood et al, 2013). The ultimate aim of these technologies was to reduce the associated costs of and speed up the drug discovery process. CADD techniques are commonly utilised in the early stages of the drug discovery process, facilitating the identification of hit compounds whose activity is then confirmed in *in vitro* screening assays (A. Wadood et al, 2013) (Figure 1). A variety of CADD techniques have been developed and include Virtual High Throughput Screening techniques (vHTS) consisting of both ligand-based and structure-based virtual screening approaches, chemoinformatic and bioinformatics approaches (G. Sliwoski et al, 2014; A. Wadood et al, 2013), which are discussed further in this chapter.

The availability of proteomic and genomic data in recent times has allowed various computational tools to be developed and has opened up other fruitful avenues in drug development, one of which includes drug repurposing (S. C. Gupta et al, 2013). In an effort to combat the low number of novel therapeutics reaching the clinics and the timely and expensive process of drug development, drug repurposing has become an attractive drug development strategy in anticancer research. It is the process of finding new indications for existing drugs or newly identified drugs that are used for the treatment of diseases other than the drugs intended disease (J. S. Shim et al, 2014). Compared to *de novo* drug development techniques, drug repurposing can be timeously performed in a cost effective manner as the safety, pharmacokinetic and pharmacodynamic properties of most repurposing candidates are

already known (J. S. Shim et al, 2014; J. Li et al, 2015). For example, the angiogenesis inhibitor Beracizumab, a monoclonal antibody developed to treat metastatic colon cancer is now used to treat or reverse the abnormal vascularisation of the retina, a symptom of exudative mascular degeneration (J. S. Shim et al, 2014). Bioinformatic tools have become widely popular to integrate pharmacological, genetic, chemical and clinical data to identify drug targets and to find new indications for existing drugs, through the use of connectivity maps (cMAPs) (J. S. Shim et al, 2014). cMAPs have been constructed with the aim of providing a detailed map for functional diseases, genetic alterations and drug actions, for example the construction of large scale gene expression profiles from human cancer cell lines treated with different drugs under different conditions (J. S. Shim et al, 2014). By integrating cMAPs with other functional genomic databases, drug repurposing is commonly performed using two approaches. The first approach consists of determining inverse drug-disease relationships by comparing the drugs gene expression profile and the disease gene expression profile. This approach is also known as 'signature reversion' and is used to identify which drugs are able to produce an inverse disease gene expression (J. S. Shim et al, 2014). This approach was used for the repurposing of Pentamidine isethionate, an anti-protozoal agent commonly used to treat Pneumocystis *carinii* pneumonia (A. P. McGrath et al, 2009; R. R. Tidwell et al, 1990). Pentamidine isethionate was found to produce an inverse gene expression in cancer cells lines through the use of cMAPs and has been demonstrated to have anticancer activity in clear cell renal cell carcinoma (ccRCC) (L. F. Zerbini et al, 2014). Another useful approach in utilising cMAPs for drug repurposing includes 'guilt by association' which identifies drugs that invoke a similar transcriptional response and thus sharing a similar mechanism of action (J. S. Shim et al, 2014). Publicly available repositories or databases of drug and disease gene expression profiles have provided the opportunity for many research institutions and academia to take advantage of computational drug repurposing tools in a timely and cost effective manner (J. Li et al, 2015).

In the context of this study, commonly used CADD techniques were utilised to identify compounds that act similarly to two previously identified compounds that have been reported in literature to have antitumor activity. These compounds include the commercially available Pentamidine Isothionate and N4- [2-(4-phenoxyphenyl) ethyl]-4, 6-quinazoline diamine (QNZ). Pentamidine is an anti-protozoal compound that has been identified to have antitumor activity (L. F. Zerbini et al, 2014; L. M. Scott et al, 2010; M. K. Pathak et al, 2002). QNZ is a potent inhibitor of the transcription factor Nuclear Factor-κB (NF-κB),

which plays a key role in the development and progression of cancer (M. Tobe et al, 2003; M. Tobe et al, 2003; B. Rayet et al, 1999; M. Karin, 2009; B. Hoesel et al, 2013). In addition, QNZ had been shown to induce apoptosis in ovarian cancer cell lines and when used in combination with Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) such as Diclofenac, an enhanced apoptotic effect is induced (L. F. Zerbini et al, 2011). The CADD techniques that were utilised in this study to find similar inhibitors to Pentamidine and QNZ are discussed further in this chapter.

## 1.2 Virtual High-Throughput Screening

vHTS has become a popular computational tool that aims to mimic the *in vitro* high throughput screening (HTS) of large chemical databases. vHTS is a powerful technique which reduces the associated costs and time involved in HTS, by selecting compounds that are predicted to be active and excluding those that are predicted to be inactive which are subsequently tested for *in vitro* activity (G. Sliwoski et al, 2014). A variety of computer algorithms have been developed to exploit available information to efficiently screen large chemical databases and identify compounds that are predicted to illicit the desired biological response. This has resulted in the development of two commonly used vHTS techniques which include ligand-based virtual screening (LBVS) and structure-based virtual screening (SBVS) (A. Wadood et al, 2013; G. Sliwoski et al, 2014). Certain limitations of computational methods such as their inability to identify the toxicities associated in a complete biological system still remains. Although certain tools such as filters, can be applied to identify and remove toxic groups or compounds from a database. Filters may also be applied to remove compounds which are not considered to be 'drug-like' or have poor ADME properties (S. Subramaniam et al, 2008). Furthermore, depending on the task at hand, computational studies may be time consuming and computationally exhausting but nevertheless, they provide useful insights to drug design (G. Klebe, 2006). These techniques are discussed in more detail below.

### 1.2.1 Ligand-Based Virtual Screening

Ligand Based Virtual Screening (LBVS) methods retrieve compounds that are closely related to known biologically active ligands (A. S. Reddy et al, 2007; P. Willet, 2006). LBVS is commonly performed when there is no structural data available or the intended target of an active ligand has not been identified. LBVS includes several methodologies such as the popular similarity search, pharmacophore search, sub-structure search, clustering methods and quantitative structure activity relationships (QSAR) (J. Bajorath, 2002).

The similarity search approach is based on the similarity property principle, which states that similar ligands will have similar biological activity and physicochemical properties (M. A. Johnson et al, 1990). Similarity is either computed based on structural or topological similarity or pharmacophoric similarity and therefore can be based on either 2 Dimensional (2D) or 3 Dimensional (3D) descriptors (P. Willet, 2006). A similarity measure incorporates 3 components, a representation to characterise molecules, a weighing scheme to assign the different degrees of importance to the components represented and a coefficient to determine the degree to relatedness between two structural representations (P. Willet, 2006; J. D. Holliday et al, 2002). Over the years, many types of molecular representations have been developed to compute molecular similarity, which include 2D fingerprints, molecular graphs, topological indices and molecular fields (P. Willet, 2006).

From the plethora of 2D molecular descriptors, molecular fingerprints such as the Extended Connectivity Finger Print (ECFP) remains highly popular due to their demonstrated effectiveness, rapid calculation and ability to represent a wide number of molecular features (D. Rogers et al, 2010; P. Willet, 2006; J. Hert et al, 2004; Y. Hu et al, 2009). In addition, molecular fingerprints are the method of choice to implement when only a single weak lead is available for a starting point (J. Bajorath, 2002). ECFPs are bit-binary strings that can represent an infinite number of molecular features including their stereo-chemical information. The ECFP fingerprint records the absence or presence of molecular features that are predicted to be responsible for molecular or drug activity (D. Rogers et al, 2010; M. Awale et al, 2014). They were initially developed for and are still widely used to compute the presence of sub-structures but are now applied in similarity searches, clustering and molecular classifications (D. Rogers et al, 2010). Their increased popularity is partly due to their implementation in the Pipeline Pilot software, where the ECFP_4 fingerprint is

reportedly a sufficient tool for a similarity search (Y. Hu et al, 2009). The ECFP_4 fingerprint, which was utilised in this study, describes both the identity and connectivity of an atom in a molecule and searches for similar descriptors within a diameter of 4 bonds radiating from each heavy atom (D. Rogers et al, 2010; Y. Hu et al, 2009).

Various similarity coefficients have been developed to quantify the degree of relatedness between two structural features and include but are not limited to the Cosine, Hamming, Russell-Rao, Forbes and the Tanimoto coefficient (A. S. Reddy et al, 2007; P. Willet, 2006). The Tanimoto Coefficient is widely used and is reported to provide a reliable and an unbiased measure of similarity (P. Willet, 2006; M. Syuib et al, 2013; J. D. Holliday et al, 2002). Furthermore, the Tanimoto coefficient is an appropriate coefficient to quantify fingerprint similarity searching (Y. Hu et al, 2009). The Tanimoto coefficient has been utilised to quantify similarity in both 2D and 3D descriptor space and is defined as follows:

$$\frac{c}{a+b-c}$$

**Figure 2: The equation used to calculate the Tanimoto Score. Where *a* and *b* represent bits in the bit-binary strings of 2 independent molecules and c represents the bit-binary string that is present in the fingerprint of both molecules (P. Willet, 2006).**

The 3D descriptors are considered more advanced and complex, which include shape-matching algorithms, shape-based fingerprints, 3D feature representations of molecular conformations, molecular field descriptions and pharmacophore fingerprints (H. Eckert et al, 2007). The application of 3D descriptors is critically dependent on the accurate generation of the conformation of a compound. The bioactive conformation of a compound is ordinarily assumed from crystal structures, but in the case of a database, where active conformations are generally not known, the accuracy in predicting these conformations plays a crucial role in the success of the 3D virtual screening (H. Eckert et al, 2007). The generation of conformations is a computationally exhausting task and the prediction of the correct bioactive conformation of compounds remains to be the bottleneck associated with the 3D similarity search (H. Eckert et al, 2007; K. Dobi et al, 2014). This adds to the debate for the preference of 2D descriptor over 3D shape-based descriptors (V. Venkatraman et al, 2010; M. Thimm et al, 2004; J. Bajorath, 2002).

3D shape based matching algorithms remain highly popular and several algorithms have been developed including Rapid Overlay of Chemical Structures (ROCS) (OpenEye Scientific Software Inc.), Cat-Shape (CATALYST) (Accelrys, San Diego), Phase-Shape (Schrodinger, LLC) etc. ROCS is highly preferred as its performance is reportedly not affected by the conformation of the query molecule (A. Hamza et al, 2012). Furthermore, ROCS incorporates a functional group or pharmacophoric feature matching, which alleviates the reliance on the correct bioactive conformation of the query molecule (A. Hamza et al, 2012).

ROCS performs a pair-wise comparison of the molecular shape and chemical features between the query compound and those compounds included in a database. Molecular structures on ROCS are represented as atom centered Gaussian functions, where their gradient based optimisation allows for the overlap of the rigid query and database molecules (S. M. Kearnes et al, 2014; P. C. D. Hawkins et al, 2007). An alignment with the highest shape (volume) overlap is retained and a ShapeTanimoto score is calculated using the equation depicted in Figure 2. In addition, a colour force field is implemented to determine the chemical features, where the interaction between the same types of features is considered an attractive interaction and will therefore form an overlap of colour force fields (C. Sotriffer, 2011). The chemical groups and functionalities identified on ROCS include hydrogen bond donors, acceptors, charged atoms, rings and hydrophobic regions (S. M. Kearnes et al, 2014). A ColourTanimoto score is calculated based on the weight and strength of these interactions, using the calculation in Figure 2. These Tanimoto scores calculated by ROCS can be analysed individually (ShapeTanimoto or ColourTanimoto, with a maximum score of 1) or in combination as a sum of the individual calculations (TanimotoCombo score, with a maximum score of 2) (C. Sotriffer, 2011).

The 3D pharmacophore search is another powerful technique that is widely used in drug discovery and can either be structure-based or ligand-based (V. Vyas et al, 2008). Pharmacophore virtual screening is a simple and vigorous methodology to rapidly identify active compounds (A. Wadood et al, 2013). The conventional ligand-based pharmacophore is generated on a set of known active ligands, in the absence of structural data. In pharmacophore mapping, it is hypothesised that the set of known active ligands will form similar interactions with the protein target and therefore captures the 3D electrostatic and geometric arrangement of the essential pharmacophore features within a molecular framework (V. Vyas et al, 2008; A. Wadood et al, 2013). Pharmacophore virtual screening

has the advantage of identifying biologically active compounds of novel chemical scaffolds (O. Dror et al, 2009; A. Wadood et al, 2013; V. Vyas et al, 2008).

Ligand-based pharmacophores may either be generated manually or automatically. Manual generation is preferred when there is experimental evidence that certain functional groups are important for biological activity (V. Vyas et al, 2008). This is in contrast to the alignment of a set of active compounds from which a pharmacophore can be automatically derived through the use of computational software (V. Vyas et al, 2008). The automatic generation of pharmacophores, such as the common feature pharmacophore, requires the correct conformational generation and alignment of a set of ligands, which can be time consuming and is further subjected to the quality of conformation generation (V. Vyas et al, 2008). The functional features that are captured in a pharmacophore include hydrogen bond donors, hydrogen bond acceptors, acidic groups, basic groups, partial charges, aliphatic hydrophobic and aromatic hydrophobic moieties and furthermore, distance and locational restraints can be applied to define their spatial arrangement (G. Sliwoski et al, 2014).

### 1.2.2 Structure-Based Virtual Screening

Structure Based Virtual Screening (SBVS) methods primarily include molecular docking calculations. There are certain pre-requisites for docking studies, which include a well-defined protein structure and the identification of a binding site (C. Hetenyi et al, 2002). In the case where there is no crystal structure available, homology modelling of the protein structure may provide an alternative methodology (G. Sliwoski et al, 2015).

Docking algorithms aim to predict the correct orientation and position of a ligand in the active site of a protein in addition to predicting the binding interactions between the ligand and protein and ranking them in a suitable order. Docking calculations that are performed in a specified binding site are known as focus docking calculations (D. Ghersi et al, 2009). This is different to an alternative methodology known as blind docking which utilises a ligand to scan the entire protein surface to identify suitable binding sites on the protein, as such, blind docking is commonly performed when the binding site on a protein is unknown (C. Hetenyi et al, 2002). Focused docking calculations are reported to predict the correct docking pose more frequently and hence complements a blind docking analysis (D. Ghersi et al, 2009).

There are many molecular docking tools available including Autodock4 (Scripps Research Institute), Glide (Schrodinger, LLC), Dock (UCSF) etc. (I. Wandzik, 2006; J. B. Cross et al, 2009). Plenty of comparative studies on the available docking tool have been reported in literature, which suggest that Glide is one of the superior docking tools available to perform docking calculations and for virtual screening. Glide is reportedly able to reproduce the binding pose of ligands in co-crystallised protein structures and is able to accurately predict and rank the binding energies of known ligands and their protein targets from a virtual screening (I. Wandzik, 2006; J. B. Cross et al, 2009; E. Kellenberger et al, 2004; E. Perola et al, 2004).

Glide's superiority in obtaining the accurate binding pose and binding affinities is thought to be a consequence of the series of hierarchical filters used to progressively search for a ligand's position and orientation with respect to the conformation of the protein. The shape and properties of the protein binding site are represented on a grid field which adds to the accuracy of the ligand pose generated (R. A. Friesner et al, 2004). A pre-screening of a ligand's orientation and position is performed, reducing the binding space in which the ligand is further minimised using the OPLS-AA force field in conjunction with a distance-dependant dielectric model (R. A. Friesner et al, 2004). The lowest energy ligand poses are selected and further subjected to a Monte Carlo procedure, which correctly orientates side chains and the internal torsion angles before the final GlideScore is calculated. Glide utilises a modified version of the ChemScore to calculate binding affinities (R. A. Friesner et al, 2004). Two forms of the GlideScore are calculated in Glide, the GlideScore SP (Standard Precision) and GlideScore XP (Extra Precision). The GlideScore SP function is considered to be a 'softer' function which is adapted to identify ligands which have a reasonable propensity to bind, regardless of imperfections in their binding pose and aims to minimise false negatives (R. A. Friesner et al, 2004). The equation used to calculate the GlideScore SP is summarised below in Figure 3.

$$\Delta G_{\text{bind}} = C_{\text{lipo-lipo}} \Sigma f(r_{\text{lr}}) + C_{\text{hbond-neut-neut}} \Sigma g(\Delta r)\, h(\Delta\alpha) + C_{\text{hbond-neut-charged}} \Sigma g(\Delta r)\, h(\Delta\alpha) + C_{\text{hbond-charged-charged}} \Sigma (\Delta r)\, h(\Delta\alpha) + C_{\text{max-metal-ion}} \Sigma f(r_{\text{lm}}) + C_{\text{rotb}} H_{\text{rotb}} + C_{\text{polar-phob}} V_{\text{polar-phob}} + C_{\text{coul}} E_{\text{coul}} + C_{\text{vdW}} E_{\text{vdW}} + \text{solvation terms}$$

**Figure 3: The GlideScore is a modified extension of the ChemScore, used to calculate the binding energy in the SP docking calculations (R. A. Friesner et al, 2004).**

The GlideScore XP function is considered to be a 'harder' function which scores penalties for poses that violate basic physical chemistry principles such as the solvent exposure of charged and polar groups. In addition to applying desolvation penalties, the Glide XP function is able to identify structural motifs such as the formation of salt bridges, which contribute to binding affinities (R. A. Friesner et al, 2006). The GlideScore XP equation is summarised in Figure 4 below.

$$\text{XP GlideScore} = E_{\text{coul}} + E_{\text{vdW}} + E_{\text{bind}} + E_{\text{penalty}}$$

Where,

$$E_{\text{bind}} = E_{\text{hyd\_enclosure}} + E_{\text{hb\_nn\_motif}} + E_{\text{hb\_cc\_motif}} + E_{\text{PI}} + E_{\text{hb\_pair}} + E_{\text{phobic\_pair}}$$

$$E_{\text{penalty}} = E_{\text{desolv}} + E_{\text{ligand\_strain}}$$

**Figure 4: The GlideScore XP equation used to calculate the ligand binding affinities (R. A. Friesner et al, 2006).**

AutoDock4 has been widely used in many molecular docking studies to identify potential anticancer compounds (T. Usha et al, 2014; B. Banerji et al, 2013; S. Sreenivasan et al, 2014). In addition to rapidly and accurately predicting the binding pose and binding energies of ligands and their protein targets, AutoDock4 utilises a grid base methodology which

allows for searching the large conformational space available to a ligand around a protein and is therefore a useful tool to perform blind docking calculations (G. M. Morris et al, 2009; G. M. Morris et al, 1998; D. S. Goodsell et al, 1996; C. Hetenyi et al, 2002). In AutoDock4, conformational searching is performed utilising the Lamarckian Genetic Algorithm (LGA) (G. M. Morris et al, 2009). The LGA creates an initial population of trial conformations, where successive generations of these conformations are able to mutate, by searching the local conformational space to find their local minima before exchanging conformational parameters. Analogous to biological evolution, each conformation competes and those with the lowest binding energies are selected (G. M. Morris et al, 2009; G. M. Morris et al, 1998). To calculate the predicted binding energies of ligands to their protein targets, AutoDock4 employs the semi-empirical free energy force field, which is based on a comprehensive thermodynamic model. The adapted force field is suitable for protein-protein docking calculations and incorporates protein flexibility, in addition to intramolecular energy (R. Huey et al, 2007). The AutoDock4 equation used to calculate binding affinities is summarised in Figure 5 below.

$$\begin{aligned}\Delta G = {} & \Delta G_{\text{vdW}} \sum_{i,j} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) \\ & + \Delta G_{\text{hbond}} \sum_{i,j} E(t) \left( \frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} + E_{\text{hbond}} \right) \\ & + \Delta G_{\text{elec}} \sum_{i,j} \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \\ & + \Delta G_{\text{tor}} N_{\text{tor}} \\ & + \Delta G_{\text{sol}} \sum_{i,j} S_i V_j e^{(-r_{ij}^2/2\sigma^2)}\end{aligned}$$

**Figure 5: The semi-empirical free energy force field equation used to calculate ligand binding energies on AutoDock4** (http://autodock.scripps.edu/resources/science/equations).

Where, i = index of atoms in the ligand, j = index of atoms in the receptor, $W_{desolv}$ = linear regression coefficient or weight for the desolvation free energy term, $S_i$ = solvation term for atom i, $V_i$ = atomic fragmental volume of atom i, $r_{ij}$ = distance between atom i and atom j (in Å), σ = gaussian distance constant = 3.5 Å. (http://autodock.scripps.edu/resources/science/equations).

## 1.3 Biological Screening Assays

As mentioned previously the aim of virtual high-throughput screening is to reduce the cost and time associated with conventional high-throughput screening, by selecting compounds that are predicted to be active and excluding those that are predicted to be inactive (G. Sliwoski et al, 2014). A fewer number of compounds are commonly identified from the top scoring compounds resulting from a LBVS and/or SBVS, for *in vitro* screening assays to confirm their biological activity (G. Sliwoski et al, 2014). In a target-driven drug discovery approach, the primary and secondary screening utilised to confirm *in vitro* activity of compounds are usually performed in either target-based or cell-based assays. Screening assays are generally sought to be low in cost, reproducible and provide good target specificity and sensitivity (A. S. Verkman, 2003; J. Major, 1998). The different types of screening assays are discussed in more detail below.

### 1.3.1 Target-Based Assays

Target-based assays are conducted *in vitro* on purified proteins or macromolecules and therefore require that the intended target has been well established and defined. Target-based assays are easier to perform if the target of interest can be easily expressed and purified in large quantities, making them more amenable for HTS and therefore they are commonly performed in industry (F. Sams-Dodd, 2005; J. P. Hughes et al, 2011; A. M. Burger et al, 2014). Furthermore, target-based assays facilitate the structure-based drug design of novel chemotherapeutics and are useful to plan SAR studies (V. Khurana et al, 2015).

A major limitation with target-based assays occurs in translating *in vitro* efficacy to *in vivo* efficacy as target-based assays are performed in isolation of cellular and signalling pathways. Therefore target-based assays are usually followed through with cell-based assays (V. Khurana et al, 2015). To increase the sensitivity of a target based HTS assays, fluorescent and illuminescent detection methods have been developed, for example those used in the enzyme-linked immunoadsorbent assay (ELISA) and the fluorescence resonance energy transfer (FRET) and time resolved fluorescence methods (A. M. Burger et al, 2014).

Pentamidine Isothionate has been shown to possess anticancer activity (L. F. Zerbini et al, 2014) and a number of protein targets have been identified to be inhibited by Pentamidine including the PRL family of oncogenic phosphatases (L. M. Scott et al, 2010; M. K. Pathak et al, 2002) and the serum protease coagulation factors, Factor IV and Factor Xa (A. Puckowska et al, 2012). Furthermore, Pentamidine has been identified to inhibit several enzymes involved in the polyamine biosynthetic pathway including S-Adenosylmethionine decarboxylase (AdoMet-DC) and Diamine Oxidase (DAO) by the use of target-based assays (A. P. McGrath et al, 2009; J. Stanek et al, 1993; U. Regenass et al, 1992; D. E. McCloskey et al, 2009). Target based assays have been especially useful in analysing the SAR and developing analogues of Pentamidine that have been tested for anticancer activity (A. Puckowska et al, 2012; J. Stanek et al, 1993; U. Regenass et al, 1992; D. E. McClowsky et al, 2009).

### 1.3.2 Cell-Based Assays

Without a doubt, the insights into the cellular molecular mechanisms that has made target based assays possible has become invaluable and has certainly tipped the scales in favour of the myriad of cell-based assays that have been developed to investigate targets of interest, for example the use of reporter gene assays and the second messenger assays (E. C. Butcher et al, 2004; E. A. Martis et al, 2011; R. Zang et al, 2012). Human tumour cell lines have certain characteristics such as their manageability and reproducible growth, allowing for their robust scale-up which makes them ideal systems to use from target identification and validation, primary screening, lead identification and toxicology screening (R. Zang et al, 2012; V. Khurana et al, 2015). As a result, in anticancer drug discovery using the classical monolayer cell based assay is preferred to identify quick-acting cytotoxic compounds. However, it does have some caveats that limit its use (A. M. Burger et al, 2014; W. Aherne et al, 2002). Due to the lack of blood vessels and the extracellular matrix, compounds that target these functions cannot be investigated using a monolayer cell assay (A. M. Burger et al, 2014). Furthermore, the use of cell-based assays may complicate the design of SAR studies due to issues of drug metabolism and permeability (V. Khurana et al, 2015).

In comparison to the target-based assays, cell-based assays are able to provide preliminary information on the permeability and stability of a compound when used with appropriate

controls (A. M. Burger et al, 2014). They are especially useful when the target is unknown, acting as a phenotypic screen to identify actives and later elucidate their mechanism of action, although it is reported that target elucidation can be a difficult feat (R. Zang et al, 2012). Furthermore, cell-based assays are able to discriminate between agonists, antagonists and allosteric inhibitors. These are certainly helpful properties to determine in the early stages of the drug discovery process (P. A. Johnston et al, 2002; A. M. Burger et al, 2014).

Transcriptional factors such as NF-κB have been identified as potential therapeutic targets and to gather more information on gene activation and transcription, custom made cell-based assays including the genetic reporter assays such as the luciferase assay were developed (V. Baud et al, 2009). Reporter gene assays have been tremendously useful in mapping the promoter, enhancer and silencer regulatory regions of different genes (M. F. Shannon et al, 2000). They have been useful in determining the direct binding of the different regulatory elements on the DNA, providing information on the transcriptional activation of different genes and their regulatory transcriptional factors (M. F. Shannon et al, 2000). The effects of inhibiting NF-κB have been previously investigated using custom made cell-based assays. For instance, the Interleukin (IL)-6 promoter region has been shown to be regulated by NF-κB and AP-1 protein, where point mutations on the NF-κB and AP-1 binding site resulted in reduced IL-6 promoter activity (L. F. Zerbini et al, 2003). In addition, cell-based assays were utilised to determine the inhibitory action of QNZ towards NF-κB (M. Tobe et al, 2003).

Cell-based assays were also utilised in this study to evaluate the *in vitro* cytotoxic activity of the compounds that were identified from the virtual screening techniques performed. Furthermore, modified cell-based assays such as the luciferase assay, were utilised to confirm the mechanism of action of the active compound in the secondary screening assays described further in Chapter 4, section 4.3.

## 1.4 Aims and Objectives

### 1.4.1 Study Aims

The aim of this study was to identify similar acting compounds to Pentamidine Isethionate and QNZ, through the use of computational techniques including LBVS and SBVS.

'Similarity' was measured by the *in vitro* biological activity of the compounds that were identified from the virtual screening in several cancer cell lines, including prostate, ovarian and renal cancer cells.

### 1.4.2 Study Objectives

1. To perform a hierarchical virtual screening of the ZINC database including of a similarity search and docking calculations, to identify similar compounds to Pentamidine and QNZ based on their similarity scores and predicted binding affinities.
2. To identify the binding site of QNZ utilising a blind docking protocol on the available human crystal structures of the NF-κB pathway proteins.
3. The selection of compounds identified from the ZINC database included those of good, intermediate and poor predicted binding energies to determine if any correlation to their experimentally derived *in vitro* half maximal Inhibitory Concentrations ($IC_{50}$) values could be made.
4. To perform a similarity search on the NCI database to identify similar compounds to QNZ, based on their similarity scores and previously reported SAR data extracted from literature.
5. To develop a pharmacophore hypothesis based on previously reported SAR data identified from literature on quinazoline inhibitors of NF-κB and to perform a pharmacophore VS on the NCI database to identify similar compounds to QNZ.
6. To perform a substructure search on the NCI database to identify similar acting functional groups to the benzamidine moiety of Pentamidine.
7. To determine a binding hypothesis of Pentamidine in the AdoMet-DC enzyme, a key therapeutic target in the polyamine biosynthetic pathway.
8. To determine the *in vitro* biological activity of the compounds identified from the NCI and ZINC database in prostate, renal and ovarian cancer cell lines.
9. To confirm that the active, most potent compound identified from the NCI database acts similarly to its parental compound QNZ, by inhibiting the transcription factor NF-κB, *in vitro.*

## 1.5 References


1. E. Ravina, The Evolution of Drug Discovery, From Traditional Medicines to Modern Drugs, WILY-VCH Verlag & Co., ISBN: 978-3-527-32669-3, 2011.
2. S. Nobili, D. Lippi, E. Witort, M. Donnini, L. Bausi, E. Mini and S. Capaccioli, Natural Compounds For Cancer Treatment and Prevention, Pharmacological Research, Vol. No. 59, Pp. 365-378, 2009.
3. M. Stiefel, A. Shaner and S. D. Schaefer, The Edwin Smith Papyrus: The Birth of Analytical Thinking in Medicine and Otolaryngology, The Laryngoscope, Vol. No. 116, Issue No. 2, Pp. 182-188, 2006.
4. J. Drews, Drug Discovery: A Historical Perspective, Drug Discovery Review, Science Magazine, Vol. No. 287, Pp. 1960-1964, 2000.
5. V. T. DeVita Jr. and E. Chu, A history of Cancer Chemotherapy, Cancer Research, Vol. No. 68, Issue No. 21, Pp. 8643-8653, 2008.
6. B. C. Baguley, A Brief History of Anticancer Chemotherapy, Chapter 1 in Anticancer Drug Development (Eds: B. C. Baguley and D. J. Kerr), Academic Press, Pp. 111, 2002.
7. M. Arruebo, N. Vilaboa, B. Saez-Gutierrez, J. Lambea, A. Tres, M. Valladares and A. Gonzalez-Fernandez, Assesment of the Evolution of Cancer Treatment Therapies, Cancers, Vol. No. 3, Pp. 3279-3330, 2011.
8. E. Ratti and D. Trist, The Continuing Evolution of the Drug Discovery Process in the Pharmaceutical Industry, Il Farmaco, Vol. No. 56, Pp. 13-19, 2001.
9. M. Suggitt and M. C. Bibby, 50 Years of Preclinical Anticancer Drug Screening: Empirical to Target-Driven Approaches, Clinical Cancer Research, Vol. No. 11, Pp. 971-981, 2005.
10. J. N. Latosińska and M. Latosińska. Anticancer Drug Discovery--From Serendipity to Rational Design. INTECH Open Access Publisher, 2013.
11. J. P. Kennedy, L. Williams, T. M. Bridges, R. N. Daniels, D. Weaver and C. W. Lindsley, Application of Combinatorial Chemistry Science on Modern Drug Discovery, Journal of Combinatorial Chemistry, Vol. No. 10, Issue No. 3, Journal of Combinatorial Chemistry, Vol. No. 10, Issue No. 3, Pp. 345-354, 2008.
12. R. Goulding and E. Marden, Setting the Context: An Overview of Drug Discovery and Drug Dveleopment, OHI Deliverables, Vol. No. 1, 2009.

13. R. Borchardt, E. Kerns, M. Hageman, D. Thakker and J. Stevens, Optimising the 'Dug-Like' Properties of Leads in Drug Discovery, Biotechnology and Pharmaceutical Aspects, Volume 4, Springer Science & Business Media, 2007. ISBN: 0387449612.
14. A. S. Reddy, S. P. Pati, P. Kumar, H. N. Pradeep and G. N. Sastry, Virtual Screening in Drug Discovery – A Computational Perspective, Current Protein and Peptide Science, Vol. No. 8, Issue No. 4, Pp. 329-351, 2007.
15. S. A. Patil, Role of Medicinal Chemist in the Modern Drug Discovery and Development, Organic Chemistry Current Research: Editorial, Vol. No. 1, Issue No. 3, 2012.
16. J. G. Lombardino and J. A. Lowe III, The Role of the Medicinal Chemist in Drug Discovery – Then and Now, Nature Reviews Drug Discovery, Vol. No. 3, Pp. 853-862, 2004.
17. A. Wadood, N. Ahemd, L. Shah, A. Ahmad, H. Hassan and S. Shams, In-silico Drug Design: An Approach which Revolutionaised the Drug Discovery Process, OA Drug Design & Delivery, Vol. No. 1, Issue No. 3, 2013.
18. G. Sliwoski, S. Kothiwale, J. Meiler and E. W. Lowe, Jr., Computational Methods in Drug Discovery, Pharmacological Reviews, Vol. No. 66, Issue No. 1, Pp. 334-395, 2014.
19. S. C. Gupta, B. Sung, S. Prasad, L. J. Webb and B. B. Aggarwal, Cancer Drug Discovery by Repurposing: Teaching New Tricks to Old Dogs, Trends in Pharmacological Science, Vol. No. 34, Issue No. 9, Pp. 508-517, 2013.
20. J. S. Shim and J. O. Liu, Recent Advances in rug Repositioning for the Discovery of New Anticancer Drugs, International Journal of Biological Sciences, Vol. No. 10, Issue No. 7, Pp. 654-663, 2014.
21. J. Li, S. Zheng, B. Chen, A. J. Butte, S. J. Swamidass and Z. Lu, A Survey of Current Trends in Computational Drug Repositioning, Briefings in Bioinformatics, 2015.
22. A. P. McGrath, K. M. Hilmer, C. A. Collyer, E. M. Shepard, B. O. Elmore, D. E. Brown, D. M. Dooley and J. M. Guss, Structure and Inhibition of Human Diamine Oxidase, Biochemistry, Vol. No. 48, Issue No. 41, Pp. 9810-9822, 2009.
23. R. R. Tidwell, S. K. Jones, J. D. Geratz, K. A. Ohemeng, C. A. Bell, B. J. Berger and J. E. Hall, Development of Pentamidine Analogues as New Agents for the Treatment of Pneumocystic *carinii* Pneumonia, Annals of the New York Academy of Science, AIDS: Anti-HIV Agents, Therapies and Vaccines, Vol. No. 616, Pp. 421-441, 1990.

24. L. F. Zerbini, M. K. Bhasin, J. F. de Vasconcellos, J. D. Paccez, X. Gu, A. L. Kung and T. A. Libermann, Computational Repositioning and Preclinical Validation of Pentamidine for Renal Cell Cancer, Molecular Cancer Therapeutics, Vol. No. 13, Issue No. 7, Pp. 1929-1941, 2014.
25. L. M. Scott, H. R. Lawrence, S. M. Sebti, N. J. Lawrence and J. Wu, Tageting Protein Tyrosine Phosphotases for Anticancer Drug Discovery, Current Pharmaceutical Design, Vol. No. 16, Issue No. 16, Pp. 1843-1862, 2010.
26. M. K. Pathak, D. Dhawan, D. J. Lindner, E. C. Borden, C. Farver, Pentamidine is an Inhibitor of PRL Phosphatases with Anticancer Activity, Molecular Cancer Therapeutics, Vol. No. 1, Pp. 1255-1264, 2002.
27. M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, H. Takahashi, T. Fukazawa and H. Hayashi, Discovery of Quinazolines as a Novel Structural Class of Potent Inhibitors of NF-κB Activation, Bioorganic & Medicinal Chemistry, Vol. No. 11, Pp. 383-391, 2003.
28. M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, H. Takahashi and H. Hayashi, A Novel Stuctural Class of Potent Inhibitors of NF-κB Activation: Structure-Activity Relationships and Biological Effects of 6-Aminoquinazoline Derivatives, Bioorganic & Medicinal Chemistry, Vol. No. 11, Pp. 3869-3878, 2003.
29. B. Rayet and C. Gelinas, Aberrant rel/nfκb Genes and Activity in Human Cancers, Oncogene, Vol. No. 19, Pp. 6938, 6947, 1999.
30. M. Karin, NF-κB as a Critical Link Between Inflammation and Cancer, Cold Spring Harbour Perspective in Biology, Vol. No. 1, 2009.
31. B. Hoesel and J. A. Schmid, The Complexity of NF-κB Signalling in Inflammation and Cancer, Molecular Cancer, Vol. No. 12, Issue No. 86, 2013.
32. L. F. Zerbini, R. E. Tamura, R. G. Correa, A. Czibere, J. Cordeiro, M. Bhasin, F. M. Simabuco, Y. Wang, X. Gu, L. Li, D. Sarkar, J. Zhou, P. B. Fisher and T. A. Libermann, Conbinatorial Effects of Non-Steroidal Anti-inflammatory Drugs and NF-κB Inhibitors in Ovarian Vancer Therapy, PLoS one, Vol. No. 6, Issue No. 9, 2011.
33. S. Subramaniam, M. Mehrotra and D. Gupta, Virtual High Throughput Screening (vHTS) – A Perspective, Bioinformation, Vol. No. 3, Issue No. 1, Pp. 14-17, 2008
34. G. Klebe, Virtual Ligand Screening: Strategies, Perspectives and Limitations, Drug Discovery Today, Vol. No. 11, Issue No. 13/14, 580-594, 2006.

35. A. S. Reddy, S. P. Pati, P. Kumar, H. N. Pradeep and G. N. Sastry, Virtual Screening in Drug Discovery – A Computational Perspective, Current Protein and Peptide Science, Vol. No. 8, Issue No. 4, Pp. 329-351, 2007.
36. P. Willet, Similarity-Based Virtual Screening using 2D Fingerprints, Drug Discovery Today, Vol. No. 11, Issue No. 23/24, Pp. 1046-1053, 2006.
37. J. Bajorath, Integration of Virtual and High Throughput Screening, Nature Review Drug Discovery, Vol. No. 1, Issue No. 11, Pp. 882-894, 2002.
38. M. A. Johnson and G. M. Maggiora, Concepts and Applications of Molecular Similarity, John Wiley, 1990.
39. J. D. Holliday, C. Y. Hu and P. Willet, Grouping of Coefficients for the Calculation of Inter-Molecular Similarity and Dissimilarity Using 2D Fragment Bit-Strings, Combinatorial Chemistry and High Throughput Screening, Vol. No. 5, Issue No. 2, Pp. 155-166, 2002.
40. D. Rogers and M. Hahn, Extended-Connectivity Fingerprints, Journal of Chemical Information and Modelling, Vol. No. 50, Issue No. 5, Pp. 742-754, 2010.
41. J. Hert, P. Willet, D. J. Wilton, P. Acklin, K. Azzaoui, E. Jacoby and A. Schuffenhauer, Comparison of Topological Descriptors for Similarity-Based Virtual Screening using Multiple Bioactive Reference Structures, Organic & Biomolecular Chemistry, Vol. No. 2, Issue No. 22, Pp. 3256-3266, 2004.
42. Y. Hu, E. Lounkine and J. Bajorath, Improving the Search Performance of Extended Connectivity Fingerprints through Activity-Oriented Feature Filtering and Application of a Bit-Density-Dependent Similarity Function, ChemMedChem, Vol. No. 4, Issue No. 4, Pp. 540-548, 2009.
43. M. Awale and J. Reymond, A Multi-Fingerprint Browser for the ZINC Database, Nucleic Acid Research, 2014.
44. M. Syuib, S. Arif and N. Malim, Comparison of Similarity Coefficients for Chemical Database Retrieval, 2013 1st International Conference on Artificial Intelligence, Pp. 129-133, ISBN: 978-1-4799-3250-4, 2013.
45. H. Eckert and J. Bajorath, Molecular Similarity Analysis in Virtual Screening: Foundations, Limitations and Novel Approaches, Drug Discovery Today Review, Vol. No. 12, Issue No. 5/6, Pp. 225-233, 2007.
46. K. Dobi, I. Hajdu, B. Flachner, G. Fabo, M. Szaszko, M. Bogner, C. Magyar, I. Simon, D. Szisz, Z. Lorincz, S. Cseh and G. Dorman, Combination of 2D/3D Ligand-Based Similarity Search in Rapid Virtual Screening from Multimillion Compound

Repositories. Selection and Biological Evaluation of Potential PDE4 and PDE5 Inhibitors, Molecules, Vol. No. 19, Issue No. 6, Pp. 7008-39, 2014.

47. V. Venkatraman, V. I. Perez-Nueno, L. Mavridis and D. W. Ritchie, Comprehensive Comparison of Ligand-Based Virtual Screening Tools against the DUD Data set Reveals Limitations of Current 3D Methods, Journal of Chemical Information and Modelling, Vol. No. 50, Issue No. 12, Pp. 2079-2093, 2010.
48. M. Thimm, A. Goede, S. Hougardy and R. Preibner, Comparison of 2D and 3D Superposition. Application to Searching a Conformational Drug Database, Journal of Chemical Information and Computer Science, Vol. No. 44, Pp. 1816-1822, 2004.
49. A. Hamza, N. N. Wei and C. G. Zhan, Ligand-based Virtual Screening Approaches Using a New Scoring Function, Journal of Chemical Information and Modelling, Vol. No. 52, Issue No. 4, Pp. 963-974, 2012.
50. S. M. Kearnes, I. S. Haque and V. S. Pande, SCISSORS: Practical Considerations, Journal of Chemical Information and Modelling, Vol. No. 54, Issue No. 1, Pp. 5-15, 2014.
51. P. C. D. Hawkins, A. G. Skillman and A. Nicholls, Comparison of Shape-Matching and Docking as Virtual Screening Tools, Journal of Medicinal Chemistry, Vol. No. 50, Issue No. 1, Pp. 74-82, 2007.
52. C. Sotriffer, Virtual screening: Principles, challenges and practical guidelines, John Wiley & Sons, 2011.
53. V. Vyas, A. Jain, A. Jain, A. Gupta, Virtual Screening: A Fast Tool for Drug Design, Scientia Pharmaceutica, Vol. No. 76, Pp. 333-360, 2008.
54. O. Dror, D. S. Duhovny, Y. Inbar, R. Nussinov and H. J. Wolfson, A Novel Appraoch for Efficient Pharmacophore-based Virtual Screening: Methods and Applications, Journal of Chemical Information and Modelling, Vol. No. 49, Issue No. 10, Pp. 2333-2343, 2009.
55. C. Hetenyi and D. Van Der Spoel, Efficient docking of peptides to proteins without prior knowledge of the binding site, Protein Science, Vol. 11, Pg. No. 1729-1737, 2002.
56. D. Ghersi and R. Sanchez, Improving accuracy and efficacy of blind protein-ligand docking by focusing on predicted binding sites, Proteins, Vol. No. 74, Pp. 417-424, 2009.
57. I. Wandzik, MATCH: Communication In Mathematical and Computer Chemistry, Vol. No. 55, Pp. 271-278, 2006.

58. J. B. Cross, D. C. Thompson, B. K. Rai, J. C. Baber, K. Y. Fan, Y. Hu and C. Humblet, Comparison of Several Moelcular Docking Programs: Pose Prediction and Virtual Screening Acuuracy, Journal of Chemical Information and Modelling, Vol. No. 49, Vol. No. 6, Pp. 1455-1474, 2009.
59. E. Kellenberger, J. Rodrigo, P. Muller and D. Rogan, Comparative Evaluation of Eight Docking Tools for Docking and Virtual Screening Accuracy, PROTEINS: Structure, Function and Bioinformatics, Vol. No. 57, Pp. 225-242, 2004.
60. E. Perola, W. P. Walters and P. S. Charifson, A Detailed Comparison of Current Docking and Scoring Methods on Systems of Pharmaceutical Relevance, PROTEINS: Structure, Function and Bioinformatics, Vol. No. 56, Pp. 235-249, 2004.
61. R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, Glide: A New Approach for Rapid, Accurate Docking and Scoring 1. Method and Assessment of Docking Accuracy, Journal of Medicinal Chemistry, Vol. No. 47, Issue No. 7, Pp. 1739-1749, 2004.
62. R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes, Journal of Medicinal Chemistry, Vol. No. 49, Issue No. 21, Pp. 6177-6169, 2006.
63. T. Usha, S. K. Middha, A. K. Goyal, M. Karthik, D. A. Manoj, S. Faizan, P. Goyal, H. P. Prashanth and V. Pande, Molecular Docking Studies of Anti-Cancerous Candidates in *Hippophae rhamnoides* and *Hippophae salicifolia,* Journal of Biomedical Research, Vol. No. 28, Issue No. 5, Pp. 406-415, 2014.
64. B. Banerji, S. K. Pramanik, U. Pal and N. C. Maiti, Potent Anticancer Activity of Cysteine-Based Dipeptides and their Interaction with Serum Albumins, Chemistry Central Journal, Vol. No. 7, Issue No. 91, 2013.
65. S. Sreenivasan, M. Alameen, S. Krishnakumar and U. Vetrivel, Computaional Docking Studies Reveal the Most Potential Inhibitory Binding Mode of Demethoxylated and Reduced Curcumin Compounds on Lung Resistance Protein, International Journal of Pharmacy and Pharmaceutical Sciences, Vol. No. 6, Issue No. 5, Pp. 234-238, 2014.
66. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, AutoDock4 and AutoDockTools4: Automated Docking with Selective

Receptor Flexibility, Journal of Computational Chemistry, Vol. No. 30, Issue No. 16, Pp. 2785-2791, 2009.

67. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olsen, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, Journal of Computational Chemistry Vol. No. 19, Issue No. 14, Pp. 1639-1662, 1998.
68. D. S. Goodsell, G. M. Morris and A. J. Olsen, Automated Docking of Felxible Ligands: Applications of AutoDock, Journal of Molecular Recognition, Vol. No. 9, Issue No. 1, Pp. 1-5, 1996.
69. R. Huey, G. M. Morris, A. J. Olsen and D. S. Goodsell, A Semiempirical Free Energy Force Field with Charge-Based Desolvation, Journal of Computational Chemistry, Vol. No. 28, Issue No. 6, Pp. 1145-52, 2007.
70. A. S. Verkman, Drug Discovery in Academia, American Journal of Physiology Cell Physiology, Vol. No. 286, Pg. No. C465-C474, 2004.
71. J. Major, Challenges and opportunities in high throughput screening: implications for new technologies, Journal of Biomolecular Screening, Vol. 3, Issue No. 1, Pg. No. 13-17, 1998.
72. F. Sams-Dodd, Target-based Drug Discovery: Is Something Wrong? Drug Discovery Today Targets, Vol. No. 10, Issue No. 2, 2005.
73. J. P. Hughes, S. Rees, S. B. Kalindjian and K. L. Phillpott, Principles of Early Drug Discovery, British Journal of Pharmacology, Vol. No. 162, Issue No. 6, Pp. 1239-1249, 2011.
74. A. M. Burger and H. H. Fiebig, Preclinical Screening for New Anticancer Agents, 2014. In: M. A. Rudek et al, Handbook of Anticancer Pharmacokinetics and Pharmacodynamics, Cancer Drug Discovery and Development, Sprnger Science and Business Media, 2014.
75. V. Khurana, D. F. Tardiff, C. Y. Chung and S. Lindquist, Towards Stem Cell-Based Phenotypic Screens for Neurodegenrative Diseases, Nature Reviews Neurology, Vol. No. 11, Pp. 339-350, 2015.
76. L. F. Zerbini, M. K. Bhasin, J. F. de Vasconcellos, J. D. Paccez, X. Gu, A. L. Kung and T. A. Libermann, Computational Repositioning and Preclinical Validation of Pentamidine for Renal Cell Cancer, Molecular Cancer Therapeutics, Vol. No. 13, Issue No. 7, Pp. 1929-1941, 2014.

77. L. M. Scott, H. R. Lawrence, S. M. Sebti, N. J. Lawrence and J. Wu, Tageting Protein Tyrosine Phosphotases for Anticancer Drug Discovery, Current Pharmaceutical Design, Vol. No. 16, Issue No. 16, Pp. 1843-1862, 2010.
78. M. K. Pathak, D. Dhawan, D. J. Lindner, E. C. Borden, C. Farver, Pentamidine is an Inhibitor of PRL Phosphatases with Anticancer Activity, Molecular Cancer Therapeutics, Vol. No. 1, Pp. 1255-1264, 2002.
79. A. Puckowska, D. Drozdowska, M. Rusak, T. Bielawski, I. Bruzgo and K. M. Nowaczek, Amino and Chlorambucil Analogues of Pentamidine – Synthesis and Biological Examination, Acta Poloniae Pharmaceutica- Drug Research, Vol. No. 69, Issue No. 1, Pp. 63-73, 2012.
80. J. Stanek, G. Caravatti, H. G. Capraro, P. Furet, H. Mett, P. Schneider and U. Regenass, S-Adenosylmethionine Decarboxylase Inhibitors: New Aryl and Heteroaryl Analogues of Methylglyoxal Bis (guanylhydrazone), Journal of Medicinal Chemistry, Vol. No. 36, Issue No. 1, Pp. 46-54, 1993.
81. U. Regenass, G. Caravatti, H. Mett, J. Stanek, P. Schneider, M. Muller, A. Matter, P. Vertino and C. W. Porter, New S-Adenosylmethionine Decarboxylase Inhibitors with Potent Antitumor Activity, Cancer Research, Vol. No. 52, Pp. 4712-4718, 1992.
82. D. E. McClowsky, S. Bale, J. A. Secrist, III, A. Tiwari, T. H. Moss, III, J. Valiyaveettil, W. H. Brooks, W. C. Guida, A. E. Pegg and S. E. Ealick, New Insights into the Design of Inhibitors of Human S-Adenosylmethionine Decarboxylase: Studies of Adenine $C^8$ Substitution in Structural Analogues of S-Adenosylmethionine, Journal of Medicinal Chemistry, Vol. No. 52, Issue No. 5, Pp. 1388-1407, 2009.
83. E. C. Butcher, E. L. Berg and E. J. Kunkel, Systems biology in drug discovery, Nature Biology Perspective, Vol. No. 22, Issue No. 10, 2004.
84. E. A. Martis, R. Radhakrishnan, R. R. Badve, High Through-put Screening: The Hits and Leads of Drug Discovery- An Overview, Journal of Applied Pharmaceutical Science, Vol. No. 01, Issue. 01, Pg. 02-10, 2011.
85. R. Zang, D. Li, I. Tang, J. Wang and S. Yang, Cell-Based Assays in High-Throughput Screening for Drug Discovery, International Journal of Biotechnology for Wellness Industries, Vol. No. 1, Issue No. 1, Pg. No. 31-51, 2012.
86. W. Aherne, M. Garrett, T. McDonald and P. Workman, Mechanism-based high-throughput screening for novel anticancer drug discovery, Anticancer Drug Development, Academic Press, Pg. No. 249-267, 2002.

87. P. A. Johnston and P. A. Johnston, Cellular Platforms for HTS: Three Case Studies, Drug Discovery Today, Vol. No. 7, Issue No. 6, Pp. 353-363, 2002.
88. V. Baud and M. Karin, Is NF-κB a Good Target for Cancer Therapy? Hopes and Pitfalls, Nature Reviews Drug Discovery, Vol. No. 8, Issue No. 1, Pp. 33-40, 2009.
89. M. F. Shannon, S. R. Himes, Assays for Transcriptional Activity Based on the Luciferase Reporter Gene, Methods in Molecular Biology, Vol. No. 130, Pp. 165-174, 2000.
90. L. F. Zerbini, Y. Wang, J. Cho and T. A. Libermann, Constitutive Activation of Nuclear Factor κB p50/p65 and Fra-1 and JunD is Essential for Deregulated Interleukin 6 Expression in Prostate Cancer, Cancer Research, Vol. No. 63, Pp. 2206-2215, 2003.

# Chapter 2: Virtual Screening of the ZINC Database

## 2.1 Introduction

The virtual screening techniques are primarily performed as a cost effective alternative to HTS (S. Subramaniam et al, 2008). Taking this into consideration, a combination study of both ligand and structure based virtual screening techniques were performed in this study, to screen the ZINC database for potential anti-cancer drugs. Combination studies are said to have the advantage of taking all available information into account, i.e. active ligands and structural data of the targets making them more robust despite their predictive element, and thus increase the hit rate (V. Kumar et al, 2014; M. N. Drwal et al, 2013). The aims of this study to perform the knowledge driven hierarchical virtual screening of the ZINC database are summarised in Figure 1.

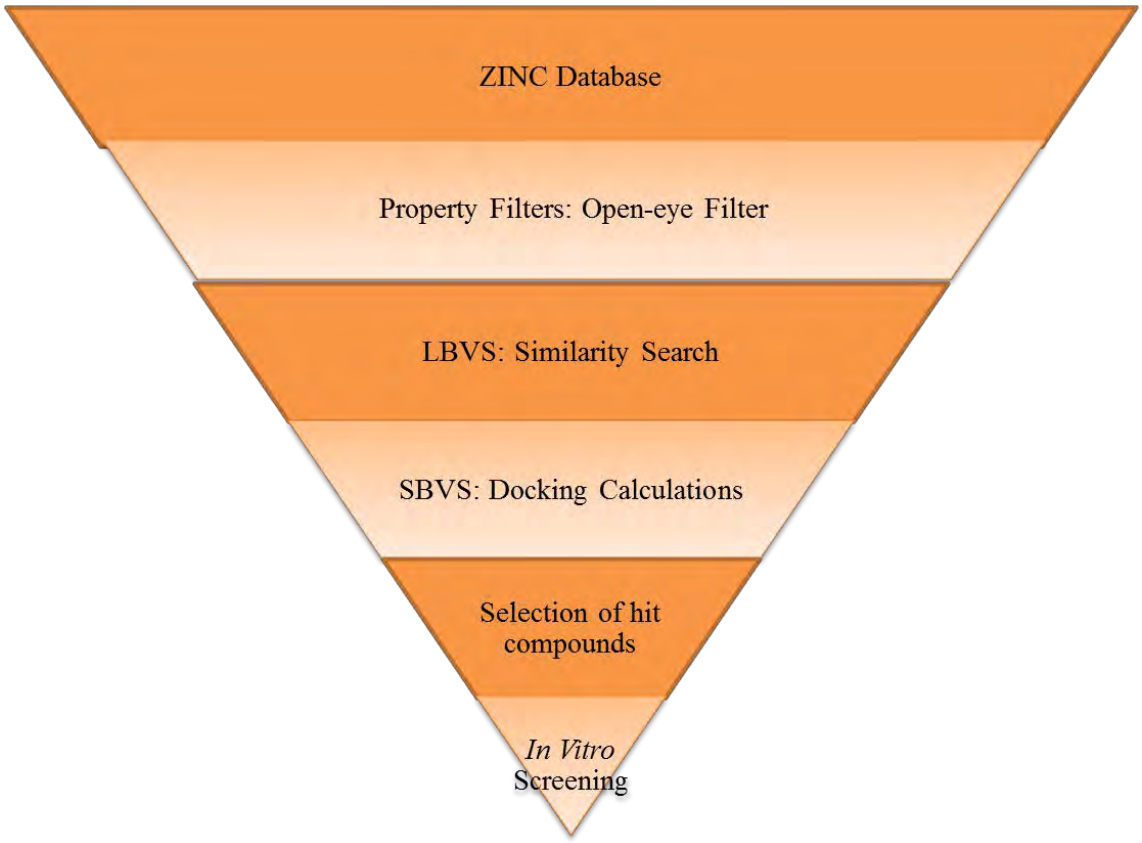



**Figure 1: The aims of the combined hierarchical ligand-based and structure-based virtual screening performed on the ZINC database.**

The success of a virtual screening study not only depends upon the quality and content of the structural and chemical information of a protein target and ligand, but also on the quality and content of a screening database (E. Beilska et al, 2011). A common strategy used to enrich a screening database with drug-like and desirable compounds is to apply a drug-like filter before a virtual screening study is embarked upon (E. Beilska et al, 2011). This strategy also reduces the millions of compounds contained in a database to a computationally manageable number that can be screened in a reasonable amount of time. Furthermore, it reduces the burden of selecting compounds with desirable properties in the final stages of the virtual screening (T. Cheng et al, 2012).

The ECFP fingerprint is reported to be a sufficient tool to use in a 2D similarity search, where it records the absence or presence of molecular features predicted to be responsible for molecular or drug activity and was therefore used in this study (D. Rogers et al, 2010; M. Awale et al, 2014). As mentioned previously, the ECFP_4 fingerprint describes both the identity and connectivity of an atom within a diameter of 4 bonds radiating from each heavy atom within a molecule (S. R. Langdon et al, 2013; D. Rogers et al, 2010; Y. Hu et al, 2009), where similar fingerprints are searched for in a database.

The 3D similarity search utilising 3D descriptors is considered more advanced, describing the spatial arrangement of atoms in a molecule with the aim of characterizing the molecular surfaces and volumes of compounds. The application of 3D descriptors is critically dependent on the accurate generation of the conformation of the compounds, which is a computationally exhausting task and remains to be the bottleneck associated with the 3D similarity search (H. Eckert et al, 2007). This adds to the debate for the preference of 2D descriptor over 3D shape-based descriptors (V. Venkatraman et al, 2010; M. Thimm et al, 2004; J. Bajorath, 2002). In view of this, the aim was to make the screening protocol more robust by including both a 2D and 3D similarity search.

The final step of the hierarchical study included the SBVS which consisted of the docking calculations. The crystal structure of Pentamidine found in complex with the DAO enzyme (A. P. McGrath et al, 2009), met the pre-requisites to perform the docking calculations, as the protein structure and binding site are well defined (C. Hetenyi et al, 2002). This allowed for a more straight-forward methodology which included a comparison of the docking pose in the crystal structure to that obtained from the docking calculations, to calculate the root mean

square deviation (RMSD) between the two poses (Z. Li et al, 2013). In the case of QNZ, a different strategy to identify the binding site was considered.

There is no crystal structure of the NF-κB proteins in complex with an inhibitor reported in literature. Therefore, the energetically most favourable binding site of QNZ was sought amongst the NF-κB pathway proteins, using a technique known as blind docking. Blind docking has been adopted and tested in identifying binding sites of small molecules, which gave the confidence to apply the technique in this study (C. Hetenyi et al, 2006; D. Ghersi et al, 2009). A cluster analysis on AutodockTool4 was applied to determine the most populated sites identified through the blind docking, which were further investigated by performing focused docking calculations on Glide5.7. Control compounds, which include the NF-κB inhibitors JSH-23 and BAY11-7085, were included in the blind docking calculations since their mechanisms of action have been experimentally elucidated to some extent and reported on in literature (A. Kumar et al, 2011; J. W. Pierce et al, 1997). The predicted binding sites identified in this study for the control compounds provided a reasonable correlation to the experimental data reported in literature, which provided the confidence in the binding sites identified for QNZ.

In this chapter, the methodology used to perform the 2D and 3D similarity search is described. The top ranked compounds, which have the highest similarity to QNZ and Pentamidine, were further enriched by performing the docking calculations described in this chapter. Good, intermediate and poor scoring compounds, in terms of their predicted binding energies, were selected for *in vitro* testing, with the goal of determining any correlation between their predicted binding energies and *in vitro* $IC_{50}$ values.

## 2.2 Methodology of the vHTS performed with QNZ

### 2.2.1 Preparation of ZINC Database

Structure information for the 14,322,885 compounds was previously downloaded from the ZINC website (http://zinc.docking.org/browse/subsets/). In order to remove compounds with known toxic functional groups or toxicophores and undesirable compounds such as dyes and non-drug like compounds, the ZINC database was filtered with an OpenEye Scientific

Software drug-like Filter (Filter, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2010) (A. Lavecchia et al, 2013). Default parameters were kept except for the following changes primarily made to the Lipinski rules (T. I. Oprea, 2002; C. Abad-Zapatero, 2007), which are summarised in Table 1 below.

**Table 1: The changes made to the OpenEye Drug-like Filter (Filter, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2010).**

| Parameter | Value |
|---|---|
| Min. molecular weight | 100 |
| Max. molecular weight | 500 |
| Min. H-bond donors | 0 |
| Max. H-bond donors | 6 |
| Min. H-bond acceptors | 0 |
| Max. H-bond acceptors | 5 |
| Min. xlogP | -5.0 |
| Max. xlogP | 5.0 |
| Min. 2D-PSA | 75.0 |
| Max. 2D-PSA | 140.0 |

These changes were made to take into consideration the drug-like properties of both parental compounds, QNZ and Pentamidine and to acquire a drug-like library within a reasonable and appropriate chemical space (A. Lavecchia et al, 2013). A simple script was written to efficiently filter the database. The complete Filter can be found in Appendix 1, Table 1.

### 2.2.2. Generation of Ligand and Database Conformations

In preparation for the 3D similarity search and the SBVS, conformations of the parental and ZINC database compounds were generated using the Confgen, an application on the Schrodinger suite (ConfGen, version 5.6, Schrödinger, Inc., New York, NY, 2009). The standard protocol on ConfGen was selected and default parameters were maintained. The fast search protocol was utilised, where conformers are considered redundant at a threshold of an RMSD of 1Å (K. S. Watts et al, 2010). The input and output structures were minimised.

### 2.2.3 2D Similarity Search

The Discovery studios 'Find similar molecules by fingerprints' protocol on Pipeline Pilot was used to conduct the 2D similarity search (Accelrys Software Inc., *Discovery Studio Modelling Environment, 4.5, San Diego: Accelrys Software Inc., 2007*). The input ligands included the filtered ZINC database and the reference ligand included QNZ. The minimum similarity was set to 0.50 and the Tanimoto coefficient was selected with a pre-defined property of the ECFP_4 fingerprint. The top 2000 compounds that scored a minimum similarity of 0.50 were retained for the SBVS that followed.

### 2.2.4 3D Similarity Search

The simple run on the ROCS interface (vROCS) was employed to perform the 3D similarity search (ROCS, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2010). Default parameters were maintained for the inputs, where the colour force field was kept at the default Implicit Mills-Dean. The Structure Data File (sdf) for QNZ (query) and the 3D conformations of the filtered ZINC database were uploaded for screening. The number of best hits to retain was increased to 1000 compounds.

### 2.2.5 Blind Docking Calculations

The available crystal structures of the NF-κB pathway proteins selected from the Protein Database Bank (PDB) website include the NF-κB (p65/p50) complex bound to DNA (1VKX.pdb) (F. E. Chen et al, 1998), the NF-κB/Inhibitory-κBα protein (IκBα) complex (1NFI.pdb) (M. D. Jacobs et al, 1998), and the NF-κB Essential Modulator (NEMO) subunit of the IκB Kinase (IKK) complex (3BRT.pdb) (M. Rushe et al, 2008). Taking into consideration the unavailability of suitable mammalian derived crystal structures for the other NF-κB subunits namely, RelB, c-Rel, p105 and p52/p100, they were excluded from the blind docking. Furthermore, it is reported in literature that the most prevalent NF-κB heterodimer found in cancer cells is the (RelA) p65/p50 dimer, which gave the confidence to proceed

using the selected structures (B. Rayet et al, 1999; A. Oeckinghaus et al, 2011; R. G. Uzzo et al, 2006).

AutoDock4 and AutoDock4Tools (G .M. Morris et al, 1998; G. M. Morris et al, 2009) were used to scan the entire extracellular surface of the selected NF-κB pathway protein structures with the ligands QNZ, JSH-23 and BAY11-7085, to identify their suitable binding sites. The water molecules were removed from the PDB structure files and hydrogen atoms were added using AutoDock4Tools. Default parameters were kept in processing the protein structures. Ligand preparation was performed keeping default parameters, where the numbers of rotatable bonds were identified for flexible docking. The Autogrid4 was generated to include the whole protein complex and the grid parameters used between each successive trail were kept constant. A grid spacing of approximately 0.4 Å was maintained, ensuring that the search methodology scans the entire extracellular surface for suitable binding sites.

To identify the correct search parameters using the Genetic Algorithm (GA) on AutoDock4, C. Hentenyi et al, 2002 determined that the number of evaluations and the population size has an influence on the probability of finding the correct binding site and mode. It is recommended that maintaining the energy evaluations above 10 million and the population size above 50, are ideal parameters to identify the correct binding sites on macromolecules (C. Hetenyi et al, 2002). Taking this into consideration, successive calculations were performed adjusting the number of evaluations from 5 million to 10 million, to determine which binding sites are still being identified at a 10 million number of evaluations. The population size was retained at 150 and the remaining parameters were kept at default. A cluster analysis on AutoDockTools4 was performed to identify the most populated binding sites for QNZ, JSH-23 and BAY11-7085, which were further investigated using Glide5.7.

### 2.2.6 Focused Docking Calculations

The identification of 2 favourable binding sites for QNZ on the NF-κB/IκBα complex, allowed for the focused docking calculations which were performed on Glide5.7 (Glide, version 5.7, Schrödinger, Inc., New York, NY, 2009). There are several processing steps to perform a docking study on Glide5.7, which include protein and ligand preparation, receptor grid generation and finally the ligand docking. The conformations of QNZ and the ZINC database were further processed using the LigPrep application on Maestro (Maestro, version

5.6, Schrödinger, Inc., New York, NY, 2009). Their isomers and tautomers were generated using default parameters. The ionisation states were generated at a default pH of 7 ± 2 and the partial atomic charges were assigned using the OPLS_2005 force field. The 1NFI.pdb structure that was identified as the target for QNZ through the blind docking was prepared using the Protein Preparation Wizard on Maestro (Maestro, version 5.6, Schrödinger, Inc., New York, NY, 2009). A pre-processing step using default parameters preceded the hydrogen bond assignment using PROPKA. A restrained minimisation step was performed using the OPLS_2005 force field. The residues included in the 2 binding sites identified on the NF-κB/IκBα complex were specified to generate the grid file, using the Receptor Grid Generation application on Glide (Binding site 1: GLN29/p65, GLN220/p65, LYS221/p65; Binding site 2: PHE142/IκBα and TYR351/p50). Default parameters were maintained on the Glide SP docking protocol which was used to carry out the docking calculations.

## 2.3 QNZ Results

### 2.3.1 QNZ Similarity Search

The Tanimoto Coefficient was evaluated to measure similarity. The TanimotoCombo score (maximum score of 2) generated by vROCS is the sum of the individual ShapeTanimoto (volume) and ColourTanimoto (functional group similarity) scores (C. Sotriffer, 2011). The Tanimoto score generated on Pipeline Pilot is based on the 2D molecular descriptor, the ECFP_4 fingerprint and may have a maximum similarity score of 1 (Y. Hu et al, 2009), which is described in more detail in Chapter 1, section 1.2.1.

The TanimotoCombo scores obtained from vROCS were evaluated and over 1000 compounds that were 50% or more similar to QNZ were retrieved. The top-ranking compound retrieved by both vROCS and Pipeline Pilot was ZINC32089624 (Figure 2) with a TanimotoCombo and Tanimoto score respectively of 1.46 and 0.71. Furthermore, it was the only compound identified by vROCS that possessed a quinazoline ring. The compound ZINC32089624 (Figure 2) was previously reported in literature as part of a quinazoline series of NF-κB inhibitors, that had *in vitro* antitumor activity with an $IC_{50}$ value of 12nM in human Jurkat cells (M. Tobe et al, 2003). A large majority of the remaining compounds possessed a 5- and 6- membered ring system with varying lengths of alkyl chains and side chain

substituents attached to them. The top 20 hits identified from the 3D vROCS similarity search are included in Appendix 1, Table 2.

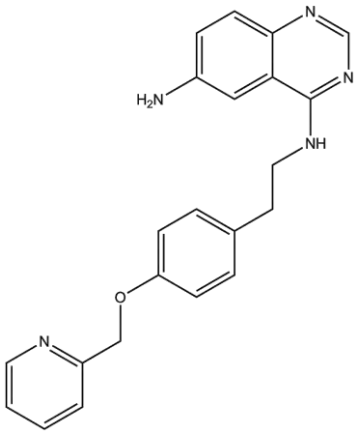



**Figure 2: The compound ZINC32089624 was identified as the top ranking compound on vROCS and Pipeline pilot, with a similarity score of 1.46 and 0.71 respectively.**

The compounds identified on Pipeline Pilot using the ECFP_4 search methodology resulted in a fewer number of compounds (63) with a 50% or greater similarity to QNZ. The 2nd top ranking compound retrieved on Pipeline Pilot was ZINC32089602 (Figure 3), with a Tanimoto score of 0.68, was also identified to be part of the series of quinazoline inhibitors of NF-κB reported in literature (M. Tobe et al, 2003). ZINC32089624 is more potent than ZINC32089602, which had a higher $IC_{50}$ value of 8271nM in human jurkat cells (M. Tobe et al, 2003).

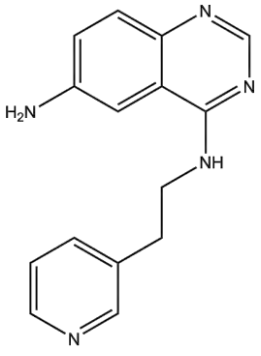



**Figure 3: The compound ZINC32089602 was identified on Pipeline Pilot and ranked the 2nd most similar compound to QNZ with a Tanimoto score of 0.68.**

The top 20 compounds identified from the 2D Pipeline Pilot similarity search are shown in Appendix 1, Table 3.

### 2.3.2 QNZ SBVS

#### 2.3.2.1 Blind Docking Calculations

The binding site of QNZ had to first be identified in order to perform the SBVS. Blind docking was performed on the different crystal structures of the NF-κB pathway proteins, with each of the inhibitors identified including QNZ, JSH-23 and BAY11-7085. A cluster analysis on AutoDockTools4 allowed for the identification of clusters that were highly populated, for each inhibitor. The binding scores obtained at each populated binding site were further ranked after the focused docking calculations performed on Glide5.7 (E. Kellenberger et al, 2004). The binding energies obtained from the blind docking performed with each compound on each structure are tabulated in Table 2 below.

**Table 2: A table of the binding energies obtained from the blinding docking calculation performed for each compound on the different NF-κB crystal structures. A cluster analysis was performed on AutoDockTools4 to determine the highly populated clusters at the 10 million energy evaluation. The binding energies obtained at each cluster from the focused docking calculations performed using Glide5.7 are shown in red. The cluster that obtained the lowest binding energy identified on Glide5.7 for each ligand was selected as the final binding site (bold typeface). LBE = Lowest Binding Energy (kcal/mol).**

| **Structure** | **Compounds** | | |
|---|---|---|---|
| **1NFI.pdb** | **QNZ (kcal/mol)** | **JSH-23 (kcal/mol)** | **BAY11-7085 (kcal/mol)** |
| **100runs/5 million evaluations** | LBE= -4.25 | LBE= -4.63 | LBE= -5.88 |
| **100runs/10 million evaluations** | LBE= -5.64/-7.070 | LBE= -6.52/-7.568 | LBE= -7.55/-5.150 |
| **Cluster analysis at 100runs/10 million evaluations** | **Cluster 1: -5.64/-7.070**<br>Cluster5: -4.94/-6.342<br>Cluster16: -3.76/-5.972 | **Cluster1:-6.52/-7.568**<br>Cluster4:-5.28/-4.901 | Cluster1:-7.55/-5.155<br>Cluster20:-5.63/-3.182 |
| | | | |
| **1VKX.pdb-p50** | | | |
| **100runs/5 million evaluations** | LBE= -5.12 | LBE= -5.22 | LBE= -5.56 |
| **100runs/10 million evaluations** | LBE= -5.97/-6.663 | LBE= -4.17/-3.967 | LBE= -3.72/-5.102 |
| **Cluster analysis at 100runs/10 million evaluations** | Cluster 1: -5.97/-6.663<br>Cluster25:-5.33/-6.844 | Cluster1:-4.17/-3.967<br>Cluster2:-5.78/-7.382 | Cluster1:-3.72/-5.102<br>Cluster2:-5.62/-3.446 |
| | | | |
| **1VKX.pdb-p65** | | | |
| **100runs/5 million evaluations** | LBE= -4.48 | LBE= -5.14 | LBE= -5.35 |
| **100runs/10 million evaluations** | LBE= -4.60/-5.582 | LBE= -4.74/-6.622 | LBE= -4.88/-3.622 |
| **Cluster analysis at 100runs/10 million evaluations** | Cluster 1: -4.60/-5.582<br>Cluster4:-5.14/-5.663<br>Cluster10:-4.92/-5.591 | Cluster1:-4.74/-6.622<br>Cluster4:-3.57/-4.044 | Cluster1:-4.88/-3.622<br>Cluster2:-5.88/-4.050 |
| | | | |
| **3BRT.pdb** | | | |
| **100runs/5 million evaluations** | LBE= -5.19 | LBE= -5.33 | LBE= -5.4 |

| | | | |
|---|---|---|---|
| **100runs/10 million evaluations** | LBE= -5.07/-4.994 | LBE= -5.89/-7.238 | LBE= -5.89/-5.218 |
| **Cluster analysis at 100runs/10 million evaluations** | Cluster 1: -5.07/-4.994<br>Cluster2:-5.33/-5.340 | Cluster1:-5.89/-7.238<br>Cluster16:-5.2/-5.168 | **Cluster1:-5.89/-5.218**<br>Cluster2:-5.46/-4.943<br>Cluster5:-5.33/-3.919 |

The refined binding energies obtained from the focused docking calculations performed on Glide5.7 (shown in red in Table 2) were evaluated to identify the final binding site for each compound. The binding site for JSH-23 was predicted to be between the p65 subunit and IκBα subunit, shown in Figure 4 below. It is in close proximity to the p65 Nuclear Localisation Signal (NLS) region which directs the nuclear translocation of the p65 subunit. JSH-23 is reported in literature to inhibit the nuclear translocation of the p65 subunit without inhibiting the degradation of IκBα (H. Shin et al, 2005; A. Kumar et al, 2011; D. Kesanakurti et al, 2013), which would suggest that it either binds directly to the p65 subunit or to the NF-κB/IκBα complex. This hypothesis would be plausible considering the binding site identified for JSH-23 in this study. On the other hand, despite reasonably poor binding energies obtained from the blind docking for BAY11-7085, its binding site was identified on the NEMO subunit of the IKK complex (Figure 4). The relatively high binding energies of BAY11-7085, compared to the other binding energies obtained for the other compounds may suggest that an alternative binding site may exist. The reported mechanism of action of BAY11-7085 indicates that it may bind to the IKK complex which inhibits the phosphorylation of IκBα without affecting the constitutive phosphorylation of IκBα (J. W. Pierce et al, 1997; C. L. Scaife et al, 2002; R. Tiwary et al, 2010) and despite correlating with our findings, it is still possible that BAY11-7085 may bind to either the IKKα or IKKβ subunits which were not included in the blind docking calculations performed (Figure 4).

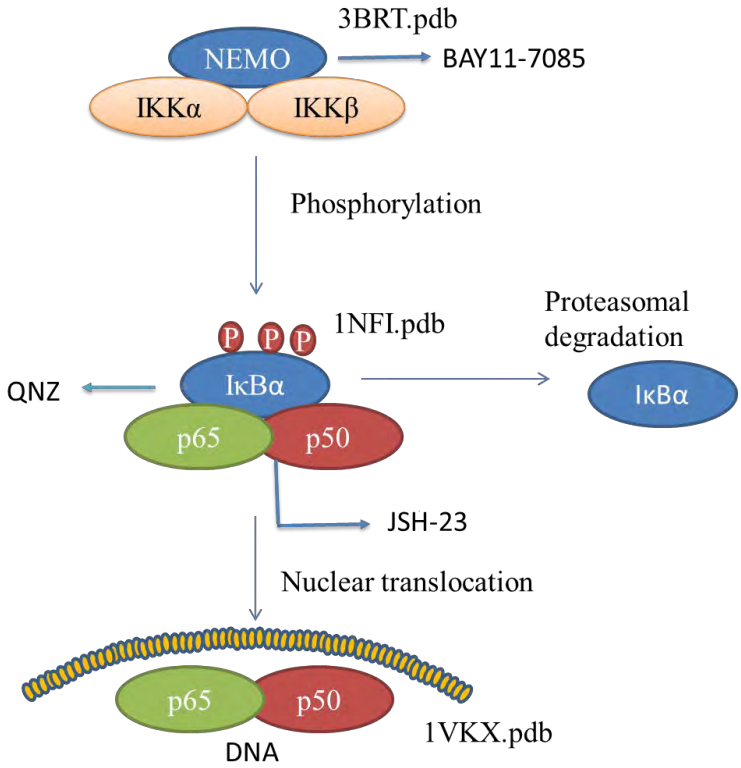



**Figure 4: A simplified representation of the NF-κB signalling pathway. Briefly, NF-κB activation is achieved by the phosphorylation of the IκBα subunit by the IKK complex. The IκBα subunit is subsequently ubiquitinated for proteasomal degradation. This allows for the nuclear translocation of the NF-κB (p65/p50) complex, where it binds to its target genes. The diagram illustrates the complexes on which the lowest energy binding sites for QNZ, JSH-23 and BAY11-7085 were identified. The protein data bank numbers of the crystal structures, on which the blind docking was performed, are also indicated on the figure.**

The predicted binding site identified in this study for QNZ, binding site 1 (ΔG = -7.070 kcal/mol), is in a hydrophobic pocket formed by the IκBα PEST region and p65. The p65 phosphorylation site at Serine (Ser) 276 is in close proximity to the binding site. The predicted binding mode of QNZ, suggests that the lipophilic end, the phenoxy phenyl region, is buried in the hydrophobic pocket formed by IκBα PEST region with the electron dense quinazoline core forming hydrogen bonds with the polar side chains of the Glutamine (Gln) 29 and Glutamine (Gln) 220 residues on the p65 subunit (Figure 5). The sidechain hydroxyl group of Gln29 forms a hydrogen bond with the amine group on the alkyl chain of QNZ. The polar Lysine (Lys) 221 residue on the p65 subunit was predicted to form π-cation interactions with the phenyl aromatic ring. The hydrophobic portion of QNZ is nestled in the hydrophobic pocket whereas the quinazoline core is solvent exposed and lies in the hydrophilic region.

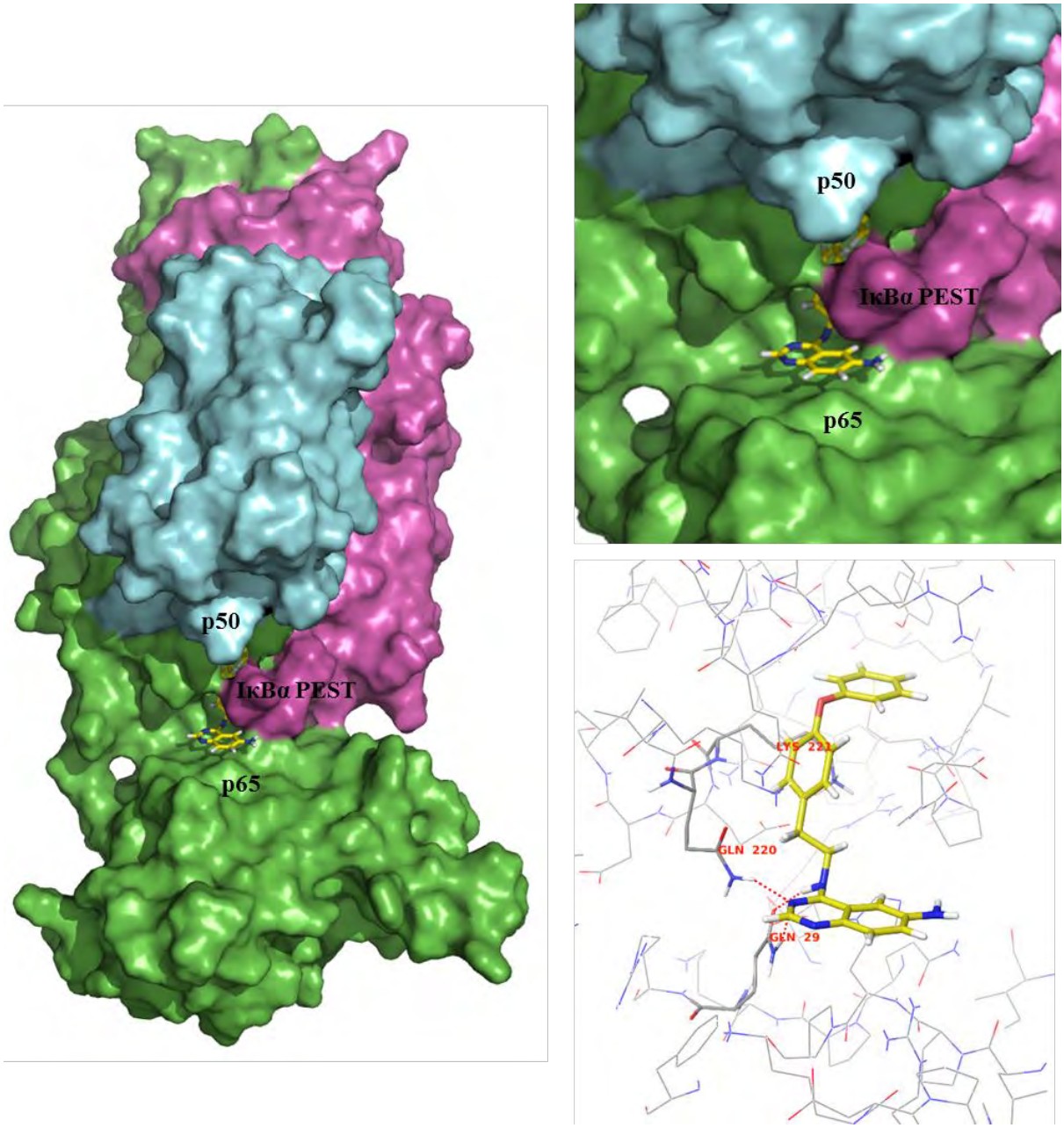



**Figure 5: The binding site 1 identified for QNZ on the NF-κB/IκBα crystal structure (1NFI.pdb) (left). The binding site is formed between the PEST domain of IκBα and the p65 subunit. The hydrophobic portion of QNZ is buried in the pocket formed by the PEST domain of IκBα and the quinazoline core is exposed to the solvent (top right). The predicted binding interactions at site 1 include hydrogen bonds between the side chains of Gln29 and Gln220 residues on the p65 subunit and the amino group on the quinazoline core and the alkyl amino group of QNZ (bottom right).**

Bearing in mind that a different binding site on the same structure was identified for JSH-23, the binding energy of QNZ was determined at this site. A lower binding energy for QNZ was obtained at the JSH-23 binding site (binding site 2), with a binding energy of -8.168 kcal/mol.

The predicted binding pose of QNZ in the binding site 2 pocket (Figure 6), shared similar binding interactions as JSH-23, forming a hydrogen bond between the backbone hydroxyl group of the Phenylalanine (Phe) 142 residue on the IκBα subunit and the amino group on the alkyl chain of QNZ. Furthermore, π-π stacking interaction between the Tyrosine (Tyr) 351 residue on the p50 subunit and the quinazoline core were predicted (Figure 6).

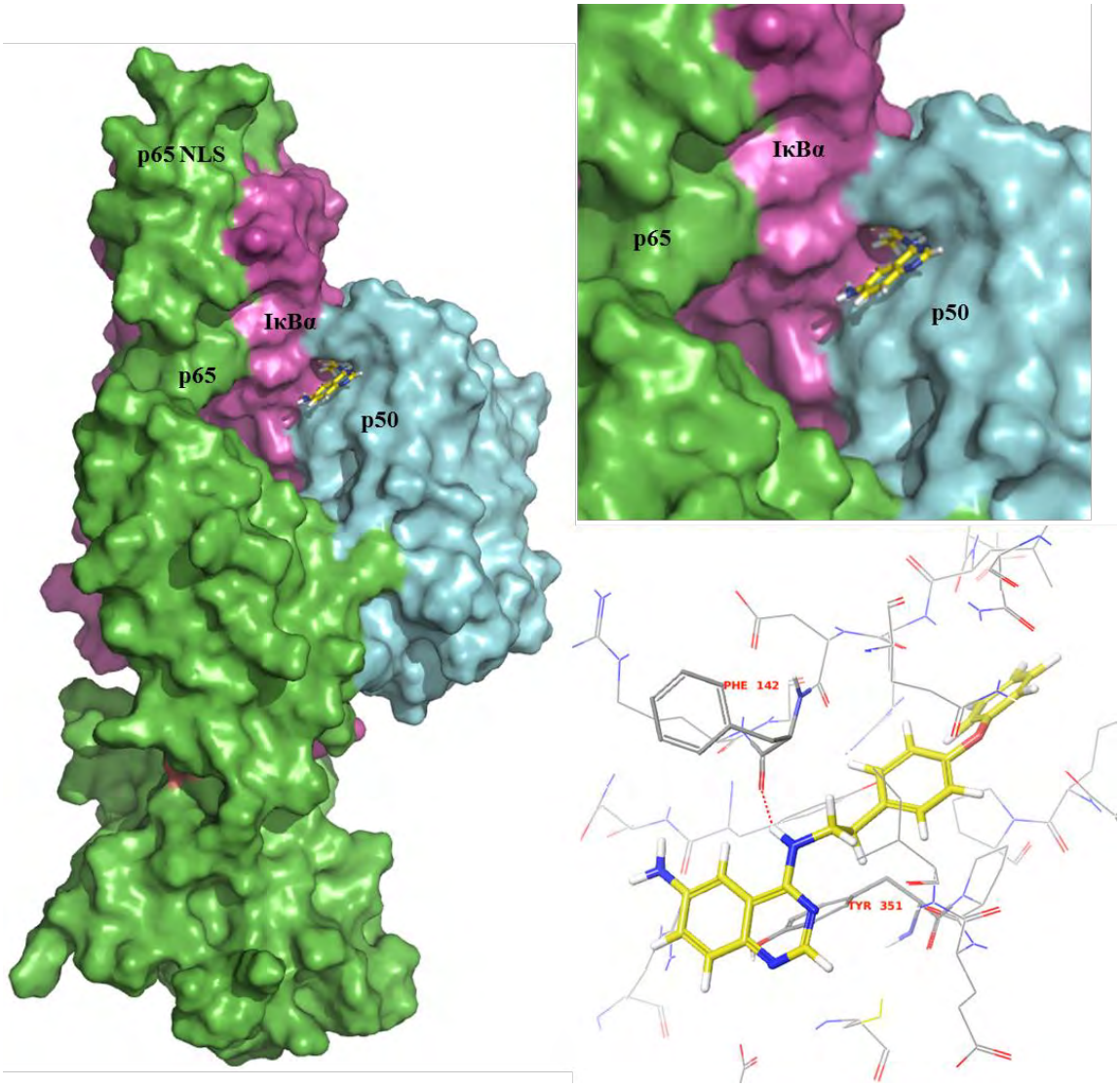



**Figure 6: The binding site 2 identified for JSH-23 on the NF-κB/IκBα crystal structure (1NFI.pdb), between IκBα and p50 subunit (left). QNZ was found to bind at this binding site with a lower binding energy compared to the binding energy obtained at binding site 1. The hydrophobic portion of QNZ is buried in the pocket formed by IκBα and p50 and the quinazoline core is exposed to the solvent (top right). The binding interactions include a hydrogen bond between the backbone hydroxyl of the Phe142 residue on the IκBα subunit and the alkyl amino group of QNZ.**

#### 2.3.2.2 Focused Docking Calculations

The focused docking calculations were performed using Glide 5.7, where the docking score or Glide score was considered when evaluating the docking calculations. QNZ was included in the docking calculations to allow for a comparison of binding energies to those obtained for the compounds in the ZINC database. There was no similarity of compounds identified with low binding energies between binding sites 1 and 2.

The docking calculations revealed a lower binding energy with the ZINC compounds at site 1 compared to site 2. The binding energies obtained from the compounds identified by the 2D similarity search were comparably similar to those compounds obtained from the 3D similarity search. At both binding sites, the ZINC compounds had lower binding energies than QNZ, which theoretically suggest that they would have better *in vitro* binding affinities, although there is reportedly a poor correlation between experimental and predicted binding affinities (R. Kim et al, 2008).

The compounds from the SBVS were classified as good, intermediate and poor binding compounds relative to QNZ, based on their predicted binding energies. This classification was performed to determine if any correlation existed between the predicted binding energies from the docking calculations and the *in vitro* activity data, and was performed despite the poor correlation reported in literature (R. Kim et al, 2008). Final compounds were selected for procurement based on their commercial availability, binding scores and chemical tractability. A total of 21 compounds identified from the docking calculations were procured for *in vitro* screening (Appendix 1, Table 4).

## 2.4 QNZ Discussion

It was evident through the similarity search performed, that the ZINC database did not have a sufficient number of quinazoline containing compounds. A maximum similarity of approximately 70% was achieved for the highest scoring compound in the ZINC database. The correct identification and ranking of the top 2 scoring compounds by the 2D similarity search, which form part of the SAR series reported in literature (M. tobe et al, 2003), may support the preference of 2D descriptors over 3D descriptors (V. Venkatraman et al, 2010; M.

Thimm et al, 2004; J. Bajorath, 2002). The compounds that were scored with a poor similarity had a striking commonality of the presence of a sulphonamide group. The sulphonamide group was either placed as a substituent at the terminal aromatic group or on the alkyl chain (Appendix 1, Table 2 and 3).

The correlation of the binding sites identified for the control compounds with the experimental data reported in literature, instilled confidence in the blind docking calculations. The blind docking methodology allowed for the consideration of the different proteins involved in the NF-κB pathway and also took into consideration the different conformational states of the NF-κB subunits, p65 and p50. This was with the exception of IKKα and IKKβ, having not being included due to the unavailability of a suitable human crystal structure. Interestingly the binding site of QNZ on the p50 subunit which was reported by K. C. Tsai et al, 2009 was also identified in the blind docking performed in this study. This binding site was not as populated as other binding sites and as it had a comparably higher binding energy to the other more populated binding sites; it was not included in this study. Taking into consideration that these findings are not confirmed by experimental assays, this does not rule out the possibility of QNZ binding at this site.

The binding sites that were identified for QNZ are located on or near areas on the NF-κB/IκBα structure that have been described in literature as suitable 'hot spots' or residues that are important for, and contribute to, the binding energy at the protein-protein interface (S. Bergqvest et al, 2008). The first hot stop is described as the 'PEST hot spot', which includes the PEST domain and the surrounding region making contact with the dimerization domain of the p65 subunit (S. Bergqvest et al, 2008). This is the same region found to be the energetically most favourable binding site for QNZ through the blind docking, binding site 1. Bergqvest et al, 2008 report that the deletion or mutation of the actual PEST domain of IκBα (residues 276-287) decreased binding of IκBα to p65 by 500 fold. Taking this into account, it can be postulated that the binding of QNZ results in an increased binding interaction between IκBα and p65, keeping the complex tightly bound together. Furthermore, binding site 1 is in close proximity to the p65 phosphorylation site, Ser276, which when phosphorylated by Protein Kinase A (PKA) increases the transcriptional activity of p65 (T. Okazaki et al, 2003; P. Arun et al, 2009; A. Spooren et al, 2010; R. Moreno et al, 2010). Taking this into consideration, it could be hypothesised that binding of QNZ may confer inhibition by inducing a conformational change that masks the Ser276 phosphorylation site, keeping the complex bound together. The second hot spot identified by S. Bergqvest et al, 2008, is

located between the p65 NLS region and the 1st ankyrin repeat of IκBα, which is in close proximity to binding site 2 or the JSH-23 binding site. In its natural state, the 1st ankyrin repeat of IκBα is reported to sequester the NLS of p65, maintaining the inactive conformation and keeping it localised in the cytoplasm. The binding studies reported in literature indicate that mutating the specific residues 305 and 321 in the p65 NLS region, significantly weakened the binding interactions of the NLS which would support the disassociation of IκBα from the p65 NLS. (S. Bergqvest et al, 2008). Considering that JSH-23 was found to bind here and that it's reported mechanism of action results in the decreased nuclear localisation of p65, it would suggest that binding of JSH-23 strengthens the interactions, keeping the complex intact. This may imply that QNZ has a similar effect upon binding.

Interestingly, lower predicted binding energies were obtained for ZINC32089602 and ZINC32089624 at binding site 1 compared to binding site 2, which may suggest their preference for binding at site 1. This is unlike for QNZ, which obtained a lower binding energy at binding site 2. A comparison of the binding interactions at both binding sites reveals a greater number of predicted binding interactions at binding site 1 (Figure 7), which would explain the lower binding energies obtained. An interesting observation is the different binding pose predicted for QNZ compared to the other inhibitors at both binding sites. The predicted binding pose or orientation of QNZ at the binding site 2 (Figure 8), indicates that the lipophilic region, the phenoxy phenyl moiety, is completely buried in the binding pocket, unlike at binding site 1 where it is solvent exposed (Figure 7). This may suggest the presence of increased hydrophobic interactions which may explain the lower binding energy obtained at this binding site. It has been reported in literature that optimised hydrophobic interactions, including hydrogen bonding, efficiently stabilises a ligand which leads to an increase in binding affinity and drug efficiency (R. Patil et al, 2010).

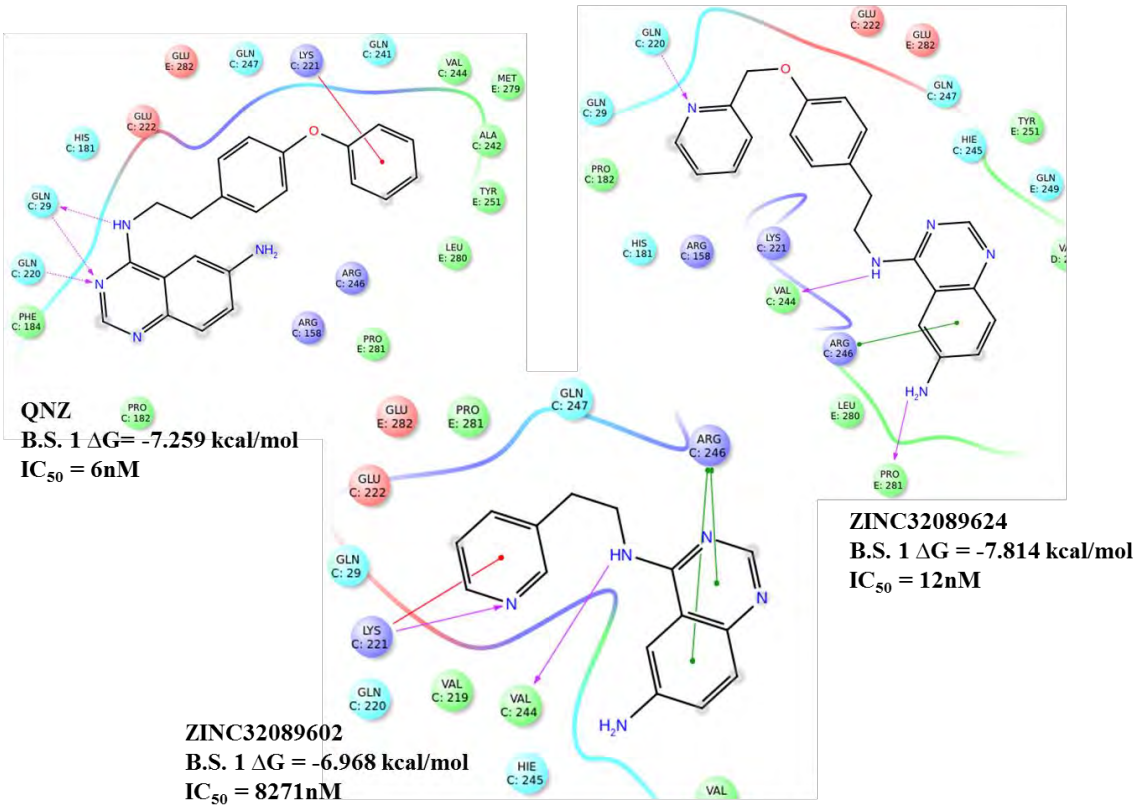



**Figure 7: The binding interactions predicted at binding site 1 for QNZ (top left), ZINC32089624 (top right) and ZINC32089602 (bottom). The red lines represent π-cation bonds; the purple solid lines represent back-bone hydrogen bonds; the green lines represent π-π stacking interactions. The grey shaded circles represent areas of the ligand that are exposed to the solvent. The predicted binding energies at the binding site 1 and the experimental $IC_{50}$ values reported in literature are depicted (M. Tobe et al, 2003).**

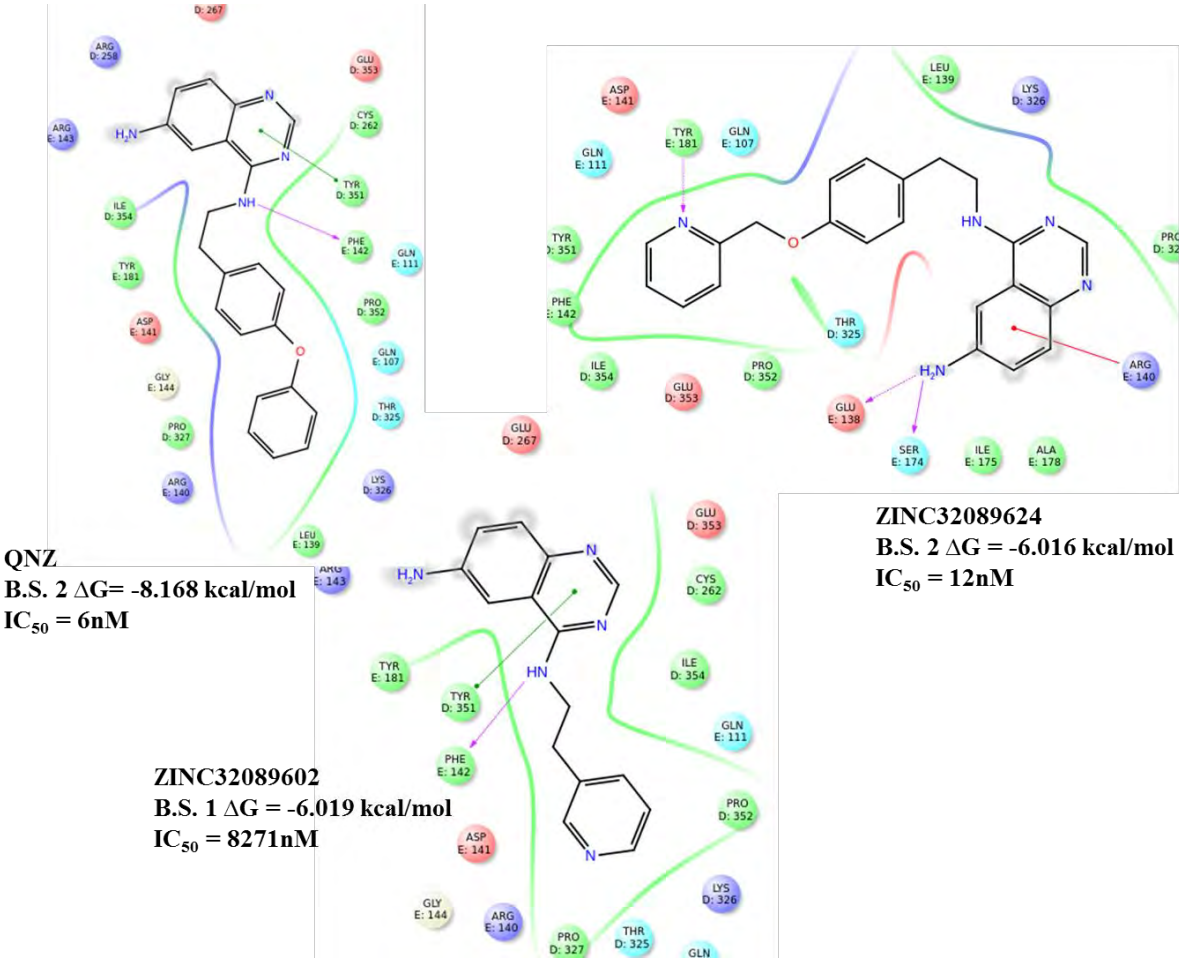



**Figure 8: The binding interactions predicted at binding site 2 for QNZ (top left), ZINC32089624 (top right) and ZINC32089602 (bottom). The red lines represent π-cation bonds, the purple solid lines represent back-bone hydrogen bonds; the green lines represent π-π stacking interactions. The grey shaded circles represent areas of the ligand that are exposed to the solvent. The predicted binding energies at the binding site 2 and the experimental $IC_{50}$ values reported in literature are depicted (M. Tobe et al, 2003).**

The successful identification of the binding site for QNZ was followed by the docking calculations of the compounds in the ZINC dataset. Overall, the binding energies obtained at sites 1 and 2 with the ZINC database compounds were comparable, within a range of -8 to -9 kcal/mol. The selection of the top-ranked compounds for *in vitro* screening was hampered due to their commercial unavailability. Therefore, the final selection of good scoring compounds was derived from further down the list of the top-ranked compounds.

## 2.5 Methodology of the vHTS performed with Pentamidine

### 2.5.1 2D Similarity Search

The Discovery studios 'Find similar molecules by fingerprints' protocol on Pipeline Pilot was used to conduct the similarity search for Pentamidine (Accelrys Software Inc., *Discovery Studio Modeling Environment, 4.5*, San Diego: Accelrys Software Inc., 2007). The input ligands included the filtered ZINC database and the reference ligand, Pentamidine. The minimum similarity was set to 0.50 and the Tanimoto coefficient was selected with a pre-defined property of the ECFP_4 fingerprint. The top 2000 compounds that scored a minimum similarity of 0.50 were retained for the SBVS.

### 2.5.2 3D Similarity Search

The Zinc database was filtered and the 3D generation of the conformation was performed for Pentamidine, as described above (section 2.2.2). The simple run on the ROCS interface was employed to perform the 3D similarity search for Pentamidine (ROCS, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2010). Default parameters were maintained for the inputs, where the colour force field was kept at the default Implicit Mills-Dean. The structure data file for Pentamidine and the 3D conformations of the filtered ZINC database were uploaded for screening. The number of best hits to retain was increased to 1000 compounds.

### 2.5.3 Docking Calculations on Glide5.7

The methodology for the docking calculations used for Pentamidine was much simpler than that used for QNZ. This was primarily due to the crystal structure of the human DAO enzyme found in complex with Pentamidine (3HII.pdb) (A. P. McGrath et al, 2009). To validate the binding site and docking parameters, Pentamidine was removed and re-docked to determine if the same binding orientation and interactions could be achieved with Glide5.7 (Glide,

version 5.7, Schrödinger, Inc., New York, NY, 2009). Pentamidine was included in the docking calculations to allow for a comparison of the predicted binding scores to those obtained for the compounds in the ZINC database.

The conformations of Pentamidine and the ZINC database were further processed using the LigPrep application on Maestro (Maestro, version 5.6, Schrödinger, Inc., New York, NY, 2009). Their isomers and tautomers were generated using default parameters. The ionisation states were generated at a default pH of 7 ± 2 and partial atomic charges were assigned using the OPLS_2005forcefield. The 3HII.pdb structure was prepared using the Protein Preparation Wizard on Maestro (Maestro, version 5.6, Schrödinger, Inc., New York, NY, 2009). A pre-processing step using default parameters preceded the hydrogen bond assignment using PROPKA. A restrained minimisation step was performed using the OPLS_2005 force field. The Receptor Grid Generation application on Glide5.7 was used to identify the binding site by selecting the ligand, Pentamidine, in the 3HII.pdb structure. The Glide SP docking protocol was employed to carry out the ligand docking calculations. Pentamidine was removed from the 3HII.pdb file in order to perform the docking calculations with the compounds contained in the ZINC database. Default parameters were maintained.

## 2.6 Pentamidine Results

### 2.6.1 Pentamidine Similarity Search

The similarity search performed on Pipeline Pilot identified compounds with a maximum Tanimoto score of 0.5625 or a 56% similarity to Pentamidine (Figure 9). Pipeline Pilot retrieved 30 compounds with a 50% or more similarity to Pentamidine. The top 20 scoring compounds identified from the 2D similarity search are tabulated in Appendix 1, Table 5. None of the compounds identified had a bis-benzamidine moiety similar to Pentamidine.

A total number of 675 compounds that were 50% or more similar were identified from vROCS and the maximum TanimotoCombo similarity score was 1.146, which translates to a 57% similarity to Pentamidine. None of the compounds identified had a bis-benzamidine moiety that was similar to Pentamidine, although varying lengths of the alkyl chain between aromatic moieties was a common feature (Figure 9). The 20 top scoring compounds

identified from the vROCS similarity search are tabulated in Appendix 1, Table 6. Consistent with the aims of this study to perform a hierarchical screening, the top scoring compounds of 50% or greater similarity were selected to perform the SBVS.

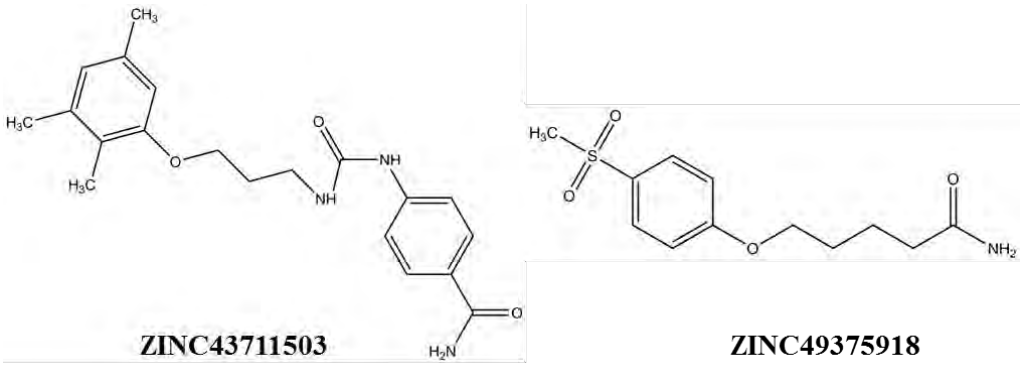



**Figure 9: The compound ZINC43711503 was identified from the 3D similarity search with the highest TanimotoCombo score of 1.146 and compound ZINC49375918 identified from the 2D similarity search with a Tanimoto score of 0.5625. Both compounds lack the bis-benzamidine moiety, compared to Pentamidine.**

### 2.6.2 Pentamidine SBVS

The docking calculations were performed using the Glide5.7 SP docking protocol. Pentamidine was re-docked in the DAO crystal structure to validate the docking calculation parameters. A RMSD was calculated of the crystal structure pose of Pentamidine and of the re-docked pose to be 1.25Å (Figure 10), which indicated the regeneration of the crystal structure conformation.

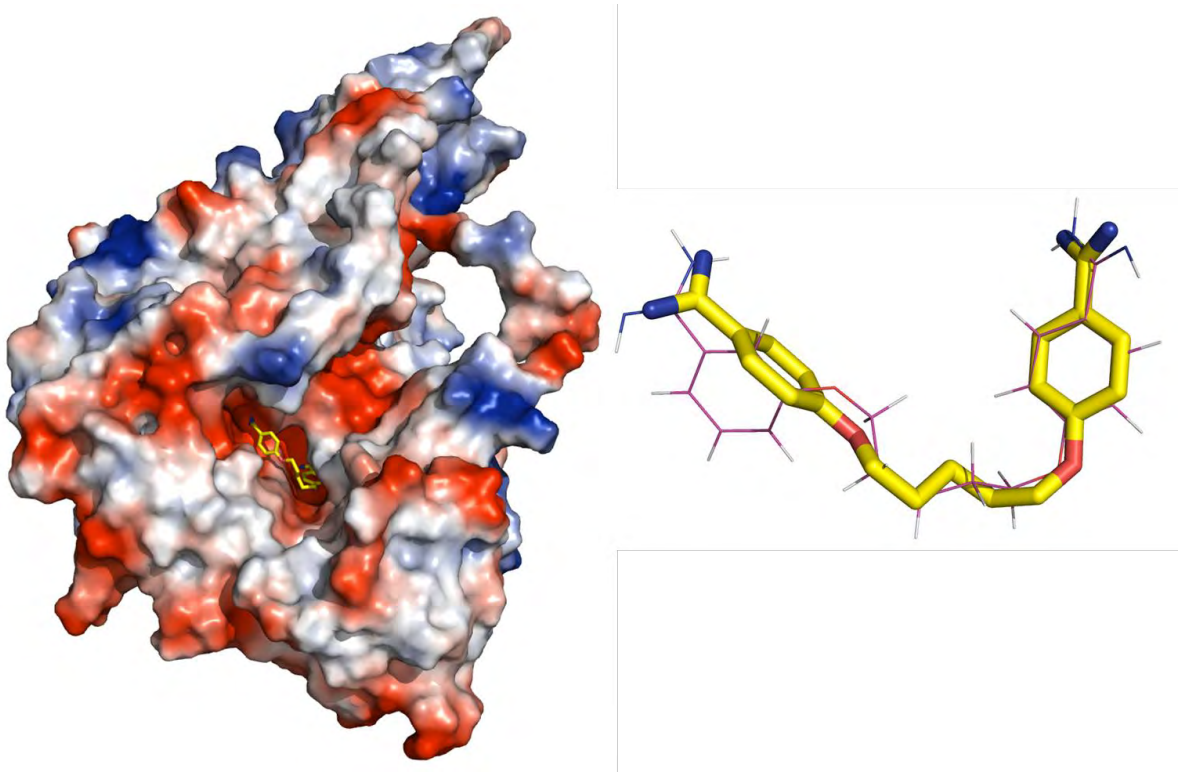


**Figure 10: The DAO crystal structure in complex with Pentamidine (left). The electrostatic map potential illustrates the negatively charged buried binding pocket, thought to interact with the positively charged substrates (A. P. McGrath et al, 2009). The re-produced Pentamidine pose (shown on the right in line view) overlaid on the crystal structure pose (in stick view and yellow). The RMSD was 1.2525Å, which was considered an acceptable value.**

Pentamidine assumes a 'horse-shoe' conformation which is consistent with the pose reported in the crystal structure (A. P. McGrath et al, 2009) (Figure 10). Certain key interactions reported in literature include, a hydrogen bond to the conserved Aspartic acid (Asp) 373 residue. The Asp373 residue reportedly forms the base of the catalytic site in the DAO structure, which when bound to an inhibitor; it inhibits the natural substrates from accessing the catalytic site. Consistent with the binding interactions observed in the crystal structure, the hydrogen bond between Pentamidine and the Asp373 residue was re-produced. The second buried amidine group is predicted to form a hydrogen bond with the carboximide group of the Asparagine (Asn) 460 residue and the trihydroxyphenylalanine quinone (TPQ) cofactor (Figure 11). The DAO crystal structure is reportedly modelled in the 'on-copper' conformation, which is the inactive conformation of the enzyme, where the TPQ cofactor coordinates a copper ion in the active site, which effectively blocks the access of the natural substrate molecule to the active site (A. P. McGrath et al, 2009). The predicted interaction of Pentamidine with the TPQ cofactor is plausible since other inhibitors such as 2-hydrazino-

pyridine, have been shown to interact with the TPQ cofactor where they mimic the formation of the Schiff base intermediate and inhibit the natural substrate molecule from accessing the active site. This interaction between the TPQ cofactor and Pentamidine was not modelled in the crystal structure of DAO in complex with Pentamidine (A. P. McGrath et al, 2009). None of the ZINC compounds were predicted to interact with the TPQ cofactor. The Aspartic acid (Asp) 186 residue, has been identified as the key residue responsible for substrate specificity and reportedly interacts with the natural ligands such as histamine and putrescine as well as other potent inhibitors of DAO. Pentamidine does not form this interaction with the conserved Asp186 residue compared to other inhibitors as it lacks a positively charged group on the central linker (A. P. McGrath et al, 2009). This explains the reduced potency of Pentamidine compared to other DAO inhibitors despite the fact that Pentamidine does not fully occupy the active site of the enzyme (A. P. McGrath et al, 2009). Instead, due to its increased length, it is able to form a hydrogen bond between the amidine group that extends out of the active site and the Serine (Ser) 380 residue of DAO. The hydrogen bond between the solvent exposed amidine group and Ser380 was not re-produced but instead the amidine group was predicted to form a hydrogen bond with the carboxyl group of the Tryptophan residue (Trp) 376 (Figure 11). Other interactions reported in the crystal structure include π-π stacking interactions between the buried amidine group and Tyrosine residue (Tyr) 371 and the Tryptophan (Trp) 376 residue of DAO (A. P. McGrath et al, 2009). The re-produced docking pose predicted a π-π stacking interaction between the buried phenyl ring and the Tyr371residue of DAO. Figure 11 below, depicts the binding interactions of Pentamidine obtained from the re-docking of Pentamidine in the DAO crystal structure along with the top scoring compounds from the docking calculations. The compound ZINC22536415 was identified from the 2D similarity search, which obtained a binding energy higher than Pentamidine.

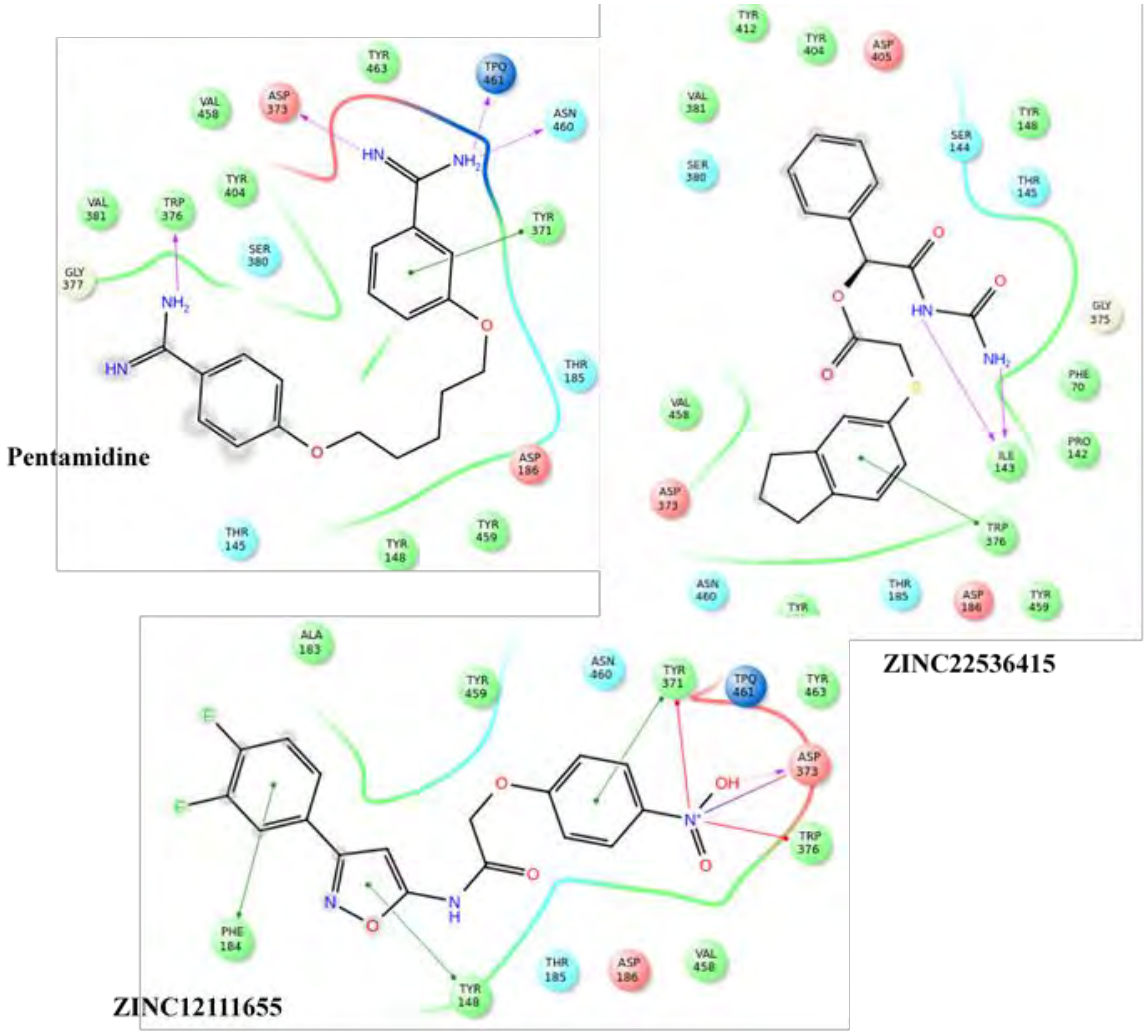



**Figure 11: The predicted binding interactions of Pentamidine (left), ZINC22536415 (right) and ZINC12111655 (bottom) obtained from the docking calculations performed on Glide5.7. The key interaction with the Asp373 residue at the catalytic base is made with both Pentamidine and the compound ZINC12111655, which was identified from the 3D similarity search. The compound ZINC22536415, identified from the 2D similarity search did not form the key interaction with Asp373, which may be reflected in its predicted high binding energy. All compounds including Pentamidine were not predicted to form a hydrogen bond with the conserved residue Asp186, which is responsible for the increased binding affinity and ligand specificity (A. P. McGrath et al, 2009).**

The compounds identified from the 2D similarity search were predicted to have higher binding energies compared to Pentamidine and the compounds identified from the 3D similarity search. Their docking poses reveal that they don't form a hydrogen bond with the conserved residue Asp373, which forms part of the catalytic base (A. P. McGrath et al, 2009) which may be a consequence of their higher binding energies. As can be seen in Figure 11,

the interaction with the Asp186 residue, which is reported to be responsible for ligand specificity and increased binding affinity is absent in all predicted docking poses, including Pentamidine. The final compounds selected for *in vitro* screening were identified based on their commercial availability, binding scores and chemical tractability. A total of 14 compounds identified from the docking calculations were procured for *in vitro* screening (Appendix 1, Table 7).

## 2.7 Pentamidine Discussion

No compounds were identified with the benzamidine moiety, which may be seen as a limitation of the ZINC database. A maximum similarity of approximately 57% was achieved with both the 2D and 3D similarity search methods. A larger number of similar compounds to Pentamidine were identified from the 3D similarity search method compared to the 2D similarity search method.

There was confidence in the docking protocol employed in this study based on the acceptable RMSD of less than 2Å, between the crystal structure pose of Pentamidine and the pose obtained from the re-docking of Pentamidine (R. A. Friesner et al, 2004; Z. Li et al, 2013). None of the compounds that were procured for *in vitro* screening were seen to interact with the conserved residue Asp186, which was identified to be responsible for substrate specificity and also responsible for the increased potency of other inhibitors of DAO (A. P. McGrath et al, 2009). Despite this, contact with the Asp373 residue of the catalytic base of DAO was made with the compounds, which would suggest that an inhibitory action is most likely to occur (A. P. McGrath et al, 2009). The compounds identified from the 2D similarity search were not predicted to interact with the Asp373 residue of DAO. Furthermore, these compounds were predicted to have higher binding energies compared to Pentamidine and the compounds from the 3D similarity search, which may be a consequence of not interacting with the catalytic base of DAO.

Compounds that had good, intermediate and poor docking scores were identified for *in vitro* screening. As with QNZ, the selection of the top-ranked compounds for *in vitro* screening was hampered due to their commercial unavailability. Therefore, the good scoring compounds were derived from further down the list of top-ranked compounds.

## 2.8 Conclusion

A hierarchical screening consisting of a ligand based and structure based virtual screening was performed for QNZ and Pentamidine on the ZINC database. A blind docking protocol was performed on the NF-κB proteins to identify the energetically most favourable binding site of QNZ. This revealed 2 potential binding sites for QNZ on the NF-κB/IκBα complex, which correlated with 2 'hotspots' previously reported in literature. The SBVS performed on the DAO crystal structure revealed that the ZINC compounds did not form a binding interaction with the conserved Asp186 residue of DAO but despite this they do form an interaction with the catalytic base of the DAO enzyme. A total number of 35 compounds were procured for *in vitro* testing in ovarian, renal and prostate cancer cell lines, described in Chapter 4.

## 2.9 References


1. S. Subramaniam, M. Mehrotra and D. Gupta, Virtual High Throughput Screening (vHTS) – A Perspective, Bioinformation: Views & Challenges, Vol. No. 3, Issue No. 1, Pp. 14-17, 2008.
2. V. Kumar, S. Krishna, M. I. Siddiqi, Virtual screening strategies: recent advances in the identification and design of anti-cancer agents, Elsevier Methods, 2014.
3. M. N. Drwal and R. Griffith, Combination of ligand- and structure- based methods in virtual screening, Drug Discovery Today: Technologies, Vol. 10, Issue No. 3, Pg. No. 395-401, 2013.
4. E. Bielska, X. Lucas, A. Czerwoniec, J. M. Kasprzak, K. H. Kaminska and J. M. Bujnicki, Virtual screening strategies in drug design – methods and applications, Journal of Biotechnology, Computational Biology and Bionanotechnology, Vol. No. 92, Issue No. 3, Pp. 249-264, 2011.
5. T. Cheng, Q. Li, Z. Zhou, Y. Wang and S. H. Bryant, Structure-based virtual screening for drug discovery: a problem-centric review, The AAPS Journal, Vol. No. 14, Issue No. 1, Pp. 133-141, 2012.
6. P. Willet, Similarity-Based Virtual Screening using 2D Fingerprints, Drug Discovery Today, Vol. No. 11, Issue No. 23/24, Pp. 1046-1053, 2006.
7. D. Rogers and M. Hahn, Extended-connectivity Fingerprints, Journal of Chemical Information and Modelling, Vol. 50, Issue No. 5, Pg. No. 742-754, 2010.
8. M. Awale and J. Reymond, A Multi-Fingerprint Browser for the ZINC Database, Nucleic Acid Research, 2014.
9. S. R. Langdon, I. M. Westwood, R. L. M. van Montfort, N. Brown and J. Blagg, Scaffold-focused virtual screening: Prospective application to the discovery of TTK inhibitors, Journal of Chemical Information and Modelling, Vol. 53, Pg. No. 1100-1112, 2013.
10. Y. Hu, E. Lounkine and J. Bajorath, Improving the search performance of extended connectivity fingerprints through activity-orientated feature filtering and application of a bit-density-dependent similarity function, ChemMedChem, Vol. 4, Pg. No. 540-548, 2009.

11. H. Eckert and J. Bajorath, Molecular Similarity Analysis in Virtual Screening: Foundations, Limitations and Novel Approaches, Drug Discovery Today Review, Vol. No. 12, Issue No. 5/6, Pp. 225-233, 2007.
12. V. Venkatraman, V. I. Perez-Nueno, L. Mavridis and D. W. Ritchie, Comprehensive Comparison of Ligand-Based Virtual Screening Tools against the DUD Data set Reveals Limitations of Current 3D Methods, Journal of Chemical Information and Modelling, Vol. No. 50, Issue No. 12, Pp. 2079-2093, 2010.
13. M. Thimm, A. Goede, S. Hougardy and R. Preibner, Comparison of 2D and 3D Superposition. Application to Searching a Conformational Drug Database, Journal of Chemical Information and Computer Science, Vol. No. 44, Pp. 1816-1822, 2004.
14. J. Bajorath, Integration of Virtual and High Throughput Screening, Nature Review Drug Discovery, Vol. No. 1, Issue No. 11, Pp. 882-894, 2002.
15. A. P. McGrath, K. M. Hilmer, C. A. Collyer, E. M. Shepard, B. O. Elmore, D. E. Brown, D. M. Dooley and J. M. Guss, Structure and Inhibition of Human Diamine Oxidase, Biochemistry, Vol. No. 48, Issue No. 41, Pp. 9810-9822, 2009.
16. C. Hetenyi and D. Van Der Spoel, Efficient docking of peptides to proteins without prior knowledge of the binding site, Protein Science, Vol. 11, Pg. No. 1729-1737, 2002.
17. Z. Li, X. Wang, K. Li, L. Kang and Q. Guo, A new molecular docking methods based on residue groups and PMF scoring function, Journal of Chemical and Pharmaceutical Research, Vol. No. 5, Issue No. 12, Pp. 1066-1069, 2013.
18. C. Hetenyi and D. van der Spoel, Blind Docking of Drug-Sized Compounds to Proteins with up to a Thousand Residues, FEBS Letters, Vol. No. 580, Issue No. 5, Pp. 1447-1450, 2006.
19. P. S. Kharkar, S. Warrier and R. S. Gaud, Reverse Docking: A powerful tool for drug Repositioning and Drug Rescue, Future Medicinal Chemistry, Vol. No. 6, Issue No. 3, Pp. 333-342, 2014.
20. D. Ghersi and R. Sanchez, Improving accuracy and efficacy of blind protein-ligand docking by focusing on predicted binding sites, Proteins, Vol. No. 74, Pp. 417-424, 2009.
21. R. Zheng, T. Chen and T. Lu, A Comparative Reverse Docking Strategy to Identify Potential Antineoplastic Targets of Tea Functional Components and Binding Mode, International Journal of Molecular Sciences, Vol. No. 12, Pp. 5200-5212, 2011.

22. A. Kumar, G. Negis and S. S. Sharma, JSH-23 Targets Nuclear Factor-kappa B and Reverses Various Deficits in Diabetic Neuropathy: Effect on Neuro-Inflammation and Antioxidant Defence, Diabetes Obesity and Metabolism, Vol. 13, Issue No. 8, 2011.
23. J. W. Pierce, R. Schoenleber, G. Jesmok, J. Best, S. A. Moore, T. Collins and M. E. Gerritsen, Novel inhibitors of cytokine-induced IκBα phosphorylation and endothelial cell adhesion molecule expression show anti-inflammatory effects *in vivo,* The Journal of Biological Chemistry, Vol. 272, Issue No. 34, pp. 21096-21103, 1997.
24. A. Lavecchia and C. Di Giovanni, Virtual screening strategies in Drug Discovery: a critical review, Current Medicinal Chemistry, Vol. 20, Issue No. 1, 2013.
25. T. I. Oprea, Virtual Screening in Lead Discovery: A Viewpoint, Molecules, Vol. No. 7, Pp. 51-62, 2002.
26. C. Abad-Zapatero, A Sorcerer's Apprentice and the Rule of Five: From Rule-of Thumb to Commandment and Beyond, Drug Discovery Today: Perspective, Vol. No. 12, Issue No. 23/24, Pp. 995-997, 2007.
27. K. S. Watts, P. Dalal, R. B. Murphy, W. Sherman, R. A. Friesner and J. C. Shelley, Journal of Chemical Information and Modeling, Vol. 50, Issue No. 4, Pg. No. 534-546, 2010.
28. F. E. Chen, D. B. Huang, Y. Q. Chen and G. Ghosh, Crystal structure of p50/p65 heterodimer of transcription factor NF-KappaB bound to DNA, Nature, Vol. No. 391, Pp. 410-413, 1998.
29. M. D. Jacobs, S. C. Harrison, Structure of an IκBα/NF-κB complex, Cell, Vol. No. 95, Issue No. 6, Pp. 749-758, 1998.
30. M. Rushe, L. Silvian, S. Bixler, L. L. Chen, A. Cheung, S. Bowes, H. Ceurvo, S. Berkowitz, T. Zheng, K. Guckian, M. Pellegrini and A. Lugovskoy, Structure of a NEMO/IKK-associating domain reveals architecture of the interaction site, Cell Press: Structure, Vol. No. 16, Pp. 798-808, 2008.
31. B. Rayet and C. Gelinas, Aberrant Rel/NF-Kb Genes and Activity in Human Cancer, Oncogene, Vol. No. 18, Issue No. 49, Pp. 6938-6947, 1999.
32. A. Oeckinghaus, M. S. Hayden and S. Ghosh, Crosstalk in NF-κB Signalling Pathways, Nature Immunology, Vol. No. 12, Issue No. 8, Pp. 695-708, 2011.
33. R. G. Uzzo, P. L. Crispen, K. Golovine, P. Makhov, E. M. Horwitz and V. M. Kolenko, Diverse Effects of Zinc on NF-κB and AP-1 Transcription Factors Implications for Prostate Cancer Progression, Carcinogenesis, Vol. No. 27, Issue No. 10, Pp. 1980-1990, 2006.

34. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, Autodock4 and AutoDockTools4: Automated docking with selective receptor flexibility, Journal of Computational Chemistry, Vol. No. 30. Issue No. 16, Pp. 2785-2791, 2009.
35. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olsen, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, Journal of Computational Chemistry Vol. No. 19, Issue No. 14, Pp. 1639-1662, 1998.
36. C. Sotriffer, Virtual screening: Principles, challenges and practical guidelines, John Wiley & Sons, 2011.
37. M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, H. Takahashi, T. Fukazawa and H. Hayashi, Discovery of Quinazolines as a Novel Structural Class of Potent Inhibitors of NF-κB Activation, Bioorganic & Medicinal Chemistry, Vol. No. 11, Pp. 383-391, 2003.
38. M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, H. Takahashi and H. Hayashi, A Novel Stuctural Class of Potent Inhibitors of NF-κB Activation: Structure-Activity Relationships and Biological Effects of 6-Aminoquinazoline Derivatives, Bioorganic & Medicinal Chemistry, Vol. No. 11, Pp. 3869-3878, 2003.
39. E. Kellenberger, J. Rodrigo, P. Muller and D. Rognan, Comparative evaluation of eight docking tools for docking and virtual screening accuracy, Proteins: Structure, Function and Bioinformatics, Vol. No. 57, Pp. 225-242, 2004.
40. H. Shin, B. H. Kim, E. Y. Chung, A. Jung, Y. S. Kim, K. R. Min and Y. Kim, Suppressive effect of a novel aromatic diamine compound on nuclear factor-κB-dependent expression of inducible nitric oxide synthase in macrophages, European Journal of Pharmacology, Vol. 521, Pg. No. 1-8, 2005.
41. D. Kesanakurti, C. Chetty, D. R. Maddirela, M. Gujrati and J. S. Rao, Essential Role of Cooperative NF-κB and Stat3 Recruitment to ICAM-1 Intronic Consensus Elements in the Regulation of Radiation-Induced Invasion and Migration in Glioma, Oncogene, Vol. No. 32, Issue No. 43, Pp. 5144-5155, 2013.
42. C. L. Scaife, J. Kuang, J. C. Wills, D. B. Trowbridge, P. Gray, B. M. Manning, E. J. Eichwald, R. A. Daynes and S. K. Kuwada, Nuclear Factor κB Inhibitors Induce Adhesion-Dependent Colon Cancer Apoptosis Implications for Metastasis, Cancer Research, Vol. No. 62, Pp. 6870-6878, 2002.

43. R, Tiwary, W. Yu, J. Li, S. K. Park, B. G. Sanders and K. Kline, Role of Endoplasmic Reticulum Stress in α-TEA Mediated TRAIL/DR5 Death Receptor Dependent Apoptosis, PLoS One, Vol. No. 5, Issue No. 7, 2010.
44. P. Tuffery and P. Derreumaux, Flexibility and binding affinity in protein–ligand, protein–protein and multi-component protein interactions: limitations of current computational approaches, Journal of The Royal Society Interface, Vol. No. 9, Pp. 20-35, 2012.
45. Q. V. Vuong, K. Siposova, T. T. Nguyen, A. Antosova, L. Balogova, L. Drajna, J. Imrich, M. S. Li and Z. Gazova, Binding of Glyco-Acridine Derivatives to Lysozyme Leads to Inhibition of Amyloid Fibrillisation, Biomacromolecules, Vol. No. 14, Pp. 1035-1043, 2013.
46. K. C. Tsai, L. Teng, Y. Shao, Y. Chen, Y. Lee, M. Li and N. Hsiao, The first pharmacophre model for potent NF-κB inhibitors, Bioorganic and Medicinal Chemistry Letters, Vol. 19, Pg. No. 5665-5669, 2009.
47. S. Bergqvist, G. Ghosh and E. A. Komives, The IκBα/NF-κB Complex has Two Hot Spots, One at Either End of the Interface, Protein Science, Vol. No. 17, Pp. 2051-2058, 2008.
48. T. Okazaki, S. Sakon, T. Sasazuki, H. Sakurai, T. Doi, H. Yaqita, K. Okumura and H. Nakano, Phosphorylation of Serine 276 is Essential for p65 NF-kappaB Subunit-Dependent Cellular Responses, Biochemical and Biophysical Research Communications, Vol. No. 300, Issue No, 4, Pp.. 807-812, 2003.
49. P. Arun, M. S. Brown, R. Ehsanian, Z. Chen and C. Van Waes, Nuclear NF-kappaB p65 Phosphorylation at Serine 276 by Protein Kinase A Contributes to the Malignant Phenotype of Head and Neck Cancer, Clinical Cancer Research, Vol. No. 1, Issue No. 15, Pp. 5974-5984, 2009.
50. A. Spooren, K. Kolmuc, L. Vermeulen, K. Van Wesemael, G. Haegeman and S. Gerlo, Hunting for Serine 276-Phosphorylated p65, Journal of Biomedicine and Biotechnology, Vol. No. 2010, Article ID 275892, 2010.
51. R. Moreno, J. Sobotzik, C. Schultz and M. L. Schmitz, Specification of the NF-κB Transcriptional Response by p65 Phosphorylation and TNF-Induced Nuclear Translocation of IKKε, Nucleic Acids Research, Vol. No. 38, Issue No. 18, Pp. 6029-6044, 2010.

52. R. Patil, S. Das, A. Stanley, L. Yadav, A. Sudhakar and A. K. Varma, Optimised Hydrophobic Interactions and Hydrogen Bonding at the Target-Ligand Interface Leads the Pathways of Drug-Designing, PLoS one, Vol. No. 5, Issue No. 8, 2010.
53. R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, Journal of Medicinal Chemistry, Vol. 47, Issue No. 7, pp. 1739-1749, 2004.

# Chapter 3: Virtual Screening of the NCI Database

## 3.1 Introduction

A chemo-informatic, knowledge driven virtual screening approach was used to screen the NCI database. In light of the ZINC compounds displaying no *in vitro* activity, the knowledge driven approach included only a ligand-based virtual screening of the NCI database (Figure 1). This comprised of a detailed analysis of the SAR data available in literature for both QNZ and Pentamidine. Utilising chemo-informatic tools, an analysis of the SAR data enabled us to identify which virtual screening methods would be most beneficial to perform in order to obtain similar inhibitors to QNZ and Pentamidine (Figure 1). A validation of the 2D and 3D similarity search protocol was performed in order to assess the accuracy of the screening protocols utilised on Pipeline Pilot and vROCS, respectively. The Receiver Operating Characteristic (ROC) curve analysis was used to assess the computer algorithm's ability to discriminate between active and inactive compounds (N. Triballeau et al, 2005). Hence, the sensitivity (the ability of a test to identify actives) and specificity rates (the ability of a test to discard inactives) of an algorithm are useful parameters that were calculated to determine this (N. Triballeau et al, 2005).

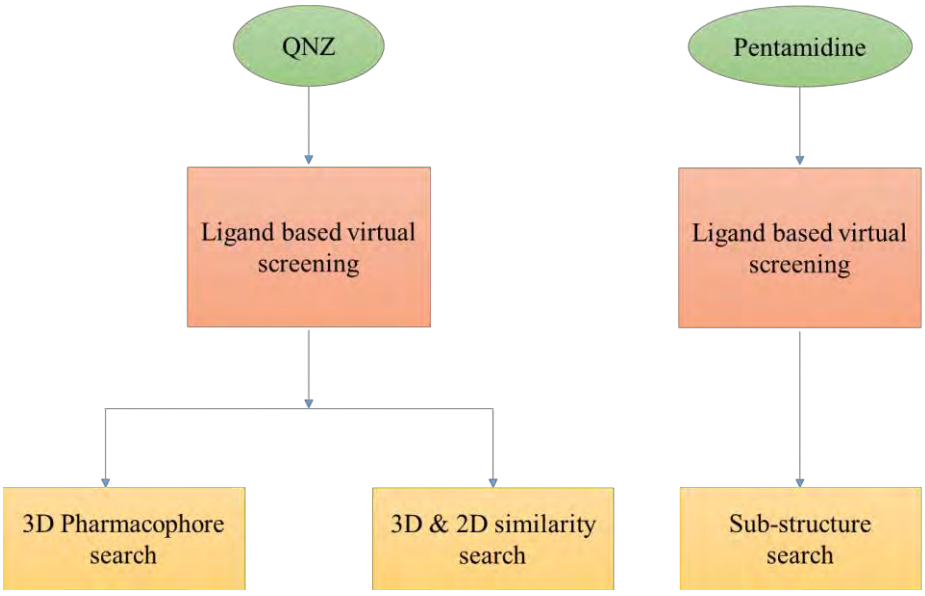



**Figure 1: The aims of the virtual screening performed on the NCI database.**

The limited amount of SAR data extracted from literature on the quinazoline inhibitors of NF-κB (M. Tobe et al, 2003; M. Tobe et al, 2003; L. Xu et al, 2013), allowed for the summary of the key structural features based on the QNZ chemical scaffold and their influence on activity. The summarised SAR guided the selection of compounds identified from the 2D and 3D similarity search, which was performed as previously described in Chapter 2. A ligand-based pharmacophore virtual screening was performed for QNZ, which consisted of developing several pharmacophores to test the hypothesis of which structural features, on the QNZ scaffold, were necessary and important for activity. A total of 4 pharmacophores were manually constructed and used to virtually screen the NCI database.

Without performing a SBVS, a binding hypothesis of Pentamidine in the AdoMet-DC enzyme was generated to provide insights not only into the binding pose of Pentamidine but also on possible ligands that may act similarly to Pentamidine. To date, there is no literature reporting on the inhibitory binding interactions of Pentamidine in the human AdoMet-DC enzyme, a more relevant and key therapeutic target in the polyamine biosynthetic pathway (A. Paasinen-Sohns et al, 2000; D. T. Koomoa et al, 2009; J. L. Ekstrom et al, 1999). The docking hypothesis generated, suggests that the benzamidine moieties and the ether oxygens of Pentamidine are involved in the binding interactions with the AdoMet-DC enzyme. Taking this into consideration, developing a pharmacophore model was not considered as defining the linker region seen in Pentamidine would be subjective. This was further substantiated by considering the variability in length of the ethyl chain and side chain substituents of the other known inhibitors of the AdoMet-DC enzyme (U. Regenass et al, 1992; U. Regenass et al, 1994; J. Stanek et al, 1993; I. Jarak et al, 2011).

An initial sub-structure search on the benzamidine moiety was performed and the limited number of compounds retrieved was indicative of the expected results that would be obtained from a similarity search. Taking this into consideration, a similarity search was not performed for Pentamidine. Bearing in mind that Pentamidine has been reported in literature to have many targets, including the minor groove of DNA and various other protein targets such as the serine protease enzymes (V. A. Johnson et al, 1976; S. J. Kempin et al, 1977), it was decided to consider the chemical space of these other protein targets. The serine-protease enzymes involved in the coagulation cascade, such as the Factor Xa and VII enzymes have been reported to be inhibited by Pentamidine (V. A. Johnson et al, 1976; S. J. Kempin et al, 1977; B. Gabriel et al, 1998; R. A. Shirk et al, 2007) and by compounds that contain a 2-aminopyridine functional group (V. Pandya et al, 2012; P. B. Choudhari et al, 2013; M.

Feroci et al, 2014; H. Nishida et al, 2002). A substructure search on the 2-aminopyridine functional group was performed to investigate if compounds containing this moiety exhibited similar inhibitory action towards the enzymes of the polyamine biosynthetic pathway.

In this chapter, the methodology used to screen the NCI database is further described in more detail. The SAR hypothesis generated on the quinazoline inhibitors of NF-κB and the docking hypothesis of Pentamidine in the AdoMet-DC crystal structure are discussed further in more detail.

## 3.2 Methodology of the LBVS performed with QNZ

### 3.2.1 Preparation of the NCI Database

The structure information for approximately 265242 compounds was downloaded from the NCI website (http://cactus.nci.nih.gov/download/nci/index.html). Taking into consideration that the NCI database contains compounds that have been previously tested for *in vitro* activity, it was not filtered with the OpenEye drug like filter (Filter, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2010).

### 3.2.2 Generation of Ligand and Database Conformations

In preparation for the 3D similarity search and pharmacophore search, conformations of QNZ and the NCI database compounds were generated using Confgen, an application in the Schrodinger suite (ConfGen, version 5.6, Schrödinger, Inc., New York, NY, 2009). The standard protocol on ConfGen was selected and default parameters were maintained. The fast search protocol was utilised, where conformers are considered redundant at a threshold of an RMSD of 1Å (K. S. Watts et al, 2010). The input and output structures were minimised.

### 3.2.3 Analysis of SAR Data from Literature

The SAR data for QNZ collected from literature included the series of quinazoline analogues generated by M. Tobe et al, 2003. An analysis of the data revealed that activity was mainly driven by log P and molecular weight and not by specific protein-ligand interactions. This was apparent when a graph of activity versus log P and molecular weight was plotted, where a general trend of increasing activity with increasing log P and molecular weight was observed (Figure 2).

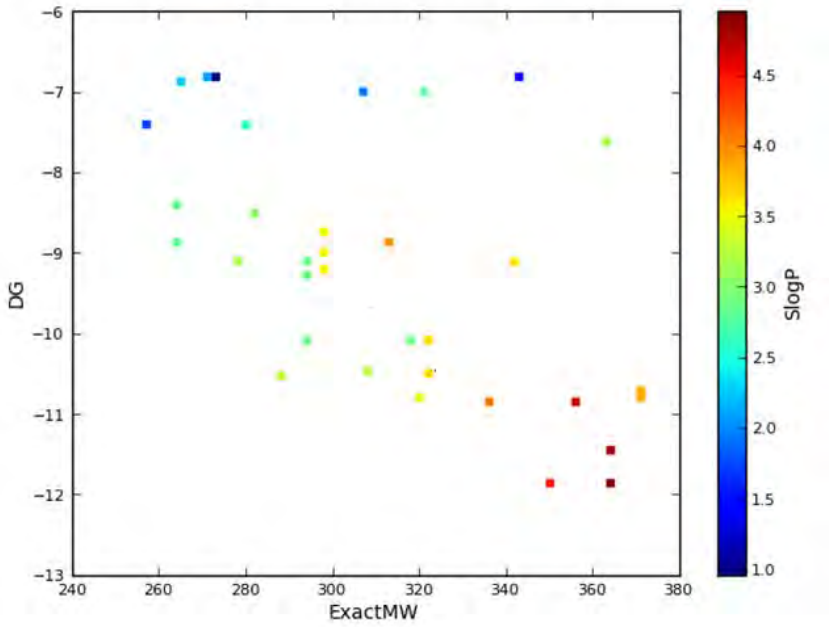



**Figure 2: A graph of Log P vs molecular weight vs the in vitro activity (DG) of the quinazoline analogues obtained from literature (M. Tobe et al, 2003) (M. Tobe et al, 2003). A general trend of increasing activity with increasing molecular weight and log P was observed. The colour coding represents the predicted atomic Log P (SlogP) values for each analogue.**

Further analysis of the SAR data revealed that a methoxy to phenoxy substitution on the terminal phenyl group, resulted in an unexpectedly low increase of activity of approximately 9 fold (M. Tobe et al, 2003), which substantiated the hypothesis that activity was not driven by specific binding interactions. A summary of the SAR data of the quinazoline inhibitors of NF-κB was generated to help guide the selection of compounds for *in vitro* testing and in formulating the hypothesis for the pharmacophore VS (Figure 3) (M. Tobe et al, 2003; M. Tobe et al, 2003).

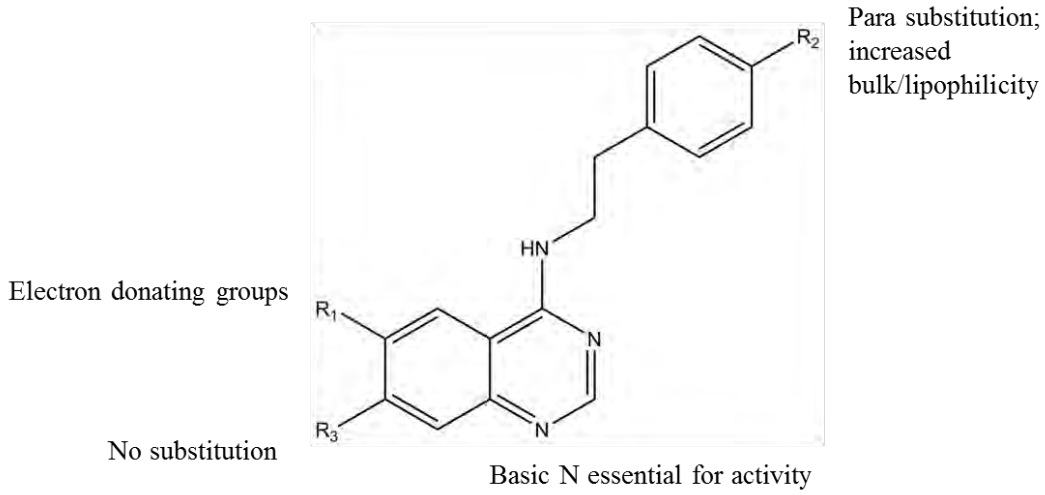



**Figure 3: A summary of the SAR data of the quinazoline inhibitors of NF-κB extracted from literature (M. Tobe et al, 2003; M. Tobe et al, 2003). The structural features and substitutions which lead to increased activity are highlighted in the diagram.**

### 3.2.4 Receiver Operating Characteristic (ROC) Curve Validation

As mentioned previously, the ROC curve analysis was performed to assess how good the protocols implemented on vROCS (ROCS, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2010) and Pipeline Pilot (Accelrys Software Inc., *Discovery Studio Modeling Environment, 4.5*, San Diego: Accelrys Software Inc., 2007) were at discriminating between the active and inactive compounds in a dataset (N. Triballeau et al, 2005). The guidelines used to perform a ROC curve analysis were reported by N. Triballeau et al, 2005 and are summarized in table 1 below.

**Table 1: The protocol used to perform the ROC curve analysis as defined by N. Triballeau et al, 2005.**

| STEP | OBJECTIVE |
|---|---|
| 1) Choice of activity cut off | Activity cut off to ensure relevance and usefulness of computer test |
| 2) Selection of a suitable sample of actives and inactives compounds | Include SAR on relevant target for assessment |
| 3) VS performed on the sample molecules | Evaluation of compounds with known activity by computer test |
| 4) ROC curve plotting and analysis & statistical evaluation | Evaluate performance of computer algorithm with and define selection threshold. |

The selection of a suitable sample of actives and inactives compounds was identified by considering the SAR data available in literature on quinazoline inhibitors of NF-κB (M. Tobe et al, 2003; M. Tobe et al, 2004). Taking into account that activity was driven by the physicochemical properties of the quinazoline series, the SAR data was analysed in a more representable and informative approach by clustering the data according to Log P, log P and molecular weight, the ECFP_4 fingerprint and by structure. This assisted in identifying the similarities between the compounds to select a representative sample of actives and inactives to perform the ROC curve validation. Within each cluster, the compounds with activity within the range of 1-1000nM were assigned as active and those with an activity in the range of 1000-10 000nM or greater, were assigned as inactive. These activity cut-off values ensured that the activity cliffs within the series of quinazoline inhibitors were assigned as inactive compounds. The 'ROC curve' protocol was selected on Discovery Studios to validate the 2D similarity search (Accelrys Software Inc., *Discovery Studio Modeling Environment, 4.5*, San Diego: Accelrys Software Inc., 2007). The 3D similarity search was validated using the 'ROCS validation run' on vROCS (ROCS, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2010). Default parameters were maintained. A ROC curve was generated for each cluster, with both the 2D and 3D protocol and a statistical validation was performed to ensure that a relationship between the structure and activity actually existed and was not due to chance (N. Triballeau et al, 2005). The randomisation trials were performed four times, due to time restraints, using the same protocols on Pipeline Pilot and vROCS, to

generate the ROC curves. The sensitivity and specificity of the computer test performance on each cluster was calculated.

### 3.2.5 2D Similarity Search

The Discovery Studios 'Find similar molecules by fingerprints' protocol on Pipeline Pilot was used to conduct the similarity search (Accelrys Software Inc., *Discovery Studio Modeling Environment, 4.5*, San Diego: Accelrys Software Inc., 2007). The input ligands included the NCI database and the reference ligand included QNZ. The minimum similarity was set to 0.50 and the Tanimoto coefficient was selected with a pre-defined property of the ECFP_4 fingerprint. The ECFP_4 fingerprint describes both the identity and connectivity of an atom in a molecule and searches for similar descriptors within a diameter of 4 bonds radiating from each heavy atom (D. Rogers et al, 2010; Y. Hu et al, 2009).

### 3.2.6 3D Similarity Search

A simple run on the ROCS interface (vROCS) was employed to perform the similarity search on the NCI database (ROCS, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2010). Default parameters were maintained for the inputs, where the colour force field was kept at the default Implicit Mills-Dean. The SDF for QNZ (query) and the 3D conformations of the NCI database were uploaded for screening. The number of best hits to retain was increased to 1000 compounds.

### 3.2.7 3D Pharmacophore Generation

The aim of the pharmacophore screening was to test the hypothesis of which structural features on the QNZ scaffold are crucial for biological activity. Therefore, the pharmacophores were manually generated using the Discovery Studios 'Create pharmacophore manually' protocol on Pipeline Pilot (Accelrys Software Inc., *Discovery Studio Modeling Environment, 4.5*, San Diego: Accelrys Software Inc., 2007). The structural

features that were selected on the QNZ scaffold to generate the 4 pharmacophore hypothesis are depicted in Figure 4, below.

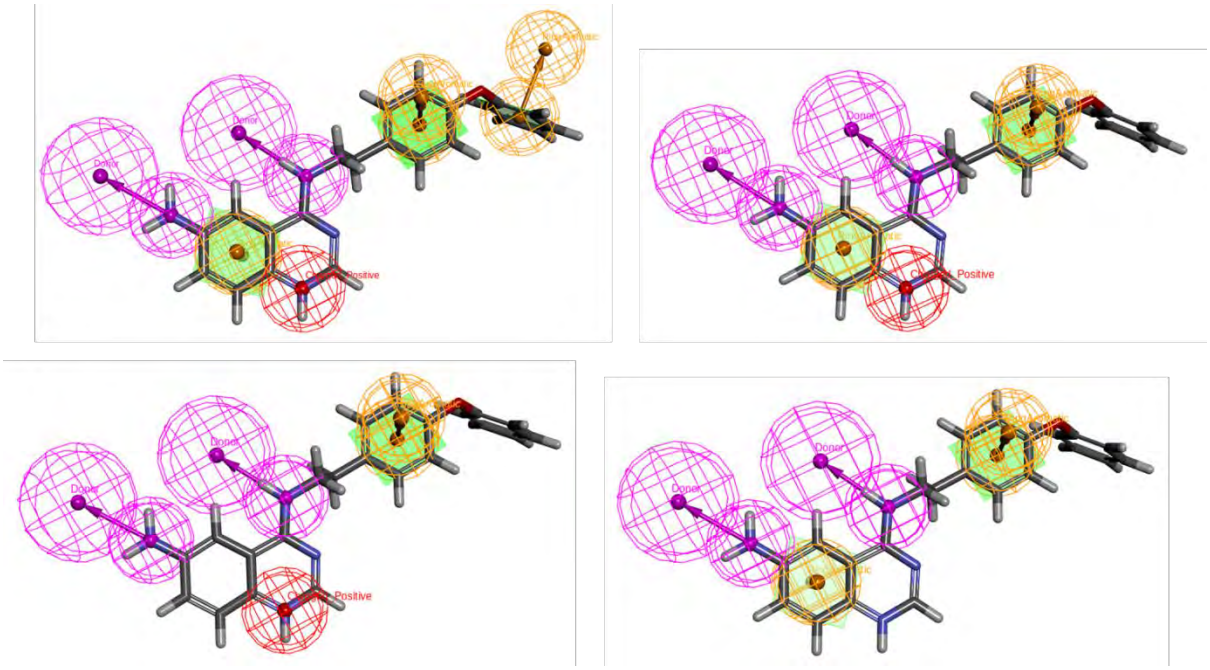


**Figure 4: The manual pharmacophores generated on Discovery Studios, based on the QNZ scaffold. The first pharmacophore (top left) identifies all the features of QNZ including the 2 hydrogen donors (pink), the aromatic moieties (orange) and the positively charged nitrogen (red). The second pharmacophore (top right) tests the importance of the terminal aromatic of the phenoxy phenyl moiety. The third pharmacophore (bottom left) excludes the aromatic moiety on the quinazoline core and the fourth pharmacophore (bottom right) excludes the positively charge atom on the quinazoline core. The phenoxy phenyl aromatic ring was omitted from the third and fourth pharmacophore, based on its poor calculated binding efficiency.**

The first pharmacophore identified all the moieties on the QNZ scaffold hypothesised to be important for activity. The second pharmacophore hypothesis excluded the second aromatic group of the phenoxy phenyl moiety, as it was predicted to have a low binding efficiency calculated on Glide5.7. This was further substantiated by the SAR data reported in literature, which indicated a higher $IC_{50}$ value for the compounds that possessed a functional moiety, other than a lipophilic group, at this position (M. Tobe et al, 2003; M. Tobe et al, 2003). The third pharmacophore excluded the aromatic moiety on the quinazoline core and the fourth pharmacophore excluded the positively charged group on the same quinazoline core. Locational constraints were added to define the relative location of these features.

### 3.2.8 3D Pharmacophore Virtual Screening

The conformations of the NCI database were uploaded on the Discovery Studios Server, using the 'Build 3D Database' protocol (Accelrys Software Inc., *Discovery Studio Modeling Environment, 4.5*, San Diego: Accelrys Software Inc., 2007). Default parameters were maintained. Each pharmacophore was used to screen the 3D database using the 'Search 3D Database' protocol on Discovery Studios (Accelrys Software Inc., *Discovery Studio Modeling Environment, 4.5*, San Diego: Accelrys Software Inc., 2007). The NCI database was selected and default parameters were maintained.

## 3.3 QNZ Results

### 3.3.1 QNZ ROC Curve Validation

The area under the curve (AUC) was determined for both the 2D and 3D similarity search protocols on Pipeline Pilot and vROCS, respectively. The specificity (true negative rate) of the 2D and 3D similarity search protocols, performed with each cluster was calculated using the equation below.

*Specificity (Sp) = number of true negatives (TN) / sum of TN + false positives (FP*

The calculated specificity of a test is the percentage of true inactives being selected and discarded and can be summarized as

$$Sp = N_{discarded\ inactives} / N_{total\ inactives}$$

The specificity of a test can range between 0-1 and a value closer to 0 indicates a high specificity which indicates that a test is better at discarding in active compounds (N. Triballeau et al, 2005).

The sensitivity (true positive rate) of the 2D and 3D similarity search protocols, performed with each cluster was calculated using the following equation below:

*Sensitivity (Se) = number of true positives (TP) / sum of TP + false negatives (FN)*

The calculated sensitivity of a test is the percentage of true actives being selected by the test and can be summarized below

$$Se = N_{selected\ active} / N_{total\ actives}$$

Similarly the sensitivity of a test can range from 0-1, where a high sensitivity is closer to 0. A high sensitivity is indicative of a test that is able to select actives from the inactive compounds (N. Triballeau et al, 2005).

The AUC, sensitivity and specificity obtained with each cluster using the 2D ROC curve protocol on Discovery Studios, are summarised in Table 2 below. The datasets, ROC curves and randomisation ROC curves, for each cluster can be found in Appendix 2.1.

**Table 2: A summary of the results obtained from the 2D ROC curve validation performed on Discovery Studios. The number of active and inactive compounds in each dataset is denoted.**

| Cluster | No. of actives | No. of in actives | AUC | Se | Sp | Statistical significance |
|---|---|---|---|---|---|---|
| **ECFP_4** | 22 | 22 | 0.711 | 0.818 | 0.681 | No |
| **Log P** | 15 | 13 | 0.855 | 0.70 | 0.444 | No |
| **Log P & Mw** | 7 | 6 | 0.762 | 0.50 | 0.714 | No |
| **Structure** | 28 | 23 | 0.860 | 0.821 | 0.304 | No |

A higher AUC indicates a higher probability that the test can discriminate between active and inactive compounds (N. Triballeau et al, 2005). The ROC curves generated with each cluster on the 2D similarity search protocol on Discovery Studios, all have good AUC but none of the clusters were found to have a statistically significant relationship between structure and activity (Table 2). The AUCs obtained for each cluster on the 3D similarity search protocol implemented on vROCS, are poor with the exception of the Log P and molecular weight cluster, although it failed to achieve statistical significance (Table 3).

**Table 3: A summary of the results obtained from the 3D ROC curve validation performed on vROCS. The number of active and inactive compounds in each dataset is denoted.**

| Cluster | No. of actives | No. of in actives | AUC | Se | Sp | Statistical significance |
|---|---|---|---|---|---|---|
| **ECFP_4** | 22 | 22 | 0.566 | 0.454 | 0.454 | Yes |
| **Log P** | 15 | 13 | 0.630 | 0.733 | 0.307 | Yes |
| **Log P & Mw** | 7 | 6 | 0.944 | 1 | 0 | No |
| **Structure** | 28 | 23 | 0.463 | 0.428 | 0.652 | No |

The ECFP_4 and Log P clusters were the only two clusters that had a statistically significant relationship between structure and activity. The ECFP_4 cluster had an equal trade-off between sensitivity and specificity. The Log P cluster has a sensitivity of approximately 73% and a specificity of 30%, which would suggest that the 3D similarity search protocol is better able to select actives than discard inactive compounds. This can be interpreted as the computer test having a reasonably high false negative rate. The datasets, ROC curves and randomisation ROC curves for each cluster can be found in Appendix 2.1.

### 3.3.2 QNZ Similarity Search

The 2D similarity search performed on Pipeline Pilot identified a total number of 8 compounds that had a quinazoline core. Despite this similarity to QNZ, these compounds were not ranked with the highest Tanimoto score. On Pipeline Pilot, the Taminoto coefficient was used to measure similarity which is described in Chapter 1, section 1.3.1. A maximum Tanimoto score of 1 can be obtained. The overall highest Tanimoto score obtained on Pipeline Pilot was 0.453, which is approximately 45% similarity to QNZ (Figure 5). The top scoring quinazoline compound, which is also depicted in Figure 5, had a Tanimoto score of 0.431. The remaining compounds had 5- and 6- membered ring systems with varied lengths of alkyl chains, with an aromatic moiety attached to them. From the final selection of top ranking compounds, a total of 8 compounds were available and requested from the NCI for *in vitro* testing, which can be found in Appendix 2.2, Table 1.

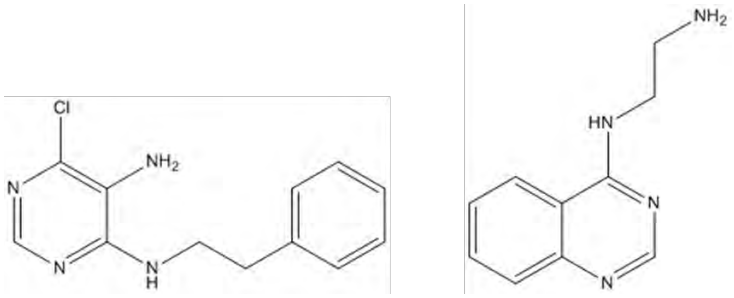



**Figure 5: The top scoring compound, NCS100126 (left) and the top scoring quinazoline containing compound, NSC403389 (right), identified from the 2D similarity search. Their Tanimoto scores are 0.453 and 0.431, respectively.**

The 3D similarity search performed on vROCS identified a fewer number of quinazoline-containing compounds compared to the 2D similarity search. On vROCS, a TanimotoCombo score (maximum score of 2) is calculated to measure similarity and is based on the sum of the individual ColourTanimoto (functional group similarity) and ShapeTanimoto (volume) scores, described in Chapter 1, section 1.2.1. The 4 compounds identified with a quinazoline core were also not assigned with the highest similarity by vROCS. The highest TanimotoCombo score obtained was 1.246, which is approximately 62% similar to QNZ (Figure 6). The top scoring quinazoline containing compound had a Tanimoto Combo score of 0.964 which is approximately 48% similar to QNZ (Figure 6).

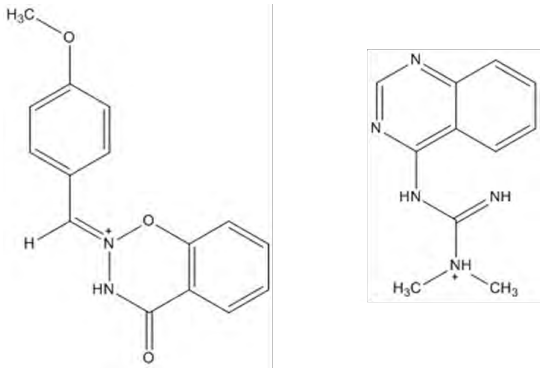



**Figure 6: The top scoring compound, NSC204382 (left), and the top scoring quinazoline containing compound, NSC403396 (right), identified from the 3D similarity search. Their TanimotoCombo scores calculated on vROCS were 1.246 and 0.964, respectively.**

The assignment of a lower TanimotoCombo score for the compound NSC403396 (Figure 6), maybe explained by the colour force field employed on vROCS. The default Implicit Mills-Dean force field utilises a simple pKa model at a pH of 7, where it identifies anions,

cations, donors and acceptors independently of their already assumed protonated or deprotonated states (OpenEye Scientific Software, 2013, https://docs.eyesopen.com/rocs/usage.html#color-force-field). Therefore, the guanidine group of NSC403398 may be preferentially protonated instead of the quinazoline core. This may further be substantiated by the calculated ColourTanimoto score obtained for NSC403396 (0.416) which was lower than the top scoring compound NSC100126 (0.536), which would suggest that the predicted chemical feature similarity of QNZ and NSC403398 was not optimal compared to NSC100126. The remaining compounds identified from the 3D similarity search consisted of 5- and 6-membered ring systems with a phenyl ethyl aromatic moiety attached to them.

A cluster analysis, based on the ECFP_4 fingerprint and on Log P and molecular weight, was performed with the compounds identified from the 3D similarity search. This and the summarised SAR were used to visually analyse and select compounds for *in vitro* screening. A total of 7 compounds were available from the NCI, which are tabulated in Appendix 2.2, Table 2.

### 3.3.3 QNZ Pharmacophore Virtual Screening

The four pharmacophores generated were used to screen the NCI database. No compounds were retrieved from the virtual screening performed with Pharmacophore 1 and 2. Pharmacophore 4 retrieved compounds that overall had better fit values than the compounds retrieved by the Pharmacophore 3 hypotheses. The fit value is calculated in Discovery Studio and consists of a geometrical fitting of the pharmacophore hypothesis to the compounds in the database. A higher fit value indicates a good overlay between the pharmacophore features and those identified on a compound.

The compounds retrieved from the pharmacophore virtual screening were structurally diverse and did not contain a quinazoline core. A visual analysis of the compounds overlaid on their respective pharmacophores was performed to identify compounds for *in vitro* screening, alongside considering their respective fit values. Furthermore, the compounds were visually analysed to determine if they contained the appropriate functional groups hypothesised to lead to increased activity (Figure 3). Those compounds that were excluded either did not contain the appropriate functional groups or they did not form a good overlay with the

pharmacophore locational constraints. A total of 4 compounds resulting from the pharmacophore 3 and 10 compounds identified from the Pharmacophore 4 virtual screening, were requested from the NCI. The structures and fit values of these compounds can be found in Appendix 2.2, Figure 1 and Figure 2.

### 3.4 QNZ Discussion

The chemoinformatic, knowledge driven virtual screening approach entailed a detailed analysis of the SAR data from literature. This led to the hypothesis that activity of the quinazoline analogues, including QNZ, was not driven by specific molecular interactions but instead by molecular weight and Log P. Furthermore, the hypothesis was consistent with the SAR data, which revealed that a lipophilic or bulky substitution at the phenylethyl aromatic ring would result in increased activity. In addition, the hypothesis also correlated with the predicted binding pose of QNZ in the binding sites identified through the blind docking described in Chapter 2. It was predicted that the lipophilic phenoxy moiety of QNZ was buried in the hydrophobic pockets of the binding sites identified, which suggests that increasing the lipophilic interactions in this region may result in increased activity.

The results obtained from the ROC curve validation were considered relatively poor and a consequence of the limited SAR data available in literature. The SAR data was limited in terms of the structural diversity of the analogues and the size (number of compounds) of the datasets used to perform the ROC curve analysis. It is recommended that datasets (of active and inactive compounds) should contain equal numbers of structurally diverse compounds, which would provide a comprehensive coverage of the chemical space regarding the SAR of a particular target (N. Triballeau et al, 2005). Nevertheless, a cluster analysis on the ECFP_4 fingerprint and Log P and molecular weight was still performed on the compounds identified from the virtual screening.

The 2D and 3D similarity search methods performed on the NCI database, did not retrieve a large number of quinazoline containing compounds. A larger number of quinazoline containing compounds were identified by the 2D methodology compared to the 3D, which may support the general preference for 2D descriptors (V. Venkatraman et al, 2010; M. Thimm et al, 2004; J. Bajorath, 2002).

The pharmacophore hypothesis aimed at rationalising the summarised SAR and took into consideration the important features that were identified (Figure 3). The pharmacophoric features of QNZ which were identified include the 2 hydrogen bond donors, 3 hydrophobic aromatic groups and 1 positively charged atom. These features differed from a previously reported pharmacophore developed from the same SAR data that was considered in this study (K. C. Tsai et al, 2009). The common feature pharmacophore, developed by K.C. Tsai et al, 2009, consisted of 1 hydrogen bond acceptor, 1 hydrophobic aromatic group and 3 hydrophobic groups. A common feature pharmacophore hypothesis is highly dependent on the training set (known actives) used to identify their common pharmacophoric features. The contents of the training set, in terms of the number of active compounds and their chemical diversity, play an important role in the final outcome of a common feature pharmacophore hypothesis (S. Y. Yang, 2010). The influence of the dataset used to generate a common feature pharmacophore hypothesis has been previously established, where using slightly different training sets on the same biological target produces a different pharmacophore hypothesis, regardless of whether or not the same programme was used to generate the pharmacophore (S. Y. Yang, 2010). Although, a common feature pharmacophore has had many successes (D. Schuster et al, 2006; A. Arooj et al, 2013), it was not an appropriate method considering the aims of this study. Furthermore, the quality of the SAR data obtained from literature was viewed as a limiting factor to produce a robust common feature pharmacophore.

## 3.5 Methodology of the LBVS performed with Pentamidine

### 3.5.1 Analysis of SAR Data from Literature

A major limitation in obtaining a sufficient and relevant amount of SAR data from literature on Pentamidine was experienced. The limitation arose in obtaining SAR that originated from cancer cells lines, with the added complexity of whether or not their activity was derived from cell-based assays versus target-based assays. Regardless of these limitations, SAR data extracted from literature on inhibitors of AdoMet-DC and DAO enzymes, which were derived from human cell and target based assays, were analysed (U. Regenass et al, 1992; U. Regenass et al, 1994; J. Stanek et al, 1993; I. Jarak et al, 2011). A cluster analysis based on

the ECFP_4 fingerprint, molecular weight and Log P, Log P and structure was performed, which revealed that Pentamidine did not share any of the similarities to the other derivatives. This was apparent as Pentamidine was clustered on its own, which furthermore ruled out the possibility of performing the ROC curve validation along with the obtaining insightful information on the SAR, to assist in selecting compounds for *in vitro* screening.

### 3.5.2 Docking of Pentamidine in the AdoMet-DC Enzyme

In order to assist in identifying similar compounds to Pentamidine, a docking hypothesis of Pentamidine in the AdoMet-DC crystal structure was generated (3H0W.pdb) (S. Bale et al, 2009). The ligand in the 3H0W.pdb file was removed and the protein structure was prepared using the Protein Preparation Wizard on Maestro (Maestro, version 5.6, Schrödinger, Inc., New York, NY, 2009), retaining the water molecules found in the PDB file. A pre-processing step using default parameters preceded the hydrogen bond assignment using PROPKA. A restrained minimisation step was performed using the OPLS_2005 force field to ensure that the hydroxyl side chains were properly orientated. This improved on the possible steric clashes and the structure was further restrained to the default RMSD tolerance (0.3Å) compared to the original protein coordinates. The Receptor Grid Generation application on Glide5.7 (Glide, version 5.7, Schrödinger, Inc., New York, NY, 2009) was used to identify the binding site by selecting the ligand, 5'-deoxy-5'-(dimethylamino)-8-methyladenosine (DMAMA), in the original 3H0W.pdb structure. The residues in the active site (Glu247, Phe223, Phe7 and Glu67) were further minimised with explicit water molecules using the MacroModel application on Maestro, to suitably re-orientate the water molecules in the binding pocket. The side chains of the serine residues (Ser229 and Ser66) and the histidine residue (His243) were rotated in order to maximise their interactions with the ether oxygens in the linker chain of Pentamidine. Furthermore, the Cysteine residue (Cys226) was deprotonated to maximise its interaction with the protonated amidine group of Pentamidine. The Glide XP docking protocol was employed to carry out the ligand docking calculations, which in addition to the empirical scoring function applies large desolvation penalties to top ligand poses by assessing the explicit water molecules. Furthermore, Glide XP is able to identify structural motifs that significantly contribute to binding affinities, such as the formation of a π-cation interaction (R. A. Friesner et al, 2006). The ligand, DMAMA, was

removed from the 3H0W.pdb file before docking Pentamidine. Default parameters were maintained. (Docking hypothesis was performed with the assistance of Dr Alexander Alex).

### 3.5.3 Sub-structure Search

A sub-structure search is based on an exact sub-structure or functional group that is hypothesised to be responsible for a desirable property or increased activity. It differs from a 2D or 3D similarity search in that a sub-structure search includes a predefined group, which is retrieved from a database (S. S. Ji et al, 2015). The sub-structure search was performed on Discovery Studios, using the 'Align to selected substructure' protocol (Accelrys Software Inc., *Discovery Studio Modeling Environment, 4.5*, San Diego: Accelrys Software Inc., 2007). Default Parameters were maintained. The sub-structures, benzamidine and 2-aminopyridine were drawn using ChemBioDraw Ultra 14, which were saved in a structure data file format.

## 3.6 Pentamidine Results

### 3.6.1 Pentamidine Docking Hypothesis in AdoMet-DC Crystal Structure

The binding hypothesis of Pentamidine in the AdoMet-DC crystal structure reveals that Pentamidine does not fully occupy the binding site which may account for its low potency, compared to the other inhibitors of AdoMet-DC (D. E. McCloskey et al, 2009). The amidine group of the buried benzamidine moiety of Pentamidine is predicted to form hydrogen bonds with the charged Cysteine (Cys) 226 residue. The second buried amidine group, which is charged, is predicted to form a hydrogen bond with the main chain carbonyl group of the Phenylalanine (Phe) 7 protein residue (Figure 7). The amidine group of the benzamidine moiety that is solvent exposed is predicted to form 2 hydrogens bonds to the Serine (Ser) 66 and Glutamic acid (Glu) 67 residue side chains (Figure 7). The ether oxygens are predicted to interact with conserved water molecules in the binding site.

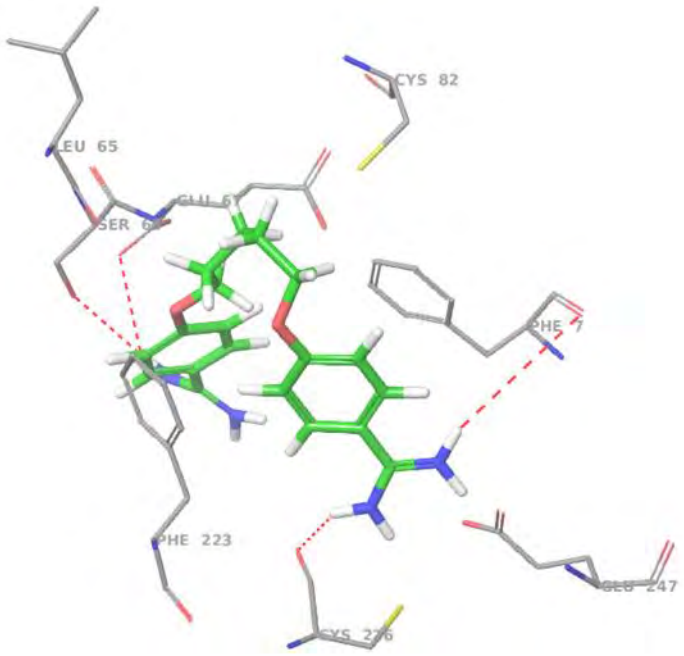



**Figure 7: The predicted binding interactions of Pentamidine in the human AdoMet-DC crystal structure (3H0W.pdb). The buried amidine groups of Pentamidine form hydrogen bonds with the Phe7 and Cys226 hydrogen bonds (hydrogen bonds shown in red). The solvent exposed amidine group of Pentamidine forms hydrogen bonds with the sidechains of Ser66 and Glu67 residues of the AdoMet-DC enzyme.**

### 3.6.2 Pentamidine Sub-structure Search

The substructure search for the benzamidine moiety retrieved only 3 compounds from the NCI database, none of which were available from the NCI for *in vitro* testing. A total of 11 compounds retrieved from the 2-aminopyridine sub-structure search were obtained from the NCI, which can be found in Appendix 2.2, Figure 3.

## 3.7 Pentamidine Discussion

An analysis of the SAR data obtained from literature, regarding Pentamidine and other inhibitors of the human AdoMet-DC and DAO enzymes proved futile. None of the analogues

shared similar properties to Pentamidine and therefore no informative conclusions could be derived from the SAR data obtained from literature.

The binding hypothesis of Pentamidine in the human AdoMet-DC crystal structure mimics the key interactions of other AdoMet-DC inhibitors that have been identified in literature (S. Bale et al, 2009; D. E. McCloskey et al, 2009). The substrate analogues and inhibitors of AdoMet-DC reportedly have a sulfonium atom which has been shown to be essential for ligand binding and inhibition. SAR studies reveal that replacing the sulfur atom with a nitrogen atom that is charged at a physiological pH also confers inhibition. Importantly, the compounds that do not have a charged group or a sulfur atom in this position do not inhibit the AdoMet-DC enzyme (S. Bale et al, 2009) (Figure 8).

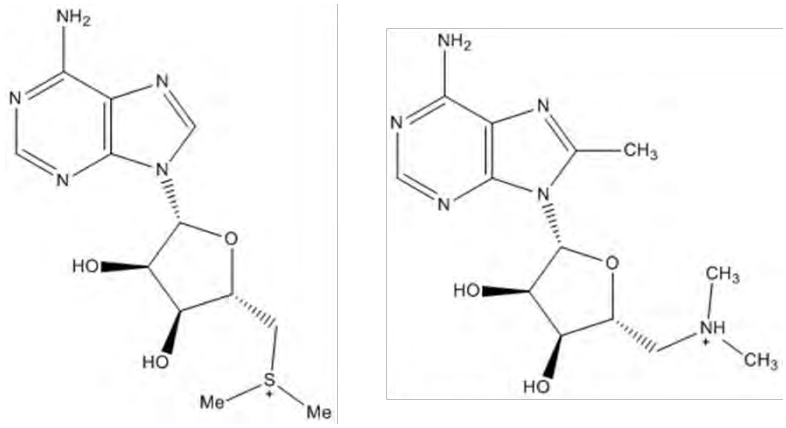



**Figure 8: The structure of 5'-deoxy-5'-(dimethylsulfonio) adenosine (MMTA) (left) and DMAMA (right). The sulfonium atom in MMTA and the positively charged nitrogen atom in DMAMA are found to be critical for inhibition of the AdoMet-DC enzyme. Compounds that do not possess a sulfur atom or charged atom at this position are found to not inhibit the AdoMet-DC enzyme (S. Bale et al, 2009).**

The role of the positively charged atom is thought to align the ligand in the favourable *syn* conformation and aid in the formation of the key π-cation interaction (S. Bale et al, 2009). The π-cation interaction has been identified to be responsible for ligand specificity, but also acts to stabilise the *syn* conformation of the ligand (S. Bale et al, 2009). The crystal structure of DMAMA in complex with the AdoMet-DC enzyme indicates that the ribose forms 2 hydrogen bonds with the Glutamic acid (Glu) residue 247. The positively charged nitrogen atom of DMAMA is at a favourable distance to stack between the Phe7 and Phe223 residues and forms the π-cation interaction, that stabilises DMAMA in the *syn* conformation (S. Bale

et al, 2009; D. E. McCloskey et al, 2009) (Figure 9). The conformation of Pentamidine in the binding hypothesis generated indicates that the buried benzamidine moiety stacks between the Phe7 and Phe223 residues. This would allow the charged amidine group to form a π-cation bond with Phe7, although a hydrogen bond formation is predicted (Figure 9). The prediction that the stabilising π-cation interaction is not formed may be consistent with the reduced potency of Pentamidine, compared to the other inhibitors of AdoMet-DC enzyme (D. E. McCloskey et al, 2009). Furthermore, it may be hypothesised that the large number of rotatable bonds in Pentamidine may not allow for the formation of the stabilising π-cation interaction, as such, the increased conformational entropy may contribute to the reduced binding affinity of Pentamidine. In addition, Pentamidine is reported in literature to have poor *in vitro* and *in vivo* permeability compared to other diamines, which may contribute to the reduced potency of Pentamidine compared to the other inhibitors of AdoMet-DC (S. Yang et al, 2014). The second amidine group of the buried benzamidine moiety is predicted to form a hydrogen bond with the Cys226 residue, which is unlike the compounds co-crystallised with the enzyme (S. Bale et al, 2009). It may be hypothesised that this amidine group is not necessary for binding affinity, in addition to the amidine group of the benzamidine moiety that is solvent exposed. The second amidine group that is solvent exposed is predicted to form a hydrogen bond with the side chains of Glu67 and Ser66. The interaction of the ether oxygens with the water molecules is plausible, as other inhibitors of AdoMet-DC have also been found to interact with conserved water molecules in the active site, stabilising their hydrogen bond complement (S. Bale et al, 2009; D. E. McCloskey et al, 2009).

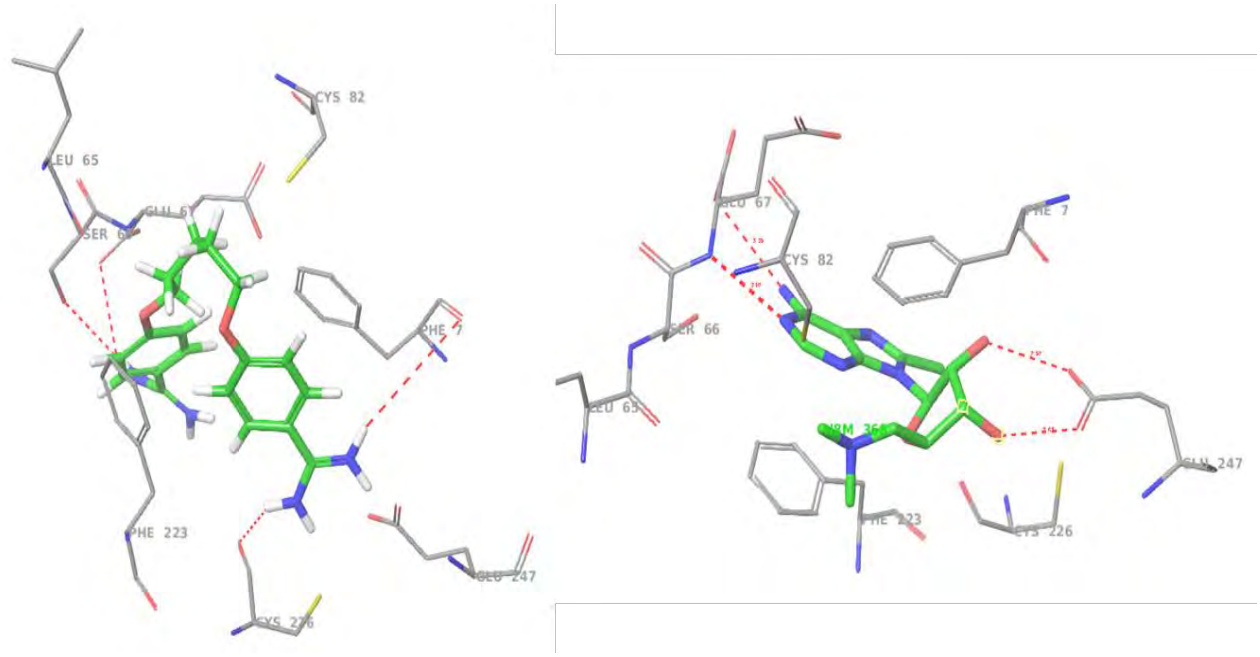



**Figure 9: The predicted binding interactions of Pentamidine in the AdoMet-DC crystal structure (left; hydrogen bonds shown in red) and the binding interaction of DMAMA co-complexed with the human AdoMet-DC crystal structure (3H0W.pdb) (right; hydrogen bonds shown in red) (S. Bale et al, 2009). The functional moieties of Pentamidine are not all predicted to interact with the enzyme and furthermore the stabilising π-cation interaction was not predicted (left). This may account for the reduced potency of Pentamidine compared to other inhibitors of the AdoMet-DC enzyme (S. Yang et al, 2014).**

A total of 11 compounds were identified from the 2-aminopyridine substructure search, for *in vitro* screening. The virtual screening workflow for Pentamidine was limited and without resorting to procuring any compounds from commercially available databases, it was concluded with the sub-structure search performed on the NCI database.

### 3.8 Conclusion

A chemo-informatic, knowledge driven virtual screening of the NCI database, consisting of a ligand based virtual screening was performed for QNZ and Pentamidine. The analysis of SAR data available in literature for QNZ, resulted in a summary of the key structural features and furthermore led to the hypothesis that activity of QNZ and the quinazoline analogues is driven by lipophilic interactions and not by specific molecular interactions. Four

pharmacophore hypotheses were developed to test the importance of the structural features of the quinazoline inhibitors of NF-κB hypothesised to be important for activity. A binding hypothesis of Pentamidine in the human AdoMet-DC crystal structure was generated, which indicated the key interactions that Pentamidine may form with the enzyme. A total number of 41 compounds were requested from the NCI database for *in vitro* testing in ovarian, renal and prostate cancer cell lines, described in Chapter 4.

## 3.9 References


1. N. Triballeau, F. Acher, I. Brabet, J. P. Pin and H. O. Bertrand, Virtual Screening Workflow Development Guided by the "Receiver Operating Characteristic" Curve Approach. Application to High-Throughput Docking on Metabotropic Glutamate Receptor Subtype 4, Journal of Medicinal Chemistry, Vol. No. 48, Issue No. 7, Pp. 2534-2547, 2005.
2. M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, H. Takahashi, T. Fukazawa and H. Hayashi, Discovery of Quinazolines as a Novel Structural Class of Potent Inhibitors of NF-κB Activation, Bioorganic & Medicinal Chemistry, Vol. No. 11, Pp. 383-391, 2003.
3. M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, H. Takahashi and H. Hayashi, A Novel Structural Class of Potent Inhibitors of NF-κB Activation: Structure–Activity Relationships and Biological Effects of 6-Aminoquinazoline Derivatives, Vol. No. 11, Pp. 3869-3878, 2003.
4. L. Xu and W. A. Russu, Molecular Docking and Synthesis of Novel Quinazoline Analogues as Inhibitors of Transcription Factors NF-κB Activation and their Anti-cancer Activities, Bioorganic & Medicinal Chemistry, Vol. No. 21, Pp. 540-546, 2013.
5. A. Paasinen-Sohns, M. Kielosto, E. Kaariainen, T. Eloranta, A. Laine, O. A. Janne, M. J. Birrer and E. Holtta, c-Jun Activation-dependent Tumorigenic Transformation Induced Paradoxically by Overexpression or Block of *S*-Adenosylmethionine Decarboxylase, The Journal of Cell Biology, Vol. No. 151, Issue No. 4, Pp. 801-809, 2000.
6. D. T. Koomoa, T. Borsics, D. J. Feith, C. C. Coleman, C. J. Wallick, I. Gamper, A. E. Pegg and A. S. Bachmann, Inhibition of S-adenosylmethionine decarboxylase by inhibitor SAM486A connects polyamine metabolism with p53-Mdm2-Akt/protein kinase B regulation and apoptosis in neuroblastoma, Molecular Cancer Therapy, Vol. No. 8, Issue No. 7, Pp. 2067-2075, 2009.
7. J. L. Ekstrom, I. I. Mathews, B. A. Stanley, A. E. Pegg, S. E. Ealick, The crystal structure of human S-adenosylmethionine decarboxylase at 2.25 Å resolution reveals a novel fold, Structure, Vol. No. 7, Issue No. 5, Pp. 583-595, 1999.

8. U. Regenass, G. Caravatti, H. Mett, J. Stanek, P. Schneider, M. Muller, A. Matter, P. Vertino and C. W. Porter, New S-Adenosylmethionine Decarboxylase Inhibitors with Potent Antitumor Activity, Cancer Research, Vol. No. 52, Pp. 4712-4718, 1992.
9. U. Regenass, H. Mett, J. Stanek, M. Mueller, D. Kramer and C. W. Porter, CGP 48664, a New S-Adenosylmethionine Decarboxylase Inhibitor with Broad Spectrum Antiproliferative and Antitumor Activity, Cancer Research, Vol. No. 54, Pp. 3210-3217, 1994.
10. J. Stanek, G. Caravatti, H. G. Capraro, P. Furet, H. Mett, P. Schneider and U. Regenass, S-Adenosylmethionine Decarboxylase Inhibitors: New Aryl and Heteroaryl Analogues of Methylglyoxal Bis (guanylhydrazone), Journal of Medicinal Chemistry, Vol. No. 36, Issue No. 1, Pp. 46-54, 1993.
11. I. Jarak, M. Marjonovic, I. Piantanida, M. Kralj, G. K. Zamola, Novel pentamidine derivatives: Synthesis, anti-tumor properties and polynucleotide-binding activities, European Journal of Medicinal Chemistry, Vol. No. 46, Pp. 2807-2815, 2011.
12. V. A. Johnson and R. L. Smith, Bovine Factor Xa and Bovine Trypsin: A Comparison with Respect to Inhibition by some Amidines and Guanidines. Archives of Biochemistry and Biophysics, Vol.; No. 176, Issue No. 1, Pp. 190-195, 1976.
13. S. J. Kempin, C. W. Jackson and C. C. Edwards, In Vitro Inhibition of Platelet Function and Coagulation by Pentamidine Isethionate, Antimicrobial Agents and Chemotherapy, Vol. No. 12, Issue No. 4, Pp. 451-454, 1977.
14. B. Gabriel, M. T. Stubbs, A. Bergner, J. Hauptmann, W. Bode, J. Sturzebecher and L. Moroder, Design of Benzamidine-Type Inhibitors of Factor Xa, Journal of Medicinal Chemistry, Vol. No. 41, Issue No. 22, Pp. 4240-4250, 1998.
15. R. A. Shirk and G. P. Vlasuk, Inhibitors of Factor VIIa/Tissue Factor, Arteriosclerosis, Thrombosis and Vascular Biology, Vol. No. 27, Pp. 1895-1900, 2007.
16. V. Pandya, M. Jain, G. Chakrabarti, H. Soni, B. Parmar, B. Chaugule, J. Patel, T. Jarag, J. Joshi, N. Joshi, A. Rath, V. Unadkat, B. Sharma, H. Ajani, J. Kumar, K. V. V. M. Sairam, H. Patel and P. Patel, Synthesis and Structure-Activity Relationship of Potent, Selective and Orally Active Anthranilamide-based Factor Xa Inhibitors: Application of Weakly Basic Sulfoximine Group as Novel S4 Binding Element, European Journal of Medicinal Chemistry, Vol. No. 58, Pp. 136-152, 2012.

17. P. B. Choudhari and M. S. Bhatia, Development of n-phenyl-3-pyridin-2-yl imino Derivatives as Anticoagulants Potential Factor VIIa Inhibitors, Journal of Chilean Chemical Society, Vol. No. 58, Issue No. 2, Pp. 1667-1673, 2013.
18. M. Feroci, I. Chiarotto, G. Forte, G. Simonetti, F. D. D'Auria, L. Maes, D. De Vita, L. Scipione, L. Friggeri, R. Di Santo and S. Tortorella, Efficient Electrochemical *N*-Alkylation of *N*-Boc-Protected 4-Aminopyridines: Towards New Biologically Active Compounds, ISRN Organic Chemistry, Hindawi Publishing Corporation, Volume 2014, Article ID 621592, 2014.
19. H. Nishida, Y. Miyazaki, T. Mukaihira, F. Saitoh, M. Fukui, K. Harada, M. Itoh, A. Muraoka, T. Matsusue, A. Okamoto, Y. Hosaka, M. Matsumoto, S. Ohnishi and H. Mochizuki, Synthesis and Evaluation of 1-Arylsulfonyl-3-piperazinone Derivatives as a Factor Xa Inhibitor II. Substituent Effect on Biological Activities, Chemical and Pharmaceutical Bulletin, Vol. No. 50, Issue No. 9, Pp. 1187-1194, 2002.
20. K. S. Watts, P. Dalal, R. B. Murphy, W. Sherman, R. A. Friesner and J. C. Shelley, ConfGen: A conformational search method for efficient generation of bioactive conformers, Journal Chemical Information and Modelling, Vol. 50, Issue No. 4, 2010.
21. D. Rogers and M. Hahn, Extended-Connectivity Fingerprints, Journal of Chemical Information and Modelling, Vol. No. 50, Issue No. 5, Pp. 742-754, 2010.Y. Hu, E. Lounkine and J. Bajorath, Improving the search performance of extended connectivity fingerprints through activity-orientated feature filtering and application of a bit-density-dependent similarity function, ChemMedChem, Vol. 4, Pg. No. 540-548, 2009.
22. V. Venkatraman, V. I. Perez-Nueno, L. Mavridis and D. W. Ritchie, Comprehensive Comparison of Ligand-Based Virtual Screening Tools against the DUD Data set Reveals Limitations of Current 3D Methods, Journal of Chemical Information and Modelling, Vol. No. 50, Issue No. 12, Pp. 2079-2093, 2010.
23. M. Thimm, A. Goede, S. Hougardy and R. Preibner, Comparison of 2D and 3D Superposition. Application to Searching a Conformational Drug Database, Journal of Chemical Information and Computer Science, Vol. No. 44, Pp. 1816-1822, 2004.
24. J. Bajorath, Integration of Virtual and High Throughput Screening, Nature Review Drug Discovery, Vol. No. 1, Issue No. 11, Pp. 882-894, 2002.
25. K. C. Tsai, L. Teng, Y. Shao, Y. Chen, Y. Lee, M. Li and N. Hsiao, The first pharmacophre model for potent NF-κB inhibitors, Bioorganic and Medicinal Chemistry Letters, Vol. 19, Pg. No. 5665-5669, 2009.

26. S. Y. Yang, Pharmacophore Modeling and Applications in Drug Discovery: Challenges and Recent Advances, Drug Discovery Today, Vol. No. 15, Issue No. 11/12, Pp. 444-450, 2010.
27. D. Schuster, E. Maurer, C. Laggner, L.G. Nashev, T. Wilckens, T. Langer and A. Odermatt, The discovery of new 11beta-hydroxysteroid dehydrogenase type 1 inhibitors by common feature pharmacophore modeling and virtual screening, Journal of Medicinal Chemistry, Vol. No. 49, Issue No. 12, Pp. 3454-3466, 2006.
28. M. Arooj, S. Sakkiah, S. Kim, V. Arulalaperumal and K. W. Lee, A combination of Receptor-based Phamracophore Modeling and QM Techniques for Identification of Human Chymase Inhibitors, PLOS one, Vol. No. 8, Issue No. 4, 2013.
29. L. F. Zerbini, M. K. Bhasin, J. F. De Vasconcellos, J. D. Paccez, X. Gu, A. L. Kung and T. A. Libermann, Computational Repositioning and Preclinical Validation of Pentamidine for Renal Cell Cancer, Molecular Cancer Therapeutics, Vol. No. 13, Issue No. 7, Pp. 1929-1941, 2014.
30. S. Bale, W. Brooks, J. W. Hanes, A. M. Mahesan, W. C. Guida and S. E. Ealick, Role of the Sulfonium Centre in Determining the Ligand Specificity of Human S-AdenisylMethionine Decarboxylase, Biochemistry, Vol. No. 48, Issue No. 27, Pp. 6423-6430, 2009.
31. S. Bale, M. M. Lopez, G. I. Makhatadze, Q. Fang, A. E. Pegg and S. E. Ealick, Structural Basis for Putrescine Activation of Human S-Adenosylmethionine Decarboxylase, Biochemistry, Vol. No. 47, Issue No. 50, Pp. 13404-13417, 2008.
32. R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes, Journal of Medicinal Chemistry, Vol. No. 49, Issue No. 21, Pp. 6177-6169, 2006.
33. S. S. Ji, H. J. Dong, X. X. Zhou, Y. M. Liu, F. X. Zhang, Q. Wang and X. A. Huang, A Structural Hierarchy Matching Approach for Molecular Similarity/Substructure Searching, Molecules, Vol. No. 20, Pp. 8791-8799, 2015.
34. D. E. McCloskey, S. Bale, J. A. Secrist III, A. Tiwari, T. H. Moss, III, J. Valiyaveettil, W. H. Brooks, W. C. Guida, A. E. Pegg and S. E. Ealick, New Insights into the Design of Inhibitors of Human S-Adenosylmethionine Decarboxylase: Studies of Adenosine $C^8$ Substitution in Structural Analogues of S-Adenosylmethionine, Journal of Medicinal Chemistry, Vol. No. 52, Pp. 1388-1407, 2009.

35. S. Yang, T. Wenzler, P. N. Miller, H. Wu, D. W. Boykin, R. Brun and M. Z. Wang, Pharmacokinetic Comparison to Determine the Mechanisms Underlying the Differential Efficacies of Cationic Diamidines Against First-and Second-Stage Human African Trypanosomiasis, Antimicrobial Agents and Chemotherapy, Vol. No. 58, Issue No. 7, Pp. 4064-4074, 2014.

# Chapter 4: Biological Screening Assays

## 4.1 Introduction

In the modern drug discovery process, the *in vitro* screening of compound libraries generally comprises of a primary and secondary screen (M. Hughes et al, 2012; A. M. Burger et al, 2014). The primary screen is performed to rapidly identify actives at a previously assigned cut-off value while the secondary screen aims to independently confirm the activity of the active compounds and to investigate the mechanism of action, which is quite frequently built upon throughout the drug discovery process (M. Hughes et al, 2012).

Assay development for compound screening is considered an important aspect as it serves to substitute clinical efficiency in the early stages of the drug discovery process (A. S. Verkman, 2003; J. Major, 1998). There are mainly 2 types of assays that are utilised in the primary and secondary biological screening, which includes the cell-based and target-based assays. As mentioned in Chapter 1, section 1.4, cell-based assays include the whole-cell assays and target-based assays are performed on the purified protein target (V. Khurana et al, 2015; A. M. Burger et al, 2014). In anticancer drug discovery, cell-based assays are favoured over the target-based assay, as the former are able to provide preliminary information on compound permeability and stability, when used with appropriate controls (A. M. Burger et al, 2014; R. Zang et al, 2012; E. A. Martis et al, 2011; E. C. Butcher et al, 2004). Cell-based assays were utilised in this study, to maintain consistency with previously reported work in literature on the parental compounds, QNZ and Pentamidine (L. F. Zerbini et al, 2011; L. F. Zerbini et al, 2014).

The primary assays in anticancer drug discovery are fundamentally aimed at establishing the activity and cytotoxicity of a compound (A. S. Narang et al, 2009). From the several cytotoxicity assays available, the 3-(4, 5-dimethylthiazol-2-yl) 2, 5 diphenyl tetrazolium bromide (MTT) assay is preferred in anticancer drug discovery as it can be used on metabolically active cells (T. L. Riss et al, 2013; T. Mosmann, 1983). The MTT assay was performed in this study as it is a well-established, rapid and precise colorimetric method which is based on the reduction of a water soluble yellow tetrazolium salt into insoluble formazan crystals, by the mitochondrial dehydrogenases (T. Mosmann, 1983). Any active

hits that have been identified and prioritized from the primary screen are further investigated in the secondary screen.

As mentioned previously, the aim of the secondary screening assays are to independently confirm the activity of the active compounds identified from the primary screen, in addition to investigating the mechanism of action of the compounds (M. Hughes et al, 2012). Cell proliferation assays are used to generate a dose-response curve, which indicates the $IC_{50}$ value. A dose response curve relays information on both the efficiency and toxicity of a compound. A further comprehensive analysis of the dose response curve could determine whether a compound has an effect on its potential target in a stoichiometric manner so as to eliminate the possibility of non-specific binding (J. P. Hughes et al, 2011). Further experiments to determine the selectivity and specificity of a compound are included in the secondary screen (M. Hughes et al, 2012). In addition to the luciferase assay, mechanism elucidation is further supported by molecular assays such as real time PCR (qPCR), which are commonly used to investigate signalling pathways (S. Hoelder et al, 2012).

In this chapter, the *in vitro* screening assays performed on the 76 compounds that were identified from the virtual screening performed on the ZINC and NCI databases (Chapter 2 and 3) are described. The primary screen, consisting of the MTT proliferation assay, was performed on prostate, renal and ovarian cancer cells lines to rapidly determine the cytotoxic activity of the compounds. The secondary screening assays were performed to independently confirm the activity of the active compound NSC727152, where a dose-response curve was generated. In addition, the mechanism of action of the most potent compound identified, NSC727152 was determined. The luciferase reporter cell assay and qPCR were performed with this compound, to determine if it acted similarly as the parental compound, QNZ, by inhibiting the transcriptional factor NF-κB in cancer cells.

## 4.2 Primary Screening Assays

The MTT assay, described above, is an established assay that is routinely performed in our laboratory to determine the *in vitro* activity of compounds. To facilitate the identification of active compounds, a more robust screening was performed, which included performing the MTT assay at a higher concentration range of between 1-1000μM. The MTT assay was

performed in prostate, renal and ovarian cancer cell lines and was repeated for reproducibility.

### 4.2.1 Materials and Methods

#### 4.2.1.1 Cell Lines

The human renal cancer cells, which includes the kidney epithelial carcinoma cell line (A498) was obtained from the American Type Culture Collection (ATCC) (Rockville, MD). The human renal clear cell (RCC4-) line is deficient of von Hippel-Lindau (VHL) tumour suppressor gene. The RCC4 plus vector (RCC4+) is encoded on the pcDNA3-VHL plasmid, conferring the VHL gene product with neomycin resistance (S. Turcotte et al, 2008) and both RCC4 cell lines were gifted by Professor Sharon Prince, University of Cape Town, South Africa. The ovarian cancer cell line OVCAR3, which is derived from an ovarian adenocarcinoma, is considered a suitable model to study drug resistance and hormonal therapy and was obtained from the American Type Culture Collection (ATCC) (Rockville, MD). PC-3 and DU145 are both derived from prostate cancer that has metastasized to the bone. PC-3 is a grade IV, adenocarcinoma that is characterized as aggressive. DU-145 is a human epithelial carcinoma cell line. Both cell lines were obtained from the American Type Culture Collection (ATCC) (Rockville, MD)

#### 4.2.1.2 Cell Culture Maintenance

RCC4 and DU-145 cell lines were grown in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, Life Technologies, USA). The A498 cell line was grown in Minimum Essential Medium (MEM) (Gibco, Life Technologies, USA). The OVCAR3 and PC-3 cell lines were grown in Roswell Park Memorial Institute-1640 (RPMI-1640) medium (Sigma-Aldrich, ST Louis, USA). All media was supplemented with10% fetal bovine serum (FBS) (Biochrom AG, Berlin) and penicillin (5000µg/ml)/streptomycin (500µg/ml) (Lonza, Walkersville MD, USA). The RCC4 media was supplemented with an additional 0.5mg/mL G418 (Gibco) to maintain for neomycin selectivity. Cells were maintained throughout in this supplemented media, referred to as complete media. The cells were incubated at 37°C, 5% $CO_2$ in a

humidified atmosphere. The medium was changed every 1 to 2 days, and the cells were sub-cultured every 2 to 4 days, or when 90% confluent.

#### 4.2.1.3 Sub-culturing Cells

Cells in a 100mm culture dish were washed with 1×phosphate buffered saline (PBS) pH 7.4 (137mM NaCl, 2.7mM KCl, 4.3mM Na2HPO4·7H2O, 1.4mM KH2PO4) and the PBS was aspirated off. Cells were trypsinised in 1 ml 0.05% Trypsin/EDTA (Gibco, Life Technologies, USA) and once they had detached, the trypsin-EDTA was inactivated by adding 5 ml of complete media. Cells were re-suspended and re-plated in a clean 100mm culture dish in complete media at a dilution of 1:7.

#### 4.2.1.4 Freezing and Thawing

A confluent dish of cells was trypsinized and neutralised with complete medium, followed by centrifugation to remove cells from the trypsinization solution. Cells were suspended in 3 ml of pre-chilled freezing medium (70 % culture medium, 20 % FBS and 10% dimethylsulfoxide (DMSO) and transferred into cryotubes, which were stored at -80°c.

To thaw cells, the frozen cell suspension was quickly thawed in a 37°C water bath and transferred to a tube containing 5ml of complete growth medium. Cells were pelleted by centrifugation at 400xg for 2 min at 4°C, were re-suspended in 10ml complete media and added to a 100 mm culture dish. Cells were incubated at 37°C, 5% $CO_2$ in a humidified atmosphere.

#### 4.2.1.5 Drugs and Inhibitors

Pentamidine isethionate salt (CAS 140-64-7) and QNZ (CAS 545380-34-5) were purchased from Sigma (Sigma Aldrich, Inc). The ZINC database compounds were purchased from the suppliers annotated on the website, with the most affordable supplier being identified. The NCI compounds were ordered through the online system on the NCI website.

All compounds were stored according to supplier's instructions upon arrival and were dissolved in DMSO, as a stock solution. Pentamidine was dissolved in molecular grade water (Gibco, CA, USA). The stock solutions were further aliquot into smaller volumes of working stocks to preserve stability between successive freeze-thaw cycles. The dissolved compounds were stored at -20°c. The molecular mass and purity of all the compounds was determined by LC/MS/MS, prior to commencing with the screening assays. Insufficient quantities were obtained from vendors to confirm their structures by Nuclear Magnetic Resonance (NMR).

#### 4.2.1.6 MTT Assay

All cell lines except for DU-145 were plated at a concentration of 5 x $10^3$ cells per well in a 96 well culture plate, at a final volume of 100µl per well. Due to the faster proliferation rate of DU-145 cells they were plated at a concentration of 3 x $10^3$ cells per well. Cells were incubated overnight at 37°C, 5% $CO_2$, in a humidified atmosphere, prior to drug treatment to allow for attachment. The next day, in a separate 96 well culture plate serial dilutions of each drug were performed in quadruplets along the divided length of the culture plate. The media from the 96 well plates containing the cells was aspirated and replaced with 100µl of each dilution in quadruplets. Cells were incubated at 37°C, 5% $CO_2$, in a humidified atmosphere. After 24 hours 10µl of MTT (0.01M in PBS) (Sigma-Aldrich, Steinheim, Germany) was added and the cells were incubated for 4 hours at 37°C, 5% $CO_2$, in a humidified atmosphere. After 4 hours, 100µl of solubilisation buffer (0.01 M *HCl*; 10% *SDS*) was added and the cells were further incubated for 16-18 hours at 37°C, 5% $CO_2$, in a humidified atmosphere. After 18 hours, the absorbance of the solubilised tetrazolium salt was read at 595nm using a plate reader.

#### 4.2.1.7 Solubility Enhancement Techniques

It was visible that certain compounds had precipitated out of solution upon adding them to the DMSO or culture media (Table 1 and 2). Solubility enhancement techniques were employed with these compounds such as the use of co-solvent, surfactants and increasing the temperature of the media in an effort to increase solubility (A. Chaudhary et al, 2012). Firstly, the effect of ethanol, polysorbate 80 or tween 80 and N, N,-Dimethylformamide (DMF) on cell proliferation was determined to establish an $IC_{50}$ concentration or a tolerable

concentration for the cancer cell lines. The compounds that precipitated out of DMSO were further dissolved in water and ethanol. Those that failed to solubilise in either of the solvents, were excluded from the primary screen. Repetitive trails were performed to solubilise the compounds in the culture media with the addition of minimal volumes of either the surfactants or co-solvents. Furthermore, samples were incubated at 50°c, in a sonicating water bath to aid in the dissolution of the compound. These techniques did not visibly improve on the solubility. The adjustment of the pH was not considered an option, as the culture conditions for cell maintenance required that they are kept at a pH of 7.

### 4.2.2 Results

#### 4.2.2.1 Compound Solubility

The solubilisation techniques were employed on the compounds that were found to be insoluble in the culture media. Although the solubilisation techniques did not visibly improve on the solubility of these compounds, they were still included in the primary screening, to determine if they had activity at a lower concentration. These compounds are listed in Table 1 and 2 found in Appendix 3.1.

#### 4.2.2.2 ZINC Database

The following results include those of the primary screening performed on the compounds identified from the ZINC database. All compounds were screened from a starting concentration of 1000μM.

##### 4.2.2.2.1 Pentamidine

A total of 17 compounds were identified from the ZINC database and tested for *in vitro* activity in the renal cancer cell line, A498. None of the compounds inhibited cell proliferation *in vitro* after 24 hours treatment. The MTT proliferation graph of the compound ZINC44920378 is depicted in Figure 1 below, which serves as an example for the rest of the inactive compounds. The MTT proliferation graphs of the remaining compounds are included in Appendix 3.2. Pentamidine was used as a positive control and was confirmed to have an $IC_{50}$ value of approximately 15μM in A498.

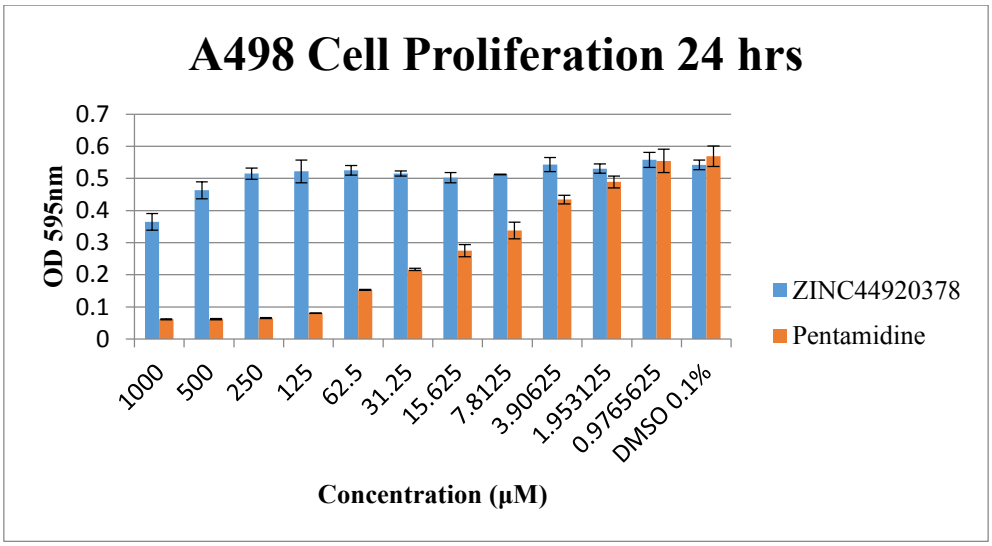



**Figure 1: The MTT proliferation graph of ZINC44920378 in A498 renal cancer cell lines, 24 hours post-treatment. Pentamidine was used as a positive control. DMSO 0.1% was used as a vehicle control. Data shown represents mean ± S.D. of three independent experiments performed in duplicate. OD = absorbance.**

The compounds that were insoluble in culture media were included in the primary screening to determine if they were active at a lower concentration. They were confirmed to be inactive and their MTT proliferation graphs are included in Appendix 3.2.

#### 4.2.2.2.2 QNZ

The effects of the 21 compounds identified from the ZINC database on the proliferation of cancer cells were determined in prostate (DU-145 and PC-3) and ovarian cancer cell lines (OVCAR3) after 24 hours of treatment. No compounds identified from the ZINC database were found to inhibit the *in vitro* cell proliferation of prostate and ovarian cancer cells. The compound QNZ was not included as a positive control in these assays, although it reportedly has an $IC_{50}$ value of 11nM, in human Jurkat T cells (M. Tobe et al, 2003).

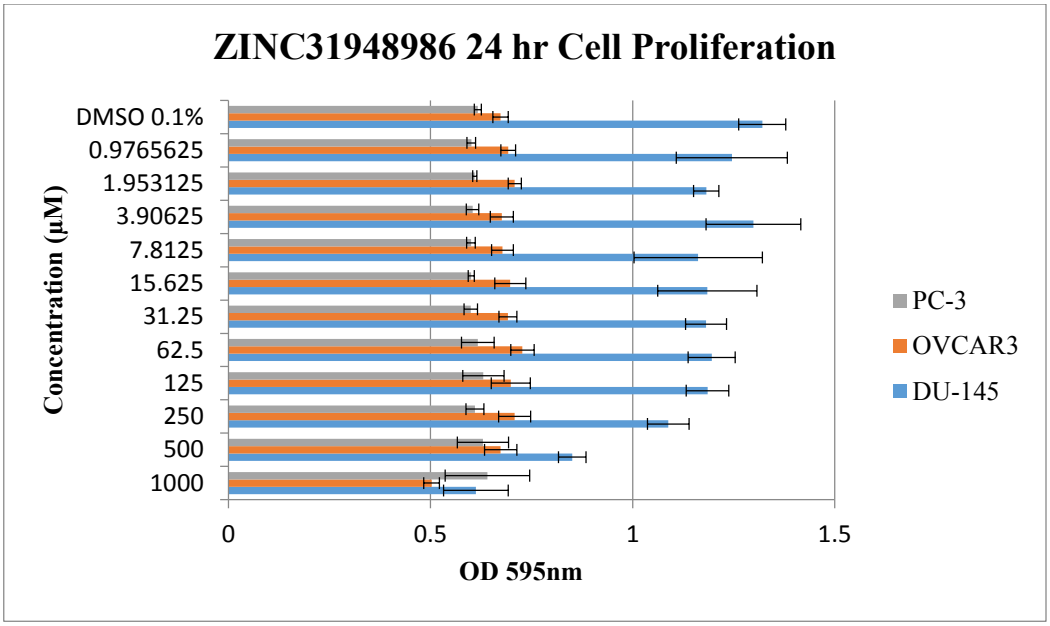



**Figure 2: The MTT proliferation graph of ZINC31948986 24 hour post-treatment in DU-145, PC3 and OVCAR3 cancer cell lines. DMSO 0.1% was used as vehicle control. Data shown represents mean ± S.D. of four independent experiments performed in duplicate. OD = absorbance.**

The proliferation graph of the compound ZINC31948986 is depicted in Figure 2 above and serves as an example of the results obtained for the rest of the inactive compounds, whose MTT proliferation graphs are included in Appendix 3.3. The insoluble compounds were included in the primary screen to rule out activity at a lower concentration. None of these compounds were confirmed with *in vitro* activity and their data is included in Appendix 3.3.

#### 4.2.2.3 NCI Database

The following results include those of the primary screening performed on the compounds identified from the NCI database.

##### 4.2.2.3.1 Pentamidine

The 11 compounds identified from the 2-aminopyridine substructure search performed on the NCI database were screened for *in vitro* activity in the renal cancer cell lines A498, RCC4+ and RCC4-. None of the compounds inhibited the *in vitro* cell proliferation of the cancer

cells. The MTT proliferation graph of the compound NSC383 is depicted in Figure 3 below and serves as an example for the rest of the inactive compounds. The MTT proliferation graphs of the remaining compounds are included in Appendix 3.4. The insoluble compounds were also included in the primary screening and their graphs are included in Appendix 3.4.

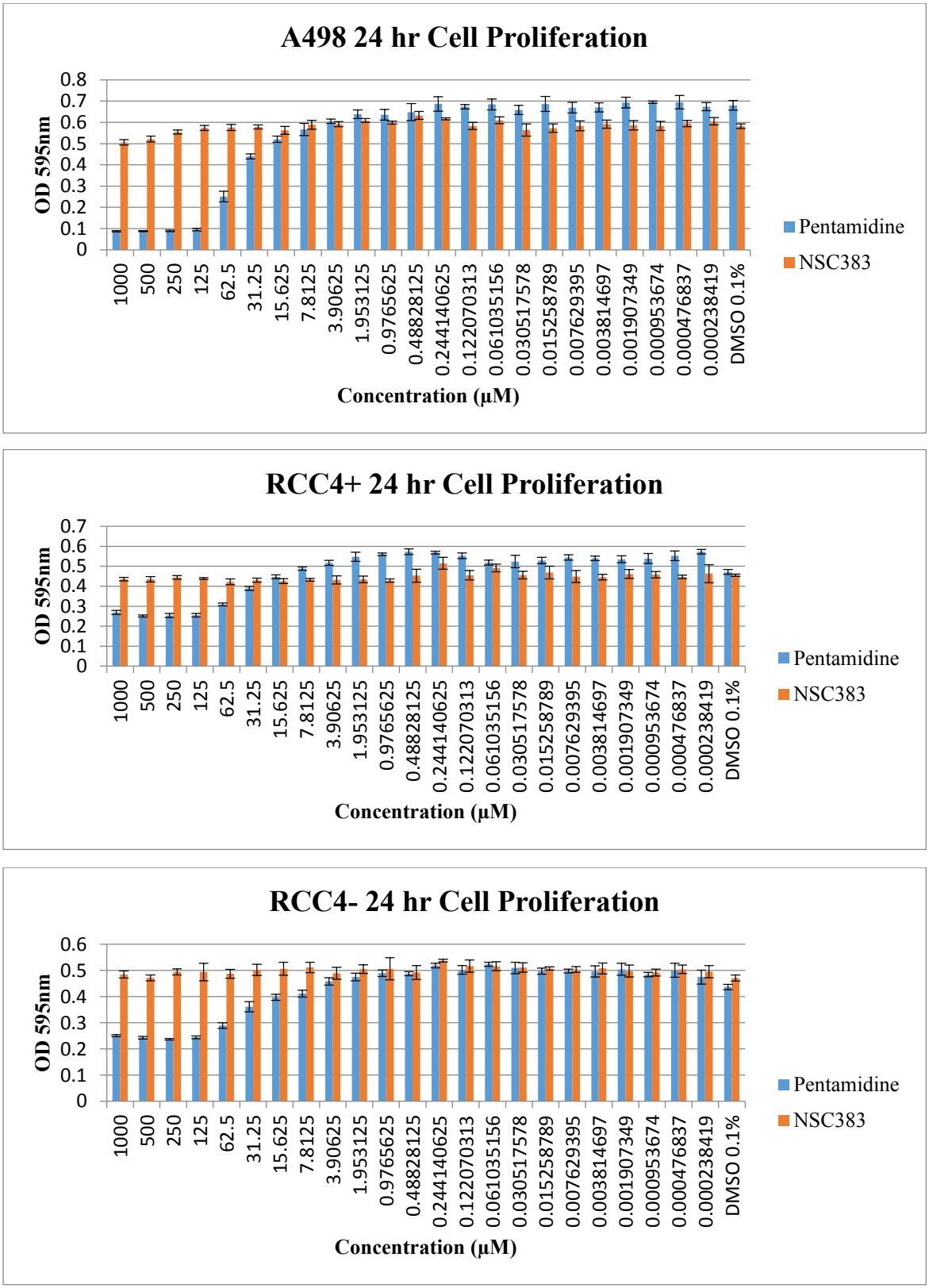



**Figure 3: The MTT proliferation graph of NSC383 in A498, RCC4+ and RCC4- cancer cell lines, 24 hours post-treatment. 0.1% DMSO used as a vehicle control and Pentamidine as the positive control. Data shown represents mean ± S.D. of four independent experiments performed in duplicates. OD = absorbance.**

#### 4.2.2.3.2 QNZ

The 30 compounds identified from the NCI database were tested for the *in vitro* inhibition of cell proliferation in prostate and ovarian cancer cell lines. Four (4) hits with *in vitro* activity in both prostate and ovarian cancer cell lines were identified. The compounds that precipitated out of the culture media could not be re-solubilised and did not show any activity at a lower concentration (data shown in Appendix 3.5).

The active hit NSC52075 was identified from the 2D similarity search that was performed in Chapter 3, section 3.2.5. The cell lines DU-145 and PC-3 showed sensitivity towards NSC52075 with an $IC_{50}$ value of 63μM and 507μM, respectively, after 24 hours of treatment (Figure 4). The other active hits include NSC727152 and NSC676169, which were identified from the virtual screening performed with the third pharmacophore hypothesis. The compound NSC329254 was identified from the virtual screening performed with the fourth pharmacophore hypothesis, both described in Chapter 3, section 3.2.8. Only DU-145 cells were sensitive towards NSC676169 with an $IC_{50}$ value of 112μM (Figure 5). DU-145 and OVCAR3 cell lines were sensitive towards NSC329254 with an $IC_{50}$ value of 137μM and 343μM, respectively (Figure 6). All 3 cancer cell lines showed sensitivity towards NCS727152 with an $IC_{50}$ value of 8μM in DU-145, 24μM in PC-3 and 7μM in OVCAR3 (Figure 7). Only the compound NSC727152 was considered for the secondary screening assays as it had most potent $IC_{50}$ value in all 3 cell lines. The compound QNZ was not included as a positive control in these assays, although it reportedly has an $IC_{50}$ value of 11nM, in human Jurkat T cells (M. Tobe et al, 2003).

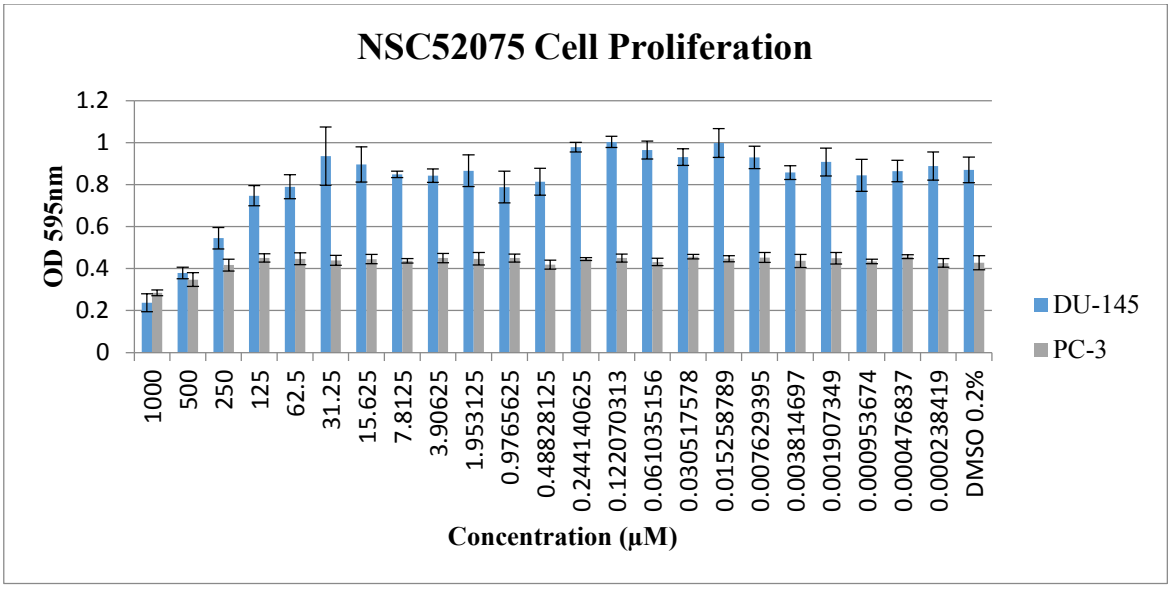



**Figure 4: The MTT proliferation graph of NSC52075 after 24 hours treatment in PC-3 and DU-145. 0.2% DMSO was used as the vehicle control. Data shown represents mean ± S.D. of 4 individual experiments performed in duplicate. OD = absorbance.**

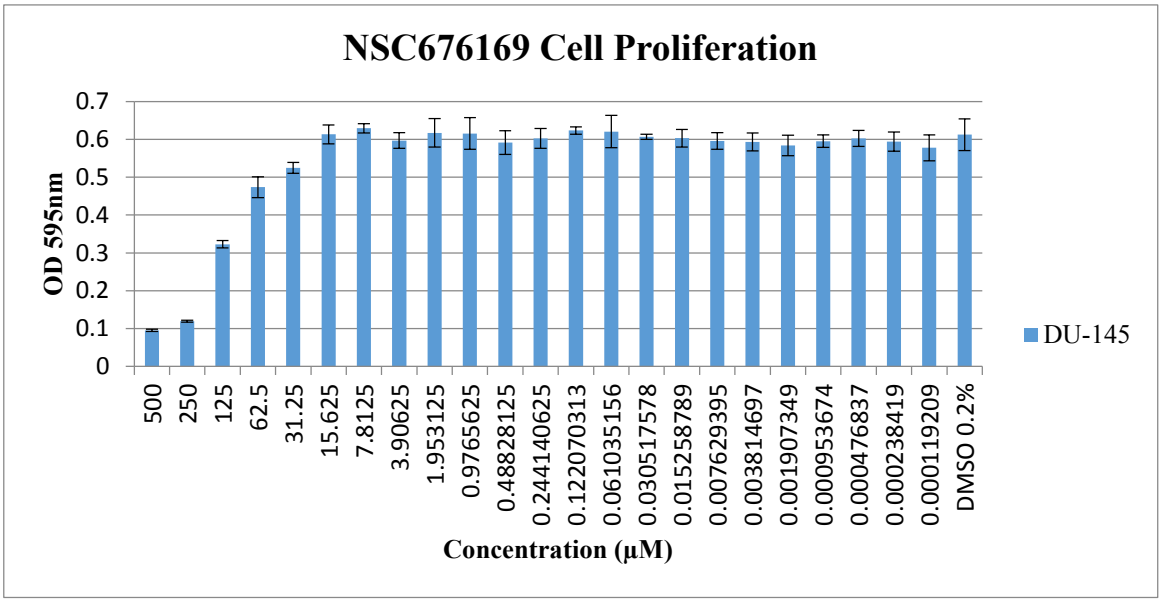



**Figure 5: The MTT proliferation graph of NSC676169 in DU-145 after 24 hours treatment. 0.2% DMSO was used as a vehicle control. Data shown represents mean ± S.D. of 4 individual experiments performed in duplicate. OD = absorbance.**

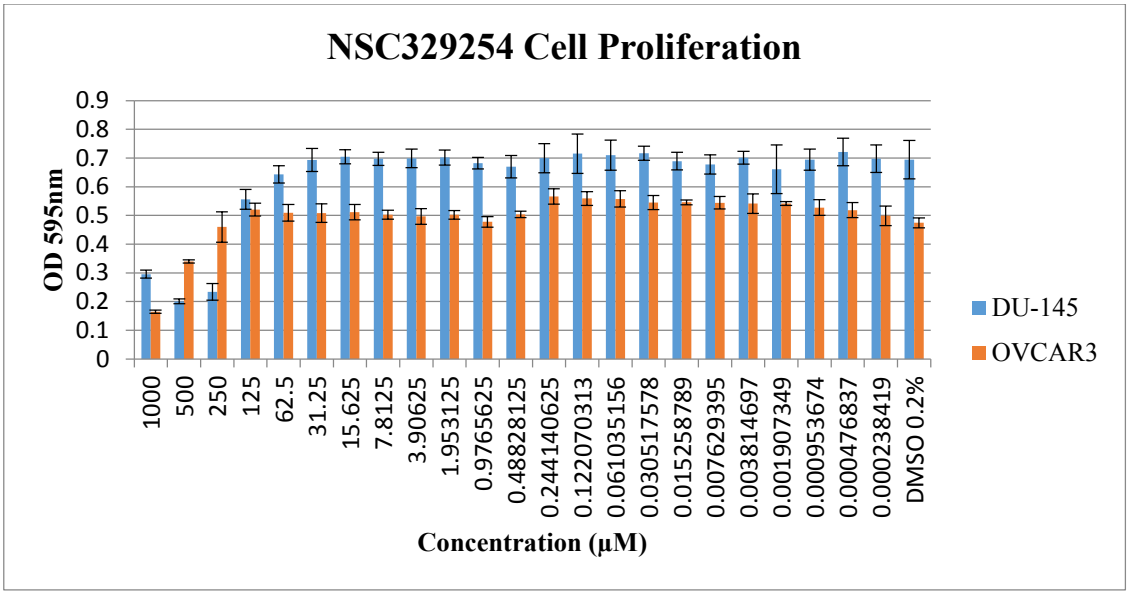



**Figure 6: The MTT proliferation graph of NSC329254 after 24 hours treatment. 0.2% DMSO was used as a vehicle control. Data shown represents mean ± S.D. of 4 individual experiments performed in duplicate. OD = absorbance.**

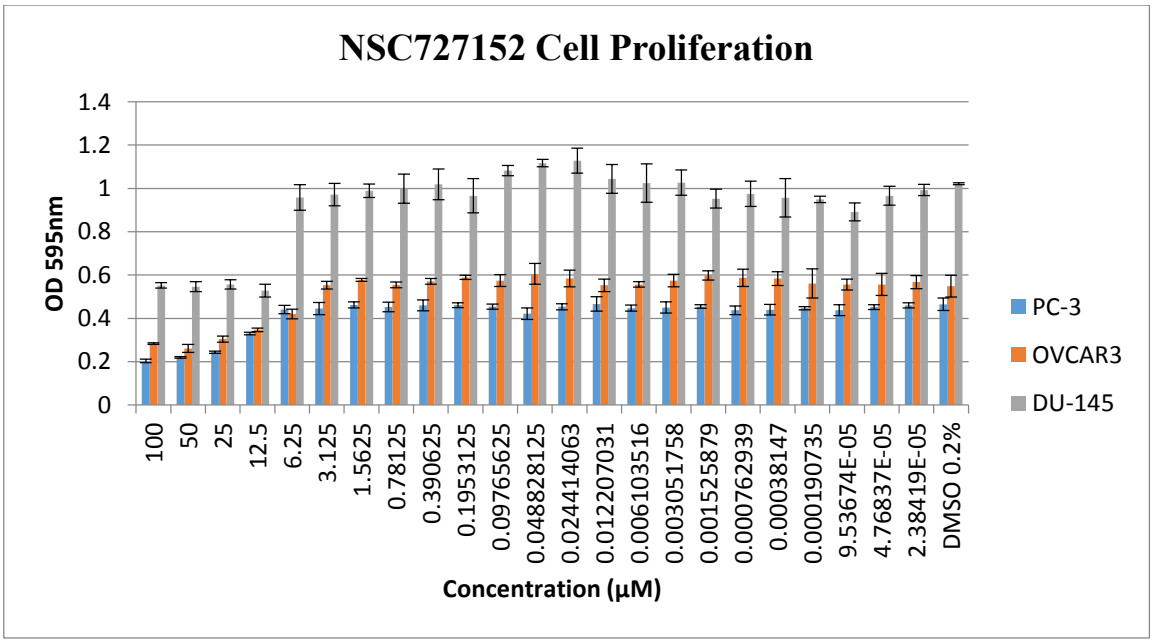



**Figure 7: The MTT proliferation graph of NSC727152 after 24 hours treatment. 0.2% DMSO was used as a vehicle control. Data shown represents mean ± S.D. of 4 individual experiments performed in duplicate. OD = absorbance.**

### 4.2.3 Discussion

#### 4.2.3.1 Pentamidine

A total of 25 hits identified from both the ZINC and NCI databases were tested for *in vitro* activity in renal cancer cell lines. None of the compounds were confirmed to inhibit the *in vitro* cell proliferation of the renal cancer cells at a starting concentration of 1000μM.

There are several possibilities which may account for the inactivity of the compounds, the most logical being that the compounds simply did not display whole-cell activity. This is a plausible explanation considering that none of the compounds identified from the ZINC database possessed a bis-benzamidine functional moiety similar to Pentamidine. Although a number of successful virtual screening studies are reported in literature (C. H. Chuang et al, 2015; H. Park et al, 2012; G. Cozza et al, 2006), the methodological limitations of a computational programme's ability to accurately predict the dynamic process of ligand-protein binding remains a major limiting factor (B. K. Shoichet, 2004).

A prerequisite to performing a successful 3D ligand-based or structure-based virtual screening is the accurate generation of the ligand and protein side chain conformations (S. Subramaniam et al, 2008; J. Bajorath, 2002; A. S. Reddy et al, 2007). Conformational search algorithms have improved dramatically to generate realistic and energetically favourable conformations, as using an inappropriate conformation will have detrimental effects in any downstream application (B. K. Shoichet, 2004; K. S. Watts et al, 2010). A major limitation still encountered in generating bioactive conformations is the associated computational cost and the large number of conformations that may be generated, increasing with the number of rotatable bonds contained in a ligand or protein (B. K. Shoichet, 2004; K. S. Watts et al, 2010). An additional complexity in generating the correct bioactive conformation arises in the correct assignment of protonation and tautomer states (B. K. Shoichet, 2004; J. C. Shelley et al, 2007). This is especially important in the case of ligands contained in a database, where bioactive conformations obtained from crystal structures are not available to be used as a reference to measure the accuracy of the generated conformation (S. Subramaniam et al, 2008; J. Bajorath, 2002). It may be hypothesised that energetically unfavourable conformations of the compounds contained in the ZINC and NCI database were predicted, which would result in inaccurate measures of the 3D similarity and binding affinities predicted in this study. In addition, this would result in the inaccurate predictions of the

binding orientation of the compounds at a physiologically relevant pH, which may not be conducive for binding to the reportedly narrow asymmetrical cone-shaped channelled conformation of the active site of the DAO enzyme. This may be a plausible explanation despite the predicted binding interaction between the ZINC compounds and the Asp373 residue of the catalytic base of the DAO residue. (A. P. McGrath et al, 2009). In addition, it could further be hypothesised that the incorrect protonation and tautomer states were predicted for the compounds in the ZINC and NCI database, taking into consideration that they possessed suitable amine groups that could interact with the negatively charged residues that line the active site of the DAO enzyme (A. P. McGrath et al, 2009).

Although incorporating protein flexibility in a virtual screening study increases the chances of a producing a reliable prediction of the binding affinities, it may be a difficult task to define the flexible protein residues in addition to being computationally exhaustive (S. Subramaniam et al, 2008; M. A. Lill, 2011). This is mainly due to the fact that the protein flexibility induced upon ligand binding may not only involve flexibility in the protein side chains but also in the main chain and loop regions of the protein and therefore would require prior knowledge on the bioactive protein and ligand conformations (M. A. Lill, 2011; M. Fischer et al, 2014). Taking this into consideration, protein flexibility was not included in the SBVS performed for QNZ and Pentamidine, which may be viewed as a limitation in this study. Although the inclusion of the different crystal structures of the NF-κB subunits in the blind docking performed in this study, namely p65/p50 either bound to IκBα or DNA, was considered to account for the observed conformational change.

Another limitation associated with computational tools is the correct scoring and ranking of ligands predicted to be active. The complex and dynamic process of ligand-protein binding, results in various parameters that need to be accounted for in scoring functions, which includes the solvation, enthalpic and entropic effects (B. K. Shoichet, 2004; D. B. Kitchen et al, 2004). These energies are very large and complex to calculate, which remains a bottle neck associated with accurately predicting binding affinities and ranking compounds accordingly (D. B. Kitchen et al, 2004; B. K. Shoichet, 2004). Although more robust scoring functions that utilise thermodynamic integration methods have been developed to address these short-comings, they are computationally expensive and still do not accurately account for a number of entropic effects which influence molecular recognition (D. B. Kitchen et al, 2004; B. K. Shoichet, 2004). Molecular Dynamic (MD) simulations have recently been developed to circumvent most of the limitations associated with LBVS and SBVS tools (S.

Subramaniam et al, 2008). Computational algorithms utilised in MD simulations are considerably more advanced and therefore are computationally more expensive. MD simulations are able to simulate a ligand-protein docking in its natural environment, to the atomic level (S. Subramaniam et al, 2008) and are frequently performed to refine docking calculations.

As mentioned previously, one of the advantages of a cell-based assay is that a preliminary indication of the compounds physicochemical properties, namely the solubility and permeability may be obtained (A. M. Burger et al, 2014). It was visually evident that a number of compounds identified from the sub-structure search performed on the NCI database were insoluble in the culture media. The solubility enhancement techniques employed were not sufficient to achieve solubility, therefore it can be said that these compounds were not present at an ideal concentration to interact with their biological targets and illicit the desired pharmacological response (J. P. Hughes et al, 2011).

In the case of this study, the protein targets (NF-κB and the enzymes of the polyamine biosynthetic pathway) are cytosolic proteins, where NF-κB may further translocate into the nucleus (A. Oeckinghaus et al, 2009; A. E. Pegg, 2009). Therefore cell membrane permeation must be achieved by the compounds, in order to interact with their respective targets. Taking into consideration that Pentamidine has a reportedly lower permeability compared to other diamines, it may be hypothesised that the soluble compounds had poor permeability and therefore were not able to permeate the cell membrane (S. Yang et al, 2014). In addition, it may be hypothesised that inaccurate predictions of Log P values for the compounds contained in the ZINC databases were obtained, which may explain the inefficiency of the drug-like filter that was applied to exclude these compounds with undesirable Log P values. An alternative assay to determine the *in vitro* inhibition of the enzymes in the polyamine biosynthetic pathway would include the use of purified enzymes in a target-based assay. The use of a target-base assay would eliminate the factors that influence cell-based assays such as permeability and those discussed below (J. P. Hughes et al, 2011; A. M. Burger et al, 2014). Several studies reported in literature have demonstrated the use of target-based assay to determine the *in vitro* activity of the series of analogues derived from Pentamidine (A. Puckowska et al, 2012; J. Stanek et al, 1993; U. Regenass et al, 1992).

Another factor that may influence the activity of compounds in a cell-based assay is the expression levels of the biological targets (Z. Zhang et al, 2012). It is reported in literature,

that cell lines may express the targets in low concentrations, requiring the over-expression of these targets (Z. Zhang et al, 2012). Furthermore, cell lines may acquire different characteristics when maintained in culture for several passages (K. Oostdik et al, 2009). Although every attempt was made to ensure that optimal tissue culture techniques were maintained in this study, it would be difficult to disregard the hypothesis that the expression levels of the target proteins in the cell lines used were not optimal. Other limitations that may influence cell-based screening assays include the non-specific binding to cellular proteins and media components such as BSA. Furthermore, a compound's stability may be compromised under differing pH and temperature conditions, which may either result in the chemical modification of the compound or in the formation of aggregates, which may contribute to their inactivity (L. Di et al, 2009).

Having said that, the MTT assay has been previously validated and commonly used in our laboratory, under the same conditions and parameters used in this study (J. D. Paccez et al, 2015). Taking this into consideration and the fact that positive and negative controls were used, any questions on the integrity and quality of the screening assay was confidently ruled out. This may further substantiate the hypothesis that the observed inactivity of the compounds is a consequences of the factors described above.

#### 4.2.3.2 QNZ

The hits identified from the ZINC and NCI database were tested for *in vitro* activity using the MTT assay. A total of 51 compounds were tested for cytotoxic activity in prostate and ovarian cancer cell lines. Of these compounds, 4 compounds identified from the NCI database were found to possess *in vitro* activity. In addition to the physicochemical properties of the ZINC and NCI compounds contributing to their inactivity, the other possibilities described above in section 4.2.3.1, may also account for the inactivity of the compounds identified from the virtual screening performed for QNZ.

In retrospective analysis it was revealed that the ZINC compounds identified from the SBVS performed on the NF-κB complex did not have the appropriate functional groups and/or substituents that were hypothesised to be necessary for activity as described in Chapter 3 and depicted in Figure 8A below. Furthermore, a large proportion of the compounds identified

from the ZINC and NCI database did not have a quinazoline core similar to QNZ. The compounds that did possess a quinazoline core include ZINC31948986 and NSC2075, where the former had no *in vitro* activity and the latter had *in vitro* activity in both DU-145 ($IC_{50}$=63μM) and PC-3 ($IC_{50}$=507μM) cancer cell lines (Figure 8B).

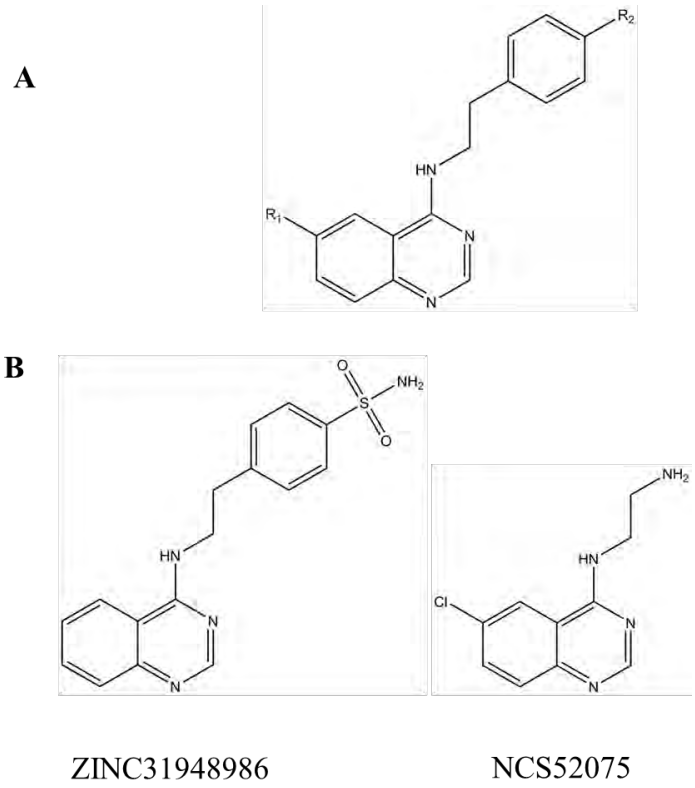



**Figure 8: A) The hypothesised SAR of the quinazoline inhibitors of NF-κB taken from literature (M. Tobe et al, 2003) (M. Tobe et al, 2003). B) The compounds identified from the virtual screening performed that possessed a quinazoline core. The compound ZINC31948986 was confirmed to have no *in vitro* activity whilst NSC52075 had *in vitro* activity in the micro-molar range.**

The compound ZINC31948986 lacked a functional group substituent at the $R_1$ position, compared to the compound NSC52075, which had an electron withdrawing, chlorine atom at this position (Figure 8). Consistent with the SAR data reported in literature an electron withdrawing group, such as chlorine at the $R_1$ position results in reduced activity compared to QNZ (M. Tobe et al, 2003), which possesses an electron donating group $NH_2$, at this position. This finding substantiates the hypothesis that an electron donating group such as $NH_2$, if not any functional moiety, at the $R_1$ position is necessary to illicit biological activity (Figure 8). Furthermore, it may be hypothesised that the terminal alkyl amino group on NSC52075 leads

to increased solubility compared to the sulphonamide group on ZINC31948986, which may also account for its activity (Figure 8B). Interestingly, the active compound NSC52075 was identified by the 2D similarity search which may add to the argument that 2D descriptors are more reliable and effective for identifying similar compounds than 3D descriptors (V. Venkatraman et al, 2010; M. Thimm et al, 2004; J. Bajorath, 2002).

The pharmacophore search was performed on the NCI database with the aim of rationalising the summarised SAR extracted from literature. Four pharmacophores were developed to test the SAR hypothesis and of these only 2 pharmacophores were able to retrieve compounds from the NCI database. Two active compounds, NCS727152 (DU145, $IC_{50}$=8μM; PC-3, $IC_{50}$=24μM; OVCAR3, $IC_{50}$=7μM) and NCS616769 (DU145, $IC_{50}$=112μM), were identified from the pharmacophore model based on the third hypothesis which determined the importance of the aromatic moiety on the quinazoline core (Figure 9). These compounds had lower $IC_{50}$ values than the compound NCS329254 (DU145, $IC_{50}$=137μM; OVCAR3, $IC_{50}$=343μM), which was identified by the fourth pharmacophore model, which determined the importance of the charged atom on the quinazoline core (Figure 9).

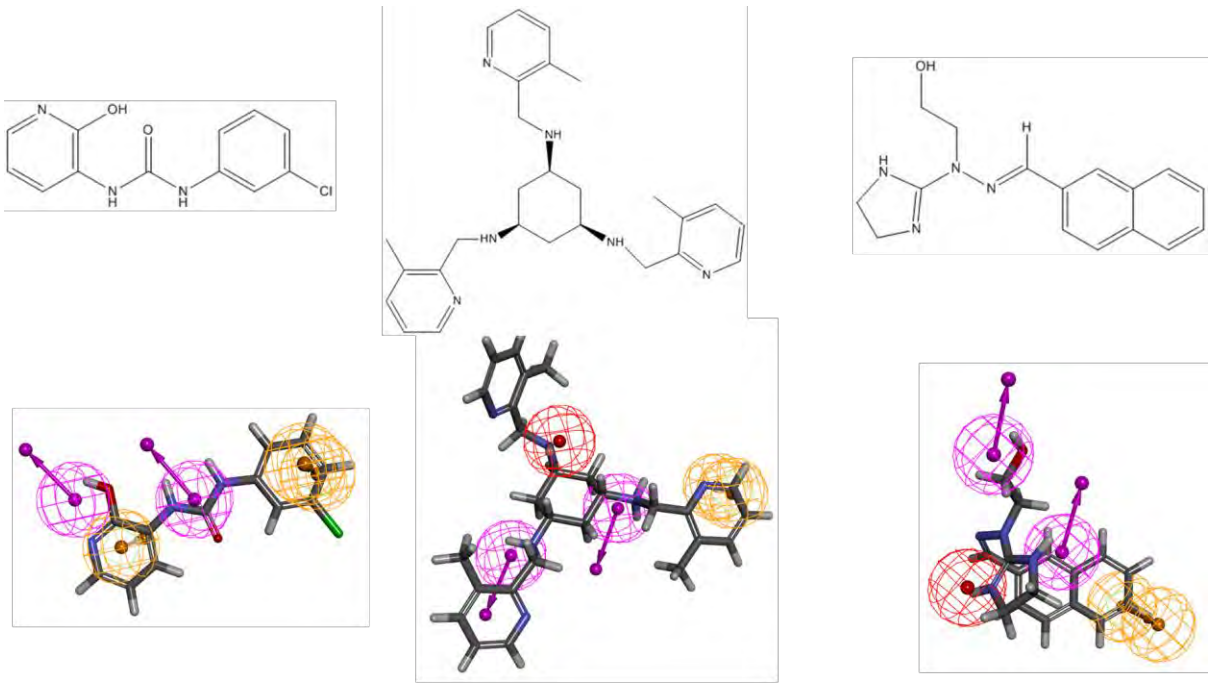



**Figure 9: The compounds identified from the pharmacophore VS performed on the NCI database and shown overlaid on their respective pharmacophores. All three compounds were confirmed to inhibit the *in vitro* cancer cell proliferation, in the micro-molar range.**

Based on the $IC_{50}$ values of the compounds identified from the pharmacophore VS, the basic nitrogen on the quinazoline core is more important for increased activity than the aromatic ring of the same core. Due to not having retrieved compounds from the virtual screening performed with pharmacophores 1 and 2, further conclusions from the hypothesis could not be made. Despite this, literature suggests that the 2 hydrogen bond donors and the lipophilic moiety at the $R_2$ position are important for activity (Figure 8) (M. Tobe et al, 2003; M. Tobe et al, 2003).

All the compounds, for both QNZ and Pentamidine, procured from the ZINC database were selected based on good, intermediate and poor binding scores. This was performed with the aim of establishing any correlation between the binding scores and the $IC_{50}$ value obtained from the biological assays. No correlation could be established as all the compounds were observed to have no *in vitro* activity, albeit literature suggests that a poor to no correlation exists between the 2 parameters (P. Tuffery et al, 2012; Q. V. Vuong et al, 2013).

## 4.3 Secondary Screening Assays

The secondary assays were performed with the aim of independently confirming or validating the active compounds identified in the primary screen and to elucidate a possible mechanism of action (M. Hughes et al, 2012). The assays serve to determine the selectivity and specificity of a compound, alongside the use of orthogonal assays that are performed to discern between a false positive and a true hit (N. Thorne et al, 2010).

This section of the chapter describes the assays used to confirm the molecular mechanism of action of NSC727152, which is hypothesised to be similar to QNZ, by acting as an inhibitor of the transcription factor NF-κB. This was performed by determining the effect of NSC727152 on the promoter activity of genes that are regulated by NF-κB, such as IL-6 and the consensus sequence κB promoter (L. F. Zerbini et al, 2004). The inhibitory effect on NF-κB target genes was confirmed by determining the effects upon NSC727152 treatment on the mRNA expression of the Growth Arrest and DNA Damage (GADD) 45 α and γ genes and IL-6. The GADD45α and GADD45γ genes, in addition to the IL-6 gene have been reported to be to be regulated by NF-κB in cancer cells (T. A. Libermann et al, 1990; L. F. Zerbini et al, 2004; L. F. Zerbini et al, 2005; L. F. Zerbini et al, 2003; M. Karin, 2009).

### 4.3.1 Materials and Methods

#### 4.3.1.1 RNA Isolation

RNA was isolated from the cell cultures using the RNeasy Mini kit, as per manufacturer's instructions. Briefly, 10µl β-mercaptoethanol (β-ME) was added for every 1 ml Buffer RLT used. Cells were scrapped using a cell scrapper and the lysate was collected in an eppendorf. Lysates were stored at -80°C, if not used immediately. The cell lysate was then transferred into a QIAshedder spin column and centrifuged at full speed for 2 minutes. An equal volume of 70% ethanol was added to the lysate and mixed by pipetting. The lysate was transferred to an RNeasy spin column and centrifuged for 15 seconds at 8000 x g. To wash the membrane 500µl of buffer RW1 was added and centrifuged for 15 seconds at 8000 x g. Three consecutive washes with Buffer RPE were performed, to ensure that no ethanol was carried over. Finally, 50µl of RNase-free water was used to elute the RNA in a clean eppendorf. The concentration of the RNA was measured using a nanodrop before storing samples at -80 °c until further use. The quality of the RNA was determined through the 260/230 nm and 260/280 nm absorbance ratios.

#### 4.3.1.2 cDNA Synthesis

Conversion of mRNA to cDNA was performed using the Roche Transcriptor First Strand cDNA synthesis kit, following manufacturer's instructions. Briefly, 1µg template mRNA, 1µl oligo-DT and $dH_2O$ up to 13µl, were added together and incubated for 10 minutes at 65°C, 60 minutes at 4°C, 5 minutes at 85°C and finally 5 minutes at 4°C. Reaction buffer, dNTPs, RNA inhibitor and reverse transcriptase were added to give a final volume of 20µl and incubated for 10 minutes at 25°C, 60 minutes at 50°C followed by 5 minutes at 85°C to inactivate the reverse transcriptase. The cDNA was stored at -20°C until needed.

#### 4.3.1.3 Quantitative real-time PCR

q-PCR was performed as described previously (L. F. Zerbini et al, 2003). Briefly, one microliter of cDNA and 20μM of the respective forward and reverse primers were added to a mixture of 1X SYBR® FAST qPCR master mix and made up to a final volume of 20μl with molecular grade water (Sigma-Aldrich, Germany). This was added to a low-profile 96 well cell culture plate (SPL Life Sciences, Korea).

The conditions for PCR were: 1 cycle of 5 minutes at 94°C; 45 cycles of 30 seconds at 95°C, 1 cycle of 30 seconds at 56°C, and 1 cycle of 30 seconds at 72°C. Melting curve genotyping: 15 seconds at 95°C, 2 minutes at 65°C, 97°C (continuous acquisition, 5 acquisitions per °C) was then performed. For each run, the human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used to normalize each sample and the primers for GAPDH were also used to quantify the amount of cDNA in a reference sample. A single point from a human GAPDH serial dilution performed previously was used as a reference sample for the absolute quantification of the amplified DNA i.e. gene expression based on a standard curve. All of the samples were run in triplicate and the final presentation of the data expressed as a ratio of the gene of interest to GAPDH. The Primers and their melting points are shown in Table 1, below.

**Table 1: The qPCR primer sequences and their melting temperatures.**

| **Primer** | **Sequence** | **Melting temperature** |
|---|---|---|
| GADD45α | 5'-GCCTGTGAGTGAGTGCAGAA-3'<br>5'-ATCTCTGTCGTCGTCCTCGT-3' | 50ºc |
| GADD45β | 5'- GAAGATCTCTATGACGCTGGAAGAGCTCGT-3'<br>5' GAAGGTACCTCAGCGTTCCTGAAGAGAGAT 3' | 50ºc |
| GADD45γ | 5'-CTGCATGAGTTGCTGCTGTC-3'<br>5'-TTCGAAATGAGGATGCAGTG-3' | 50ºc |
| IL-6 | 5'-GGGAACGAAAGAGAAGCTCT-3<br>5'-ACCAGAAGAAGGAATGCCCA-3' | 50ºc |
| GAPDH | 5'-CAAAGTTGTCATGGATGACC-3'<br>5'-CCATGGAGAAGGCTGGGG-3' | 50ºc |

#### 4.3.1.4 Luciferase Constructs

The luciferase constructs were previously designed and described elsewhere (L. F. Zerbini et al, 2004). The full length human IL-6 promoter was inserted into a pxp2 reporter gene construct containing the luciferase gene. Similarly the pxp2-κB construct was designed using a κB consensus sequence promoter (L. F. Zerbini et al, 2004).

#### 4.3.1.5 Transient Transfection

DU-145 cells were plated at a concentration of 1.5 x $10^5$ cells per well and incubated at 37°C, 5% CO2 in a humidified atmosphere, overnight. The transfection of the luciferase constructs in DU-145 cells was performed using the Invitrogen lipofectamine PLUS transfection kit, following the manufacturer procedures. Briefly, 750 ng of DNA was added to 500µl DMEM media only, to which 1.5µl of the PLUS reagent was added and incubated at room temperature for 5 minutes. 4.5µl of the lipofectamine LTX reagent was added to the DNA mixture and incubated at room temperature for a further 30 minutes. The DNA/lipofectamine mixture was added drop wise to the cells and incubated for 4 hours at 37°C, 5% $CO_2$ in a humidified atmosphere. After 4 hours the media was changed to complete media and incubated over night at 37°C, 5% CO2 in a humidified atmosphere.

#### 4.3.1.6 Luciferase Gene Reporter Assay

The next day after transfection or when the transfected cells had reach a confluency of 70-80%, they were trypsinised and plated in 96 well culture plates at a concentration of 3 x $10^5$ cells per well and incubated overnight at 37°C, 5% $CO_2$ in a humidified atmosphere. After 24 hours they were treated with the serial dilutions of NCS727152, at a starting concentration of 32µM. Appropriate controls were included such as a blank that contained no cells except for media, the drug in media only and the luciferase reagent. After 6 hours, 50µl's of the media was removed from each well and replaced with 50µl of the ONE-glo (Promega) luciferase lysis buffer. Culture plates containing the luciferase lysis buffer were kept in the dark before transferring 100µl of the media to an opaque plate that was read using a microplate luminometer (Veritas™ Microplate Luminometer, Turner Biosystems).

### 4.3.2 Results

#### 4.3.2.1 Dose Response Curve of NSC727152 in Prostate and Ovarian Cancer Cell Lines

The activity of the most potent compound NSC727152 was re-confirmed before proceeding with the secondary assays described. A dose-response curve was generated to identify its $IC_{50}$ value. The dose response curve is also used to determine the potency of the compound at inhibiting *in vitro* cell proliferation, in the different cancer cell lines. Figure 10 below illustrates the dose response curves obtained for NSC727152 in DU-145, PC-3 and OVCAR3 cancer cell lines.

The data from two independent experiments performed in quadruplet, were combined for analysis on the statistical programme GraphPad Prism Windows version 5.01 (GraphPad Software, San Diego California, USA). A sigmoidal dose response curve (variable slope) and the non-linear regression were used to analyse the data. The $IC_{50}$ values obtained from the androgen-independent prostate cancer cell lines, suggest that NSC727152 has a more potent effect in DU-145 ($IC_{50}$ = 8.7μM) cell line compared to the PC-3 cell line ($IC_{50}$ = 24.7μM). The ovarian cancer cell line, OVCAR3, was equally sensitive to NSC727152 as the DU-145 cell line, which obtained an $IC_{50}$ of 7.2μM (Figure 10).

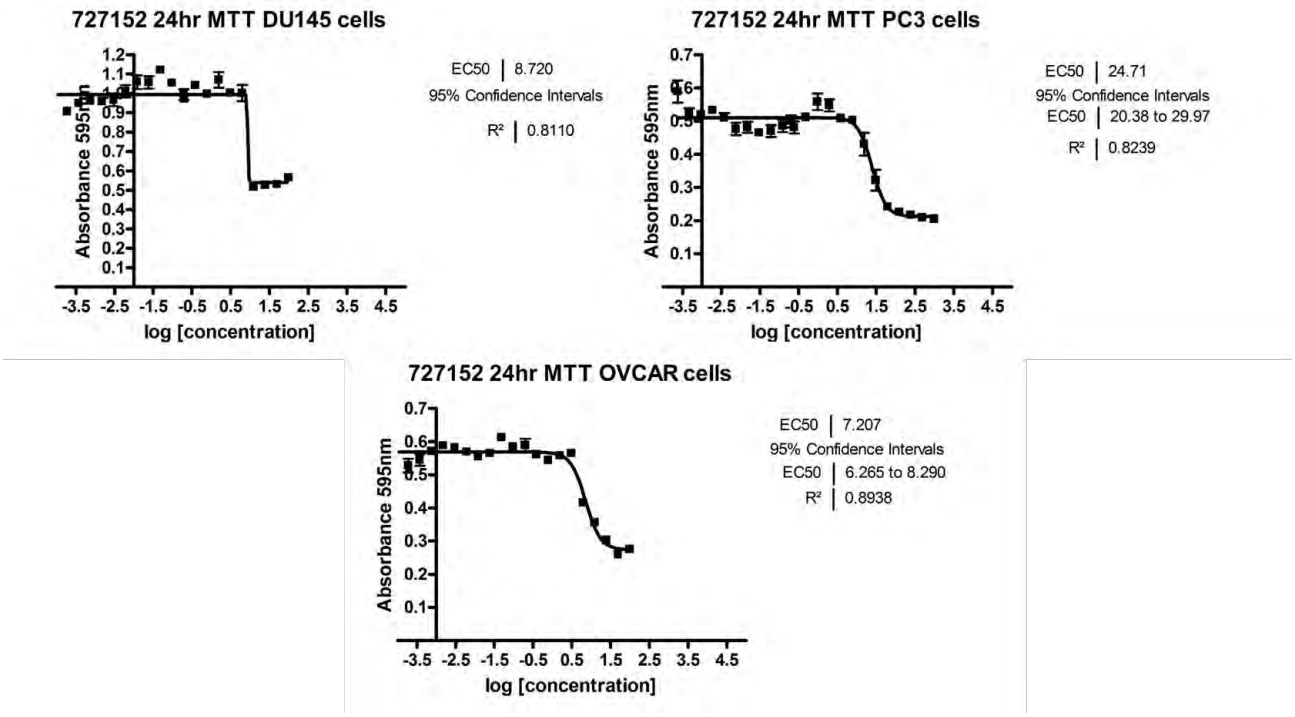



**Figure 10: The dose response curves of NSC727152 in DU-145, PC3 and OVCAR3 cancer cell lines after 24 hours treatment, generated on Prism 4.2. The compound demonstrated varying potencies in the different cancer cell lines, being most potent in the ovarian cancer cell line, OVCAR3. Data shown represents mean ± SD of 4 independent experiments performed in duplicate.**

#### 4.3.2.2 The Effect of NF-κB Inhibition with NSC727152 on the Expression of GADD45α and GADD45γ in Prostate Cancer Cells

The constitutive activation of NF-κB has been previously reported to allow cancer cells to escape programmed cell death. Furthermore, the NF-κB mediated cell survival has been shown to be completely dependent on the repression of GADD45α and GADD45γ and not GADD45β. Conversely, it has also been reported that the inhibition of NF-κB results in the up-regulation of GADD45α and GADD45γ mRNA expression (L. F. Zerbini et al, 2004). The effect of NSC727152 on the expression of GADD45α and GADD45γ was determined, to elucidate if NF-κB was being inhibited. As seen in Figure 11, the expression of GADD45α and γ were up-regulated at the mRNA level, in DU-145 cancer cells 6 hours after treatment with NSC727152. In PC-3 cells, the significant up-regulation of both GADD45α and GADD45γ mRNA was observed 6 hours post-treatment with NSC727152 (Figure 12). These results are consistent with reported findings and suggest that NSC727152 inhibits NF-κB.

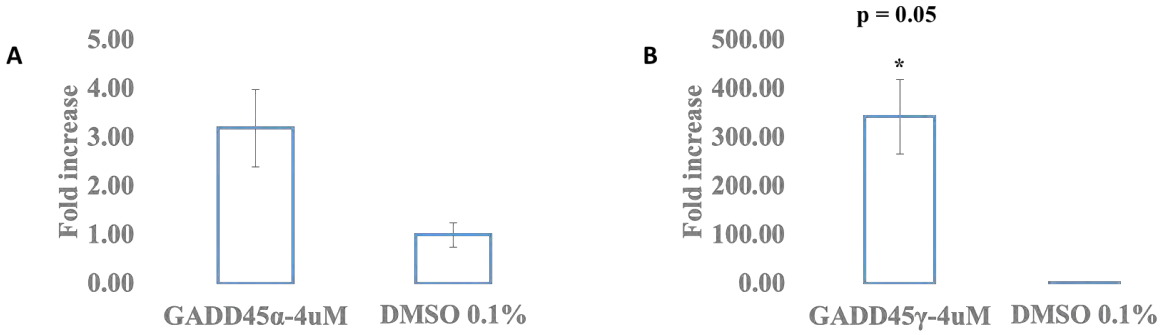



**Figure 11: Real-time PCR analysis of mRNA expression of GADD45α and γ expression in DU-145, 6 hours after treatment with NSC727152. A) GADD45α and B) GADD45γ (statistical significance p<0.05). Each RNA was normalised to GAPDH. GADD45α and γ expression levels are shown as a fold induction over the DMSO 0.1% control.**

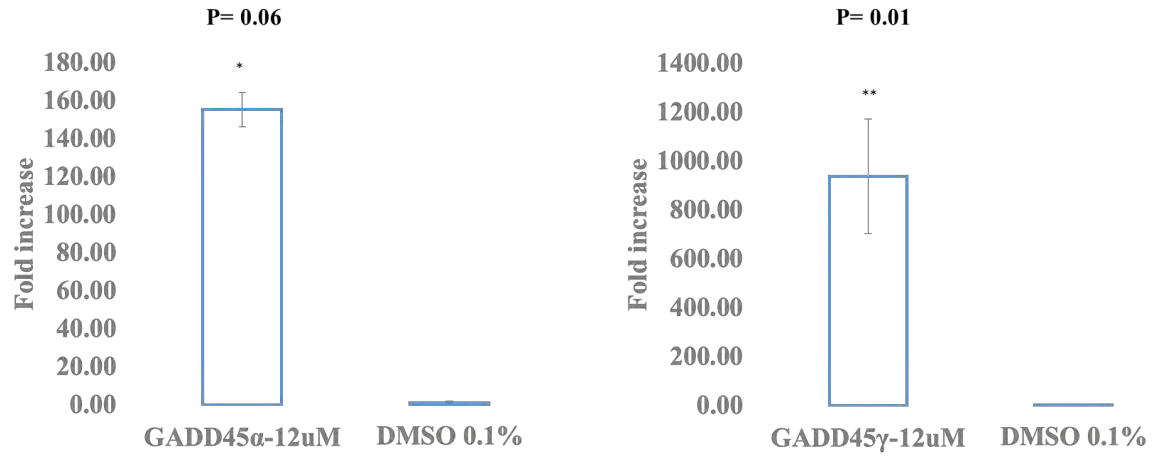



**Figure 12: Real-time PCR analysis of GADD45α and γ expression in PC-3 cells after 6 hour of treatment with NSC727152. A) GADD45α (statistical significance p<0.06) and B) GADD45γ (statistical significance p<0.01). Each RNA was normalised to GAPDH. GADD45α and γ expression levels are shown as a fold induction over the DMSO 0.1% control.**

#### 4.3.2.3 To Determine the Effect of NF-κB Inhibition with NSC727152 on the mRNA Expression of IL-6

The constitutive activation of NF-κB and AP-1 transcription factors has been reported in literature to mediate the aberrant up-regulation of IL-6 mRNA expression in both PC-3 and DU-145 cancer cell lines (L. F. Zerbini et al, 2004). The IL-6 mRNA expression is reportedly

elevated in hormone independent prostate cancer cells DU-145 and PC-3 compared to hormone dependent prostate cancer cell line LNCap and is believed to be one of the factors which promotes the progression of cancer (L. F. Zerbini et al, 2004). The inhibition of NF-κB has been previously reported to drastically reduce the expression of IL-6 mRNA in hormone independent cancer cell lines (L. F. Zerbini et al, 2004). After 6 hours of treatment with NSC727152, the expression of IL-6 in both androgen independent cancer cell lines was reduced by approximately 0.6 fold (Figure 13).

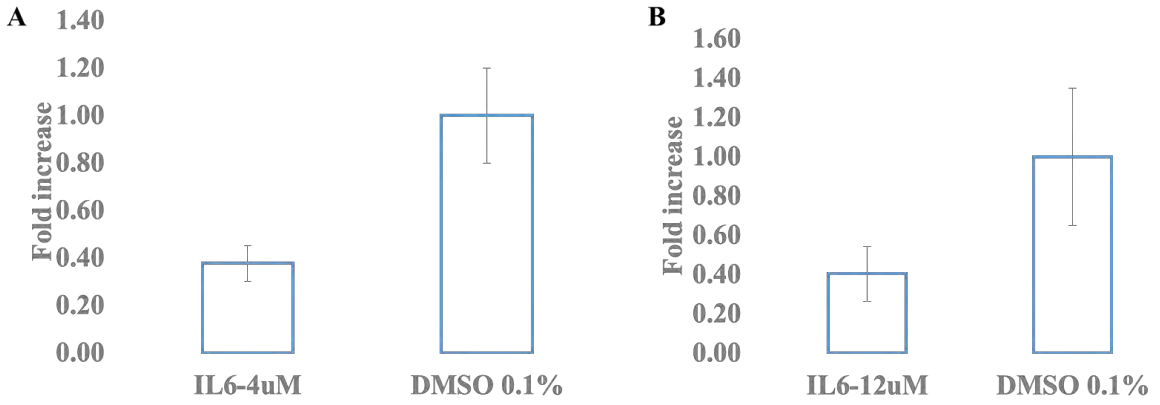



**Figure 13: Real-time PCR analysis of IL-6 expression after 6 hour treatment with NSC727152. A) In DU-145 and B) PC3 cancer cells. Each RNA was normalised to GAPDH IL-6 expression and levels are shown as a fold induction over the DMSO 0.1% control.**

#### 4.3.2.4 The Down-regulation of IL-6 mRNA Expression Correlates with its Reduced Rromoter Activity

This laboratory and others have previously used the luciferase reporter assay to establish the binding role of NF-κB on the IL-6 promoter region (L. F. Zerbini et al, 2004; T. A. Libermann et al, 1990). Through site directed point mutations of the NF-κB and AP-1 regulatory elements, a decrease in the IL-6 promoter activation was observed, further substantiating the activator and binding roles of NF-κB and AP-1 (L. F. Zerbini et al, 2004). The activation of the IL-6 promoter by the constitutively activated NF-κB and AP-1 transcription factors and the concomitant up-regulation of IL-6 mRNA expression, have been implicated in the progression of prostate cancer from a hormone dependent state to a hormone independent state (L. F. Zerbini et al, 2004). The inhibition of NF-κB in DU-145 and PC-3 cancer cell lines has been previously reported to result in the reduced promoter activity of the IL-6 gene (L. F. Zerbini et al, 2004). To determine if the reduced mRNA

expression of IL-6 correlated with the IL-6 promoter activity, the promotor activity of the IL-6 gene was measured. Approximately, a 25% reduction of the IL-6 promoter activity was observed after 6 hours of treatment with NSC727152 at a concentration of 32µM in DU-145 cells (Figure 14). To further validate the inhibition of NF-κB by NSC727152, the promoter activity of the consensus sequence κB promoter activity in DU-145 cells was measured. The reduced promoter activity that was observed after 6 hours treatment at a concentration of 4µM was approximately 25%, confirming the inhibition of NF-κB (Figure 14).

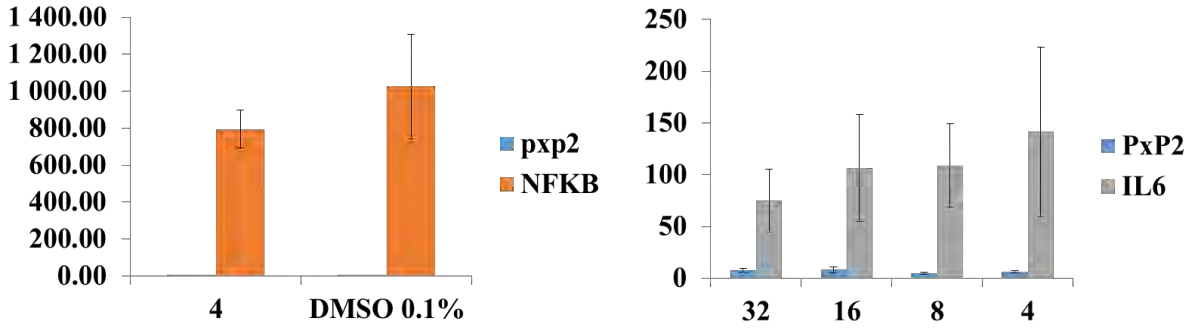



**Figure 14: The transcriptional activity of A) the consensus sequence κB promoter B) the IL-6 promoter in DU-145 prostate cancer cells after 6 hours treatment with NCS727152 and DMSO 0.1% control. Data shown represents mean +/- SD of 3 individual samples. Luciferase activity is shown as a fold induction of the parental vector, pxp2, shown on the left (blue bars).**

### 4.3.3 Discussion

The *in vitro* evaluation of NSC727152 was confirmed to inhibit cancer cell proliferation in a dose-dependent manner, in hormone independent prostate cancer cell lines including PC-3 and DU-145 and in the ovarian cancer cell line, OVCAR3. Further analysis in normal prostate and ovarian cell lines would be required to determine its selectivity for cancer cells.

A literature search has revealed that a closely related analogue to NSC727152, known as Tachpyridine (Figure 15) has been reported to have *in vitro* anti-tumour activity and is currently in pre-clinical trials (J. Turner et al, 2005). Tachpyridine acts as a metal chelator that binds divalent metals such as $Fe^{2+}$, $Zn^{2+}$, $Cu^{2+}$, $Ca^{2+}$, $Mg^{2+}$ and $Mn^{2+}$ and is reported to be more cytotoxic to cancer cells than other metal chelators such as the commercially available Desferrioxamine (DFO) (R. Zhao et al, 2004). Tachpyridine has been reported to arrest cancer cells in the $G_2$/M phase of the cell cycle in the absence or presence of p38, a mitogen-

activated protein kinase, where it activates check point kinases in the cell cycle and sensitises cells to radiation therapy, resulting in the apoptosis of cancer cells (R. D. Abeysinghe et al, 2001; R. Zhao et al, 2004). Studies reported in literature have shown that the overexpression of the Bcl-xl, a transmembrane mitochondrial protein that acts as a tumour suppressor or pro-survival protein, abrogates the effects of Tachpyridine (B. T. Greene et al, 2002). Furthermore, inhibition of caspase 9, a pro-apoptotic protein, results in cell survival, which confirms the mitochondrial apoptotic pathway induced by Tachpyridine (B. T. Greene et al, 2002).

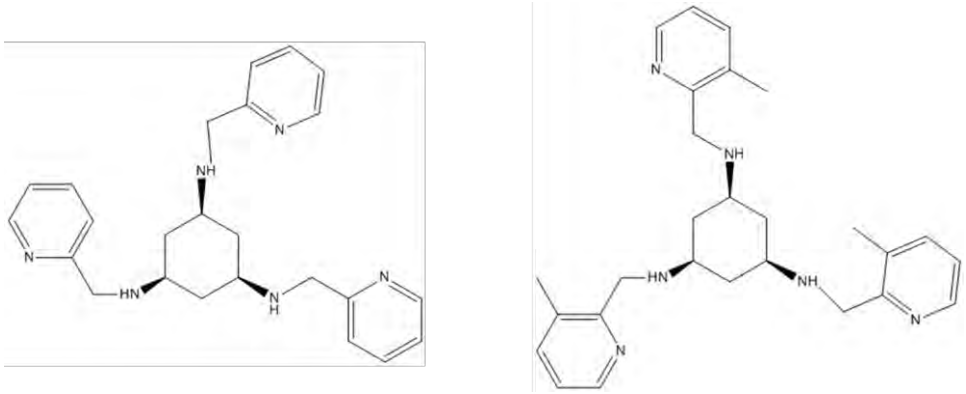



**Figure 15: The structures of Tachpyridine (left) and the analogue NSC727152 (right). NSC727152 was identified in this study to have anti-tumour activity by inhibiting NF-κB, at a low micro-molar concentration.**

Most iron chelators arrest cells in the $G_1$-S phase and the different mechanism of action of Tachpyridine is thought to be due to its $Cu^{2+}$ and $Zn^{2+}$ chelating activity (J. Turner et al, 2005). Compounds that are capable of neutralising reactive oxygen species (ROS) such as chelators or metal chelators that contain a dithiocarbamate functional group and thiol derivatives as well as several other agents have been reported to inhibit NF-κB at high micro molar concentrations (C. K. Sen et al, 1996). Interestingly, Tachpyridine has not been reported in literature to inhibit NF-κB, neither has its effects on genes that are regulated by NF-κB, such as the GADD45 proteins or IL-6, been investigated.

NF-κB has for a long time been identified as an attractive chemotherapeutic target due to its constitutive activation in many types of cancers including prostate and ovarian cancer where

it plays a key role in the development and progression of the disease (L. F. Zerbini et al, 2004; X. Dolcet et al, 2005; C. D. Chen et al, 2002; M. Karin, 2009). NF-κB is reported in literature to be responsible for the multidrug resistance observed in anticancer chemotherapy, further substantiating it as key therapeutic target (S. Papa et al, 2004; L. F. Zerbini et al, 2004). It is a transcriptional factor that has been studied extensively and has been found to bind directly to DNA and activates genes that work together to promote cancer cell growth. This is evident by the relationship between NF-κB and the GADD45 proteins, along with its activator role in the transcription of the IL-6 gene (L. F. Zerbini et al, 2004; L. F. Zerbini et al, 2003).

The induction of GADD45α and GADD45γ has been shown to lead to the inhibition of cell growth and apoptosis of cancer cells (L. F. Zerbini et al, 2004), but the role of GADD45β in response to NF-κB inhibition or activation is still not fully understood but is believed to either enhance or repress apoptosis depending on the stimulating factor and the pathway that it may activate (E. De Smaele et al, 2001; A. Amanullah et al, 2003). Taking into consideration the conflicting role of GADD45β, the effects of NSC727152 on GADD45β were not investigated. This study has demonstrated that the *in vitro* inhibition of NF-κB by NSC727152 results in the reduced cell proliferation of ovarian and prostate cancer cells. Furthermore, it was demonstrated that inhibition of NF-κB mediated by NSC727152 resulted in the up-regulation of GADD45α and GADD45γ mRNA expression.

In response to the inhibition of NF-κB, the upregulation of the GADD45 proteins have been previously shown to interact with the upstream kinase, mitogen-activated protein kinase kinase 4 (MEKK4), which activates both p38 and JNK, leading to apoptosis or inhibition of cell growth (L.F. Zerbini et al, 2005; M. Takekawa et al, 1998; M. Gupta et al, 2006). Although this analysis was not within the scope of this study, findings in literature indicate that iron chelators such as DFO have also been found to result in apoptosis of cancer cells by activation of the Mitogen-Activated Protein Kinase (MAPK) pathway (Y. Yu et al, 2011). The induced iron depletion was found to significantly increase the phosphorylation of JNK and p38 MAPKs, resulting in the phosphorylation of the downstream targets p58 and ATF-2. Taking this into consideration, it would be interesting to investigate if the upregulation of the GADD45 proteins induced by treatment with NSC727152 could result in apoptosis by activating the MAPK pathway.

Furthermore, the GADD45 proteins play an important role in cell cycle regulation (R. E. Tamura et al, 2012), where the upregulation of the GADD45 proteins in colorectal cancer cells with wild-type p53 function (a tumour-suppressor protein) was found to arrest cells in the $G_2$/M phase of the cell cycle, independent of p38 kinase activity (M. Vairapandi et al, 2002; R. E. Tamura et al, 2012). This was found to be a consequence of the inhibition of the cdc2/cyclin B1 complex by all the GADD45 proteins, where GADD45α and β were identified to disrupt the complex and not GADD45γ (M. Vairapandi et al, 2002; R. E. Tamura et al, 2012). This was an interesting finding considering that Tachypyridine, the closely related analogue of NSC727152 is also reported to arrest colorectal and ovarian cancer cells in the $G_2$/M phase of the cell cycle. The $G_2$/M cell cycle arrest observed in colorectal and ovarian cancer cell lines was found to be independent of p53 (R. D. Abeysinghe et al, 2001; R. Zhao et al, 2003). More importantly, the cytotoxic effects of Tachypyridine were observed in cell lines with both mutant and normal p53 (R. Zhao et al, 2003). In addition, Tachypyridine was found to sensitise only cancer cells and not normal cells to ionising radiation (J. Turner et al, 2005). It would certainly be interesting to investigate the effect of NSC727152 on the cell cycle, especially considering that NSC727152 induces an upregulation of the GADD45 proteins.

As mentioned previously, the constitutive activation of the IL-6 gene in hormone independent prostate cancer cell lines, which is mediated by NF-κB, plays a pivotal role in the progression of prostate cancer (L. F. Zerbini et al, 2003). IL-6 is thought to play a key role in driving the progression of prostate cancer from an androgen dependant prostate cancer to an androgen independent prostate cancer, which is more aggressive and associated with a higher death rate (L. F. Zerbini et al, 2003). This substantiates NF-κB as an attractive therapeutic target, as it binds directly on the IL-6 promoter region. In this study, the inhibition of NF-κB by NSC727152 resulted in the down-regulation of IL-6 mRNA expression and the reduced promoter activity of IL-6. The inhibition of NF-κB was further corroborated by investigating the effect of NF-κB binding on the κB consensus sequence promoter, which consistently resulted in the reduced promoter activity. The luciferase assay was performed to provide a preliminary indication of the effects of NSC727152 on the NF-κB promoter activity. Further experiments such as chromatin immunoprecipitation and/or EMSA gel shifts are necessary to better determine and substantiate the inhibitory effects of NSC727152 on the binding activity of NF-κB.

It is noteworthy that these studies reported in literature were conducted under different experimental conditions, using different cell types and taking into consideration the conflicting accounts reported in literature, it would be necessary to confirm these hypothesis under the same experimental parameters utilised in this study.

### 4.3.4 Conclusion

In conclusion, the *in vitro* activity of the four hits identified from the NCI database virtual screening was confirmed in prostate and ovarian cancer cell lines. The most potent of the 4 compounds, NSC727152 was shown to inhibit NF-κB activity, which resulted in the down regulation of IL-6 mRNA, the reduced promoter activity of the IL-6 gene and the consensus sequence κB promoter. Furthermore, as a consequence of the inhibition of NF-κB by NSC727152, an up-regulation of both GADD45α and GADD45γ mRNA in prostate cancer cell lines was observed.

### 4.3.5 References


1. M. Hughes, J. Inglese, A. Kurtz, A. Andalibi, L. Patton, C. Austin, M. Baltezor, M. Beckloff, S. Sittampalam, M. Weingarten and S. Weir, Early Drug Discovery and Development Guidelines: For Academic Researches, Collaboratoryorators and Start-up Companies, 2012. In: G. S. Sittampalam, N. Gal-Edd, M. Arkin et al, Assay Guidance Manual, Eli Lilly & Company and the National Centre for Advancing Translational Sciences, 2004. http://www.ncbi.nlm.nih.gov/boooks/NBK92015.
2. A. M. Burger and H. H. Fiebig, Preclinical Screening for New Anticancer Agents, 2014. In:M. A. Rudek et al, Handbook of Anticancer Pharmacokinetics and Pharmacodynamics, Cancer Drug Discovery and Development, Springer Science and Business Media, 2014.
3. A. S. Verkman, Drug Discovery in Academia, American Journal of Physiology Cell Physiology, Vol. No. 286, Pg. No. C465-C474, 2004.
4. J. Major, Challenges and opportunities in high throughput screening: implications for new technologies, Journal of Biomolecular Screening, Vol. 3, Issue No. 1, Pg. No. 13-17, 1998.
5. V. Khurana, D. F. Tardiff, C. Y. Chung and S. Lindquist, Towards Stem Cell-Based Phenotypic Screens for Neurodegenrative Diseases, Nature Reviews Neurology, Vol. No. 11, Pp. 339-350, 2015.
6. Z. Zhang, N. Guan, T. Li, D. E. Mais and M. Wang, Quality Control of Cell-Based High-Throughput Drug Screening, Acta Pharmaceutica Sinica B, Vol. No. 2, Issue No. 5, Pp. 429-438, 2012.
7. E. A. Martis, R. Radhakrishnan, R. R. Badve, High Through-put Screening: The Hits and Leads of Drug Discovery- An Overview, Journal of Applied Pharmaceutical Science, Vol. No. 01, Issue. No. 1, Pp. 2-10, 2011.
8. E. C. Butcher, E. L. Berg and E. J. Kunket, Systems Biology in Drug Discovery, Nature Biotechnology Perspective, Vol. No. 22, Issue No. 10, Pp.1253-1259, 2004.
9. L. F. Zerbini, R. E. Tamura, R. G. Correa, A. Czibere, J. Cordeiro, M. Bhasin, F. M. Simabuco, Y. Wang, X. Gu, L. Li, D. Sarkar, J. R. Zhou, P. B. Fisher and T. A. Libermann, Combinatorial Effect of Non-Steroidal Anti-inflammatory Drugs and NF-κB Inhibitors in Ovarian Cancer Therapy, PLoS ONE, Vol. No. 6, Issue No. 9, 2011.

10. L. F. Zerbini, M. K. Bhasin, J. F. de Vasconcellos, J. D. Paccez, X. Gu, A. L. Kung and T. A. Libermann, Computational Repositioning and Preclinical Validation of Pentamidine for Renal Cell Cancer, Molecular Cancer Therapuetics, Vol. No. 13, Issue No. 7, Pp. 1929-1941, 2014.
11. A. S. Narang and D. S. Desai, Anicancer drug development in Pharmaceutical Perspectives of Cancer Therapeutics, Ed. Y. Lu and R. I. Mahato, Springer Science and Business Media, 2009.
12. T. L. Riss, R. A. Moravec, A. L. Niles, H. A. Benink, T. J. Worzella and L. Minor, Cell Viability Assays, In: Assay Guidance Manual, Eli Lilly & Company, 2013.
13. T. Mosmann, Rapid Colorimetric Assay for Cellular Growth and Survival: Application to Proliferation and Cytotoxicity Assays, Journal of Immunological Methods, Vol. No. 65, Pg. No. 55-63, 1983.
14. J. P. Hughes, S. Rees, S. B. Kalindjian and K. L. Philpott, Principles of Early Drug Discovery, British Journal of Pharmacology, Vol. No. 162, Issue No. 6, Pp. 1239-1249, 2011.
15. S. Hoelder, P. A. Clarke and P. Workman, Discovery of Small Molecula Cnacer Drugs: Successes, Challenges and Opportunities, Molecular Oncology, Vol. No. 6, Issue No. 2, Pp. 155-176, 2012.
16. M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, H. Takahashi, T. Fukazawa and H. Hayashi, Discovery of Quinazolines as a Novel Structural Class of Potent Inhibitors of NF-κB Activation, Bioorganic & Medicinal Chemistry, Vol. No. 11, Pp. 383-391, 2003.
17. M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, H. Takahashi and H. Hayashi, A Novel Structural Class of Potent Inhibitors of NF-κB Activation: Structure–Activity Relationships and Biological Effects of 6-Aminoquinazoline Derivatives, Vol. No. 11, Pp. 3869-3878, 2003.
18. S. Turcotte, D. A. Chan, P. D. Sutphin, M. P. Hay, W. A. Denny and A. J. Giaccia, A molecule targeting VHL-deficient Renal Cell Carcinoma that induces autophagy, Cancer Cell, Vol. No. 14, Issue No. 1, 2008.
19. A. Chaudhary, U. Nagaich, N. Gulati, V. K. Sharma and R. L. Khosa, Enhancement of solubilisation and bioavailaboratoryility of poorly soluble drugs by physical and chemical modifications: A recent review, Journal of Advanced Pahrmacy Education & Research, Vol. No. 2, Issue No,. , Pg. No. 32-67, 2012. C. H. Chuang, T. C. Cheng, Y. L. Leu, K. H. Chuang, S. C. Tzou and C. S. Chen, Discovery of AKT Kinase

Inhibitors Through Structure-Based Virtual Screening and their Evaluation as Potential Anticancer Agents, International Journal of Molecular Science, Vol. No. 16, Issue No. 2, Pp. 3202-3012, 2015.

20. H. Park, S. Lee and S. Hong, Discovery of MEK/PI3K Dual Inhibitor via Structure-Based Virtual Screening, Bioorganic Medicinal Chemistry Letters, Vol. No. 22, Issue No. 15, Pp.4946-4950, 2012.
21. G. Cozza, P. Bonvini, E. Zirzi, G. Poletto, M. A. Pagano, S. Sarno, A. D. Deana, G. Zagotto, A. Rosolen, L. A. Pinna, F. Meggio and S. Moro, Identification of Ellagic Acid as Potent Inhibitors of Protein Kinase CK2: A Successful Example of a Virtual Screening Application, Journal of Medicinal Chemistry, Vol. No. 49, Issue No. 8, Pp. 2363-2366, 2006.
22. B. K. Schoichet, Virtual Screening of Chemical Libraries, Nature, Vol. No. 432, Issue No. 7019, Pp. 862-865, 2004.
23. S. Subramaniam, M. Mehrotra and D. Gupta, Virtual High Throughput Screening (vHTS) – A Perspective, Bioinformation: Views & Challenges, Vol. No. 3, Issue No. 1, Pp. 14-17, 2008.
24. J. Bajorath, Integration of Virtual and High Throughput Screening, Nature Review Drug Discovery, Vol. No. 1, Issue No. 11, Pp. 882-894, 2002.
25. A. S. Reddy, S. P. Pati, P. Kumar, H. N. Pradeep and G. N. Sastry, Virtual Screening in Drug Discovery – A Computational Perspective, Current Protein and Peptide Science, Vol. No. 8, Issue No. 4, Pp. 329-351, 2007.
26. K. S. Watts, P. Dalal, R. B. Murphy, W. Sherman, R. A. Freisner and J. C. Shelley, ConfGen: A conformational search method for efficient generation of bioactive conformers, Journal of Chemical Information and Modelling, Vol. No. 50, Issue No. 4, Pp. 534-546, 2010.
27. J. C. Shelley, A. Cholleti, L. L. Frye, J. R. Greenwood, M. R. Timlin and M. Uchimaya, Epik: A Software Program for pK (a) Prediction and Protonation State Generation for Drug-Like Molecules, Journal of Computer-Aided Molecular Design, Vol. No. 21, Issue No. 12, Pp. 681-691, 2007.
28. A. P. McGrath, K. M. Hilmer, C. A. Collyer, E. M. Shepard, B. O. Elmore, D. E. Brown, D. M. Dooley and J. M. Guss, Structure and Inhibition of Human Diamine Oxidase, Biochemistry, Vol. No. 48, Issue No. 41, Pp. 9810-9822, 2009.
29. M. A. Lill, Efficient Incorporation of Protein Flexibility and Dynamics into Molecular Docking Simulations, Biochemistry, Vol. No. 50, Issue No. 28, Pp. 6157-6169, 2011.

30. M. Fischer, R. G. Coleman, J. S. Fraser and B. K. Shoichet, The Incorporation of Protein Flexibility and Conformational Energy Penalties in Docking Screens to Improve Ligand Discovery, Nature Chemistry, Vol. No. 6, Issue No. 7, Pp. 575-583, 2014.
31. D. B. Kitchen, H. Decornez, J. R. Furr and J. Bajorath, Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications, Nature Reviews Drug Discovery, Vol. No. 3, Pp. 935-949, 2004.
32. A. Oeckinghaus and S. Ghosh, The NF-κB Family of Transcription Factors and Its Regulation, Cold Spring Harbour Perspectives in Biology, Vol. No. 1, Issue No. 4, 2009.
33. A. E. Pegg, Mammalian Polyamine Metabolism and Function, IUBMB Life, Vol. No. 61, Issue No. 9, Pp. 880-894, 2009.
34. S. Yang, T. Wenzler, P. N. Miller, H. Wu, D. W. Boykin, R. Brun and M. Z. Wang, Pharmacokinetic Comparison to Determine the Mechanisms Underlying the Differential Efficacies of Cationic Diamidines against First- and Second-Stage Human Africa Trypanosomiasis, Antimicrobial Agents and Chemotherapy, Vol. No. 58, Issue No. 7, Pp. 4064-4074, 2014.
35. A. Puckowska, D. Drozdowska, M. Rusak, T. Bielawski, I. Bruzgo and K. M. Nowaczek, Amino and Chlorambucil Analogues of Pentamidine – Synthesis and Biological Examination, Acta Poloniae Pharmaceutica- Drug Research, Vol. No. 69, Issue No. 1, Pp. 63-73, 2012.
36. J. Stanek, G. Caravatti, H. G. Capraro, P. Furet, H. Mett, P. Schneider and U. Regenass, S-Adenosylmethionine Decarboxylase Inhibitors: New Aryl and Heteroaryl Analogues of Methylglyoxal Bis (guanylhydrazone), Journal of Medicinal Chemistry, Vol. No. 36, Issue No. 1, Pp. 46-54, 1993.
37. U. Regenass, G. Caravatti, H. Mett, J. Stanek, P. Schneider, M. Muller, A. Matter, P. Vertino and C. W. Porter, New S-Adenosylmethionine Decarboxylase Inhibitors with Potent Antitumor Activity, Cancer Research, Vol. No. 52, Pp. 4712-4718, 1992.
38. K. Oostdik, A. Petterson, T. Schagat and D. Storts, Stem Cell Line Authentification and Contamination Detection, 2009, http://www.promega.com/resources/pubhub/enotes/stem-cell-line-authentication-and-contamination-detection/
39. L. Di and E. H. Kerns, Stability challenges in drug discovery, Chemistry & Biodiversity, Vol. No. 6, 2009.

40. J. D. Paccez, K. Duncan, A. Vava, R. G. Correa, T. A. Libermann, M. I. Parker and L. F. Zerbini, Inactivation of GSK3β and activation of NF-κB pathway via Axl represents an important mediator of tumorigenesis in esophageal squamous cell carcinoma, Molecular Biology of the Cell, Vol. No. 26, Issue No. 5, Pp. 821-831, 2015.
41. V. Venkatraman, V. I. Perez-Nueno, L. Mavridis and D. W. Ritchie, Comprehensive Comparison of Ligand-Based Virtual Screening Tools against the DUD Data set Reveals Limitations of Current 3D Methods, Journal of Chemical Information and Modelling, Vol. No. 50, Issue No. 12, Pp. 2079-2093, 2010.
42. M. Thimm, A. Goede, S. Hougardy and R. Preissner, Comparison of 2*D* Similarity and 3*D* Superposition. Application to Searching a Conformational Drug Database, Journal of Chemical Information and Computer Science, Vol. No. 44, Issue No. 5, Pp. 1816-1822, 2004.
43. P. Tuffery and P. Derreumaux, Flexibility and binding affinity inprotein–ligand, protein–protein and multi-component protein interactions: limitations of current computational approaches, Journal of The Royal Society Interface, Vol. No. 9, Pp. 20-35, 2012.
44. Q. V. Vuong, K. Siposova, T. T. Nguyen, A. Antosova, L. Balogova, L. Drajna, J. Imrich, M. S. Li and Z. Gazova, Binding of Glyco-Acridine Derivatives to Lysozyme Leads to Inhibition of Amyloid Fibrillization, Biomacromolecules, Vol. No. 4, Pp. 1035-1043, 2013.
45. N. Thorne, D. S. Auld and J. Inglese, Apparent activity in high-throughput screening: origins of compound-dependant assay interference, Current Opinion in Chemical Biology, Vol. No. 14, Issue No. 3, Pp. 315-324, 2010.
46. L. F. Zerbini, Y. Wang, A. Czibere, R. G. Correa, J. Cho, K. Ijiri, W. Wei, M. Joseph, X. Gu, F. Grall, M. B. Goldring, J. Zhou and T. A. Libermann, NF-κB-mediated repression of growth arrest- and DNA-damage-inducible proteins 45α and γ is essential for cancer cell survival, PNAS, Vol. No. 101, Issue No. 37, Pp. 13618-13623, 2004.
47. T. A. Libermann and D. Baltimore, Activation of interlekin-6 gene expression through NF-κB transcription factor, Molecular and Cellular Biology, Vol. 10, Issue No. 5, Pg. No. 2327-2334, 1990.
48. L. F. Zerbini, Y. Wang, J. Cho and T. A. Libermann, Constitutive Activation of Nuclear Factor κB p50/p65 and Fra-1 and JunD Is Essential for Deregulated

Interleukin 6 Expression in Prostate Cancer, American Association for Cancer Research, Vol. No. 63, Pp. 2206-2215, 2003.

49. L. F. Zerbini and T. A. Libermann, GADD45α and γ are Critical Regulators of NF-κB Mediated Escape from Programmed Cell Death, Cell Cycle, Vol. No. 4, Issue No. 1, Pp. 18-20, 2005.
50. M. Karin, NF-κB as a critical link between inflammation and cancer, Cold Spring Harbour Perspectives in Biology, 2009.
51. J. Turner, C. Koumenis, T. E. Kute, R. P. Planalp, M. W. Brechbiel, D. Beardsley, B. Cody, K. D. Brown, F. M. Torti and S. V. Torti, Tachpyridine, a metal chelator, induces G2 cell-cycle arrest, activates checkpoint kinases, and sensitizes cells to ionizing radiation, Blood Journal, Vol. No. 106, Issue No. 9, Pp. 3191-3199, 2005.
52. R. Zhao, R. P. Planalp, R. Ma, B. T. Greene, B. T. Jones, M. W. Brechbiel, F. M. Torti and S. V. Torti, Role of zinc and iron chelation in apoptosis mediated by tachpyridine, an anti-cancer iron chelator, Biochemical Pharmacology, Vol. 67, Issue No. 9, Pg. No. 1677-1688, 2004.
53. R. D. Abeysinghe, B. T. Greene, R. Haynes, M. C. Willingham, J. Turner, R. P. Planalp, M. W. Brechbiel, F. M. Torti and S. V. Torti, p53-independent apoptosis mediated by tachpyridine, ananti-cancer iron chelator, Oxford Journals, Carcinogenesis, Vol. 22, Issue No. 10, Pg. No. 1607-1614, 2001.
54. B. T. Greene, J. Thorburn, M. C. Willingham, A. Thorburn, R. P. Planalpl, M. W. Brechbiel, J. Jennings-Gee, J. WilkinsonIV, F. M. Torti and S. V. Torti, Activation of caspase pathways during iron chelator-mediated apotosis, The Journal of Biological Chemistry, Vol. 277, Issue No. 28, Pg. No. 25568-25575, 2002.
55. C. K. Sen and L. Packer, Antioxidant and redox regulation of gene transcription, The FASEB Journal, Vol. No. 10, Pg. No. 709-720, 1996.
56. X. Dolcet, D. Llobet, J. Pallares and X. Matias-Guiu, NF-κB in developemtn and progression of human cancer, Issue 446, Pg. No. 475-482, 2005.
57. C. D. Chen and C. L. Sawyers, NF-κB activates prostate-specific antigen expression and is up-regulated in androgen independent prostate cancer, Molecular and Cellular Biology, Vol. 22, Issue No. 8, Pg. No. 2862-2870, 2002.
58. S. Papa, F. Zazzeroni, C. Bubici, S. Jayawardena, K. Alvarez, S. Matsuda, D. U. Nguyen, C. G. Pham, A. H. Nelsbach, T. Melis, E. D. Smaele, W. Tang, L. D'Adamio and G. Franzoso, Gadd45β mediates the NF-κB suppression of JNK signalling by

targeting MKK7/JNKK2,Nature Cell Biology, Vol. No. 6, Issue No. 2, Pp. 146-153, 2004.

59. E. De Smaele, F. Zazzeroni, S. Papa, D. U. Nguyen, R. Jin, J. Jones, R. Cong and G. Franzoso, Induction of gadd45β by NF-κB downregulates pro-apoptotic JNK signalling, Nature, Vol. 414, Pg. No. 308-313, 2001.
60. A. Amanullah, N. Azam, A. Balliet, C. Hollander, B. Hoffman, A. Fornace Jr and D. Liebermann, Cell signalling (brief communications): cell survival and a Gadd45-factor deficiency, Nature, Issue No. 424, Pg. No. 421-422, 2003.
61. M. Takekawa and H. Saito, A family of stress-inducible Gadd45-like proteins mediate activation of the stress-responsive Mtk1/Mekk4 Mapkkk, Cell, Vol. 95, Pg. No. 521-530, 1998.
62. M. Gupta, S. K. Gupta, B. Hoffman and D. A. Liebermann, Gadd45α and Gadd45β protect hematopoietic cells from UV-induced apoptosis via distinct signalling pathways, including p38 activation and JNK inhibition, The Journal of Biological Chemistry, Vol. 281, Issue No. 26, Pg. No. 17552-17558, 2006.
63. Y. Yu and D. R. Richardson, Cellular iron depletion stimulates the JNK and p38 MAPK signalling transduction pathways, dissociation of ASK1-thioredoxin, and activation of ASK1, Journal of Biological Chemistry, Vol. No. 286, Issue No. 17, Pp. 15413-15427, 2011.
64. R. E. Tamura, J. F. de Vasconcellos, D. Sarker, T. A. Libermann, P. B. Fisher and L. F. Zerbini, GADD45 proteins: central players in tumorigenesis, Current Molecular Medicine, Vol. No. 12, Issue No. 5, Pg. No. 634-651, 2012.
65. M. Vairapandi, A. G. Balliet, B. Hoffman and D. A. Liebermann, GADD45b and GADD45g are cdc2/cyclinB1 kinase inhibitors with a role in S and G2/M cell cycle checkpoints induced by genotoxic stress, Journal of Cellular Physiology, Vol. No. 192, Issue No. 3, Pp. 327-338, 2002.

# Chapter 5: Summary and Conclusions

## 5.1 Summary

This study demonstrated that vHTS is a useful alternative to HTS in the identification of active compounds from large compound databases in a timely and cost-effective manner (S. Subramaniam et al, 2008). A virtual screening of the ZINC and NCI databases was performed to identify similar compounds to QNZ and Pentamidine. Several SBVS and LBVS techniques were utilised to identify a selected number of compounds for *in vitro* screening. The compounds were screened for cytotoxic activity in prostate, ovarian and renal cancer cell lines. Four compounds resulting from the LBVS performed for QNZ on the NCI database were confirmed to have *in vitro* activity. Furthermore, the most potent compound, NSC727152, was shown to inhibit NF-κB similarly to its parental compound, QNZ.

### 5.1.1. Virtual Screening of the ZINC Database

The hierarchical virtual screening of the ZINC database involved of performing a LBVS followed by a SBVS from which good, intermediate and poor scoring compounds were selected for *in vitro* screening. As there is no crystal structure of the NF-κB proteins in complex with an inhibitor and the target of QNZ is not known, a blind docking protocol was employed to identify the binding sites of QNZ on the NF-κB pathway proteins, in order to perform the SBVS. Two putative binding sites of QNZ were identified on the NF-κB/IκBα complex, which were found to be located in close proximity to two previously described 'hot spots' in literature (S. Bergqvest et al, 2008). QNZ was predicted to have a lower binding energy at the binding site 2, where the phenoxyphenyl moiety of QNZ is predicted to be buried in the hydrophobic pocket formed between the p50/IκBα subunits (Chapter 2, section 2.3.2). None of the compounds identified from the docking calculations performed on the NF-κB/IκBα complex were confirmed to have cytotoxic activity at a starting concentration of 1000μM.

The docking calculations performed for Pentamidine were performed using the DAO crystal structure (3HII.pdb) (A. P. McGrath et al, 2009). An analysis of the binding interactions

revealed that Pentamidine and none of the ZINC database compounds formed binding interactions with the conserved residue Asp186 of the DAO enzyme. However, Pentamidine and the ZINC compounds identified from the 3D similarity search did interact with the Asp373 residue which forms part of the catalytic base of DAO, indicating that an inhibitory action is likely to occur (A. P. McGrath et al, 2009) (Chapter 2, section 2.6.2). The binding interactions made by the ZINC compounds were similar to those made by Pentamidine but despite this, the ZINC compounds did not show *in vitro* cytotoxic activity at a starting concentration of 1000μM in renal cancer cells.

### 5.1.2. Virtual Screening of the NCI Database

A chemo-informatic approach was employed for the virtual screening performed on the NCI database. An analysis of available SAR data from literature was performed to determine the most beneficial VS techniques to utilise in order to identify similar compounds to QNZ and Pentamidine. Taking into consideration the SAR data extracted from literature for QNZ and Pentamidine, only a LBVS was performed on the NCI database.

An analysis of the SAR data available in literature for QNZ (M. Tobe et al, 2003; M. Tobe et al, 2003) resulted in the generation of 4 pharmacophore hypotheses, which aimed to rationalise the importance of the different functional groups identified for activity. Furthermore, the summarised SAR was utilised to select compounds with the desired functional groups from the similarity search that was performed (Chapter 3, section 3.2.7).

As a consequence of having no relevant SAR data for Pentamidine, a 2-aminopyridine substructure search was performed for Pentamidine. This was performed based on the hypothesis that the 2-aminopyridine functional moiety may act similarly as the benzamidine moiety of Pentamidine in conferring inhibitory effects on the enzymes of the polyamine biosynthetic pathway. None of the compounds identified from the substructure search were confirmed to have *in vitro* activity in renal cancer cells at a starting concentration of 1000μM. In addition, a binding hypothesis of Pentamidine in the crystal structure of the human AdoMet-DC enzyme was generated to characterise the binding orientation and interaction of Pentamidine, to provide insights into the possible inhibitors of the AdoMet-DC enzyme. The predicted binding pose of Pentamidine indicates that not all of the functional moieties of Pentamidine are involved in binding interactions with the AdoMet-DC enzyme. Furthermore,

the stabilising π-cation bond was not predicted in the binding interactions made between Pentamidine and the AdoMet-DC enzyme (Chapter 3, section 3.6.1). The π-cation bond is thought to stabilise the inhibitors in a favourable *syn* conformation leading to increased potency (S. Bale et al, 2009; D. E. McCloskey et al, 2009). Therefore it was hypothesised that the less favourable *anti* conformation was adopted by Pentamidine, which would account for its reduced potency compared to the other inhibitors of AdoMet-DC (S. Bale et al, 2009; D. E. McCloskey et al, 2009). In addition, the reportedly poor permeability of Pentamidine in comparison to other diamines, may contribute to its low potency compared to other inhibitors of AdoMet-DC (D. E. McClosky et al, 2009; S. Yang et al, 2014).

### 5.1.3. Biological Screening Assays

The biological assays were performed to confirm the activity of the hit compounds identified through the virtual screening techniques performed. The cytotoxic activity of the compounds was determined using the MTT cell proliferation assay.

Four novel hits identified from the LBVS performed for QNZ on the NCI database were confirmed with *in vitro* cytotoxic activity, in the micro-molar range. Only one compound was identified from the similarity search and the remaining three compounds were identified from the pharmacophore virtual screening. The active compound identified from the similarity search demonstrated the importance of a functional moiety at the $R_1$ position for activity (Chapter 4, section 4.2.2.3.2). The compounds identified from the pharmacophore hypothesis demonstrated the importance of the basic nitrogen on the quinazoline core over the aromatic ring of the same core for increased activity (Chapter 4, section 4.2.2.3.2). The most active compound NSC727152, was confirmed to have activity in the prostate cancer cell lines, DU-145 ($IC_{50}$=8μM) and PC-3 ($IC_{50}$=24μM) and in the ovarian cancer cell line, OVCAR3 ($IC_{50}$=7μM). Its mechanism of action was also confirmed to be similar to QNZ by acting as an inhibitor of NF-κB (M. Tobe et al, 2003) (M. Tobe et al, 2003). Consistent with previously reported data in literature, in response to the inhibition of NF-κB the compound NSC727152 was shown to induce an upregulation of the GADD45α and γ mRNA expression and a down-regulation of IL-6 mRNA expression (L. F. Zerbini et al, 2003; L. F. Zerbini et al, 2004). Furthermore, treatment with NSC727152 resulted in the reduced promotor activity of the IL-6 and the consensus sequence κB promotors (Chapter 4, section 4.3.2).

## 5.2 Conclusions

In conclusion, a combination of SBVS and LBVS techniques were utilised to screen the ZINC and NCI databases.

In the absence of a crystal structure of the NF-κB proteins in complex with an inhibitor, blind docking was performed to identify the lowest energy binding site of QNZ. Two putative binding sites were identified on the NF-κB/IκBα complex. The putative binding sites were found to be located in close proximity to previously identified 'hot spots' on the NF-κB/IκBα complex (S. Bergqvest et al, 2008). A binding hypothesis of Pentamidine in the AdoMet-DC enzyme has provided possible reasons for the reduced potency of Pentamidine in comparison to the other inhibitors of AdoMet-DC (D. E. McCloskey et al, 2009). The binding hypothesis revealed that the ether oxygens and the benzamidine moiety are necessary for inhibitory interactions made with the AdoMet-DC enzyme.

Biological assays were used to confirm the *in vitro* activity of the hits identified through the virtual screening performed. The compounds confirmed with *in vitro* activity resulted from the LBVS screening performed for QNZ on the NCI database. Taking into consideration the $IC_{50}$ values of these active compounds, it can be concluded that the basic nitrogen on the quinazoline core is more important for increased activity than the aromatic ring on the same core. Furthermore, a functional group at the $R_1$ position is important for activity. The most active compound, NSC727152, was confirmed to act similarly to QNZ by inhibiting NF-κB. Consistent with the data reported in literature on the inhibition of NF-κB, an up-regulation of the GADD45α and γ mRNA expression and a down-regulation of IL-6 mRNA expression was observed after 6 hours of treatment with NSC727152 in DU-145 and PC-3 cancer cell lines. Furthermore, the reduced promotor activity of the IL-6 and the consensus sequence κB promotors, confirmed the inhibitory effect of NSC727152 on NF-κB in DU-145 cancer cells.

## 5.3 Future Work

### 5.3.1 Validation of the Binding Sites Identified for QNZ through the Blind Docking Calculations.

The validation of the predicted binding sites identified for QNZ on the NF-κB/IκBα complex (Chapter 2, section 2.3.2.1) would be worthwhile, considering that they have been identified to be in close proximity to previously described 'hot spots' (S. Bergqvest et al, 2008). There are several approaches that may be used to confirm the binding sites of QNZ, including MD simulations, which aim to accurately predict the dynamics and interactions between flexible ligands and proteins and are commonly performed on the top docked calculations retrieved from a SBVS (H. Alonso et al, 2006; E. R. Lindahl, 2006). MD simulations have become invaluable in refining docking calculations in addition to providing a time-dependant, dynamic understanding of the interactions between a protein and ligand in a solvent simulating the natural environment (H. Alonso et al, 2006). MD simulations would be beneficial to perform in order to confirm the putative binding sites identified for QNZ on the NF-κB/IκBα complex before embarking on obtaining a crystal structure. The i*n vitro* verification of the predicted binding sites may be achieved by performing site directed mutagenesis (S. Ekins et al, 2005). Site directed mutagenesis is commonly used to confirm the ligand binding sites on a protein surface, by mutating protein residues hypothesised to be important for ligand interaction (S. Ekins et al, 2005). The natural function and inhibition of the protein is then confirmed to determine that the protein had still maintained its correct fold and function despite the introduced mutations and to determine whether inhibition was achieved or not with respect to the mutated residues (A. Guidmond et al, 2002; N. Vaidehi et al, 2002). Correct protein folding and site directed mutagenesis may also be explored by utilising MD simulations (S. Ekins et al, 2005; A. Alonso et al, 2006). The residues predicted to interact with QNZ at both binding site 1 and 2, may be mutated to confirm their importance in forming the binding interactions with QNZ.

### 5.3.2 Binding Hypothesis of Pentamidine in the AdoMet-DC Crystal Structure

The aim of generating the binding hypothesis of Pentamidine in the human AdoMet-DC crystal structure was to provide insightful information on the binding interactions and pose of Pentamidine (Chapter 3, section 3.6.1). From the binding hypothesis it was evident that not all of the functional moieties of Pentamidine formed binding interactions with the AdoMet-DC enzyme (Figure 1).

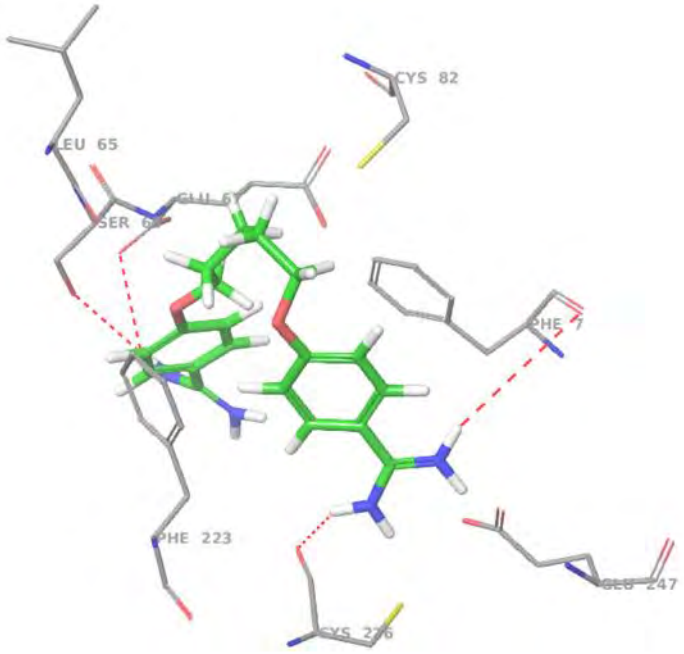



**Figure 1: The predicted binding interactions of Pentamidine in the human AdoMet-DC crystal structure (3H0W.pdb) (S. Bale et al, 2009). The buried amidine groups of Pentamidine form hydrogen bonds with the main chain carbonyl group of the Phe7 residue and the side chain of the Cys226 residue (hydrogen bonds shown in red). The solvent exposed amidine group of Pentamidine forms hydrogen bonds with the Ser66 and Glu67 residue sidechains.**

The hydrogen bond formation between Pentamidine and the Cys226 residue of AdoMet-DC is not observed with the other inhibitors of AdoMet-DC and therefore it was hypothesised that this amidine group may not be necessary for activity (S. Bale et al, 2009). In addition, the amidine group of the solvent exposed benzamidine moiety of Pentamidine is predicted to not form binding interactions with AdoMet-DC. Therefore it may be hypothesised that this moiety may also not be necessary for activity. Several series of analogues may be designed based on this hypothesis by omitting the non-interacting solvent exposed amidine group of

Pentamidine and/or the amidine that interacts with the Cys226 residue of AdoMet-DC enzyme. Replacing these moieties with other functional groups to form increased interactions with the AdoMet-DC enzyme may also be explored. Furthermore, analogues with a fewer number of rotating bonds compared to Pentamidine may be designed to decrease the conformational entropy hypothesised to contribute to the reduced potency of Pentamidine, compared to other inhibitors of AdoMet-DC. This may allow for the analogues to adopt the favourable syn conformation in the active site of the AdoMet-DC enzyme, which is responsible for increased activity compared to the less favourable *anti* conformation hypothesised to be adopted by Pentamidine (D. E. McCloskey et al, 2009). Furthermore, analogues may be designed to include a suitably positioned positively charged group to interact with the Phe7 residue to form the stabilising π-cation interaction. Literature suggests that compounds without a positively charged group fail to inhibit the AdoMet-DC enzyme (S. Bale et al, 2009; D. E. McCloskey et al, 2009).

### 5.3.3. Biological Assays to Determine the Inhibitory Effects of NSC727152 on NF-κB

The inhibitory activity of NSC727152 towards the transcription factor NF-κB may be further explored to determine the full potential of NSC727152 as an inhibitor of NF-κB. In response to the inhibition of NF-κB, the upregulation of the GADD45 proteins has been previously shown to interact with the upstream kinase, MEKK4, which activates both p38 and JNK, leading to apoptosis or inhibition of cell growth (L.F. Zerbini et al, 2004; M. Takekawa et al, 1998; M. Gupta et al, 2006). Taking into consideration that other iron chelators such as DFO have also been identified to induce the apoptosis of cancer cells via activation of the MAPK pathway, it would be worthwhile to determine whether NSC727152 induces the apoptosis of cancer cells (Y. Yu et al, 2011).

The GADD45 proteins play a central role in the cell cycle regulation, where their up-regulation has been shown to arrest cancer cells in the $G_2$/M phase of the cell cycle (R. E. Tamura et al, 2012; M. Vairapandi et al, 2002). Tachpyridine, the closely related analogue of NSC727152 is also reported to arrest colorectal and ovarian cancer cells in the $G_2$/M phase of the cell cycle. Taking into consideration that NSC727152 induced an up-regulation of the

GADD45 proteins, it would be interesting to determine the effect of NSC727152 on the cell cycle regulation.

Copper and zinc chelators have been identified to inhibit key mediators of cancer progression and migration such as the vascular endothelial growth factor, IL-8, IL-6 and NF-κB (Y. Yu et al, 2006). IL-6 has been identified to play a pivotal role in the progression of prostate cancer, contributing to its metastasis and angiogenesis (L. F. Zerbini et al, 2004; D. P. Nguyen et al, 2014) Taking into consideration that NSC727152 was shown to inhibit NF-κB in this study resulting in the reduced expression of IL-6 mRNA and its promoter activity, further analysis could include determining the effects of NSC727152 on the migration and invasion of cancer cells.

Finally, the direct inhibitory effects of NSC727152 on NF-κB determined in this study by the luciferase assays may be further substantiated by performing the chromatin immunoprecipitation assay or the EMSA gel shift assay. This would substantiate the data that NSC727152 binds directly to and inhibits NF-κB in prostate cancer cell lines.

## 5.4 References


1. S. Subramaniam, M. Mehrotra and D. Gupta, Virtual High Throughput Screening (vHTS) – A Perspective, Bioinformation, Vol. No. 3, Issue No. 1, Pp.14-17, 2008.
2. S. Bergqvist, G. Ghosh and E. A. Komives, The IκBα/NF-κB Complex has two hot spots, one at either end of the interface, Protein Science, Vol. No. 17, Pp. 2051-2058, 2008.
3. A. P. McGrath, K. M. Hilmer, C. A. Collyer, E. M. Shepard, B. O. Elmore, D. E. Brown, D. M. Dooley and J. M. Guss, Structure and Inhibition of Human Diamine Oxidase, Biochemistry, Vol. No. 48, Issue No. 41, Pp. 9810-9822, 2009.
4. M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, H. Takahashi, T. Fukazawa and H. Hayashi, Discovery of Quinazolines as a Novel Structural Class of Potent Inhibitors of NF-κB Activation, Bioorganic & Medicinal Chemistry, Vol. No. 11, Pp. 383-391, 2003.
5. M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, H. Takahashi and H. Hayashi, A Novel Structural Class of Potent Inhibitors of NF-κB Activation: Structure–Activity Relationships and Biological Effects of 6-Aminoquinazoline Derivatives, Vol. No. 11, Pp. 3869-3878, 2003.
6. S. Bale, W. Brooks, J. W. Hanes, A. M. Mahesan, W. C. Guida and S. E. Ealick, Role of the Sulfonium Centre in Determining the Ligand Specificity of Human S-AdenisylMethionine Decarboxylase, Biochemistry, Vol. No. 48, Issue No. 27, Pp. 6423-6430, 2009.
7. D. E. McCloskey, S. Bale, J. A. Secrist III, A. Tiwari, T. H. Moss, III, J. Valiyaveettil, W. H. Brooks, W. C. Guida, A. E. Pegg and S. E. Ealick, New Insights into the Design of Inhibitors of Human S-Adenosylmethionine Decarboxylase: Studies of Adenice $C^8$ Substitution in Structural Analogues of S-Adenosylmethionine, Journal of Medicinal Chemistry, Vol. No. 52, Pp. 1388-1407, 2009.
8. S. Yang, T. Wenzler, P. N. Miller, H. Wu, D. W. Boykin, R. Brun and M. Z. Wang, Pharmacokinetic Comparison to Determine the Mechanisms Underlying the Differential Efficacies of Cationic Diamidines Against First-and Second-Stage Human African Trypanosomiasis, Antimicrobial Agents and Chemotherapy, Vol. No. 58, Issue No. 7, Pp. 4064-4074, 2014.

9. L. F. Zerbini, Y. Wang, A. Czibere, R. G. Correa, J. Cho, K. Ijiri, W. Wei, M. Joseph, X. Gu, F. Grall, M. B. Goldring, J. Zhou and T. A. Libermann, NF-κB-mediated repression of growth arrest- and DNA-damage-inducible proteins 45α and γ is essential for cancer cell survival, PNAS, Vol. No. 101, Issue No. 37, Pp. 13618-13623, 2004.
10. L. F. Zerbini, Y. Wang, J. Cho and T. A. Libermann, Constitutive Activation of Nuclear Factor κB p50/p65 and Fra-1 and JunD Is Essential for Deregulated Interleukin 6 Expression in Prostate Cancer, American Association for Cancer Research, Vol. No. 63, Pp. 2206-2215, 2003.
11. H. Alonso, A. A. Bliznyuk and J. E. Gready, Combining Docking and Molecular Dynamic Simulations in Drug Design, Medicinal Research Reviews, Vol. No. 26, Issue No. 5, Pp. 531-568, 2006.
12. E. R. Lindahl, Molecular Dynamics Simulations, Molecular Modelling of Proteins: Methods in Molecular Biology, Vol. No. 443, Pp. 3-23, 2008.
13. S. Ekins, J. Mestres and B. Testa, *In Silico* Pharmacology for Drug Discovery: Applications to targets and Beyond, British Journal of Pharmacology, Vol. No. 152, Issue No. 1, Pp. 21-37, 2007.
14. A. Guimond, T. Sulea, A. Pen, P. Ear, M. D. O'Connor-McCourt, Site-Directed Mutagenesis of the Type II TGF-β Receptor Indicates a Ligand-Binding Site Disstinct form that of the Type II Activin Receptor, FEBS Letters, Vol. No. 515, Pp. 13-19, 2002.
15. N. Vaidehi, W. B. Floriano, R. Trabanino, S. E. Hall, P. Freddolino, E. J. Choi, G. Zamanakos and W. A. Goddard III, Prediction of Structure and Function of G Protein-Coupled Receptors, PNAS, Vol. No. 99, Issue No. 20, Pp. 12622-12627, 2002.
16. M. Takekawa and H. Saito, A family of stress-inducible Gadd45-like proteins mediate activation of the stress-responsive Mtk1/Mekk4 MAPKKK, Cell, Vol. 95, Pg. No. 521-530, 1998.
17. M. Gupta, S. K. Gupta, B. Hoffman and D. A. Liebermann, Gadd45α and Gadd45β protect hematopoietic cells from UV-induced apoptosis via distinct signalling pathways, including p38 activation and JNK inhibition, The Journal of Biological Chemistry, Vol. 281, Issue No. 26, Pg. No. 17552-17558, 2006.
18. Y. Yu and D. R. Richardson, Cellular iron depletion stimulates the JNK and p38 MAPK signalling transduction pathways, dissociation of ASK1-thioredoxin, and

activation of ASK1, Journal of Biological Chemistry, Vol. No. 286, Issue No. 17, Pp. 15413-15427, 2011.

19. R. E. Tamura, J. F. de Vasconcellos, D. Sarker, T. A. Libermann, P. B. Fisher and L. F. Zerbini, GADD45 proteins: central players in tumorigenesis, Current Molecular Medicine, Vol. No. 12, Issue No. 5, Pg. No. 634-651, 2012.
20. M. Vairapandi, A. G. Balliet, B. Hoffman and D. A. Liebermann, GADD45b and GADD45g are cdc2/cyclinB1 kinase inhibitors with a role in S and G2/M cell cycle checkpoints induced by genotoxic stress, Journal of Cellular Physiology, Vol. No. 192, Issue No. 3, Pp. 327-338, 2002.
21. Y. Yu, J. Wong, D. B. Lovejoy, D. S. Kalinowski and D. R. Richardson, Chelators at the Cancer Coalface: Desferrioxamine to Triapine and Beyond, American Association for Cancer Research, Vol. No. 2, Issue No. 23, Pp. 6876-6883, 2006.
22. D. P. Nguyen, J. Li and A. K. Tewari, Inflammation and Prostate Cancer: the role of Interleukin 6 (IL-6), BJU International, Vol. No. 113, Pp. 986-992, 2014.

# Chapter 6: Appendices

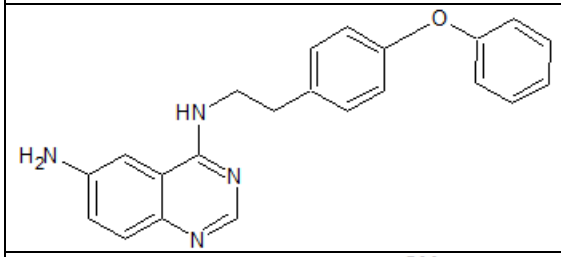
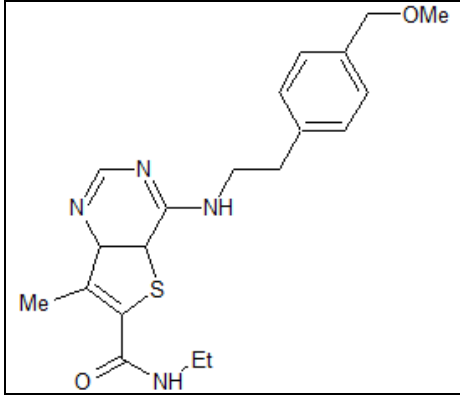
## 6.1 Appendix 1: Compounds identified from ZINC database

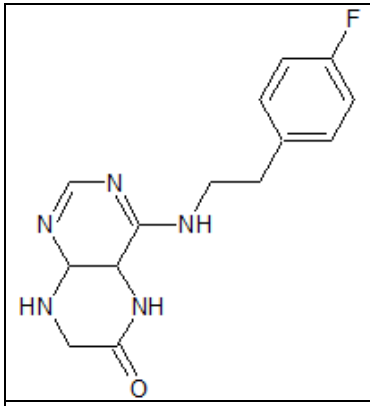
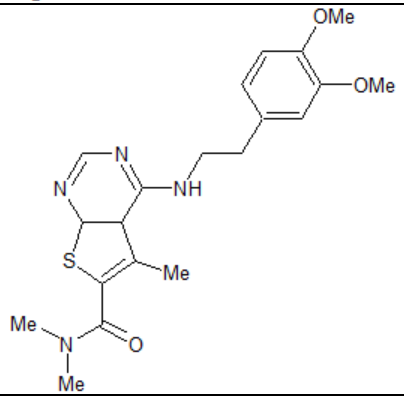
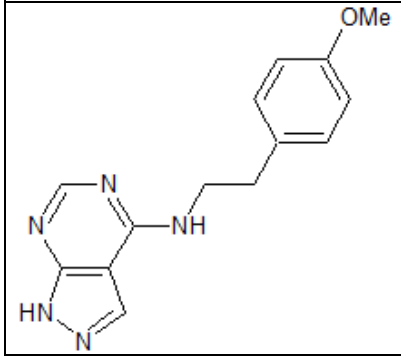
**Table 1: The complete Open-Eye drug-like filter (Filter, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2010).**

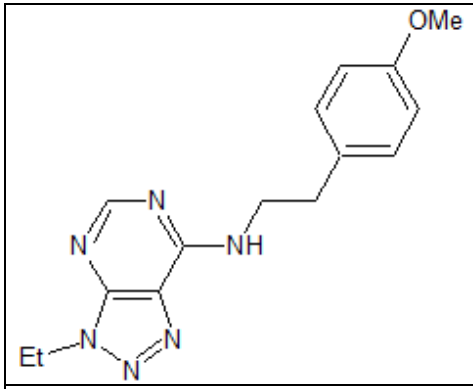
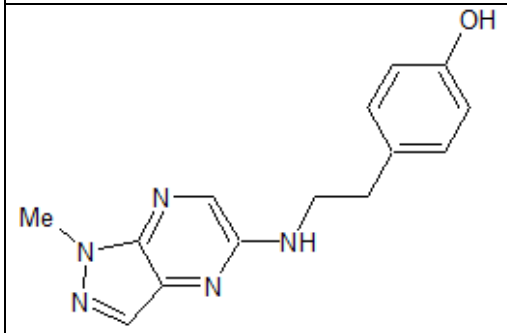
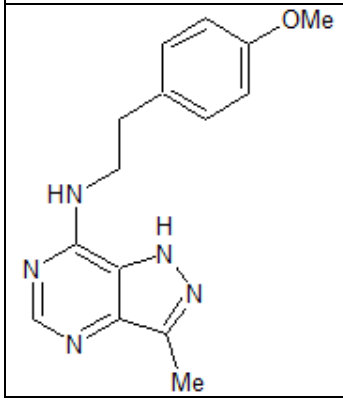
| Parameter | Value |
|---|---|
| Min. molecular weight | 100.00 |
| Max. molecular weight | 500.00 |
| Min. ring systems | 0 |
| Max. ring systems | 6 |
| Min. ring size | 0 |
| Max. ring size | 20 |
| Min. carbons | 7 |
| Max. carbons | 35 |
| Min. heteroatoms | 2 |
| Max. Heteroatoms | 20 |
| Min. het/C ratio | 0.100 |
| Max. Het/C ratio | 1.00 |
| Min. rotatable bonds | 0 |
| Max. rotatable bonds | 10 |
| Min. rigid bonds | 0 |
| Max. rigid bonds | 35 |
| Min. Con. Non_Rings | 0 |
| Max. Con. Non_Rings | 15 |
| Max. count formal charge | 0 |
| Min. count formal charge | 3 |
| Min. sum formal charge | -2 |
| Max. sum formal charge | 2 |
| Min. Hbond donors | 0 |
| Max. Hbond donors | 6 |

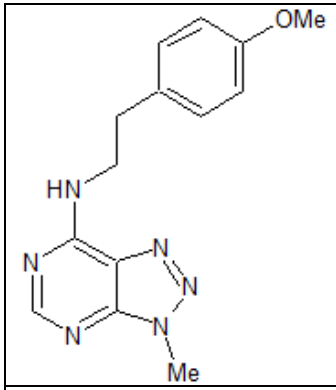
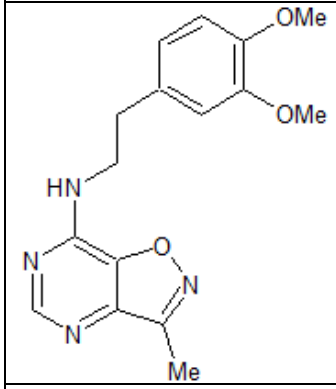
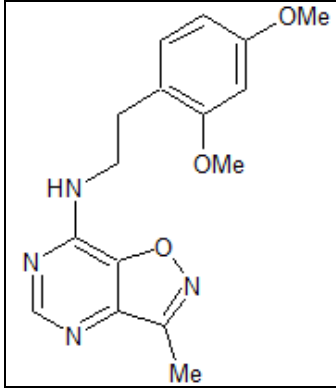
| | |
|---|---|
| Min. Hbond acceptors | 0 |
| Max. Hbond acceptors | 5 |
| Min. Lipinski donors | 0 |
| Max. Lipinski donors | 12 |
| Min. Lipinski acceptors | 0 |
| Max. Lipinski acceptors | 22 |
| Min. chiral centers | 0 |
| Max. chiral centers | 4 |
| Min. XLogP | -5.00 |
| Max. XLogP | 5.00 |
| Min. solubility | Moderately |
| Min. 2D PSA | 075.00 |
| Max. 2D PSA | 140.00 |
| Adjust rot for rings | True |
| Lipinski | Flase |
| Max. Lipinski | 3 |
| Pharmacopia | True |
| Min. Abs | 0.500 |
| Aggregators | True |
| Pred. Agg | True |

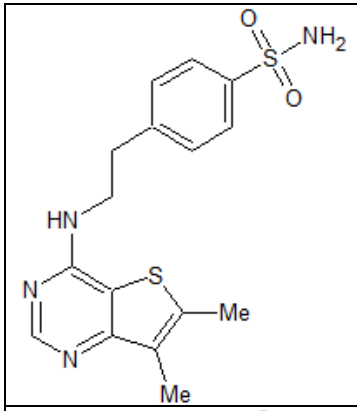
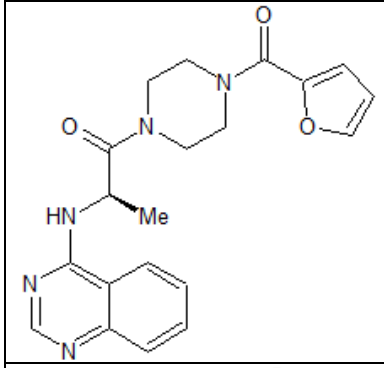
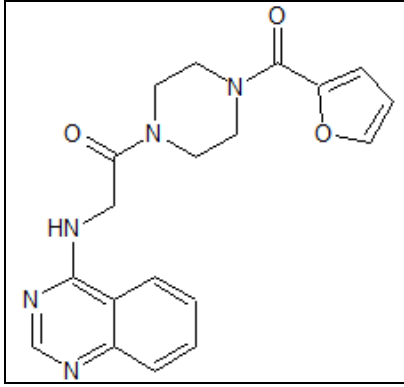
**Table 2: The top 20 hits from the 3D similarity search performed on vROCS for QNZ. The TanimotoCombo coefficient was considered to measure similarity to QNZ.**

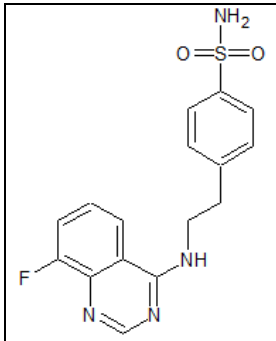
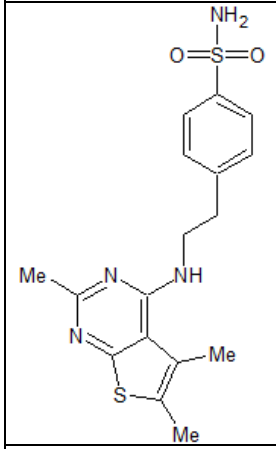
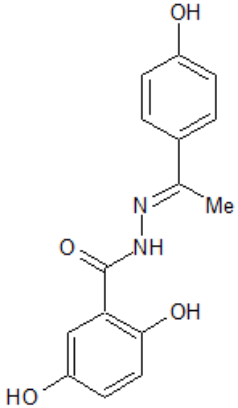
| Molecular structure | ZINC ID | Tanimoto Combo | Tanimoto shape | Tanimoto colour |
|---|---|---|---|---|
|  | QNZ | 2 | 1 | 1 |
|  | ZINC12191589 | 1.386 | 0.734 | 0.652 |

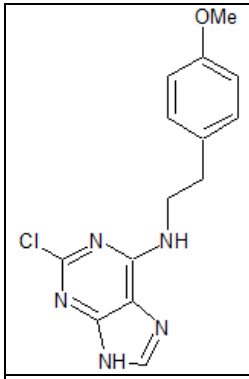
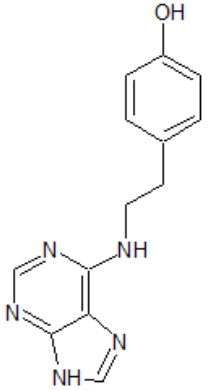
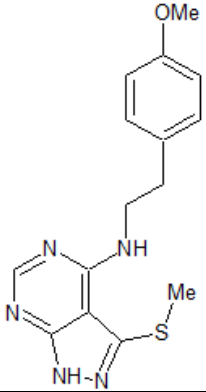
| | | | | |
|---|---|---|---|---|
|  | ZINC41121377 | 1.363 | 0.79 | 0.573 |
|  | ZINC11881002 | 1.362 | 0.711 | 0.651 |
|  | ZINC11856461 | 1.35 | 0.778 | 0.572 |

| | | | | |
|---|---|---|---|---|
|  | ZINC05250513 | 1.33 | 0.759 | 0.571 |
|  | ZINC40445373 | 1.325 | 0.753 | 0.572 |
|  | ZINC41528031 | 1.324 | 0.751 | 0.573 |

| | | | | |
|---|---|---|---|---|
|  | ZINC04941893 | 1.322 | 0.752 | 0.571 |
|  | ZINC20575137 | 1.32 | 0.749 | 0.571 |
|  | ZINC08596006 | 1.316 | 0.743 | 0.573 |

| | | | | |
|---|---|---|---|---|
|  | ZINC03428816 | 1.309 | 0.824 | 0.485 |
|  | ZINC41502800 | 1.308 | 0.814 | 0.494 |
|  | ZINC41502206 | 1.301 | 0.799 | 0.502 |

| | | | | |
|---|---|---|---|---|
|  | ZINC41471180 | 1.297 | 0.821 | 0.477 |
|  | ZINC03307744 | 1.297 | 0.812 | 0.485 |
|  | ZINC17009566 | 1.292 | 0.687 | 0.605 |

| | | | | |
|---|---|---|---|---|
|  | ZINC09419282 | 1.291 | 0.766 | 0.526 |
|  | ZINC11856480 | 1.281 | 0.752 | 0.529 |
|  | ZINC41528614 | 1.277 | 0.748 | 0.529 |

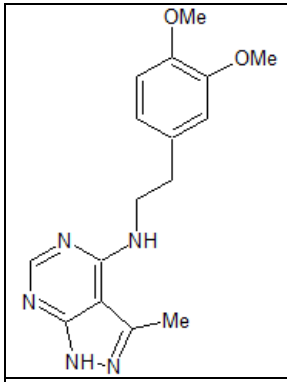
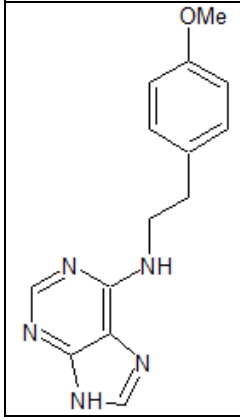
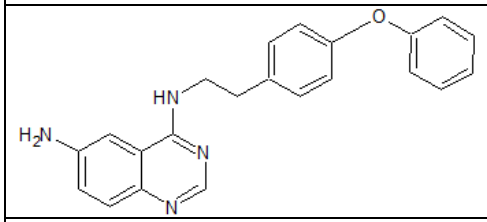
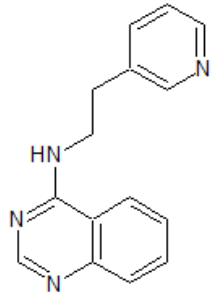
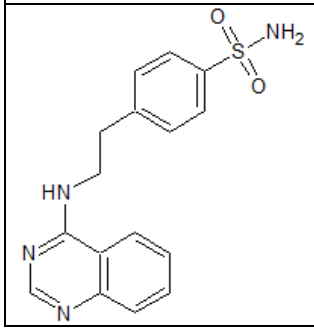
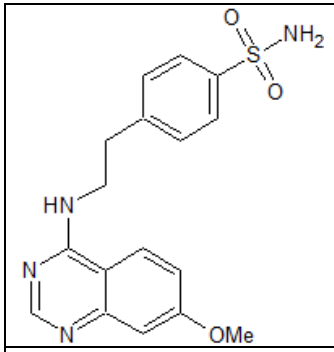
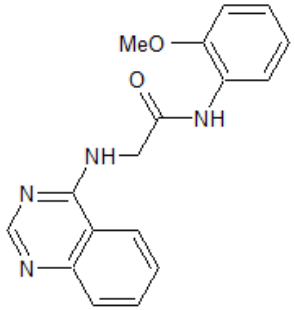
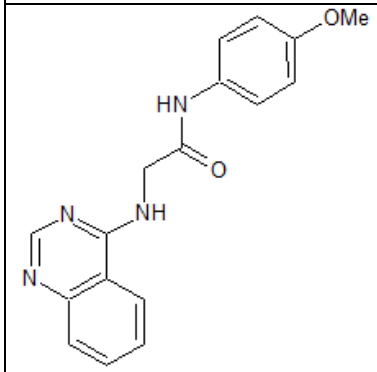
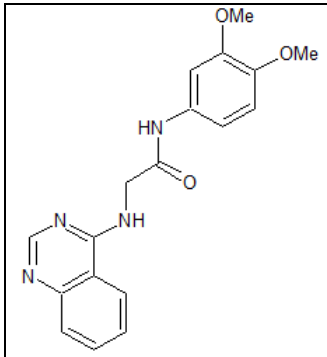
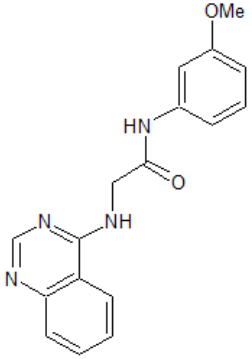
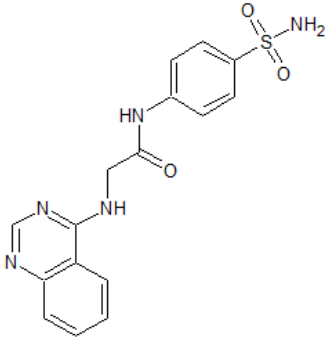
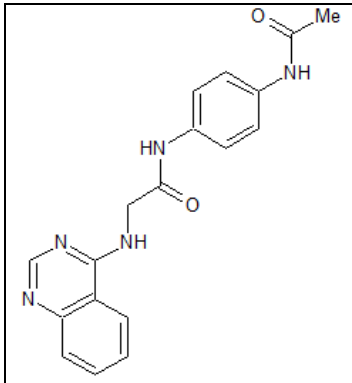
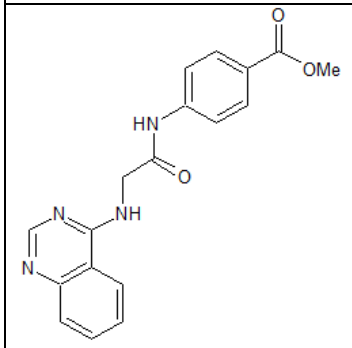
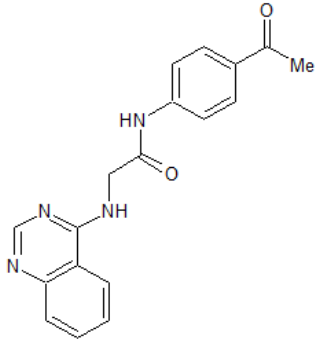
| | | | | |
|---|---|---|---|---|
|  | ZINC41528034 | 1.276 | 0.747 | 0.529 |
|  | ZINC08764584 | 1.276 | 0.75 | 0.526 |

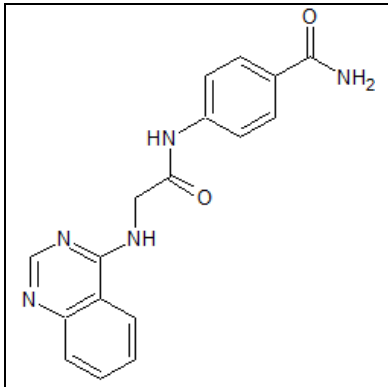
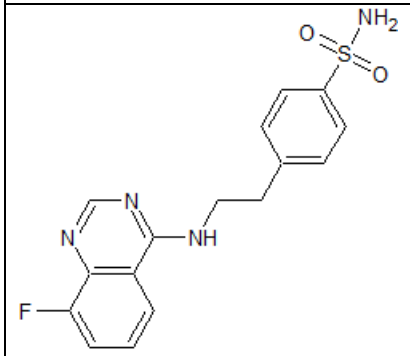
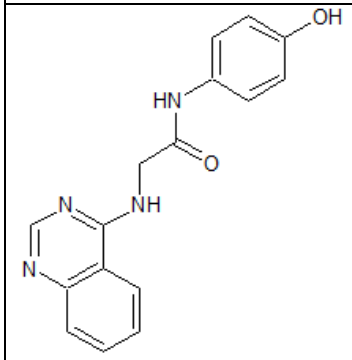
**Table 3: The top 20 compounds identified from the 2D similarity search performed On Pipeline Pilot for QNZ. Their Tanimoto scores were considered to measure similarity to QNZ.**

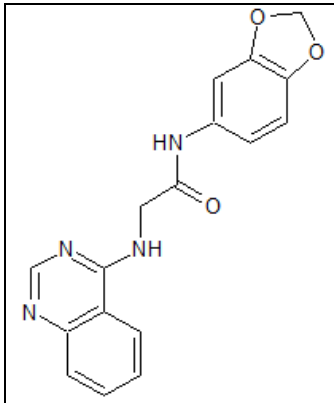
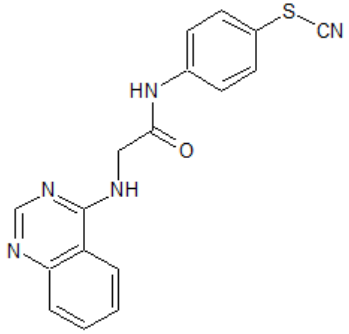
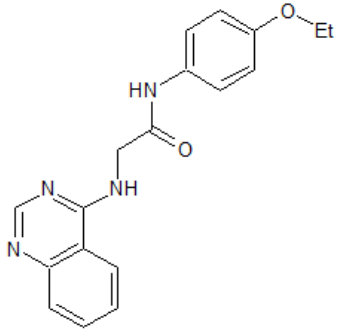
| Molecular structures | ZINC ID | Tanimoto Score |
|---|---|---|
|  | QNZ | 1 |
|  | ZINC32089602 | 0.681818182 |
|  | ZINC31948986 | 0.651162791 |

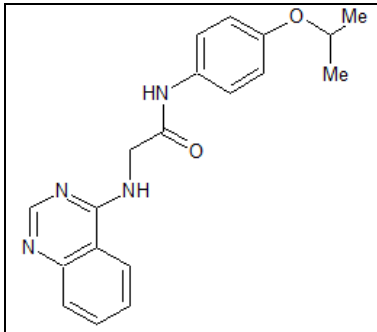
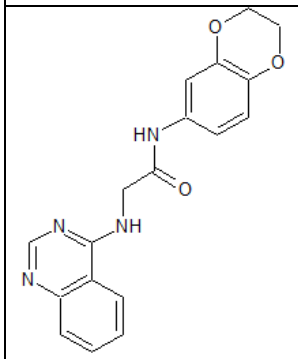
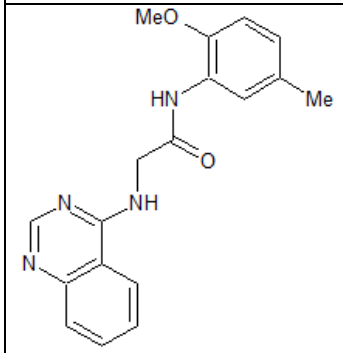
| | | |
|---|---|---|
|  | ZINC32544427 | 0.595744681 |
|  | ZINC41504168 | 0.574468085 |
|  | ZINC41504176 | 0.574468085 |

| | | |
|---|---|---|
|  | ZINC41504192 | 0.5625 |
|  | ZINC41504172 | 0.5625 |
|  | ZINC41504263 | 0.553191489 |

| | | |
|---|---|---|
|  | ZINC41504220 | 0.543478261 |
|  | ZINC41504229 | 0.541666667 |
|  | ZINC41504212 | 0.541666667 |

| | | |
|---|---|---|
|  | ZINC41504204 | 0.541666667 |
|  | ZINC41471180 | 0.541666667 |
|  | ZINC41504310 | 0.541666667 |

| | | |
|---|---|---|
|  | ZINC41504278 | 0.54 |
|  | ZINC41504342 | 0.54 |
|  | ZINC41504180 | 0.54 |

| | | |
|---|---|---|
|  | ZINC41504188 | 0.54 |
|  | ZINC41504282 | 0.54 |
|  | ZINC41504184 | 0.54 |

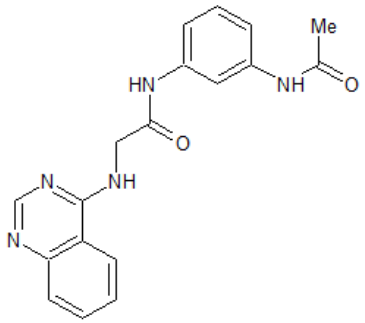
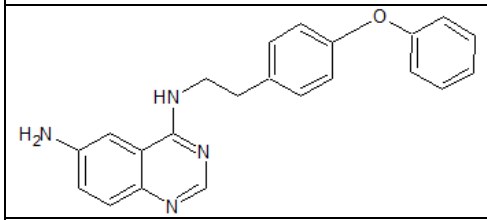
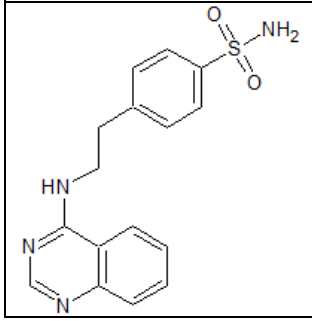
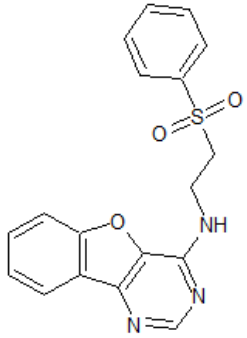
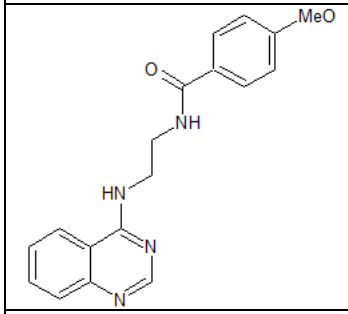
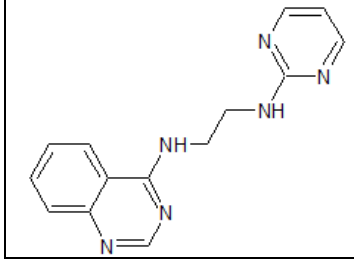
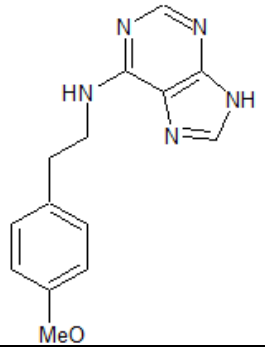
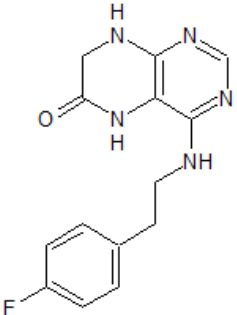
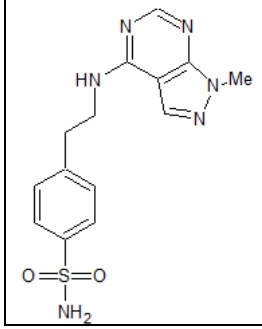
| | | |
|---|---|---|
|  | ZINC41504216 | 0.531914894 |

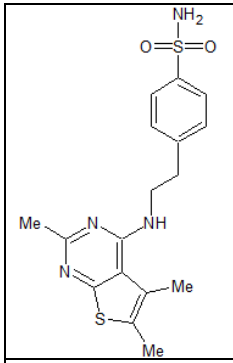
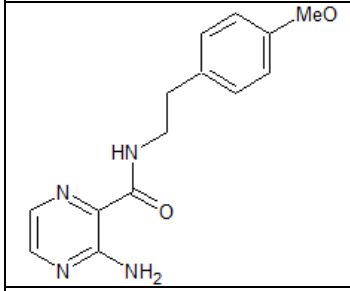
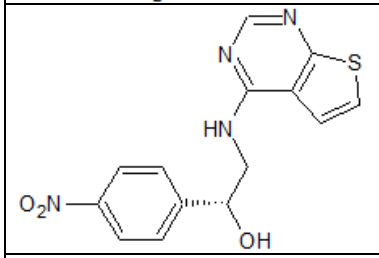
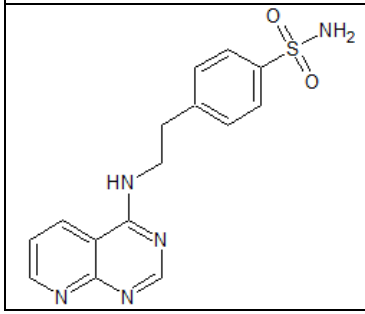
**Table 4: The compounds identified for QNZ from the docking calculations performed on Glide5.7 at the binding site 1 and 2. The compounds were ranked by good, intermediate and poor binding energies and the 2D or 3D similarity search method by which they were identified by is denoted. These compounds were procured for in vitro screening.**

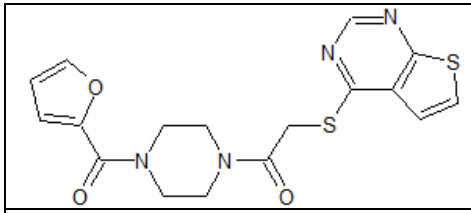
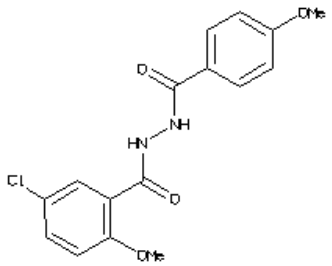
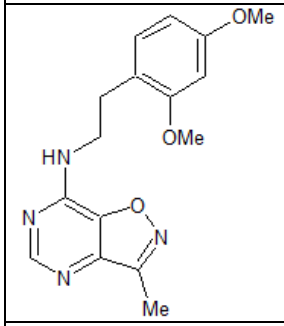
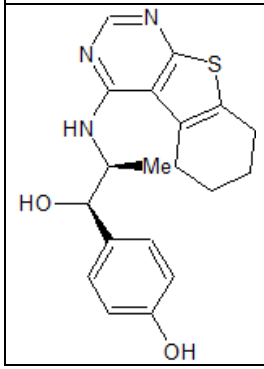
| Molecular structure | ZINC ID | Glide score | Tanimoto Score | Similarity rank | Classification |
|---|---|---|---|---|---|
|  | QNZ | -7.237354 | 1 | | |
|  | ZINC31948986 | -8.363115 | 0.651163 | 3 | good/2D/site1 |

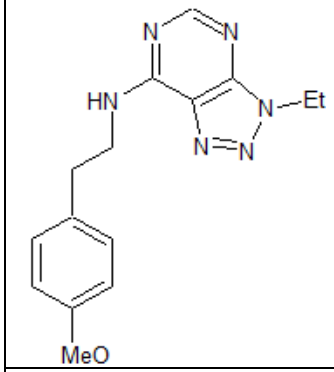
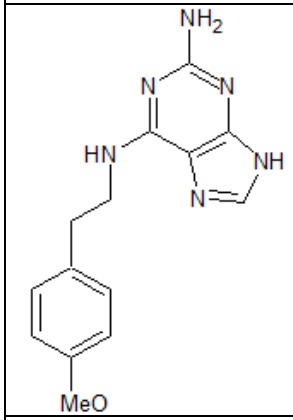
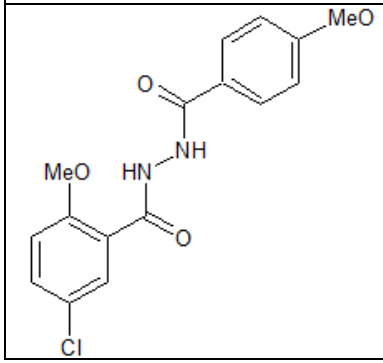
| | | | | | |
|---|---|---|---|---|---|
|  | ZINC43272460 | -7.23 | 0.431373 | 186 | intermediate/2D/site1 |
|  | ZINC52468283 | -7.156269 | 0.52 | 37 | intermediate/2D/site1 |
|  | ZINC49000199 | -7.136254 | 0.446809 | 148 | intermediate/2D/site1 |

| | | | | | | |
|---|---|---|---|---|---|---|
|  | ZINC08764584 | -6.687725 | 0.470588 | 107 | | poor/2D/site1 |

| Molecular Structure | ZINC ID | Glide score | Tanimoto Combo | Tanimoto shape | Tanimoto colour | Similarity ranking | Classification |
|---|---|---|---|---|---|---|---|
|  | ZINC41121377 | -7.60271 | 1.363 | 0.79 | 0.573 | 3 | good/3D/site2 |
|  | ZINC09058175 | -8.53897 | 1.207 | 0.772 | 0.436 | 136 | good/3D/site2 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
|  | ZINC03307744 | -8.39025 | 1.297 | 0.812 | 0.485 | 20 | good/3D/site2 |
|  | ZINC32860451 | -6.0877 | 1.232 | 0.776 | 0.456 | 73 | intermediate/3D/site2 |
|  | ZINC08147913 | -5.73493 | 1.228 | 0.765 | 0.463 | 80 | intermediate/3D/site2 |
|  | ZINC41472876 | -5.50188 | 1.274 | 0.831 | 0.443 | 32 | intermediate/3D/site2 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
|  | ZINC03417611 | -4.95 | 1.25 | 0.779 | 0.471 | 45 | poor/3D/site2 |
|  | ZINC06127758 | -4.43253 | 1.212 | 0.722 | 0.491 | 120 | poor/3D/site2 |
|  | ZINC08596006 | -6.00466 | 1.316 | 0.743 | 0.573 | 13 | poor/3D/site2 |
|  | ZINC07982118 | -7.12618 | 1.2 | 0.716 | 0.484 | 158 | good/3D/site1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
|  | ZINC05250513 | -6.98707 | 1.33 | 0.759 | 0.571 | 7 | intermediate/3D/site1 |
|  | ZINC32123960 | -7.13776 | 1.223 | 0.767 | 0.456 | 90 | intermediate/3D/site1 |
|  | ZINC23146805 | -6.95947 | 1.209 | 0.677 | 0.532 | 129 | intermediate/3D/site1 |

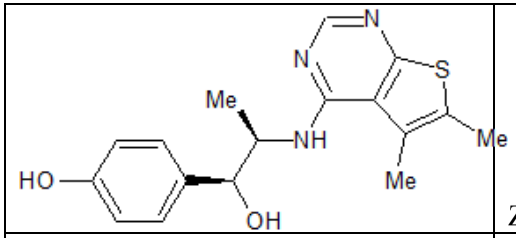
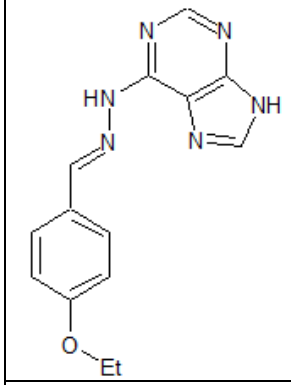
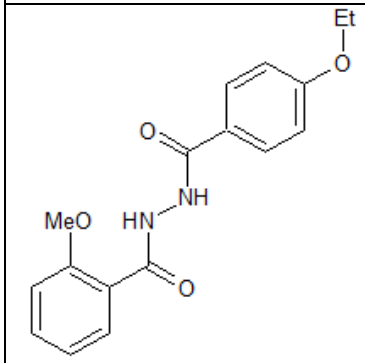
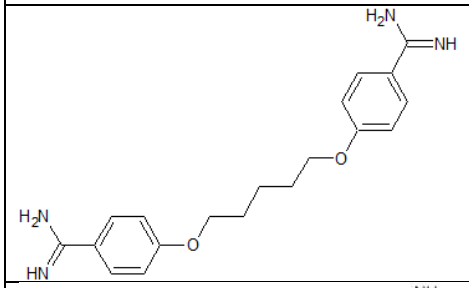
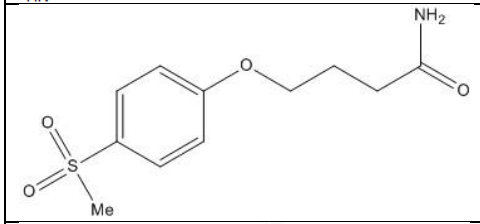
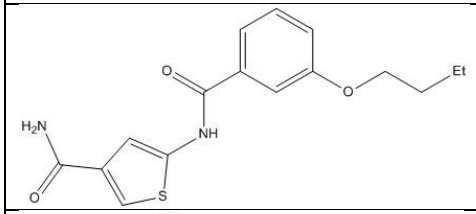
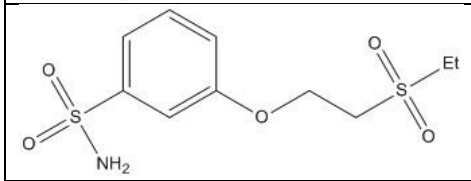
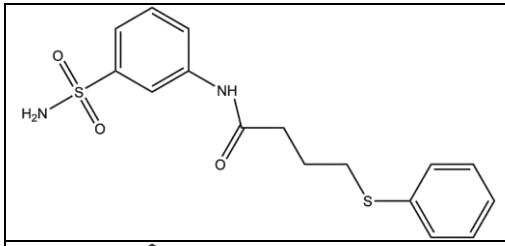
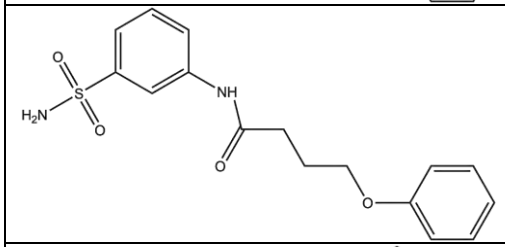
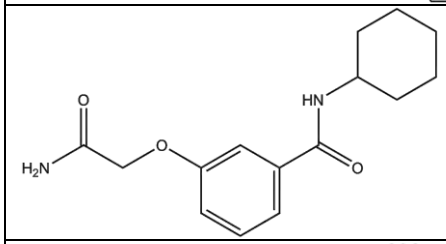
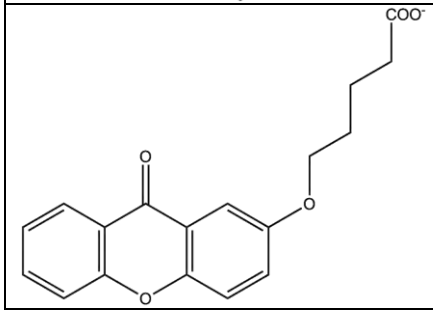
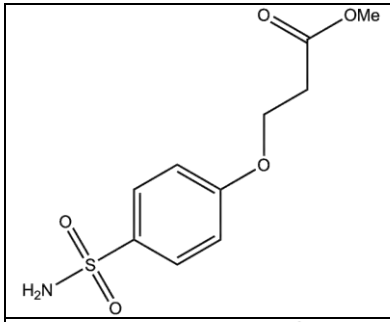
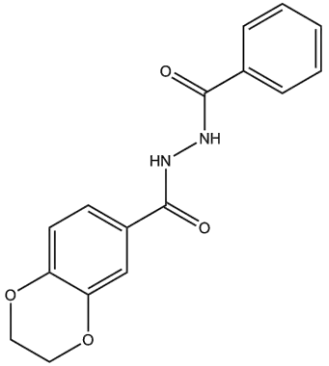
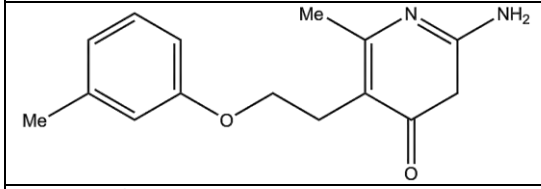
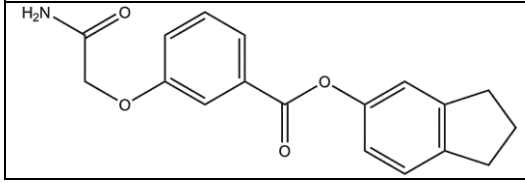
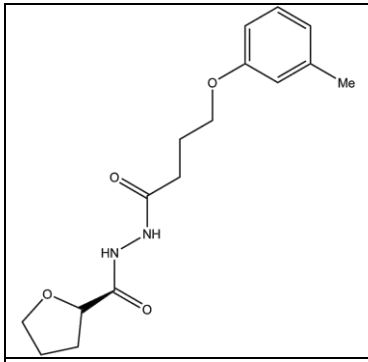
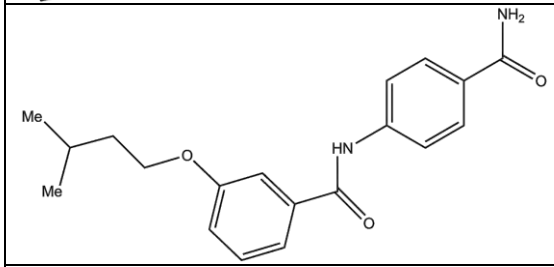
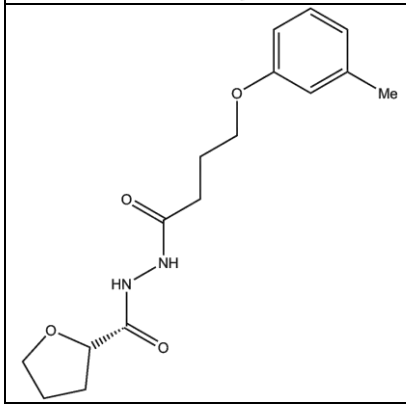
| | | | | | | | |
|---|---|---|---|---|---|---|---|
|  | ZINC07982128 | -6.7907 | 1.221 | 0.724 | 0.498 | 95 | poor/3D/site1 |
|  | ZINC04106433 | -6.64242 | 1.21 | 0.727 | 0.482 | 128 | poor/3D/site1 |
|  | ZINC05065919 | -6.36334 | 1.208 | 0.714 | 0.494 | 134 | poor/3D/site1 |

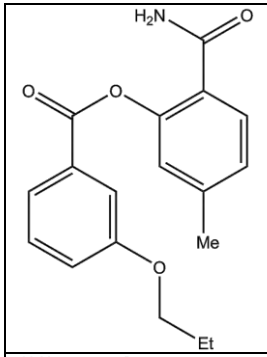
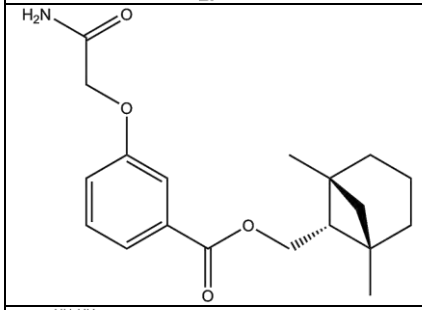
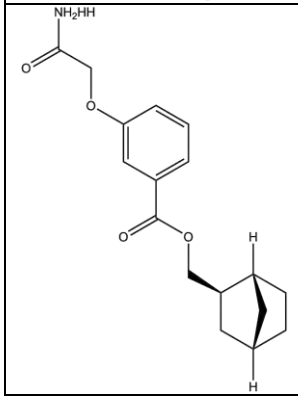
**Table 5: The top 20 hits identified from the 2D similarity search performed on Pipeline Pilot for Pentamidine. the Tanimoto score was considered to measure similarity to Pentamidine.**

| Molecular structure | ZINC ID | Tanimoto Score |
|---|---|---|
|  | Pentamidine | 1 |
|  | ZINC49375918 | 0.5625 |
|  | ZINC04281503 | 0.525 |
|  | ZINC38085888 | 0.516129 |

| | | |
|---|---|---|
|  | ZINC09473369 | 0.514286 |
|  | ZINC01748085 | 0.514286 |
|  | ZINC28264993 | 0.5 |
|  | ZINC04344168 | 0.5 |

| | | |
|---|---|---|
|  | ZINC48664715 | 0.5 |
|  | ZINC19603643 | 0.5 |
|  | ZINC04623153 | 0.5 |
|  | ZINC47382167 | 0.5 |

| | | |
|---|---|---|
|  | ZINC32262136 | 0.486486 |
|  | ZINC08567712 | 0.486486 |
|  | ZINC32262135 | 0.486486 |

| | | |
|---|---|---|
|  | ZINC49400487 | 0.486486 |
|  | ZINC50208118 | 0.486486 |
|  | ZINC50208117 | 0.486486 |

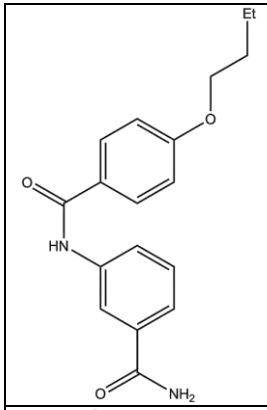
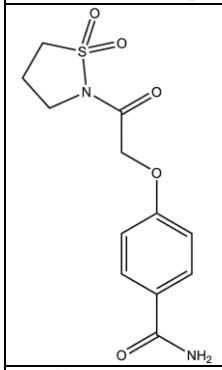
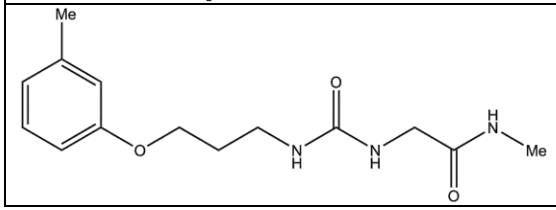
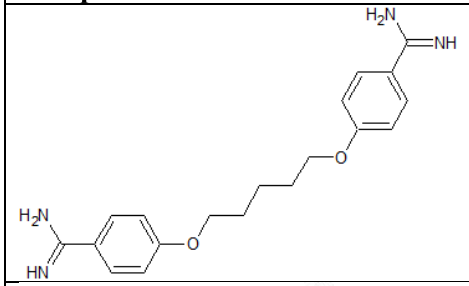
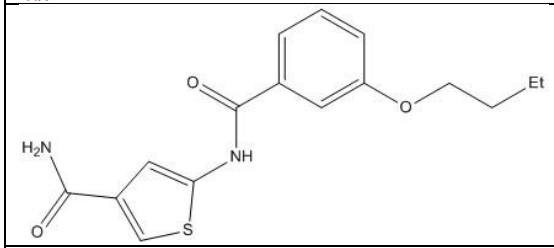
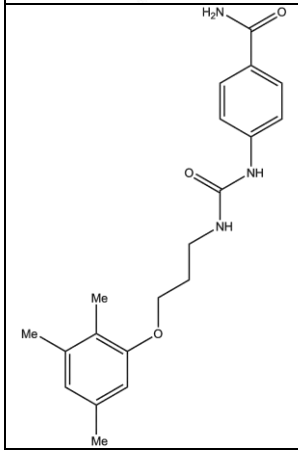
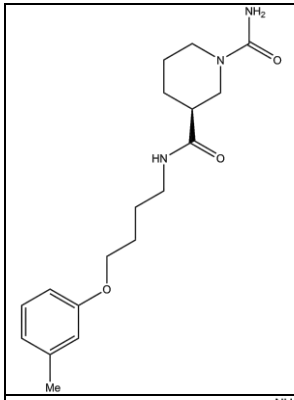
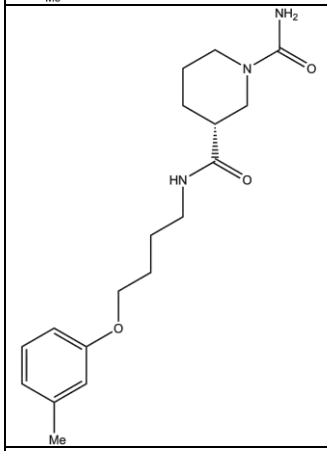
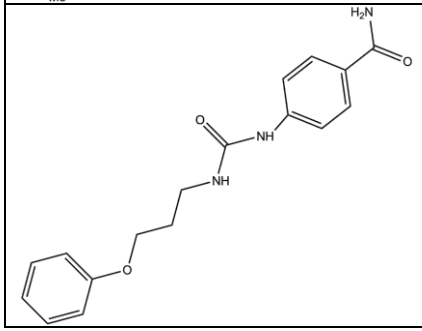
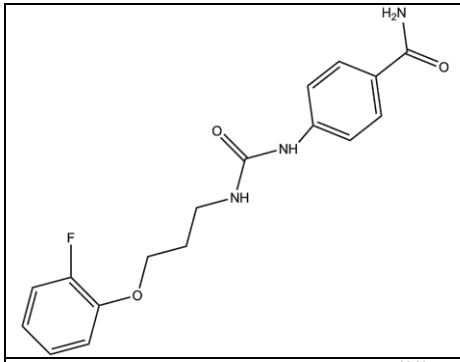
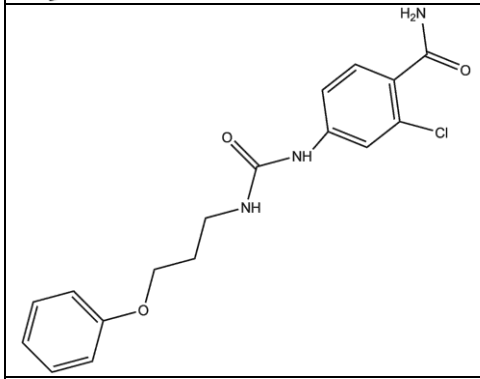
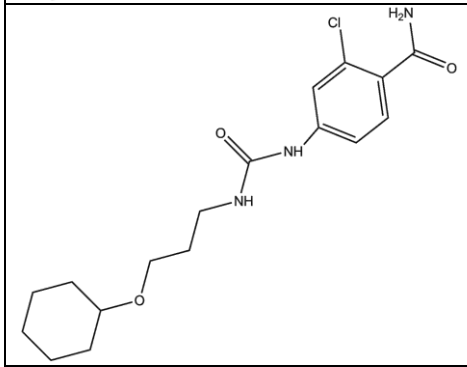
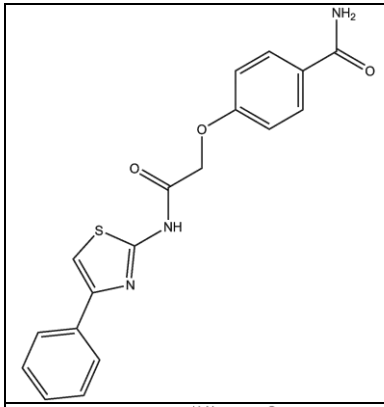
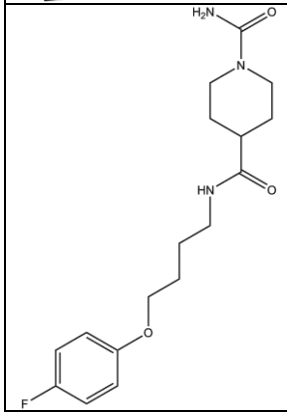
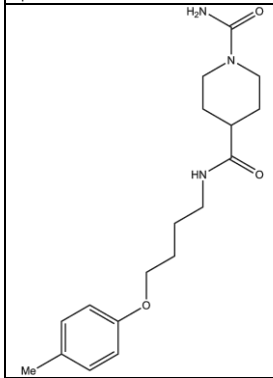
| | | |
|---|---|---|
|  | ZINC12672314 | 0.486486 |
|  | ZINC48762892 | 0.486486 |
|  | ZINC26278030 | 0.486486 |

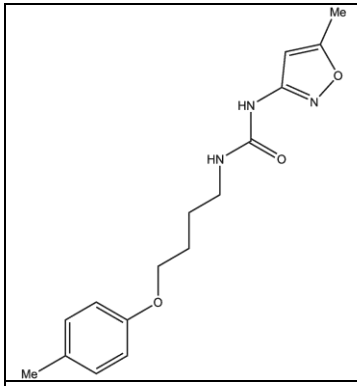
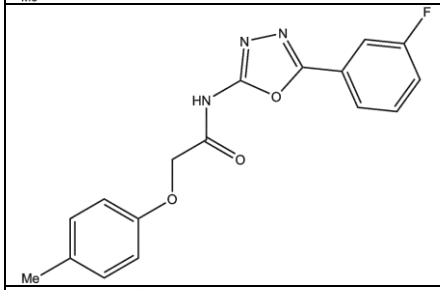
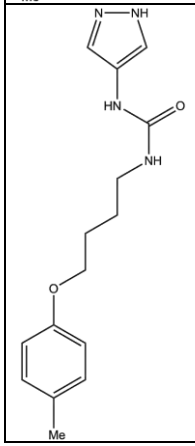
**Table 6: The top 20 hits identified from the 3D similarity search performed on vROCS for Pentamidine. The TanimotoCombo score was considered to measure similarity to Pentamidine.**

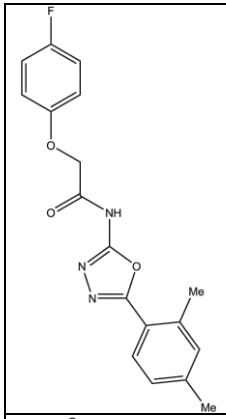
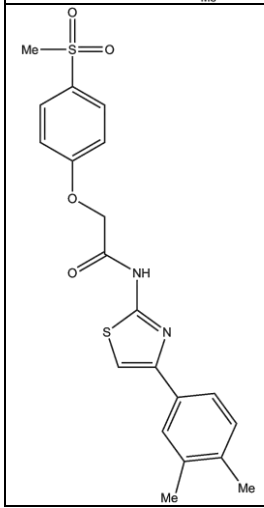
| Compound Molecular Structure | ZINC ID | Tanimoto Combo | Tanimoto shape | Tanimoto colour |
|---|---|---|---|---|
|  | Pentamidine | 2 | 1 | 1 |
|  | ZINC43711503 | 1.146 | 0.875 | 0.271 |
|  | ZINC43669978 | 1.137 | 0.867 | 0.271 |

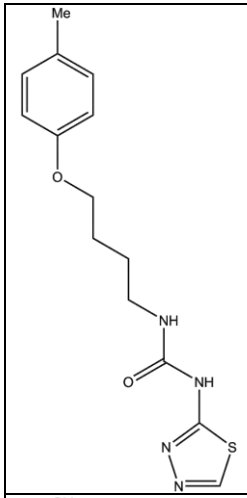
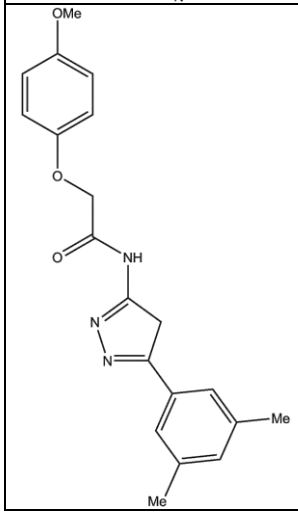
| | | | | |
|---|---|---|---|---|
|  | ZINC49324384 | 1.137 | 0.839 | 0.298 |
|  | ZINC49324382 | 1.125 | 0.819 | 0.306 |
|  | ZINC43711873 | 1.122 | 0.851 | 0.271 |

| | | | | |
|---|---|---|---|---|
|  | ZINC43713405 | 1.114 | 0.843 | 0.271 |
|  | ZINC45719659 | 1.093 | 0.828 | 0.265 |
|  | ZINC45719013 | 1.09 | 0.808 | 0.281 |

| | | | | |
|---|---|---|---|---|
|  | ZINC31673334 | 1.08 | 0.818 | 0.263 |
|  | ZINC42000823 | 1.079 | 0.779 | 0.3 |
|  | ZINC49324405 | 1.078 | 0.778 | 0.3 |

| | | | | |
|---|---|---|---|---|
|  | ZINC46785216 | 1.077 | 0.803 | 0.275 |
|  | ZINC38932342 | 1.076 | 0.858 | 0.218 |
|  | ZINC46784828 | 1.072 | 0.817 | 0.254 |

| | | | | |
|---|---|---|---|---|
|  | ZINC36682106 | 1.063 | 0.845 | 0.218 |
|  | ZINC45526358 | 1.062 | 0.87 | 0.192 |

| | | | | |
|---|---|---|---|---|
|  | ZINC44944926 | 1.058 | 0.812 | 0.246 |
|  | ZINC36682100 | 1.058 | 0.854 | 0.204 |

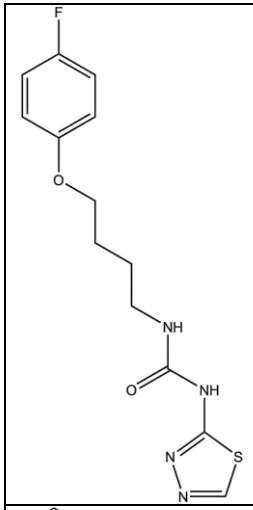
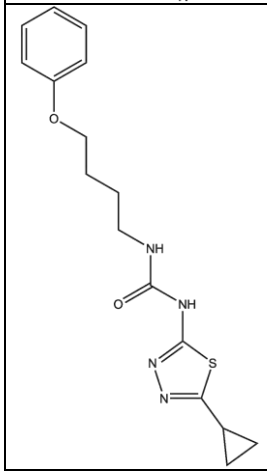
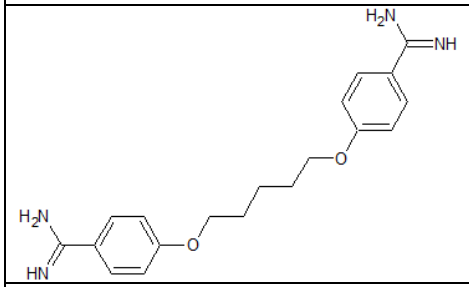
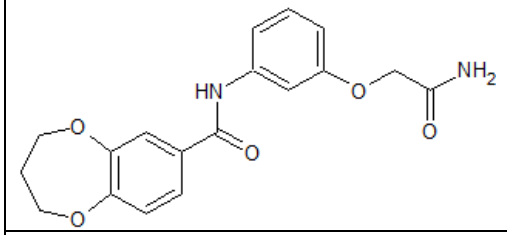
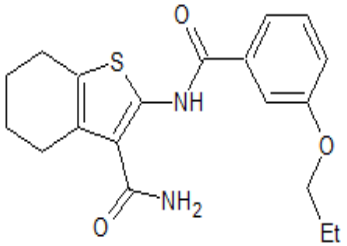
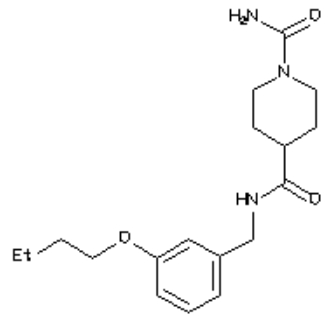
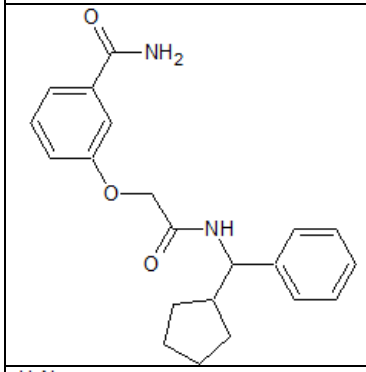
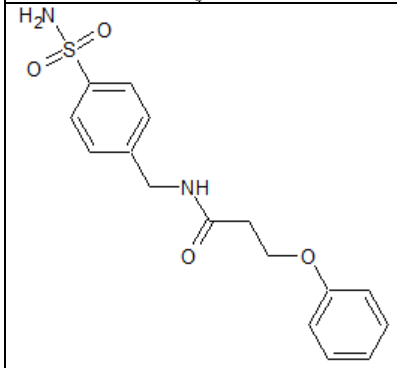
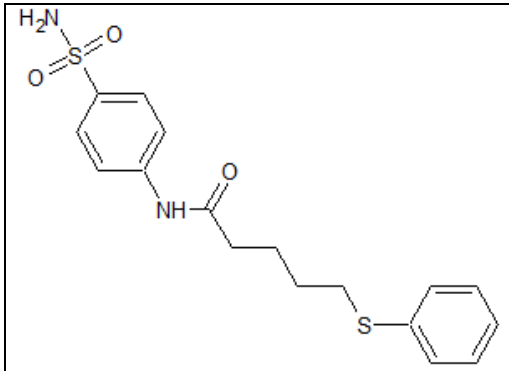
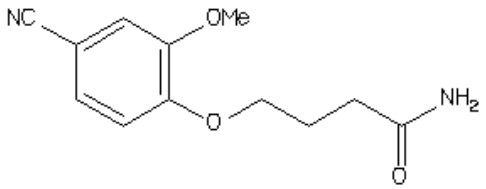
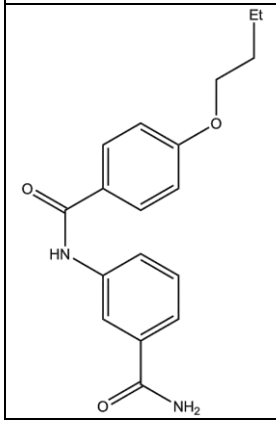
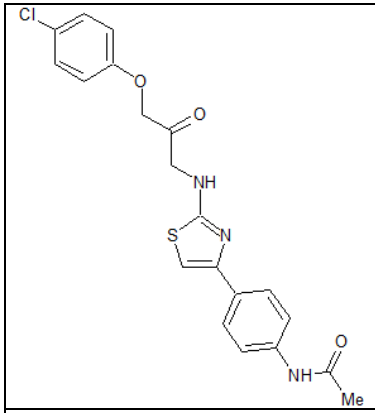
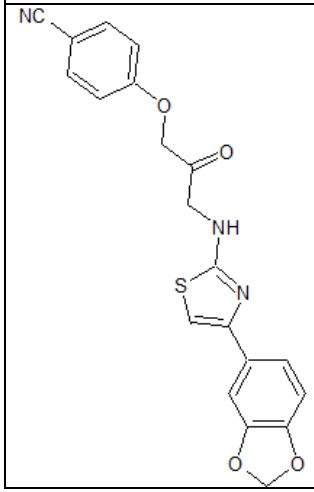
| | | | | |
|---|---|---|---|---|
|  | ZINC43128408 | 1.056 | 0.809 | 0.246 |
|  | ZINC43359379 | 1.056 | 0.866 | 0.19 |

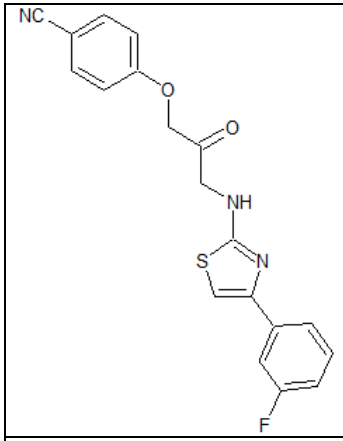
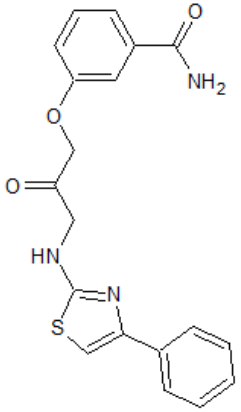
**Table 7: The compounds identified for Pentamidine from the docking calculations performed on Glide5.7. The compounds were ranked by good, intermediate and poor binding energies and the 2D or 3D similarity search method by which they were identified by is denoted. These compounds were procured for in vitro screening.**

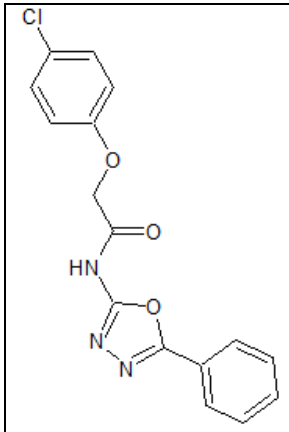
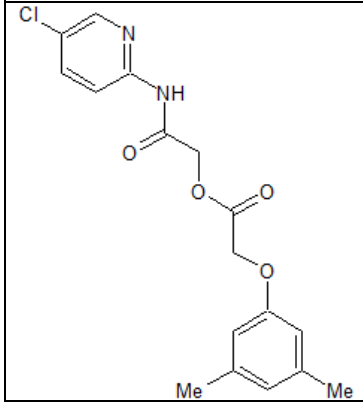
| Molecular structure | ZINC ID | Tanimoto Score | Glide score | Similarity rank | Classification |
|---|---|---|---|---|---|
|  | Pentamidine | 1 | -9.27484 | - | - |
|  | ZINC22883041 | 0.461538 | -8.65292 | 455 | Good/2D |
|  | ZINC02996814 | 0.465116 | -8.61273 | 525 | Good/2D |

| | | | | | |
|---|---|---|---|---|---|
|  | ZINC48997508 | 0.463415 | -8.50525 | 803 | Good/2D |
|  | ZINC47360766 | 0.461538 | -8.13933 | 3032 | Intermediate/2D |
|  | ZINC15613510 | 0.485714 | -8.21601 | 2492 | Intermediate/2D |

| | | | | | |
|---|---|---|---|---|---|
|  | ZINC09473369 | 0.514286 | -8.35412 | 1605 | Intermediate/2D |
|  | ZINC44920378 | 0.472222 | -7.71773 | 9959 | Poor/2D |
|  | ZINC12672314 | 0.486486 | -7.93448 | 5825 | Poor/2D |

| | | | | | |
|---|---|---|---|---|---|
|  | ZINC15991762 | | -9.55855 | 81 | Good/3D |
|  | ZINC03355852 | | -8.66227 | 1480 | Intermediate/3D |

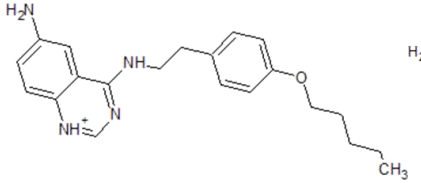
| | | | | | |
|---|---|---|---|---|---|
|  | ZINC25322020 | | -8.343 | 2212 | Intermediate/3D |
|  | ZINC31673334 | | -8.82365 | 1023 | Intermediate/3D |

| | | | | | |
|---|---|---|---|---|---|
|  | ZINC04276177 | | -7.52154 | 6789 | Poor/3D |
|  | ZINC02622770 | | -7.45961 | 7102 | Poor/3D |

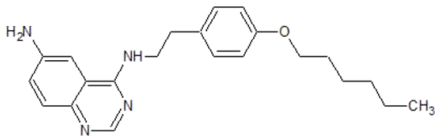
## 6.2 Appendix 2

### 6.2.1 Appendix 2.1: ROC Curve Validation

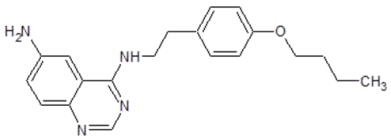
ECFP_4 Cluster: Active compounds (1-1000nM)



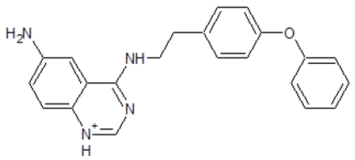
#1
$IC_{50}$ = 2nM



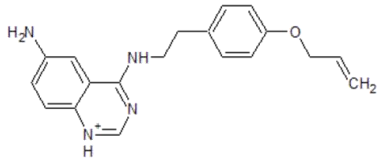
#2
$IC_{50}$ = 4nM



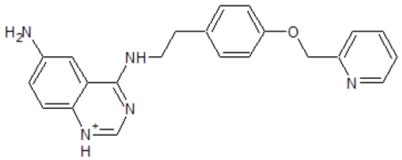
#36
$IC_{50}$ = 11nM



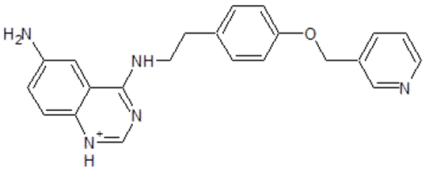
#3
$IC_{50}$ = 11nM



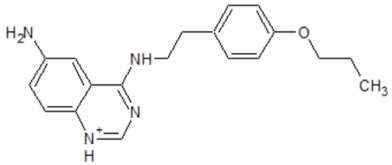
#5
$IC_{50}$ = 12nM



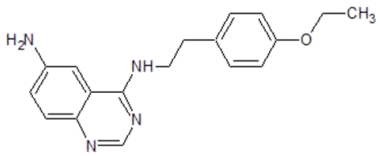
#4
$IC_{50}$ = 12nM



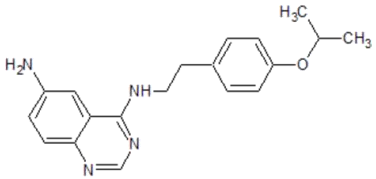
#37
$IC_{50}$ = 14nM



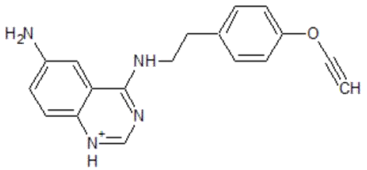
#6
$IC_{50}$ = 20nM



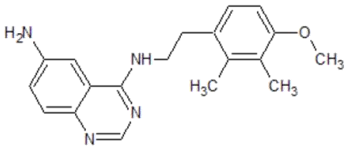
#38
$IC_{50}$ = 21nM



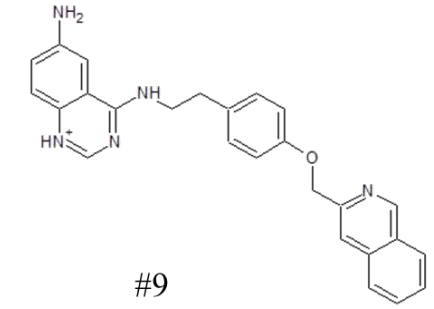
#39
$IC_{50}$ = 40nM



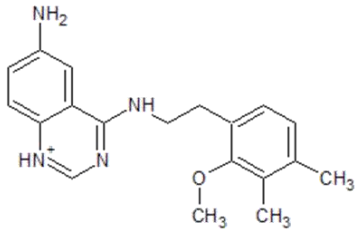
#7
$IC_{50}$ = 40nM



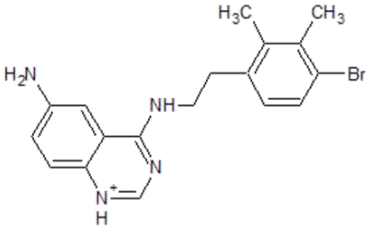
#8
$IC_{50}$ = 90nM



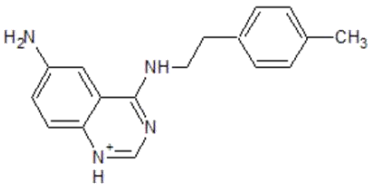
#9
$IC_{50}$ = 93nM



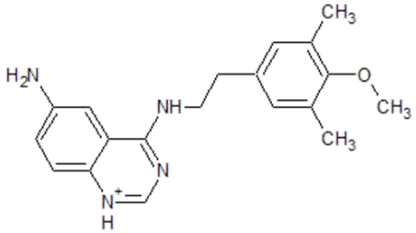
#10
$IC_{50}$ = 155nM



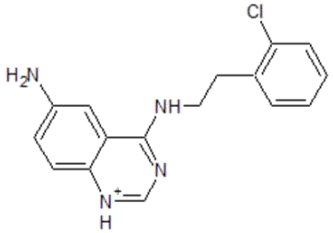
#11
$IC_{50}$ = 207nM



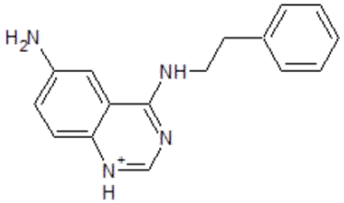
#41
$IC_{50}$ = 210nM



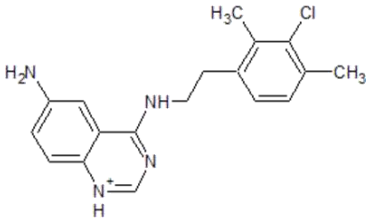
#12
$IC_{50}$ = 210nM



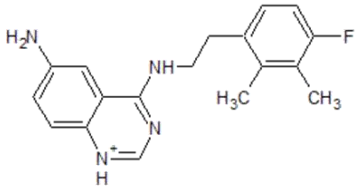
#42
$IC_{50}$ = 254nM



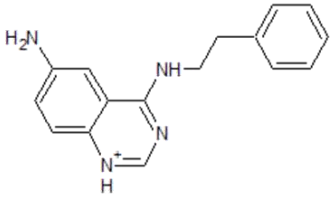
#43
$IC_{50}$ = 311nM



#15
$IC_{50}$ = 389nM



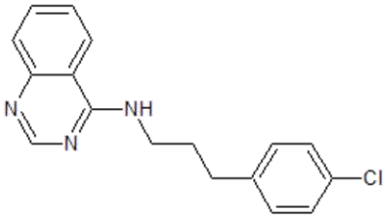
#16
$IC_{50}$ = 574nM



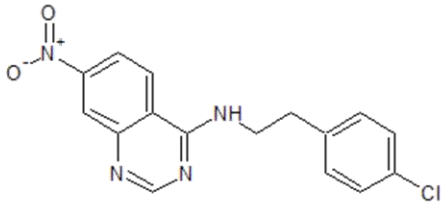
#17
$IC_{50}$ = 685nM

**Figure 2: The ECFP_4 dataset of active compounds used to perform the ROC curve validation. These compounds had activity ranging from 1-1000nM.**

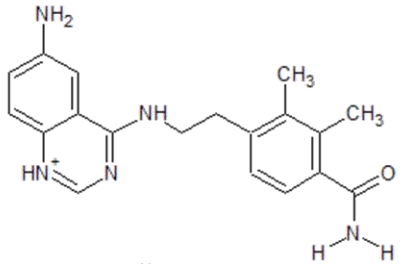
ECFP_4 Cluster: In Active compounds (1000-10000nM)



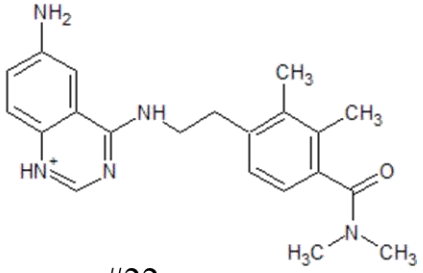
#47
$IC_{50}$ = 1401nM



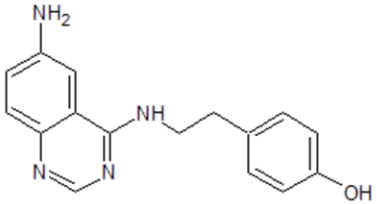
#48
$IC_{50}$ = 1765nM



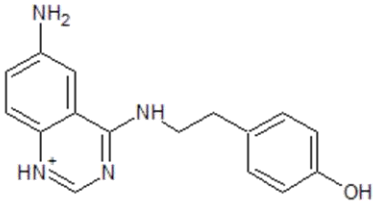
#22
$IC_{50}$ = 2594nM



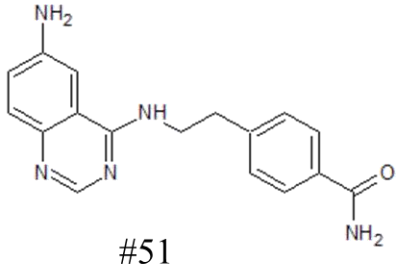
#22
$IC_{50}$ = 2594nM



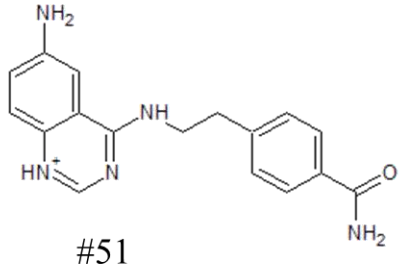
#23
$IC_{50}$ = 3645nM



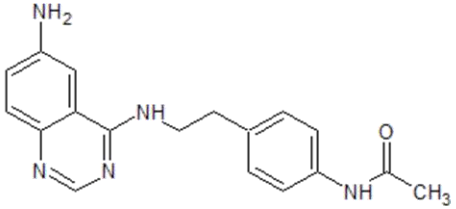
#23
$IC_{50}$ = 3645nM



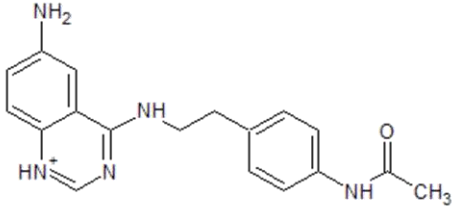
#51
$IC_{50}$ = 7333nM



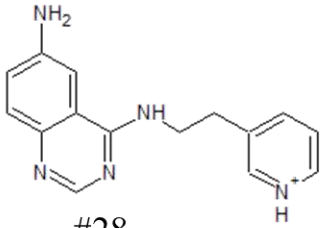
#51
$IC_{50}$ = 7333nM



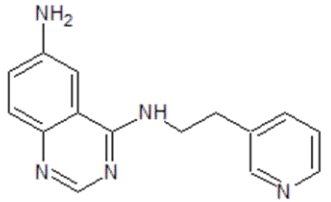
#26
$IC_{50}$ = 7340nM



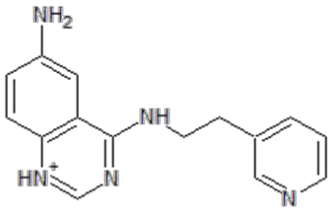
#26
$IC_{50}$ = 7340nM



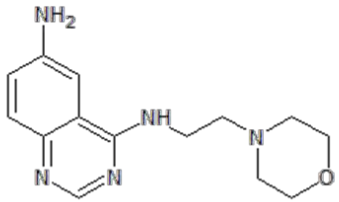
#28
$IC_{50}$ = 9060nM



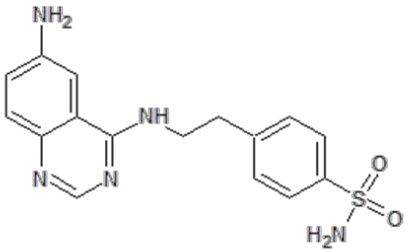
#28
$IC_{50}$ = 9060nM



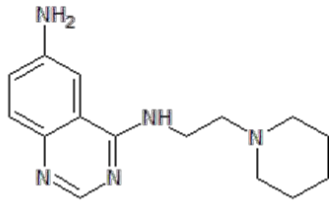
#28
$IC_{50}$ = 9060nM



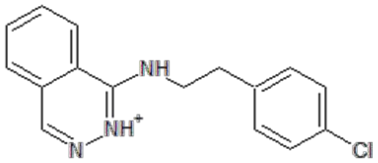
#32
$IC_{50}$ = 10000nM



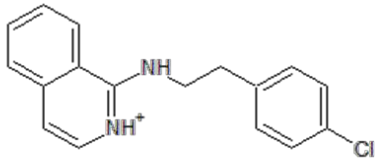
#33
$IC_{50}$ = 10000nM



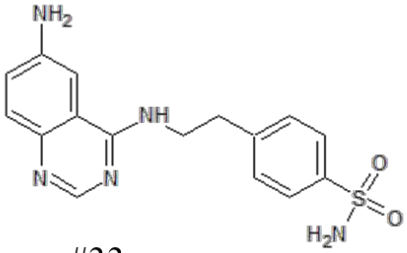
#53
$IC_{50}$ = 10000nM



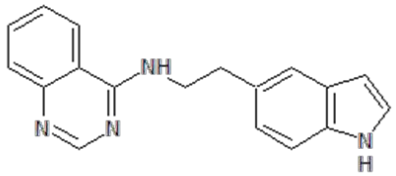
#52
$IC_{50}$ = 10000nM



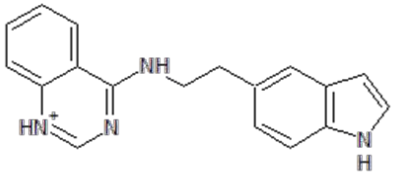
#31
$IC_{50}$ = 10000nM



#33
$IC_{50}$ = 10000nM



#2b
$IC_{50}$ = 27500nM



#2b
$IC_{50}$ = 27500nM

**Figure 3: The ECFP_4 dataset of inactive compounds used to perform the ROC curve validation. These compounds have activity ranging from 1000-10 000nM.**

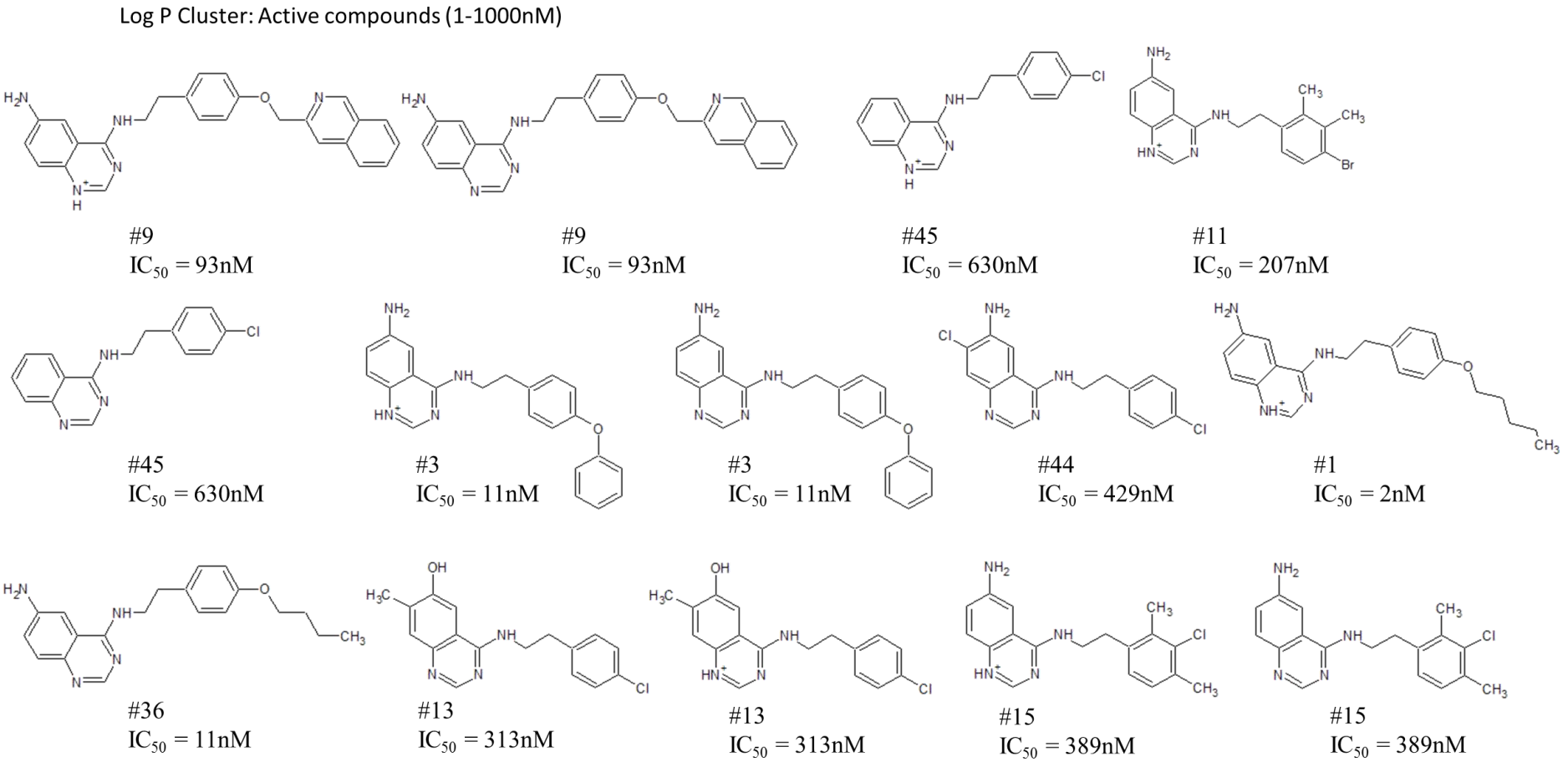
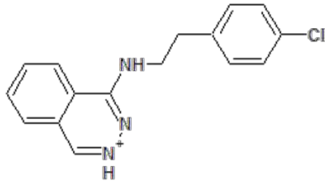
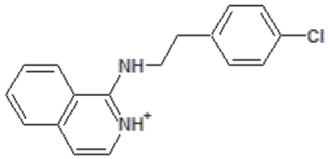



**Figure 4: The Log P dataset of active compounds used to perform the ROC curve validation. These compounds have activity ranging from 1-1000nM.**

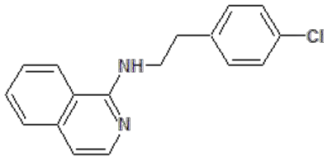
Log P Cluster: In Active compounds (1000-10000nM)



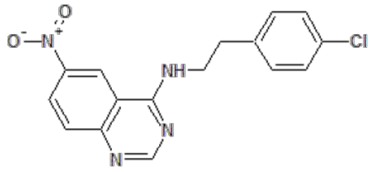
#52
$IC_{50}$ = 10000nM



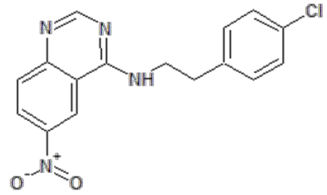
#31
$IC_{50}$ = 10000nM



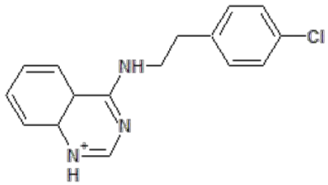
#31
$IC_{50}$ = 10000nM



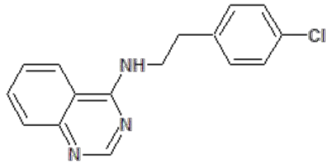
#20
$IC_{50}$ = 1515nM



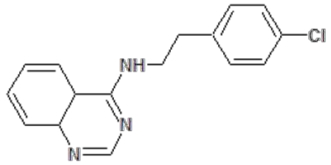
#48
$IC_{50}$ = 1765nM



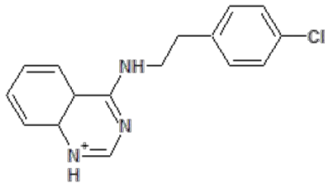
#30
$IC_{50}$ = 10000nM



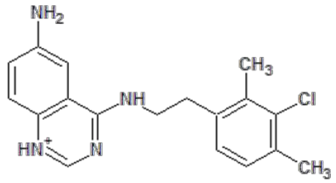
#30
$IC_{50}$ = 10000nM



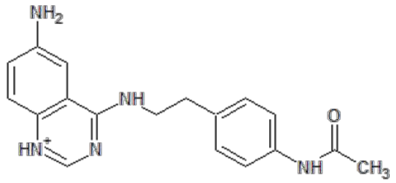
#21
$IC_{50}$ = 1823nM



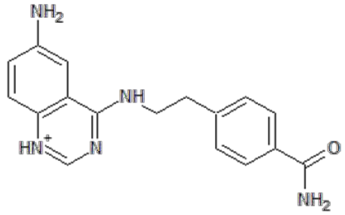
#21
$IC_{50}$ = 1823nM



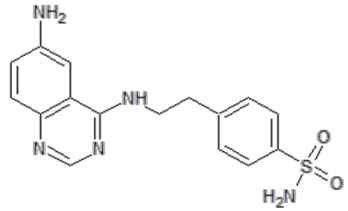
#15
$IC_{50}$ = 389nM



#26
$IC_{50}$ = 7340nM



#51
$IC_{50}$ = 7333nM



#33
$IC_{50}$ = 10000nM

**Figure 5: The Log P dataset of inactive compounds used to perform the ROC curve validation. These compounds have activity ranging from 1000-10 000nM.**

Log P & Mw Cluster: Active compounds (1-1000nM)

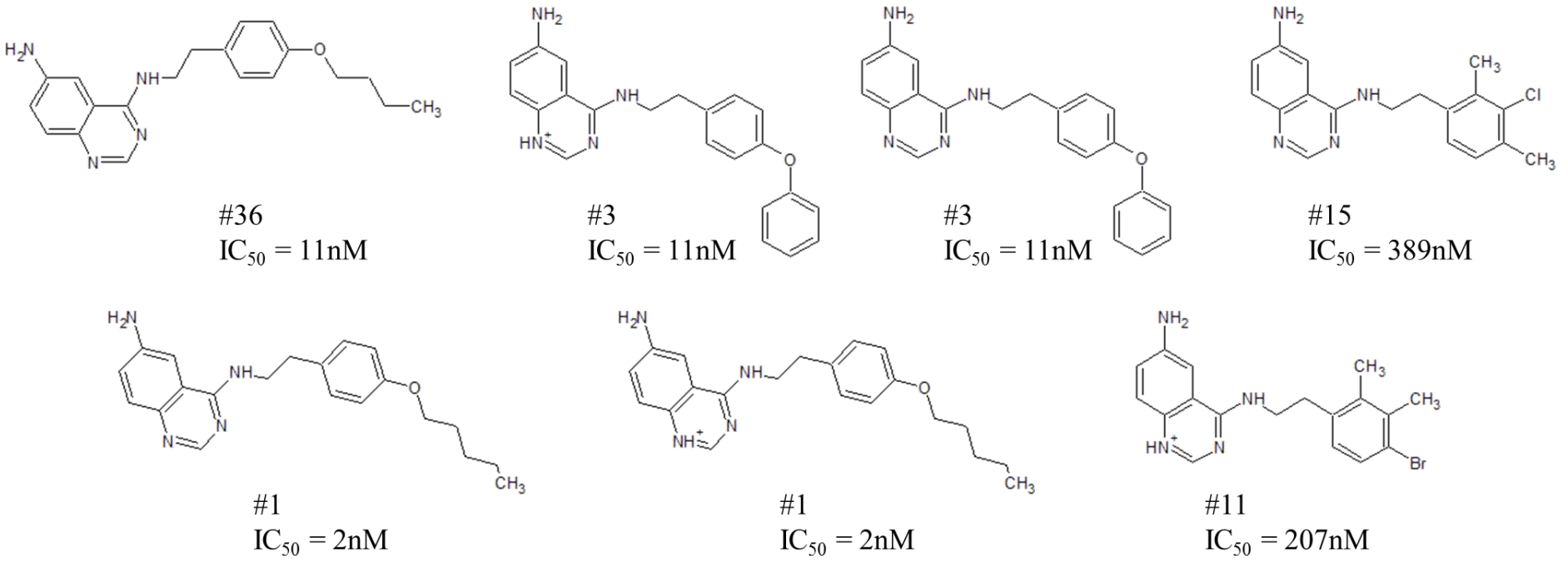
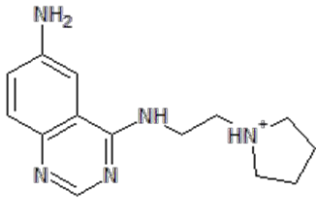
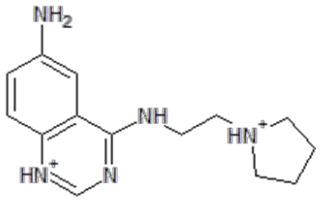



**Figure 6: The Log P & molecular weight dataset of active compounds used to perform the ROC curve validation. These compounds have activity ranging from 1-1000nM.**

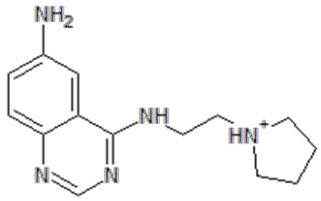
Log P & Mw Cluster: In Active compounds (1000nM-10000nM)



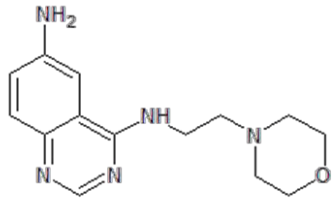
#24
$IC_{50}$ = 3694nM



#24
$IC_{50}$ = 3694nM



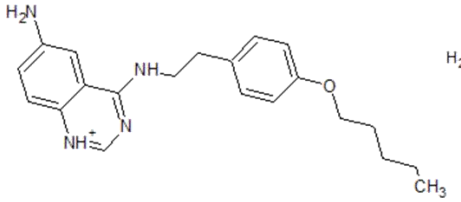
#53
$IC_{50}$ = 10000nM



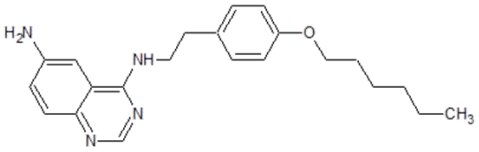
#32
$IC_{50}$ = 10000nM

**Figure 7: The Log P & molecular weight dataset of inactive compounds used to perform the ROC curve validation. These compounds have activity ranging from 1000-10 000nM.**

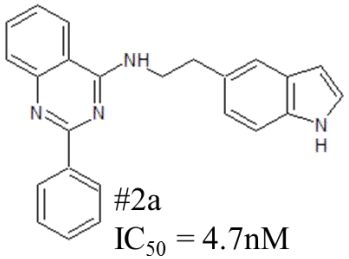
Structure Cluster: Active compounds (1-1000nM)



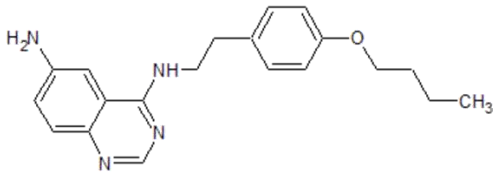
#1
$IC_{50}$ = 2nM



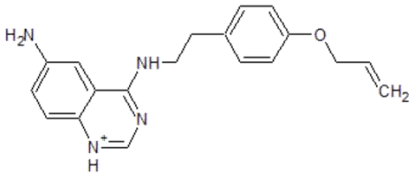
#2
$IC_{50}$ = 4nM



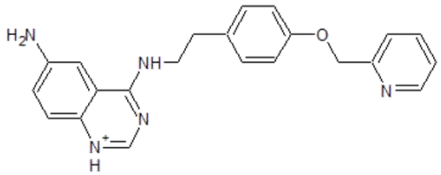
#2a
$IC_{50}$ = 4.7nM



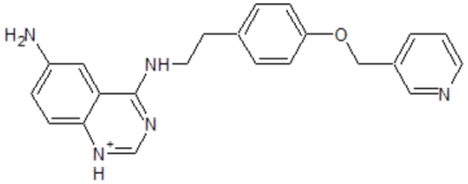
#36
$IC_{50}$ = 11nM



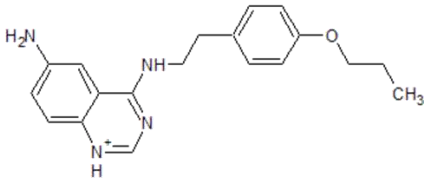
#5
$IC_{50}$ = 12nM



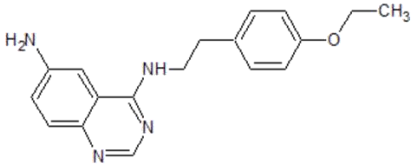
#4
$IC_{50}$ = 12nM



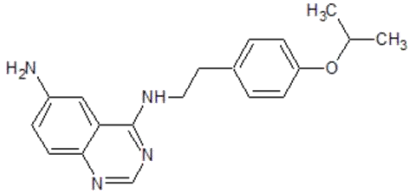
#37
$IC_{50}$ = 14nM



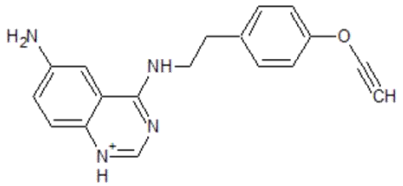
#6
$IC_{50}$ = 20nM



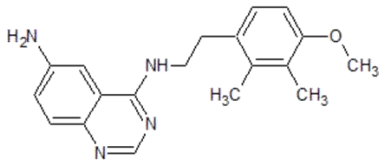
#38
$IC_{50}$ = 21nM



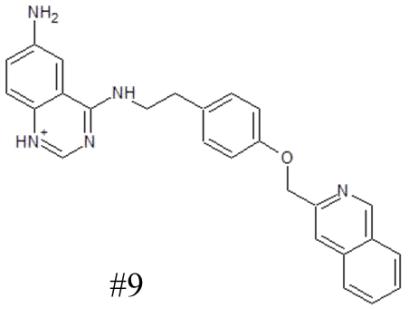
#39
$IC_{50}$ = 40nM



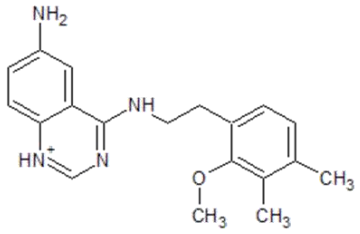
#7
$IC_{50}$ = 40nM



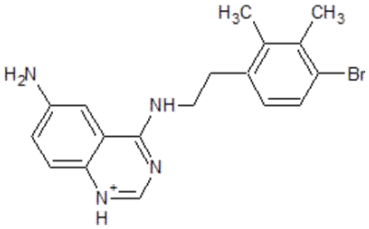
#8
$IC_{50}$ = 90nM



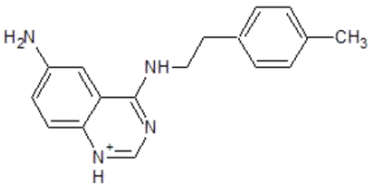
#9
$IC_{50}$ = 93nM



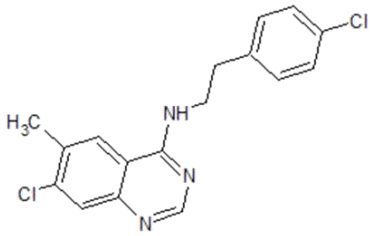
#10
$IC_{50}$ = 155nM



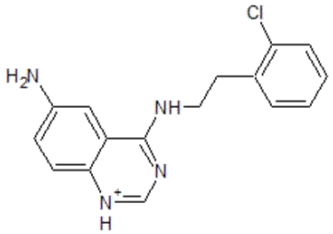
#11
$IC_{50}$ = 207nM



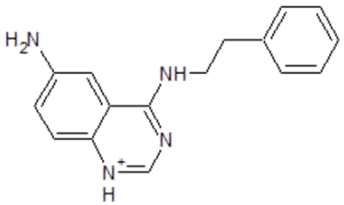
#41
$IC_{50}$ = 210nM



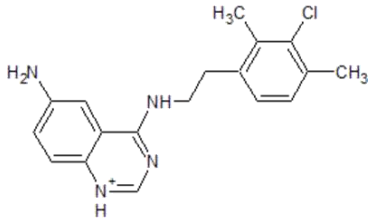
#44
$IC_{50}$ = 429nM



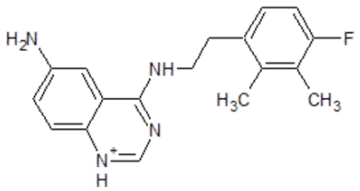
#42
$IC_{50}$ = 254nM



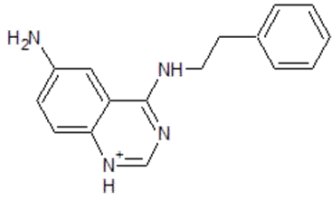
#43
$IC_{50}$ = 311nM



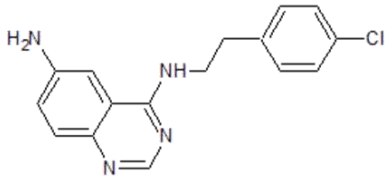
#15
$IC_{50}$ = 389nM



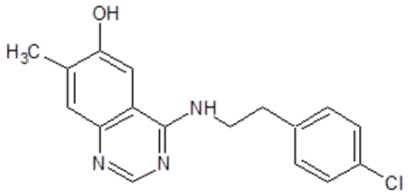
#16
$IC_{50}$ = 574nM



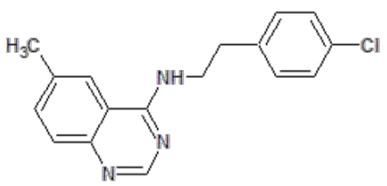
#17
$IC_{50}$ = 685nM



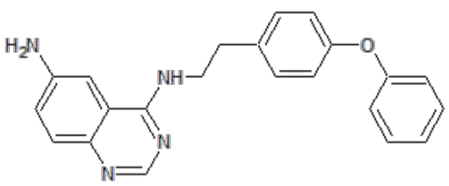
#40
$IC_{50}$ = 176nM



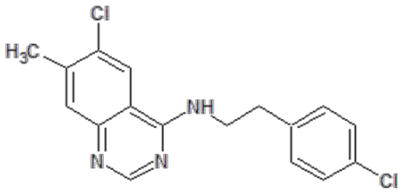
#13
$IC_{50}$ = 313nM



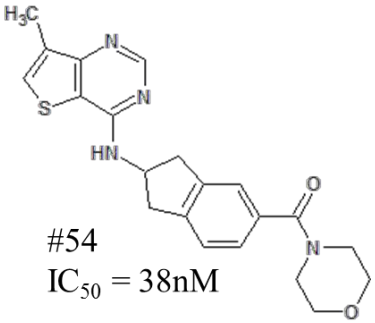
#46
$IC_{50}$ = 730nM



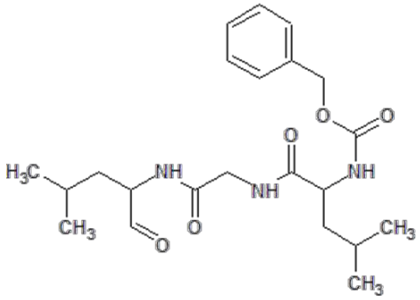
#3
$IC_{50}$ = 11nM



#14
$IC_{50}$ = 344nM



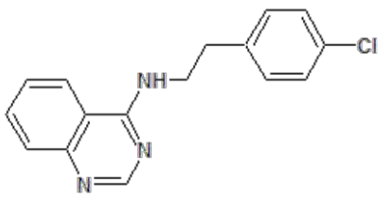
#54
$IC_{50}$ = 38nM



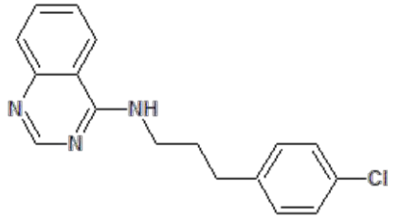
#55
$IC_{50}$ = nM

**Figure 8: The Structure dataset of active compounds used to perform the ROC curve validation. These compounds have activity ranging from 1-1000nM.**

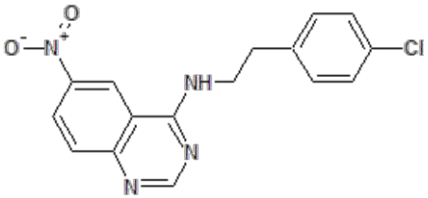
Structure Cluster: In Active compounds (1000-10000nM)



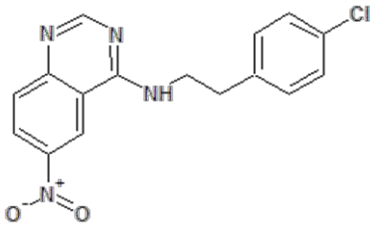
#19
$IC_{50}$ = 1392nM



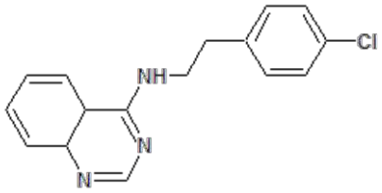
#47
$IC_{50}$ = 1401nM



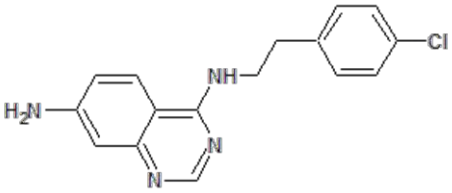
#20
$IC_{50}$ = 1515nM



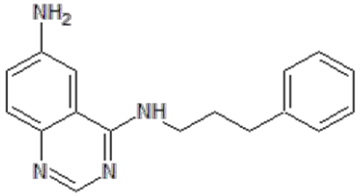
#48
$IC_{50}$ = 1765nM



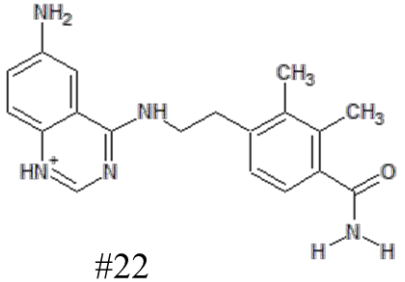
#21
$IC_{50}$ = 1832nM



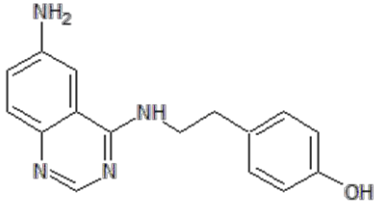
#49
$IC_{50}$ = 2202nM



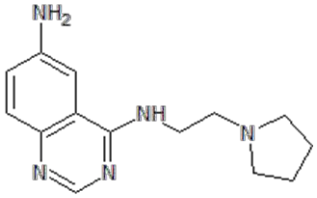
#50
$IC_{50}$ = 2381nM



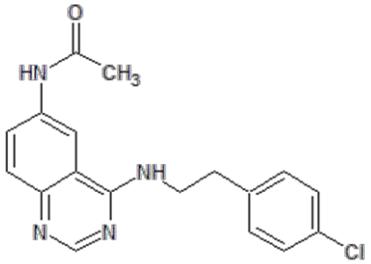
#22
$IC_{50}$ = 2594nM



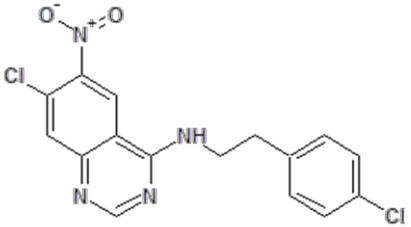
#23
$IC_{50}$ = 3645nM



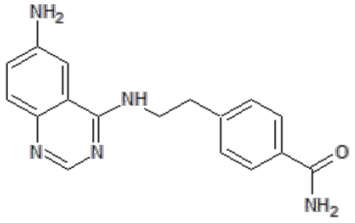
#24
$IC_{50}$ = 3694nM



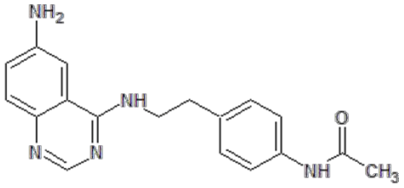
#25
$IC_{50}$ = 4562nM



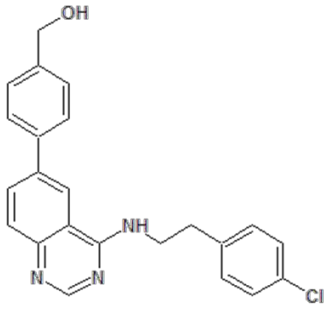
#29
$IC_{50}$ = 5979nM



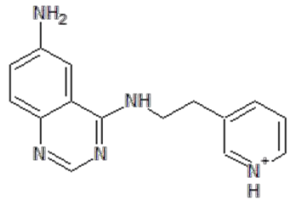
#51
$IC_{50}$ = 7333nM



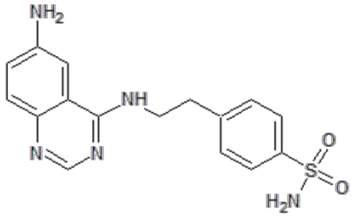
#26
$IC_{50}$ = 7340nM



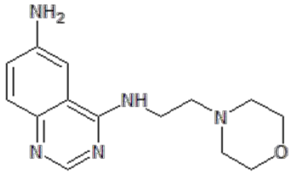
#27
$IC_{50}$ = 8271nM



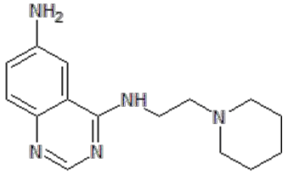
#28
$IC_{50}$ = 9060nM



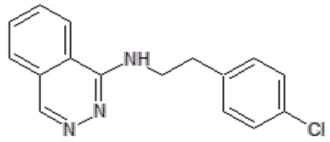
#33
$IC_{50}$ = 10000nM



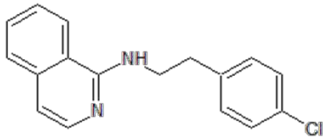
#32
$IC_{50}$ = 10000nM



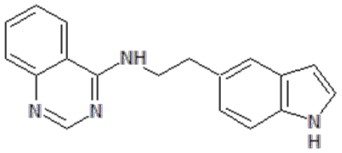
#53
$IC_{50}$ = 10000nM



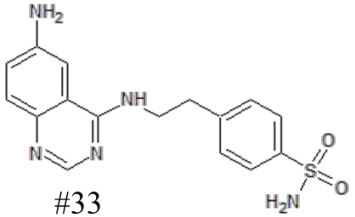
#52
$IC_{50}$ = 10000nM



#31
$IC_{50}$ = 10000nM



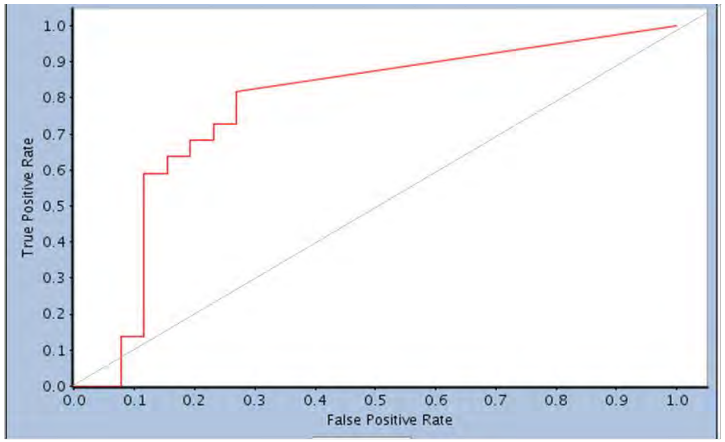
#2b
$IC_{50}$ = 10000nM

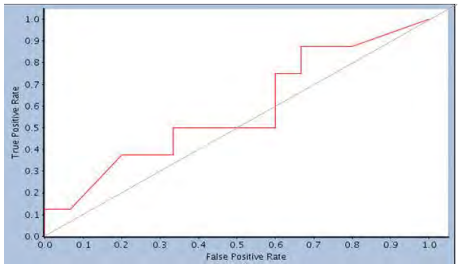


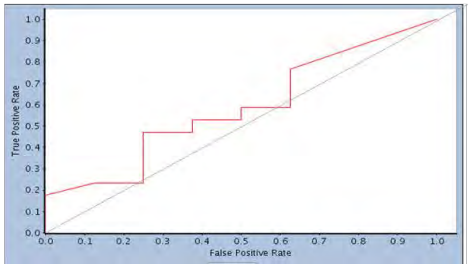
#33
$IC_{50}$ = 10000nM

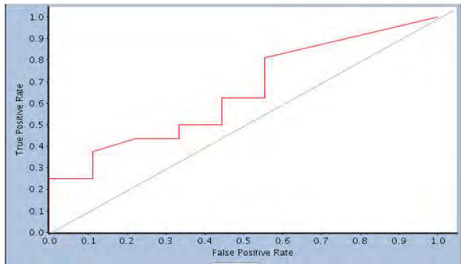
**Figure 9: The Structure dataset of inactive compounds used to perform the ROC curve validation. These compounds have activity ranging from 1000-10 000nM.**

ECFP_4 cluster 2D
ROC validation



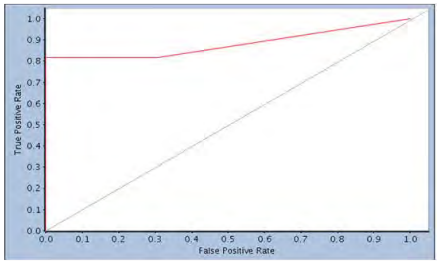
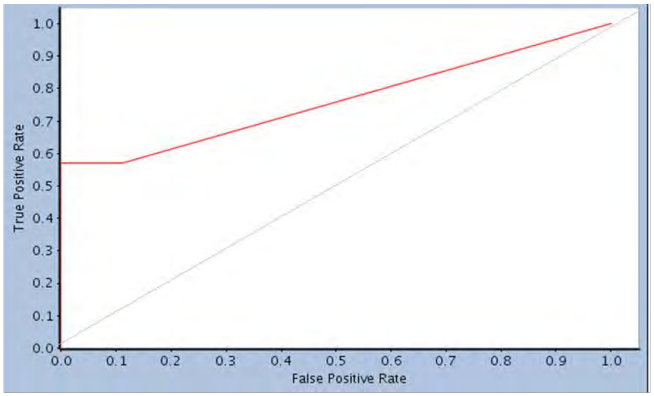




**Figure 10: The 2D ROC curve graph of the ECFP_4 cluster (top) and the four validation trials (below) . The AUC was 0.711 and no statistical relevance was found between the structure and activity of the compounds in the ECFP_4 dataset.**

Log P & Mw cluster
2D ROC validation



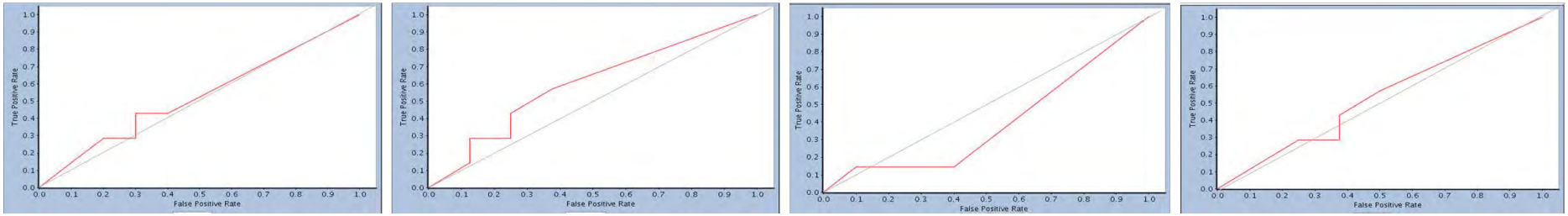




**Figure 11: The 2D ROC curve graph of the Log P & molecular weight cluster (top) and the four validation trials (below). The AUC was 0.762 and no statistical relevance was found between the structure and activity of the compounds in the Log P & molecular weight dataset.**

Log P cluster 2D ROC validation

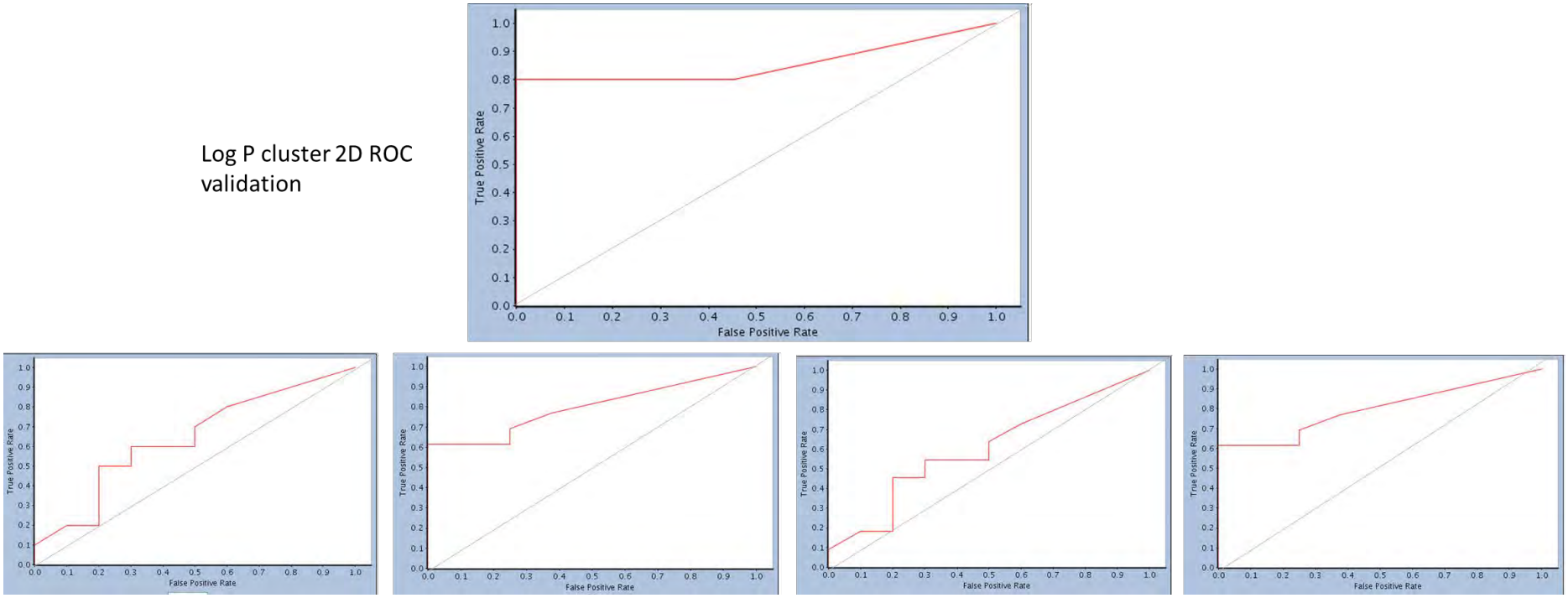



**Figure 12: The 2D ROC curve graph of the Log P cluster (top) and the four validation trials (below). The AUC was 0.855 and no statistical relevance was found between the structure and activity of the compounds in the Log P dataset.**

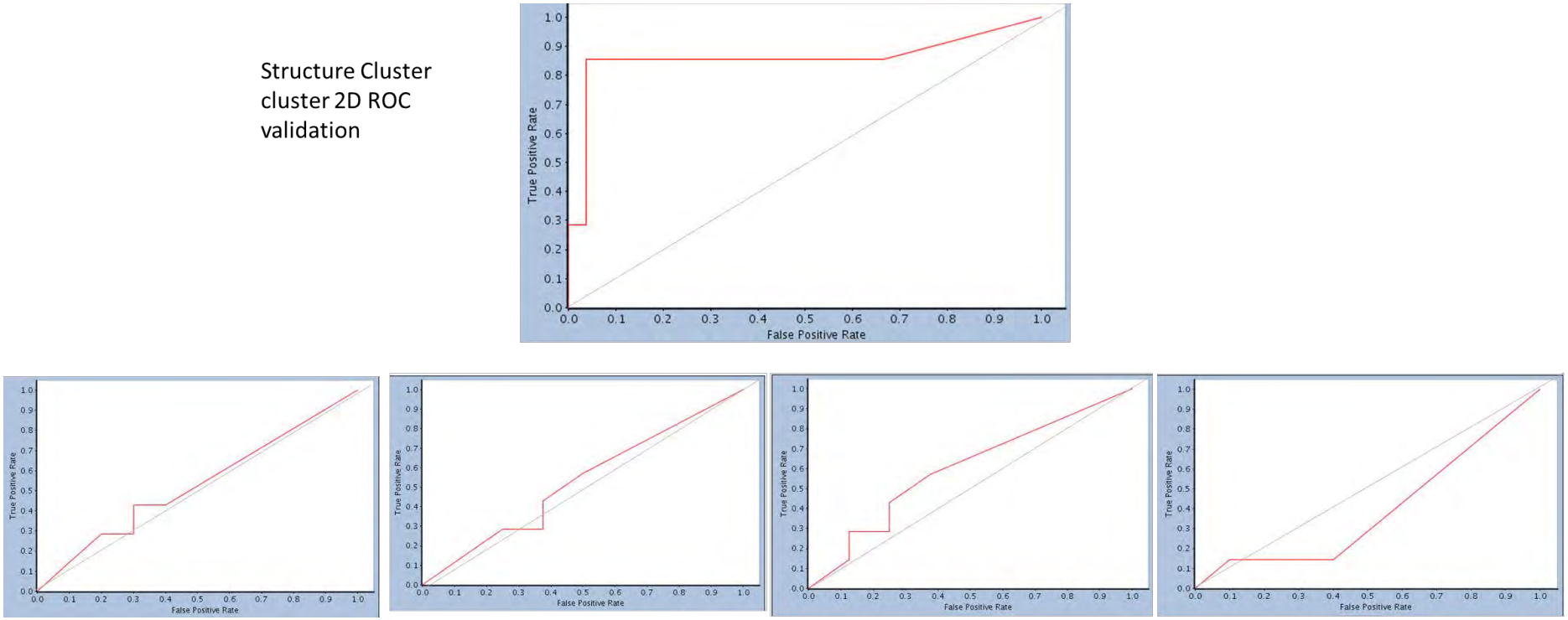
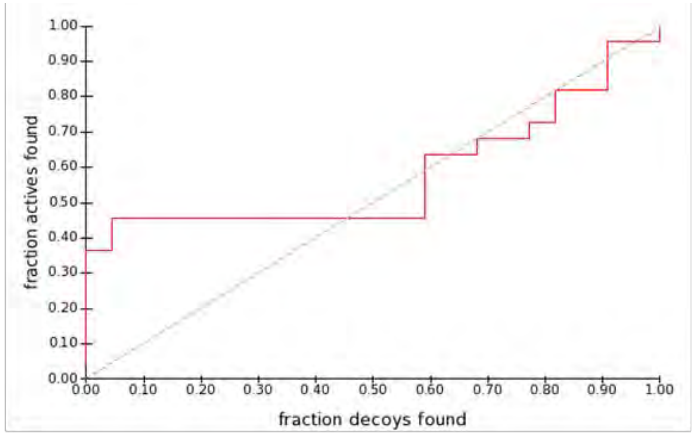



**Figure 13: The 2D ROC curve graph of the Structure cluster (top) and the four validations trials (below). The AUC was 0.860 and no statistical relevance was found between the structure and activity of the compounds in the Structure dataset.**

ECFP_4 Cluster 3D ROC curve validation



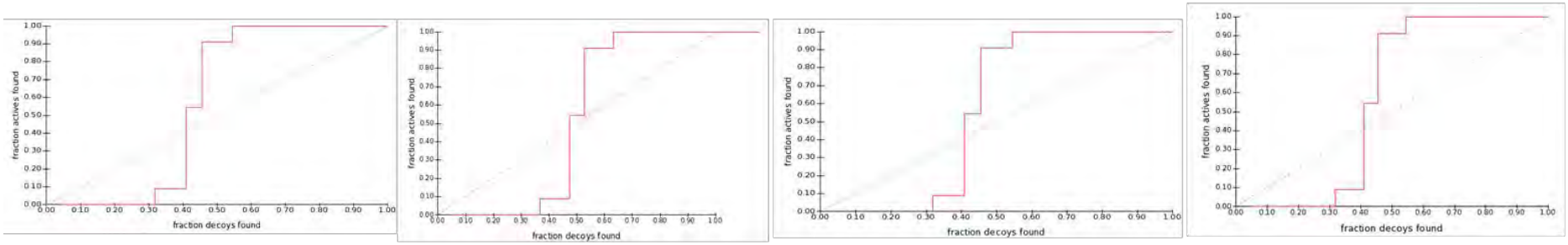




**Figure 14: The 3D ROC curve graph of the ECFP_4 cluster (top) and the four validation trials (below). The AUC was 0.566 and a statistical relevance was found between the structure and activity of the compounds in the ECFP_4 dataset.**

Log P Cluster 3D ROC curve validation

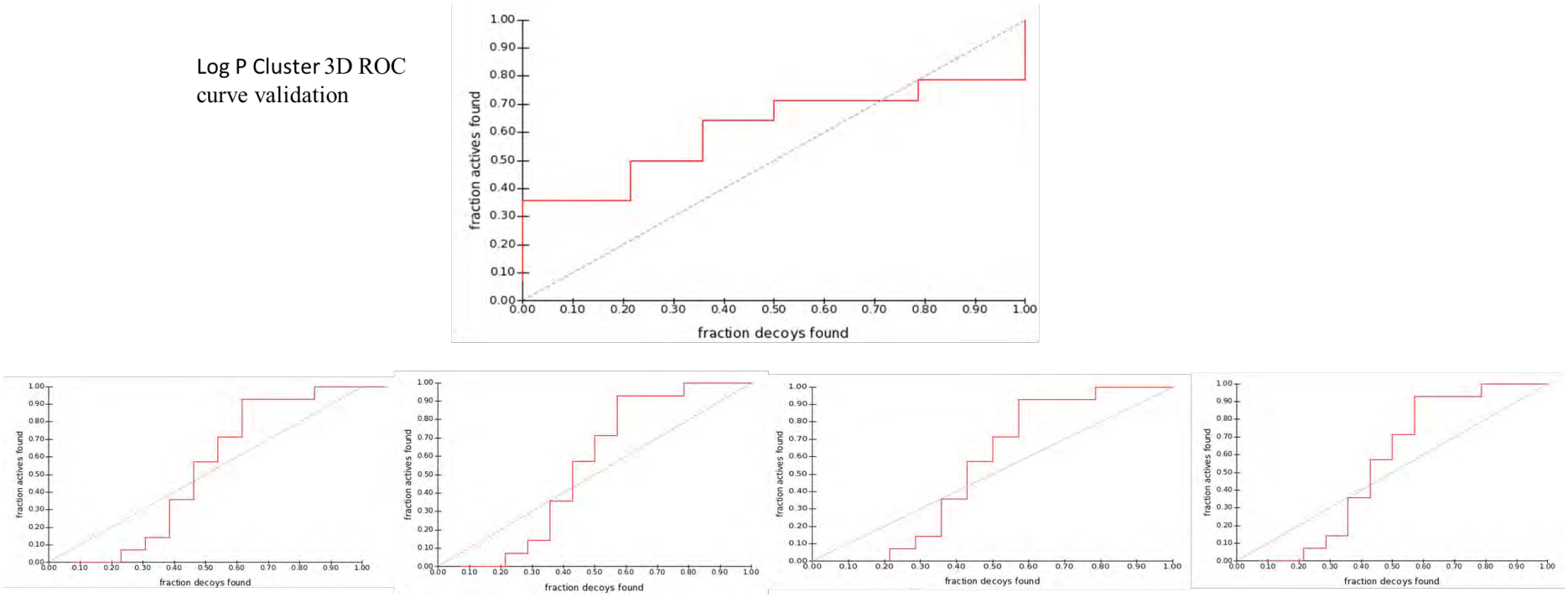
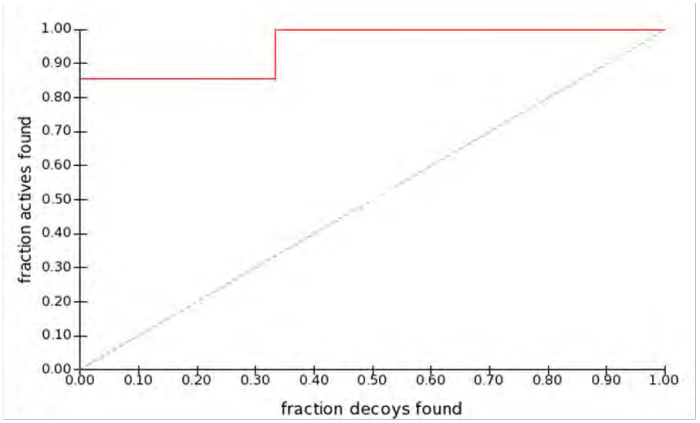



**Figure 14: The 3D ROC curve graph of the Log P cluster (top) and the four validation trials (below). The AUC was 0.630 and a statistical relevance was found between the structure and activity of the compounds in the Log P dataset.**

Log P & Mw
Cluster 3D ROC
curve validation



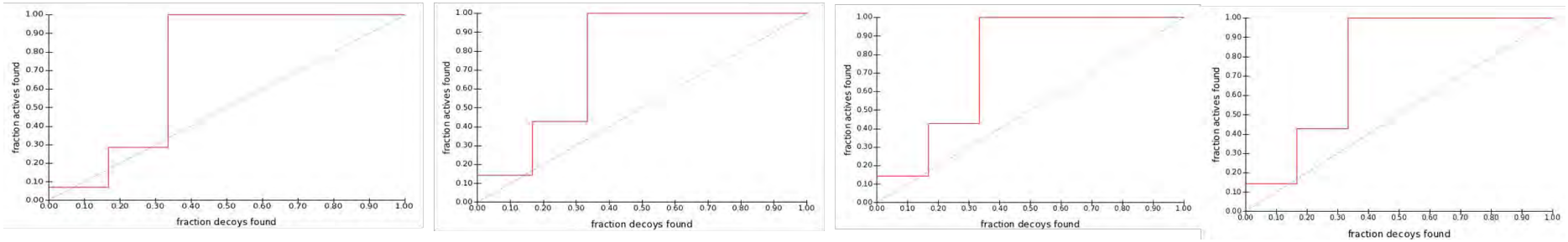
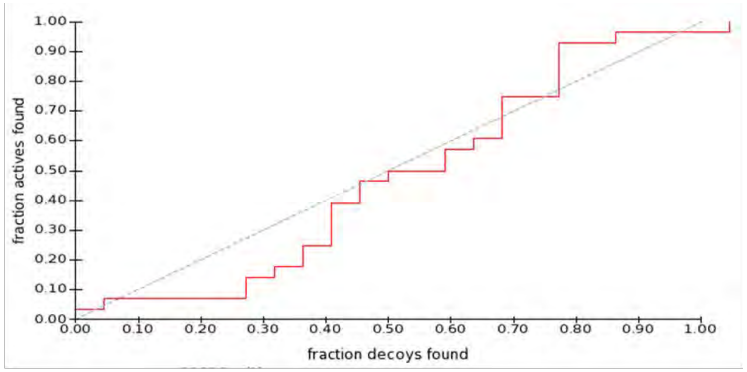




**Figure 15: The 3D ROC curve graph of the Log P & molecular weight cluster (top) and the four validation trials (below). The AUC was 0.944 and no statistical relevance was found between the structure and activity of the compounds in the Log P & Molecular weight dataset.**

Structure cluster
3D ROC curve
validation



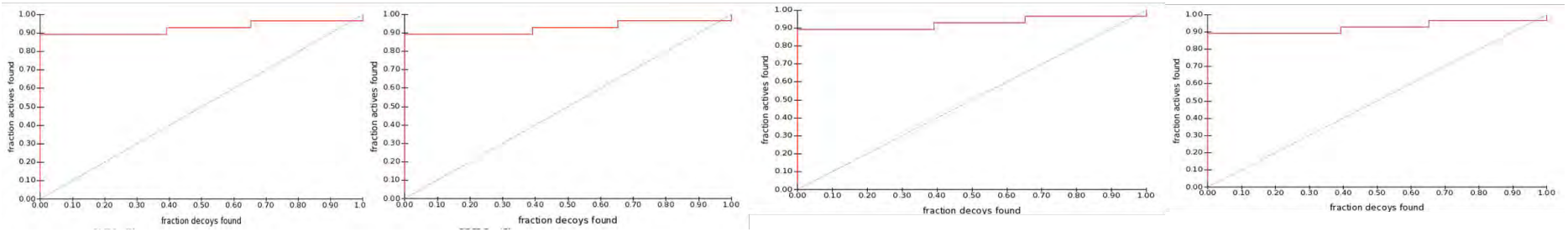
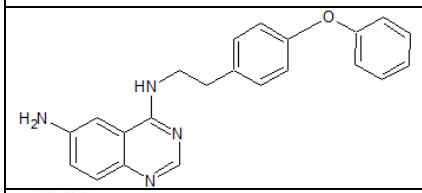
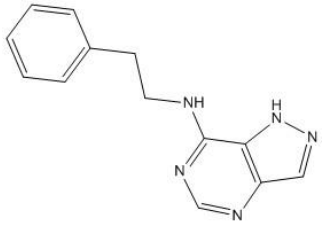
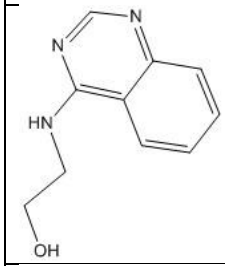
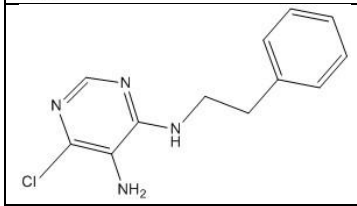




**Figure 16: The 3D ROC curve graph of the Structure cluster (top) and the four validation trials (below). The AUC was 0.463 and no statistical relevance was found between the structure and activity of the compounds in the Structure dataset.**

### 6.2.2 Appendix 2.2: Compounds identified from NCI database

**Table 1: The compounds identified from 2D similarity search performed for QNZ and their Tanimoto scores. These compounds were requested from the NCI for *in vitro* screening.**

| Molecular structures | NCI ID | Tanimoto Score |
|---|---|---|
|  | QNZ | 1 |
|  | NSC11641 | 0.421 |
|  | NSC13215 | 0.404 |
|  | NSC100126 | 0.453 |

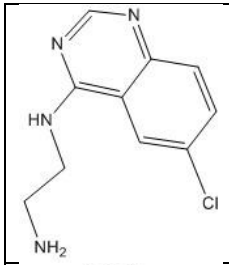
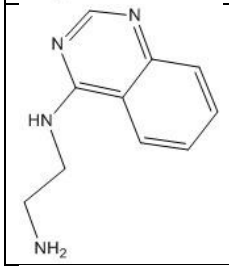
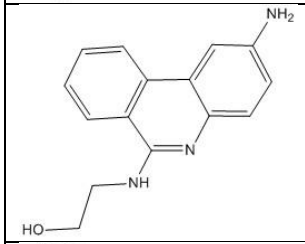
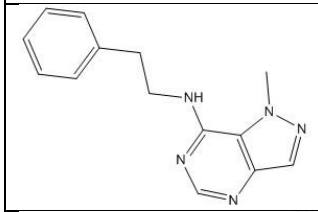
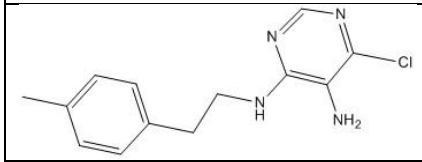
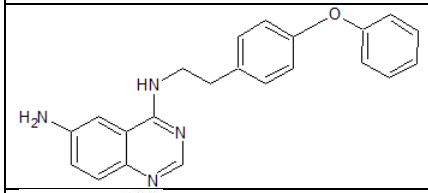
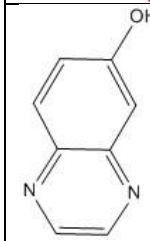
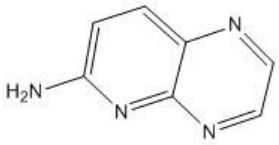
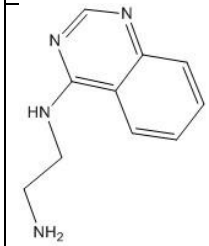
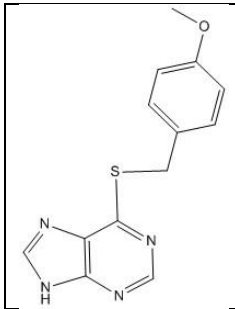
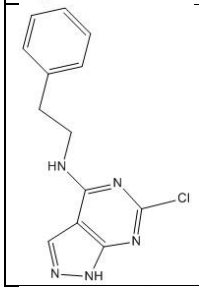
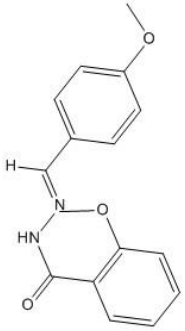
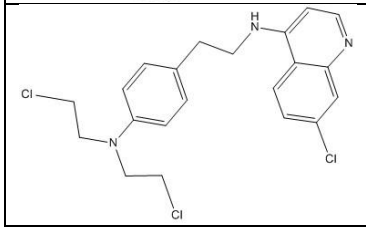
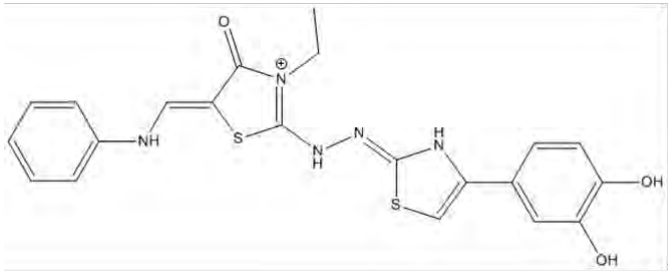
| | | |
|---|---|---|
|  | NSC52075 | 0.426 |
|  | NSC403389 | 0.431 |
|  | NSC150427 | 0.367 |
|  | NSC21105 | 0.407 |
|  | NSC94620 | 0.386 |

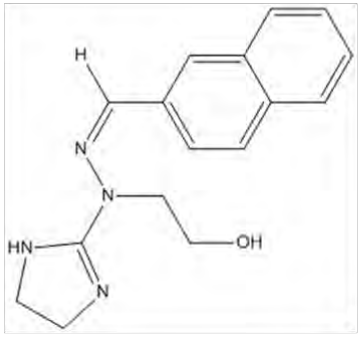
**Table 2: The compounds identified from 3D similarity search performed for QNZ and their Tanimoto scores. These compounds were requested from the NCI for *in vitro* screening.**

| Molecular structure | NCI ID | Tanimoto Combo | Tanimoto shape | Tanimoto colour |
|---|---|---|---|---|
|  | QNZ | 2 | 1 | 1 |
|  | NSC400599 | 0.899 | 0.428 | 0.47 |
|  | NSC59671 | 0.856 | 0.429 | 0.426 |
|  | NSC403396 | 0.964 | 0.548 | 0.416 |

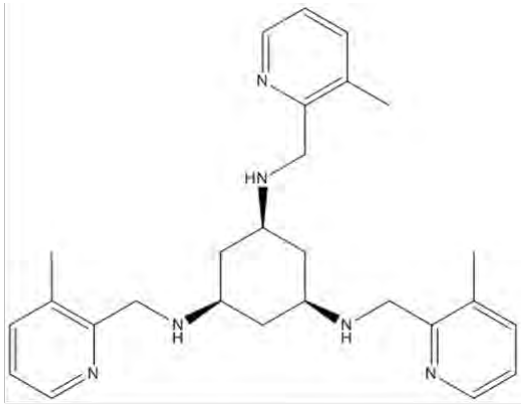
| | | | | |
|---|---|---|---|---|
|  | NSC47789 | 0.882 | 0.584 | 0.298 |
|  | NSC19136 | 1.194 | 0.702 | 0.491 |
|  | NSC204382 | 1.246 | 0.71 | 0.536 |
|  | NSC50982 | 1.241 | 0.787 | 0.454 |




NSC658798
FitValue: 1.849

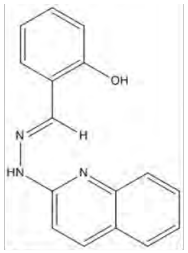



NSC676169
FitValue: 1.565

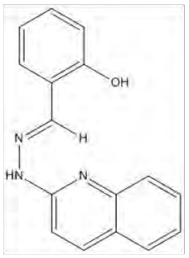



NSC727152
FitValue: 1.536

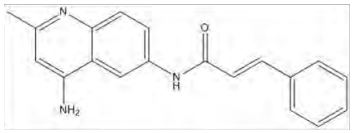
**Figure 1: The compounds identified from the virtual screening performed with the third pharmacophore hypothesis and their respective fit values. These compounds were requested from the NCI for *in vitro* screening.**



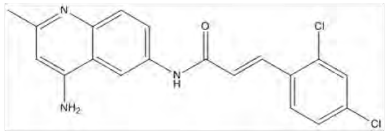
NSC181099
FitValue : 3.266



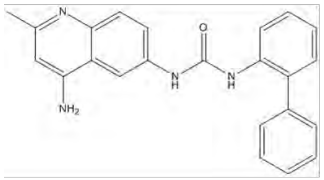
NSC89534
FitValue : 3.266



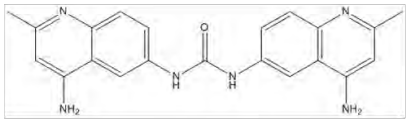
NSC33352
FitValue : 3.202



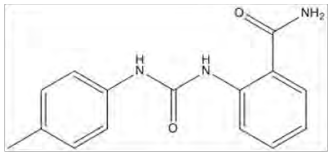
NSC33354
FitValue : 3.2



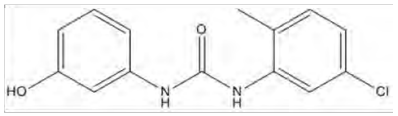
NSC33358
FitValue : 3.166



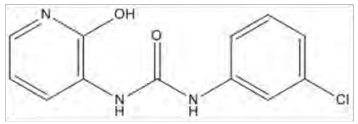
NSC12155
FitValue : 3.15



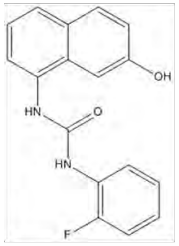
NSC153792
FitValue : 2.729



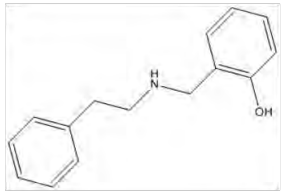
NSC205653
FitValue : 2.636



NSC329254
FitValue : 2.63

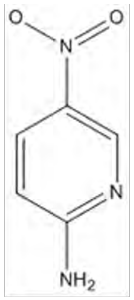


NSC214005
FitValue : 2.616

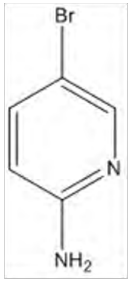


NSC13024
FitValue : 2.61

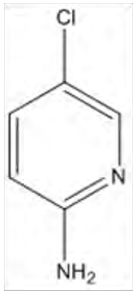
**Figure 2: The compounds identified from the virtual screening performed with the fourth pharmacophore hypothesis and their respective fit values. These compounds were requested from the NCI for *in vitro* screening.**



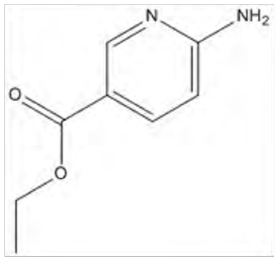

NSC23774



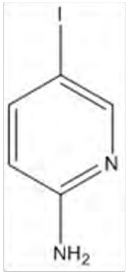

NSC26282



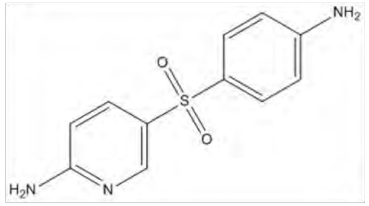

NSC26283



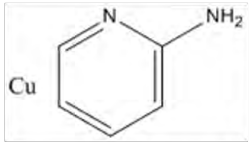

NSC37818



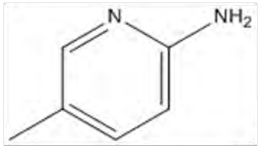

NSC31592



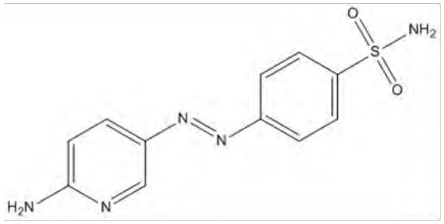

NSC33415



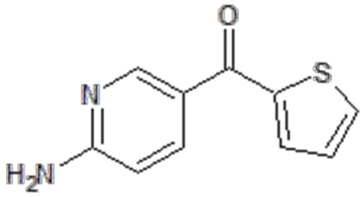

NSC403345



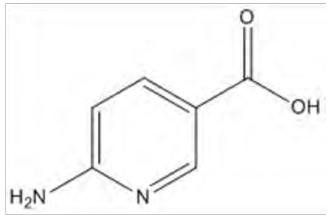

NSC96444




NSC106207




NSC343344




NSC383

**Figure 3: The compounds identified from the 2-aminopyridine substructure search performed for Pentamidine. These compounds were requested from the NCI for *in vitro* screening.**

## 6.3 Appendix 3

### 6.3.1 Appendix 3.1: Compound Solubility

**Table 1: The compounds identified from the ZINC database that were insoluble and precipitated out of the culture media. The solubilisation enhancement techniques did not improve the solubility of these compounds.**

| QNZ | Pentamidine |
|---|---|
| ZINC43272460 | ZINC02996814 |
| ZINC52468283 | ZINC48997508 |
| ZINC49000199 | ZINC09473369 |
| ZINC08764584 | ZINC15613510 |
| | ZINC47360766 |
| | ZINC12672314 |
| | ZINC31673334 |

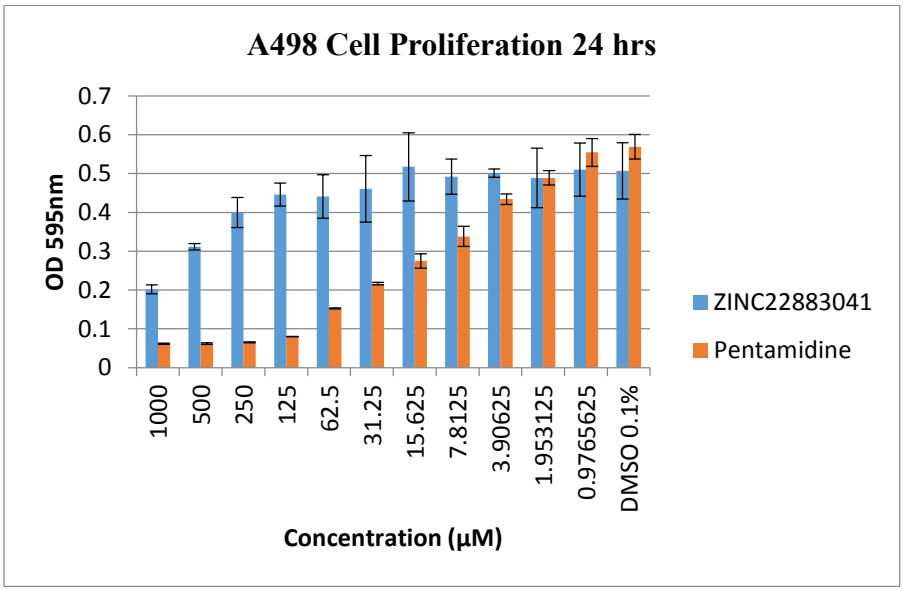
**Table 2: The compounds identified from the NCI database that were insoluble and precipitated out of the culture media. The solubilisation enhancement techniques did not improve the solubility of these compounds. The compounds denoted with an asterisk include those that were excluded from the primary screen.**

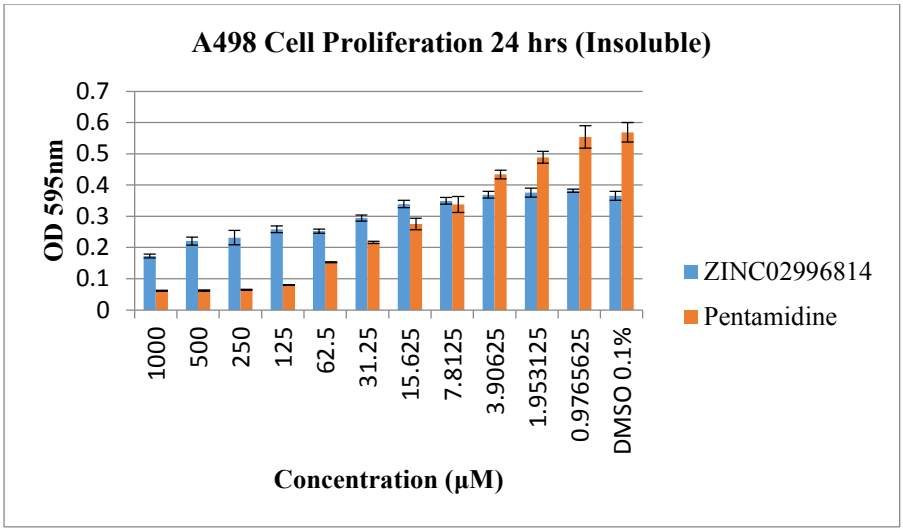
| QNZ | Pentamidine |
|---|---|
| NSC47789 | NSC403345 |
| NSC403389* | NSC334433 |
| NSC89534 | NSC383 |
| NSC12155 | NSC33415 |
| NSC153792 | NSC31592 |
| NSC33353* | NSC26282 |
| NSC150427* | NSC106207 |
| NSC21105 | |
| NSC19136 | |

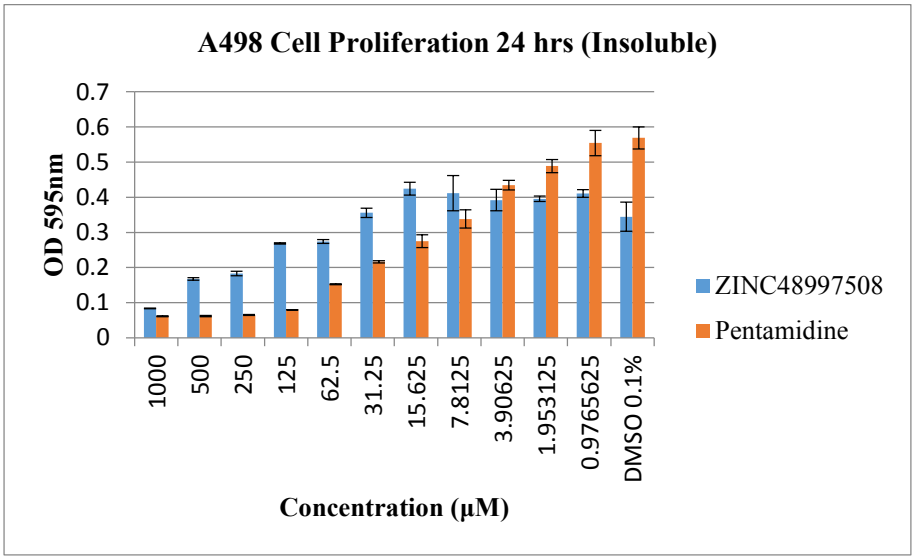
| | |
|---|---|
| NSC214005 | |
| NSC106570 | |
| NSC204382 | |
| NSC33352 | |
| NSC205653 | |
| NSC630355 | |
| NSC94620 | |
| NSC181099 | |
| NSC658798 | |
| NSC50982* | |

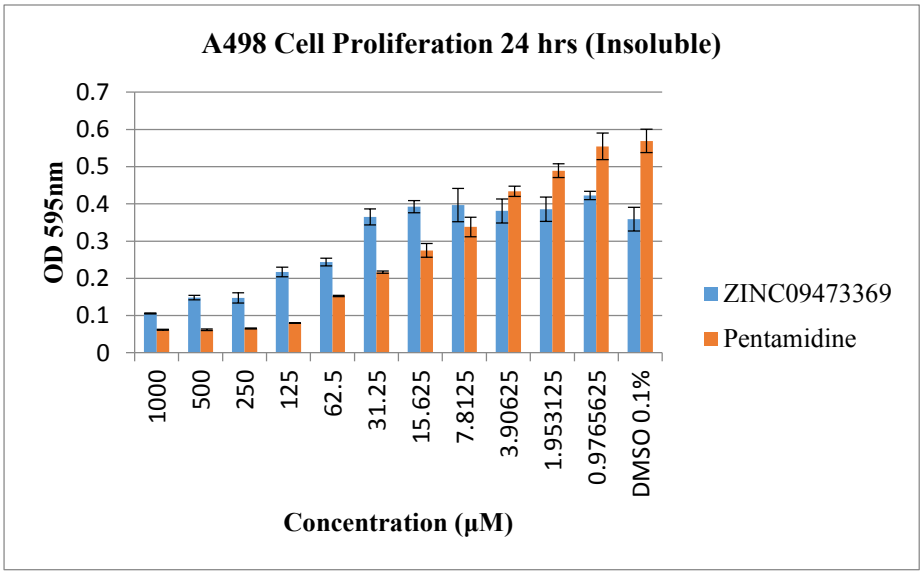
### 6.3.2 Appendix 3.2: Biological Results of ZINC Compounds Identified for Pentamidine

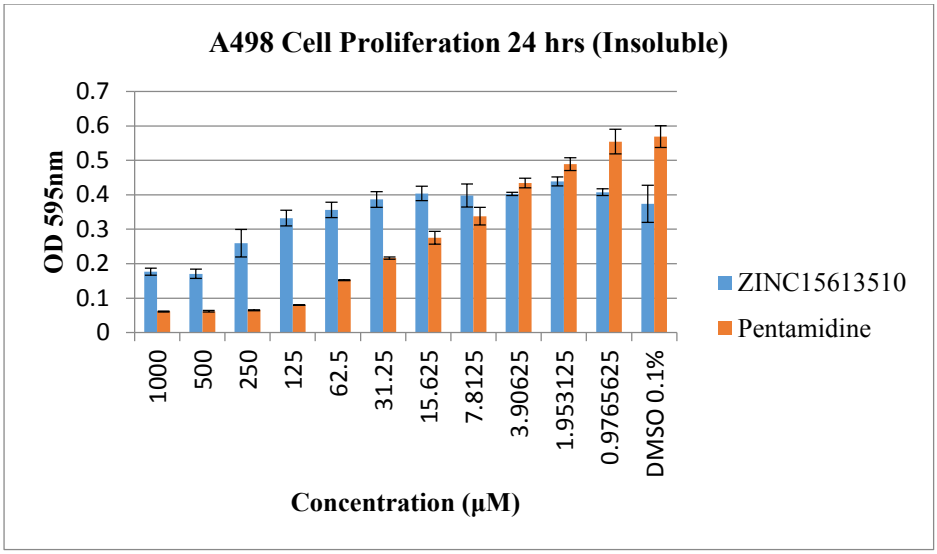
The MTT proliferation graph of ZINC compounds identified from the SBVS performed for Pentamidine in Chapter2. The graphs below include the proliferation MTT assay, 24 hour post-treatment in A498 cancer cell line. DMSO 0.1% was used as vehicle control. Data shown represents mean ± S.D. of four independent experiments performed in duplicate. OD = absorbance.

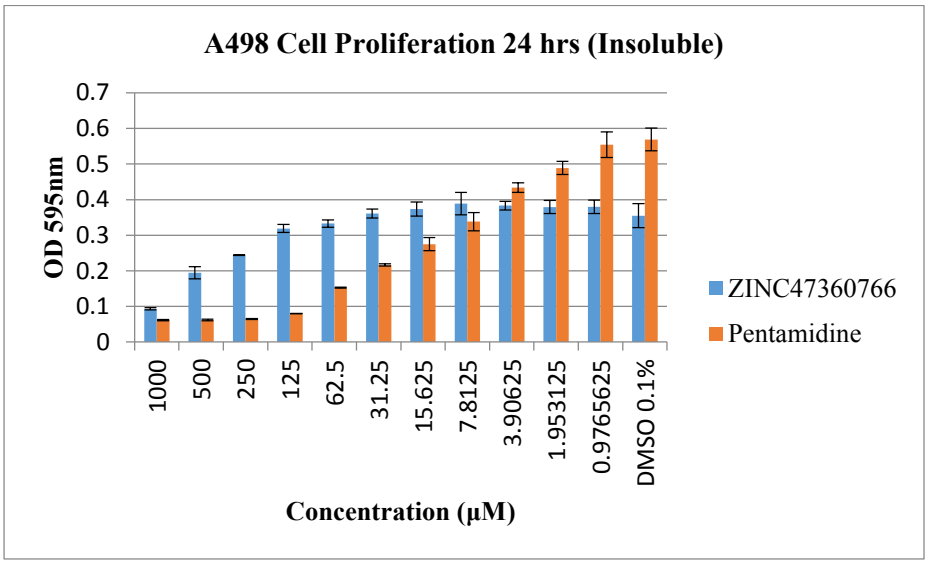


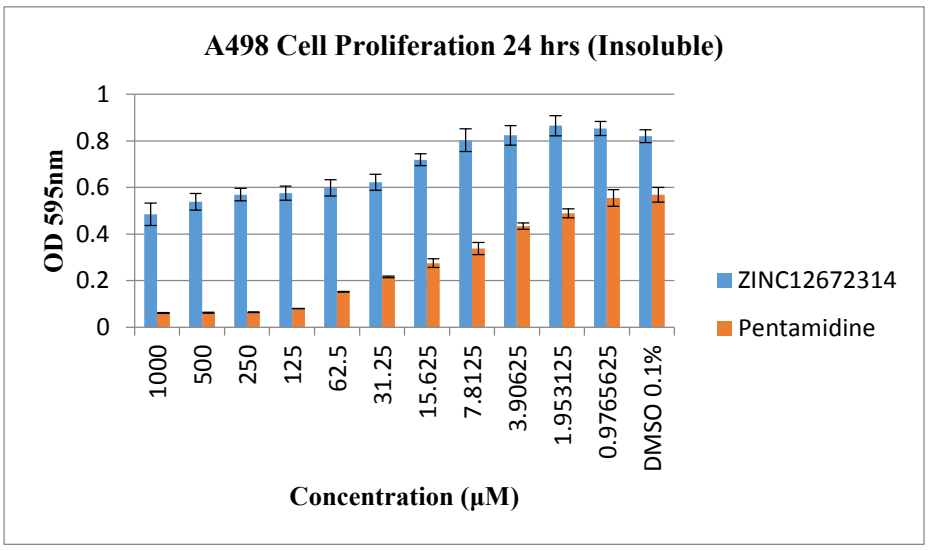

A498 Cell Proliferation 24 hrs (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC02996814
Pentamidine

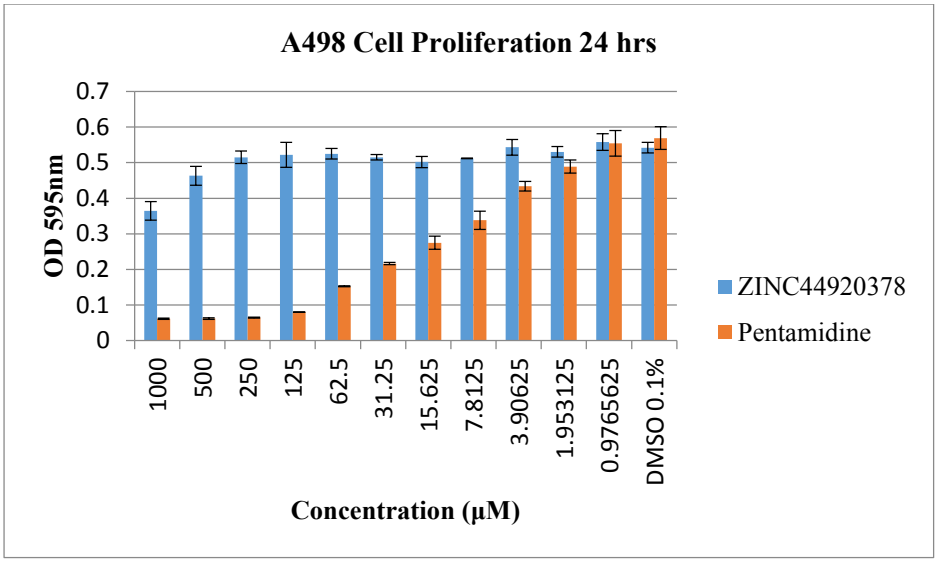

A498 Cell Proliferation 24 hrs (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC48997508
Pentamidine

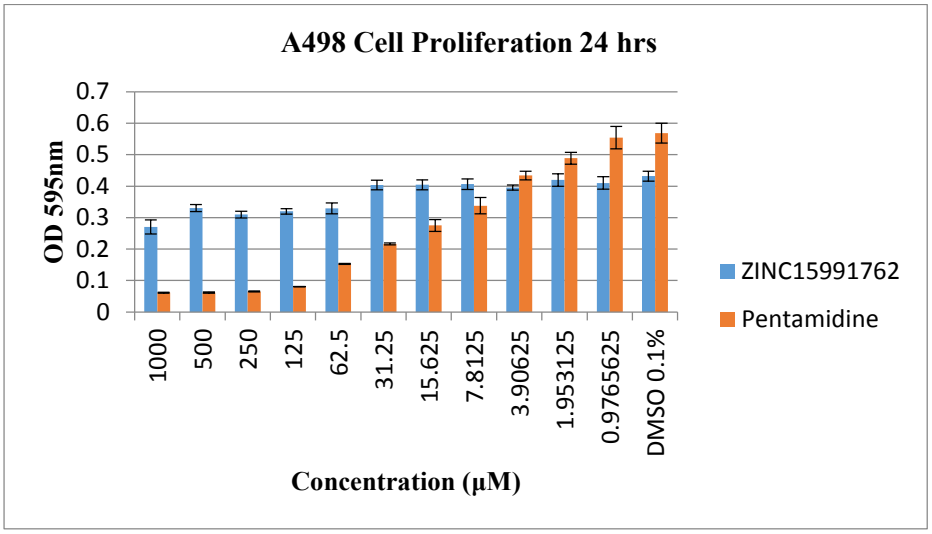

A498 Cell Proliferation 24 hrs (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC09473369
Pentamidine

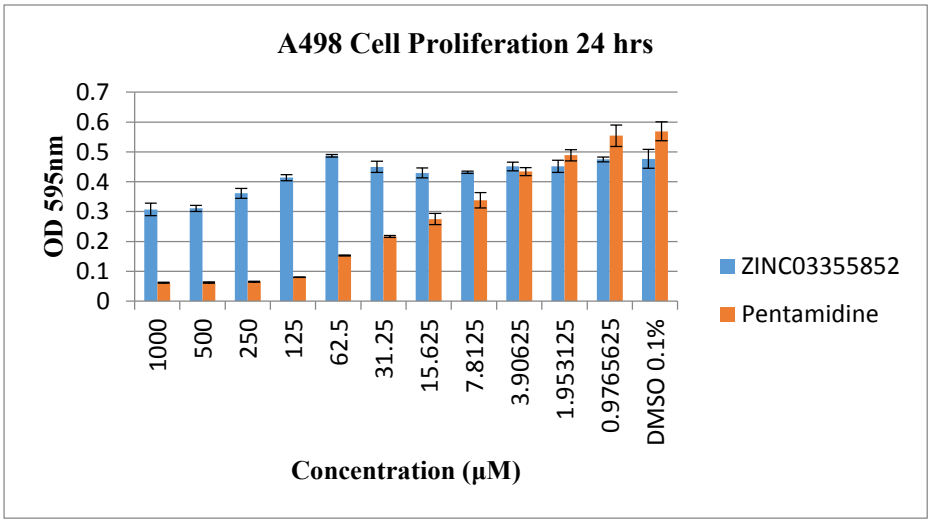

A498 Cell Proliferation 24 hrs (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC15613510
Pentamidine

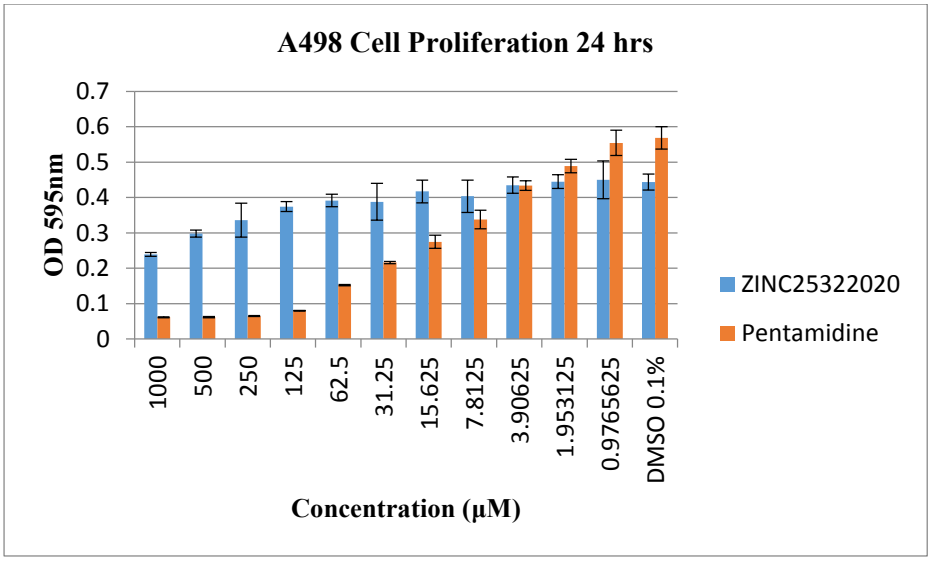

A498 Cell Proliferation 24 hrs (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC47360766
Pentamidine

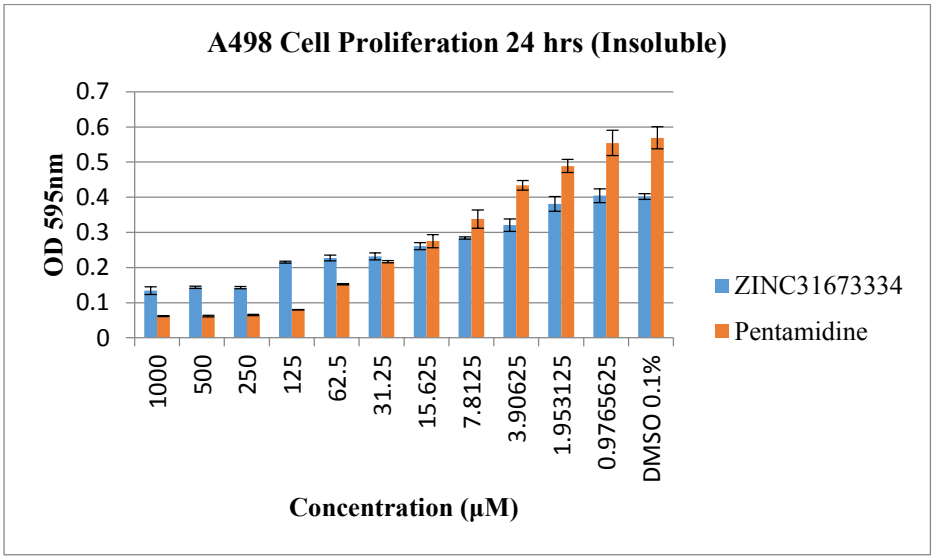

A498 Cell Proliferation 24 hrs (Insoluble)
OD 595nm
1
0.8
0.6
0.4
0.2
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC12672314
Pentamidine

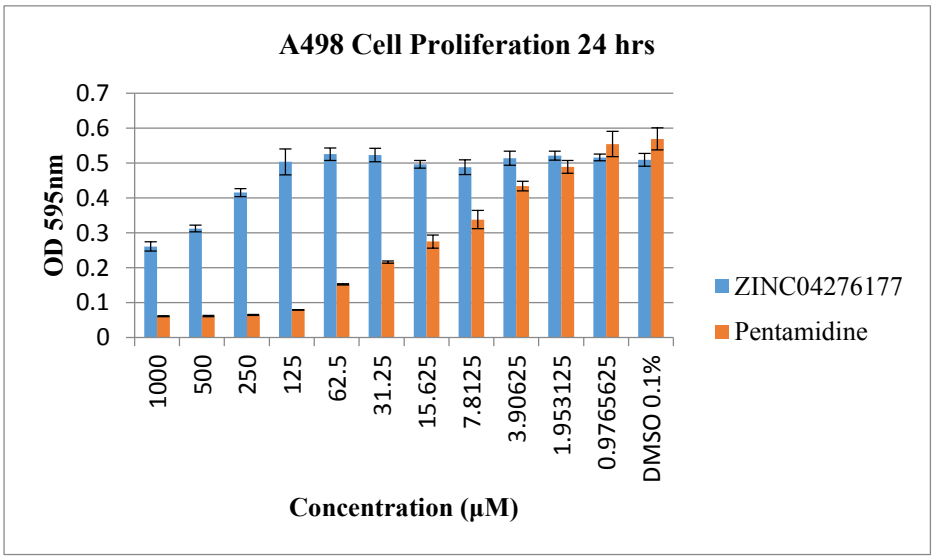

A498 Cell Proliferation 24 hrs
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC44920378
Pentamidine

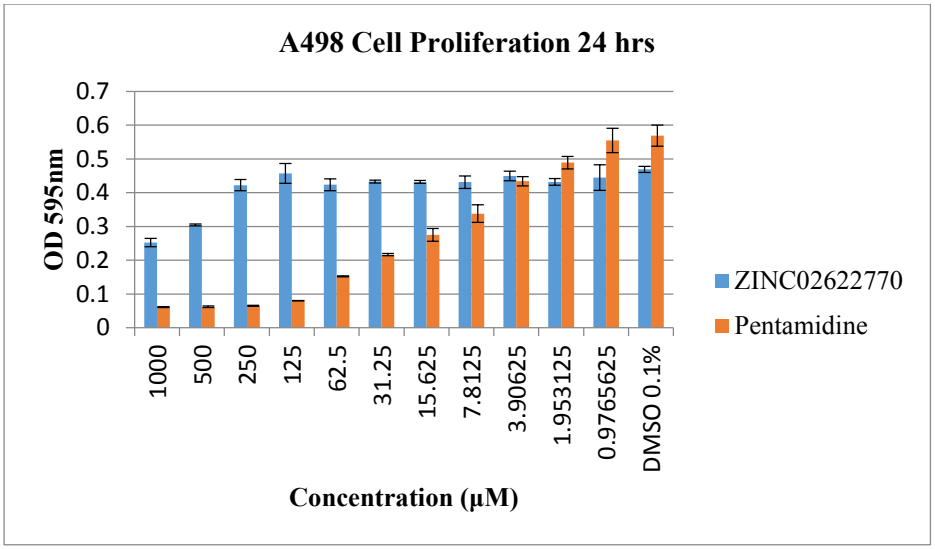

A498 Cell Proliferation 24 hrs
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC15991762
Pentamidine


A498 Cell Proliferation 24 hrs
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC03355852
Pentamidine


A498 Cell Proliferation 24 hrs
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC25322020
Pentamidine

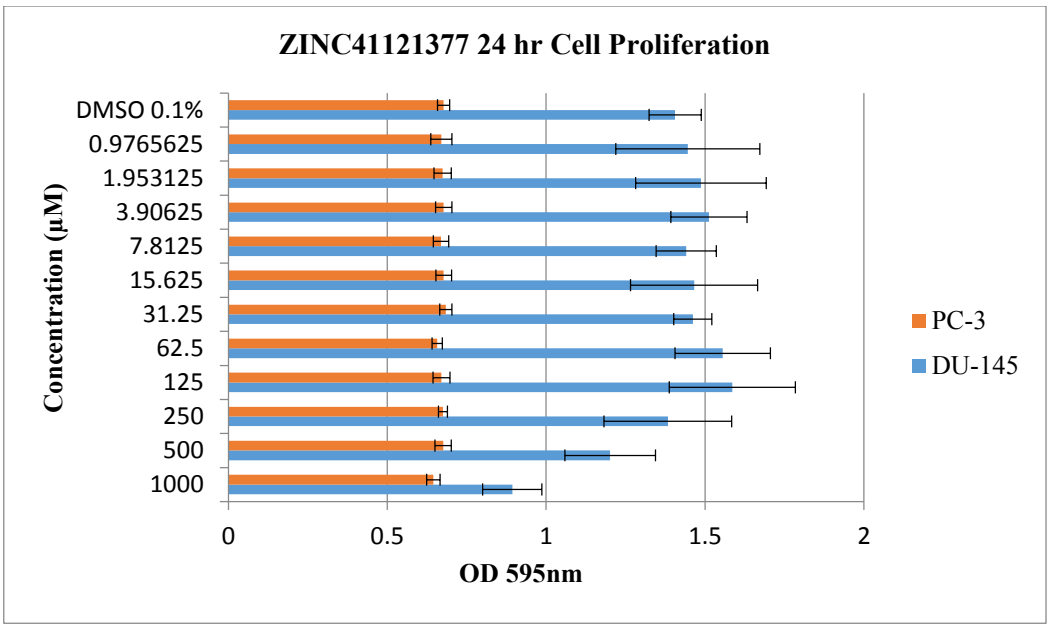

A498 Cell Proliferation 24 hrs (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC31673334
Pentamidine

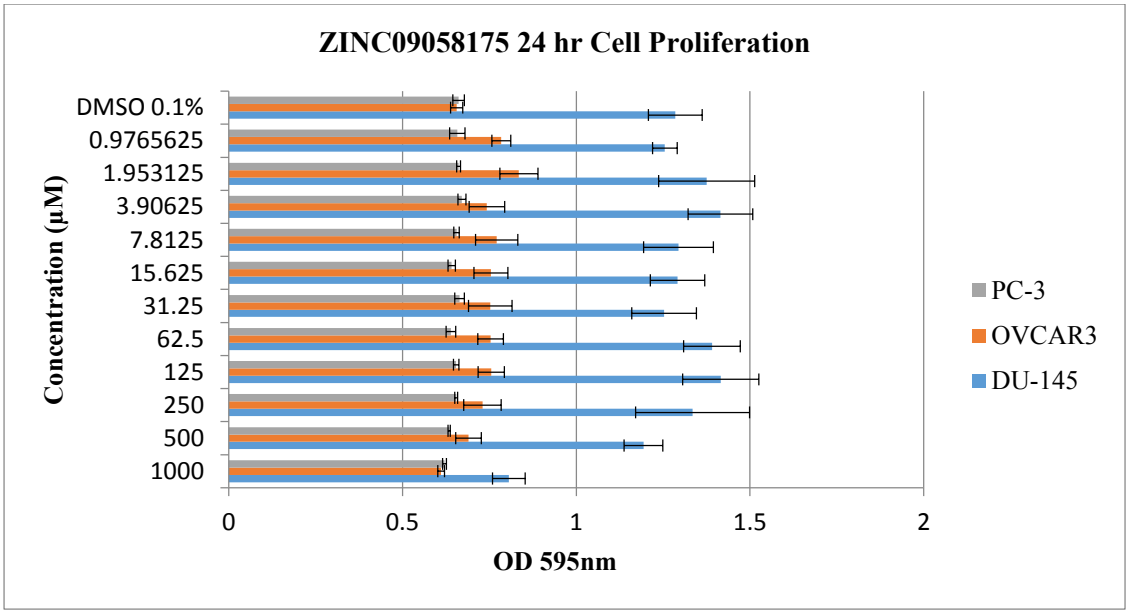

A498 Cell Proliferation 24 hrs
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
DMSO 0.1%
Concentration (µM)
ZINC04276177
Pentamidine

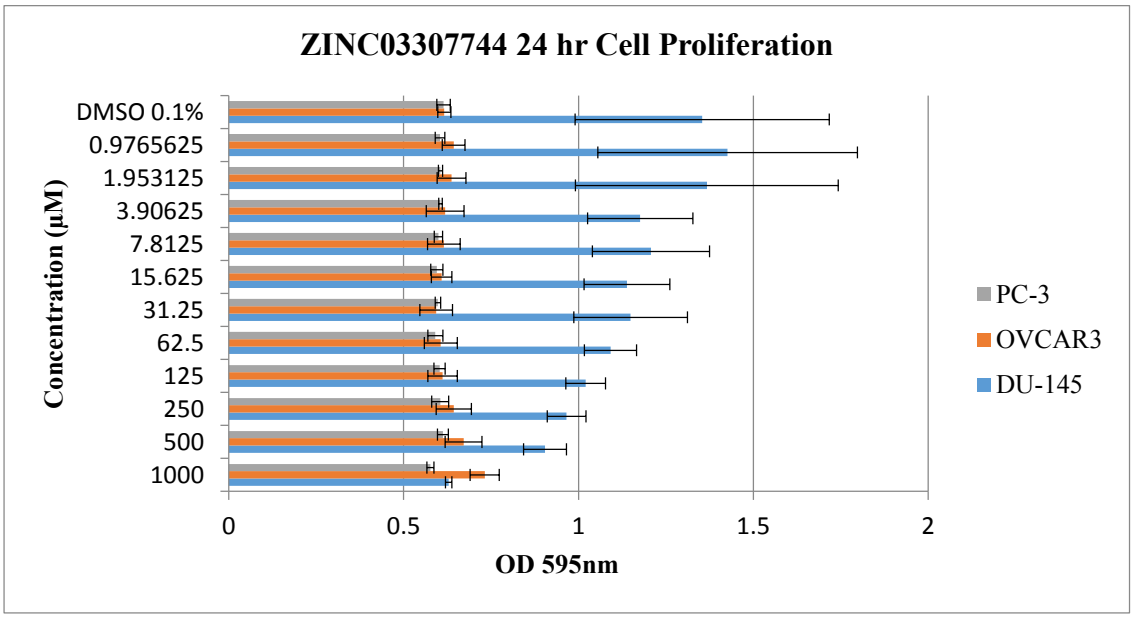


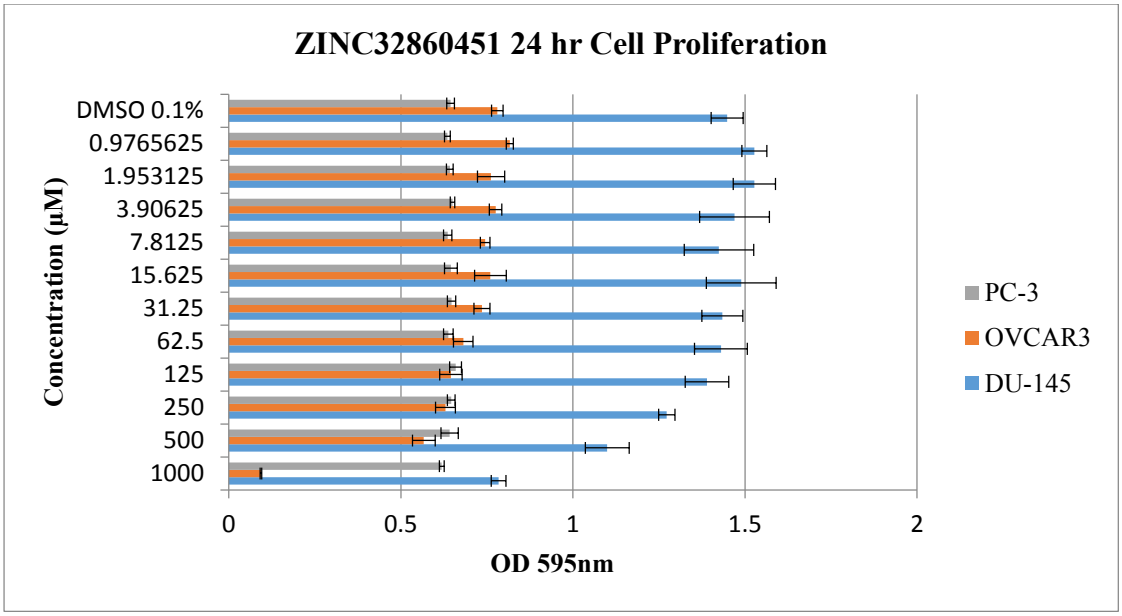

### 6.3.3 Appendix 3.3: Biological Results of ZINC Compounds Identified for QNZ

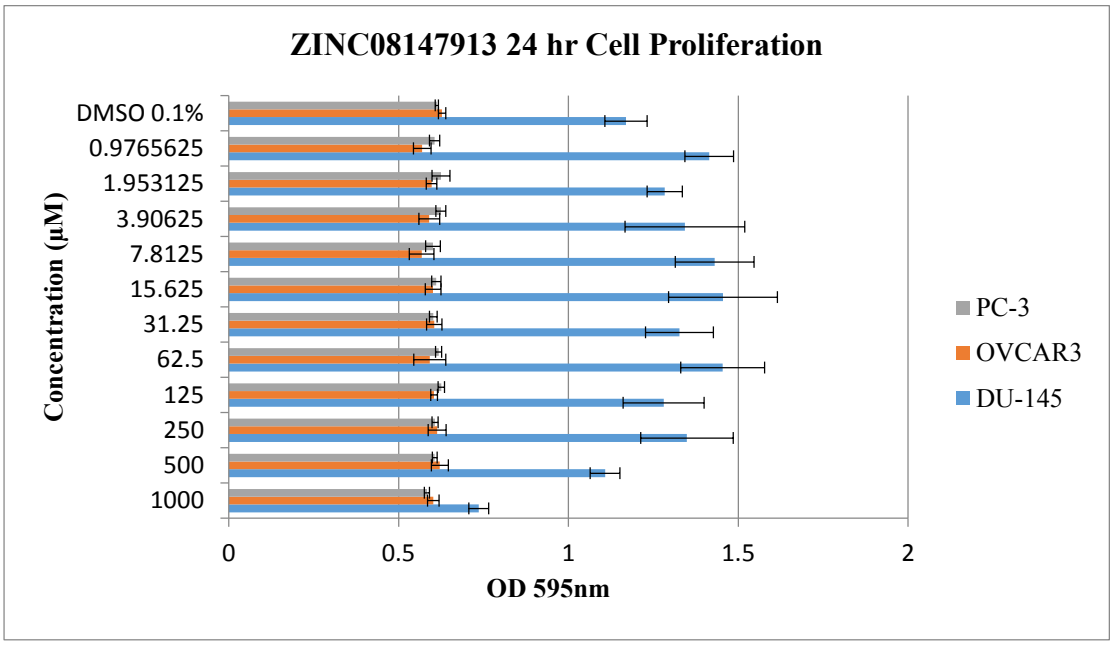
The MTT proliferation graph of ZINC compounds identified from the SBVS performed for QNZ in Chapter 2 are depicted below. The graphs include the proliferation MTT assay, 24 hour post-treatment in DU-145, PC-3 and OVCAR3 cancer cell line. DMSO 0.1% was used as vehicle control. Data shown represents mean ± S.D. of four independent experiments performed in duplicate. OD = absorbance.

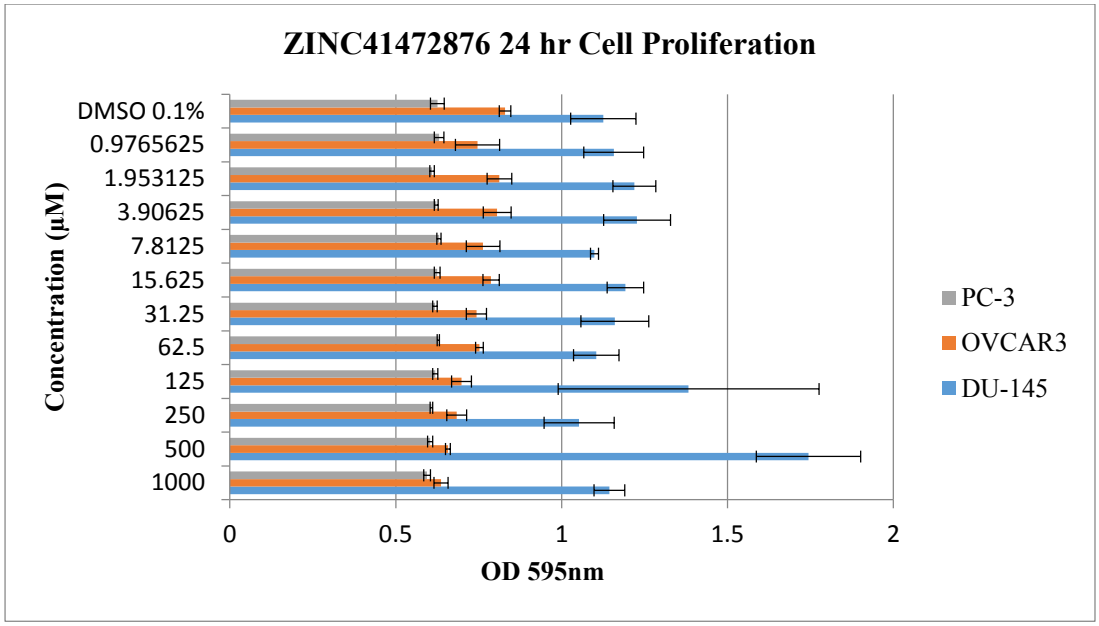


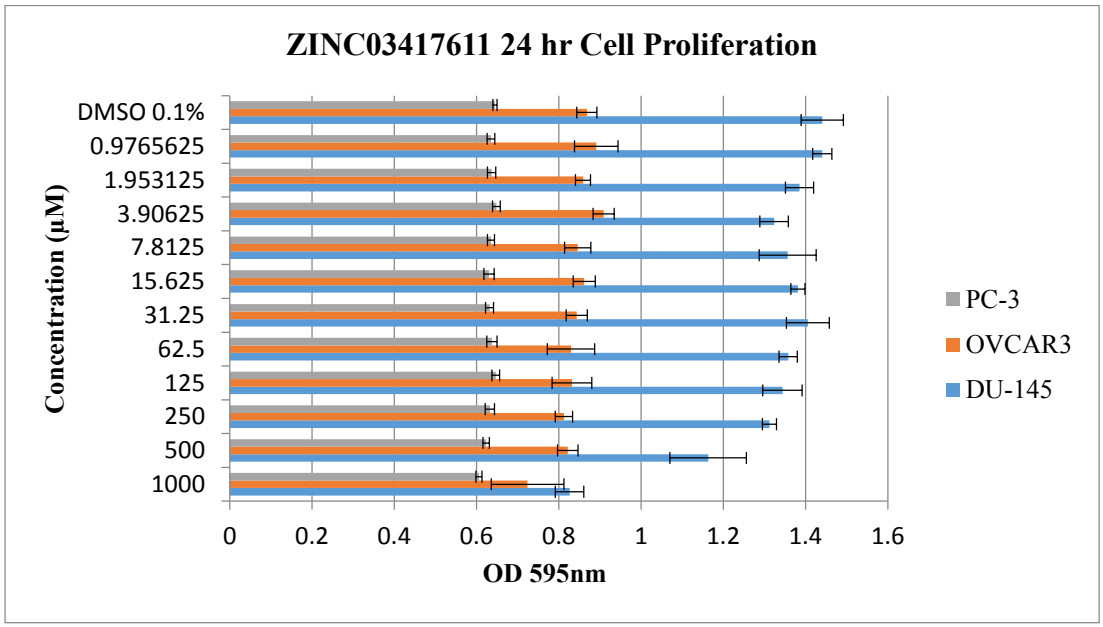

ZINC09058175 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.5
1
1.5
2
OD 595nm
PC-3
OVCAR3
DU-145

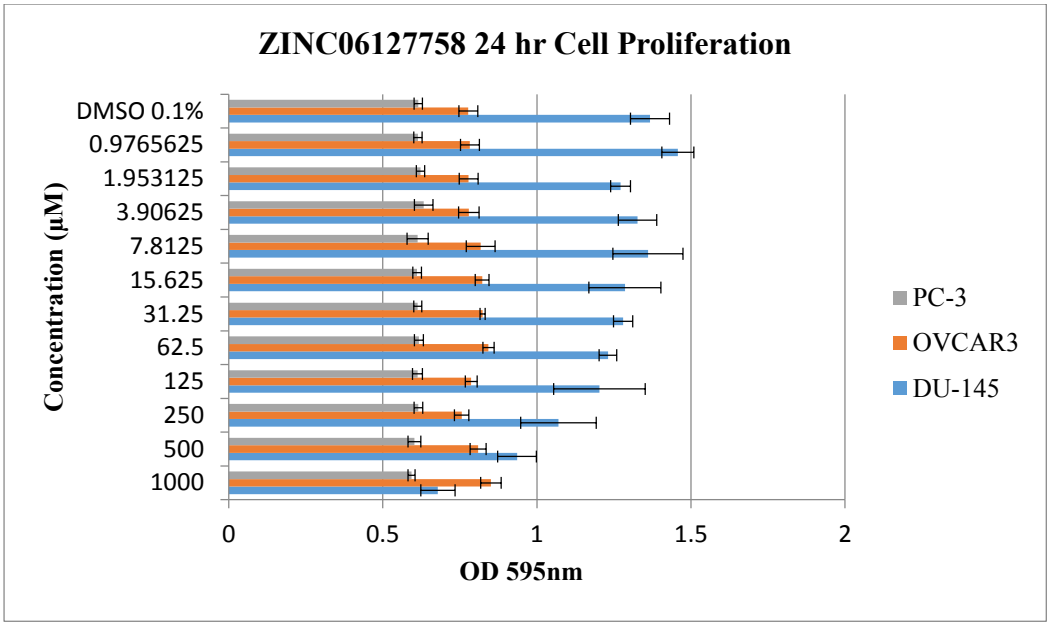

ZINC03307744 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.5
1
1.5
2
OD 595nm
PC-3
OVCAR3
DU-145

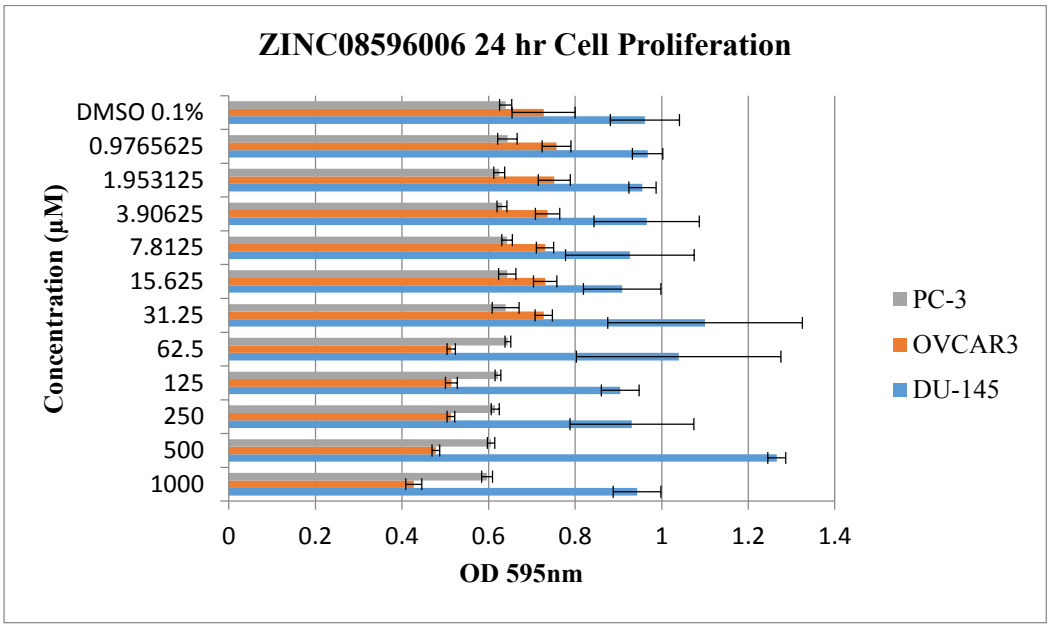

ZINC32860451 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.5
1
1.5
2
OD 595nm
PC-3
OVCAR3
DU-145

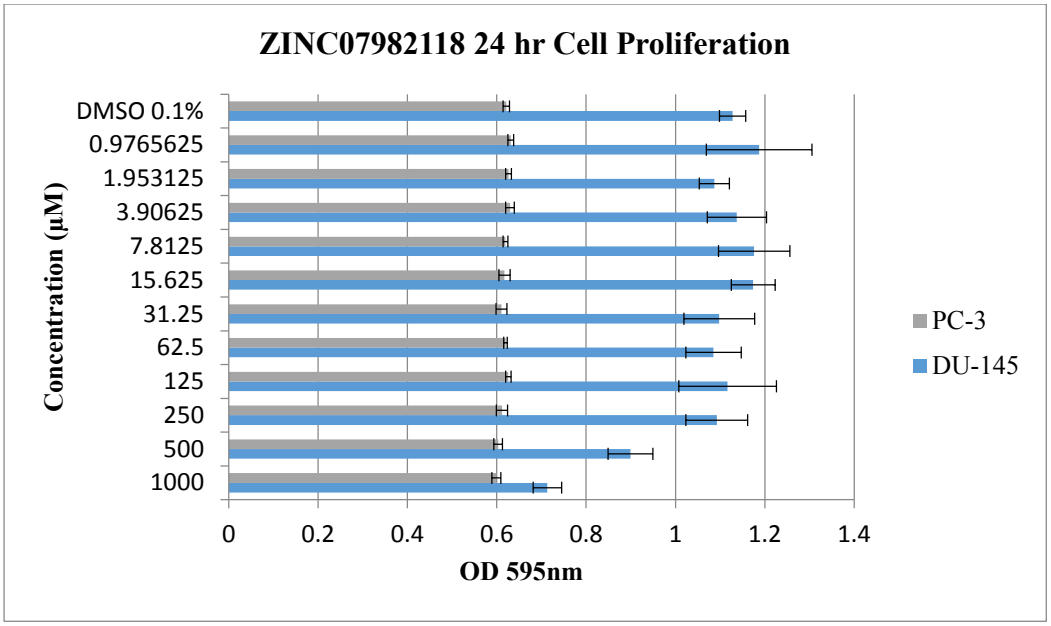

ZINC08147913 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.5
1
1.5
2
OD 595nm
PC-3
OVCAR3
DU-145

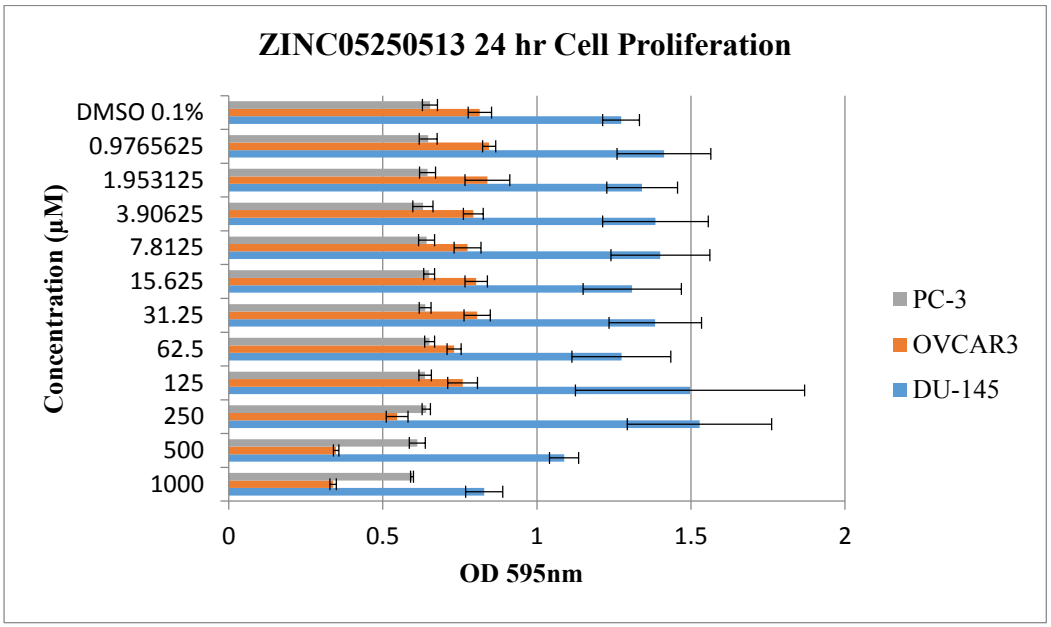

ZINC41472876 24 hr Cell Proliferation
Concentration (µM)
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
0
0.5
1
1.5
2
OD 595nm
PC-3
OVCAR3
DU-145

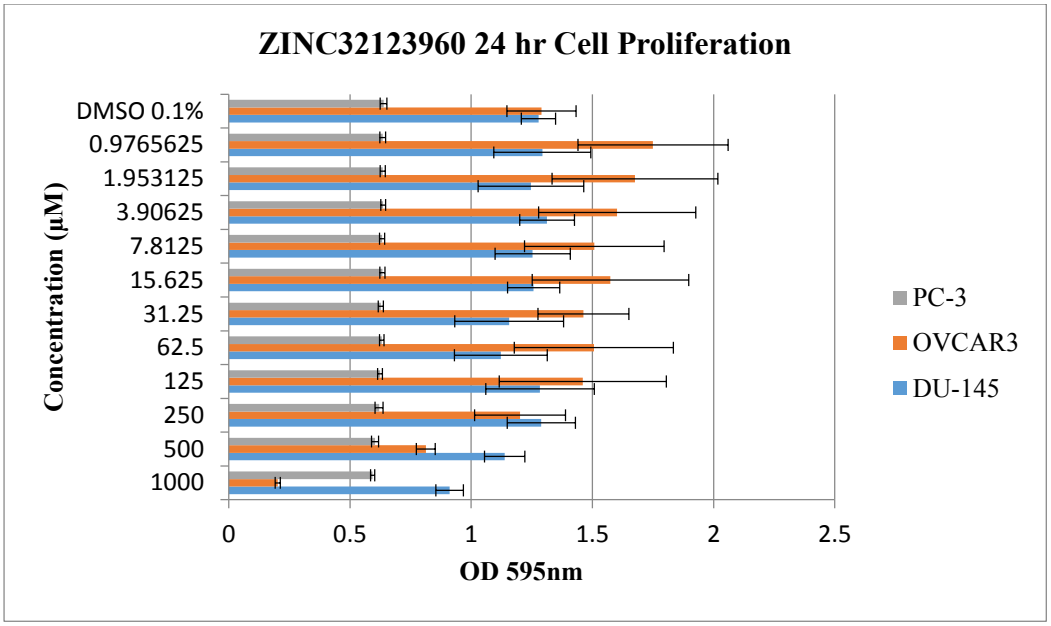

ZINC03417611 24 hr Cell Proliferation
Concentration (µM)
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
OD 595nm
PC-3
OVCAR3
DU-145

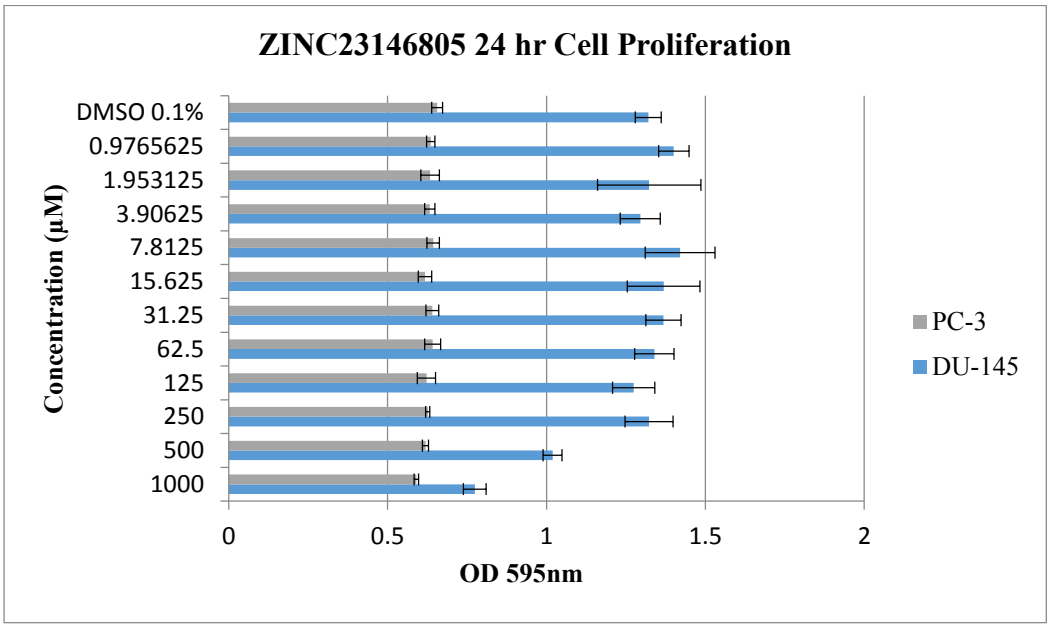

ZINC06127758 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.5
1
1.5
2
OD 595nm
PC-3
OVCAR3
DU-145

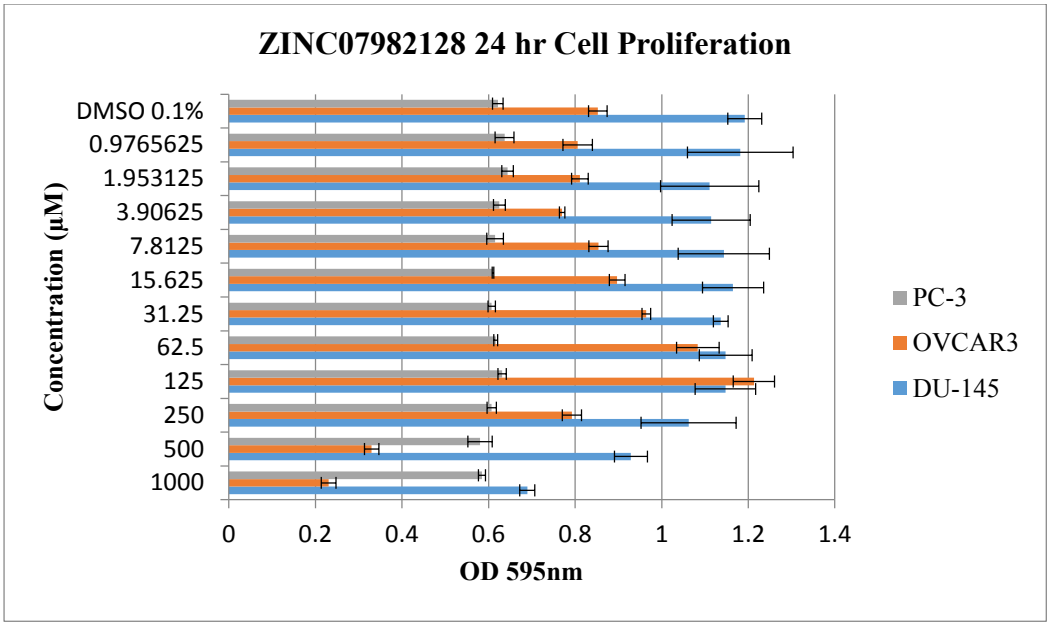

ZINC08596006 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
OD 595nm
PC-3
OVCAR3
DU-145

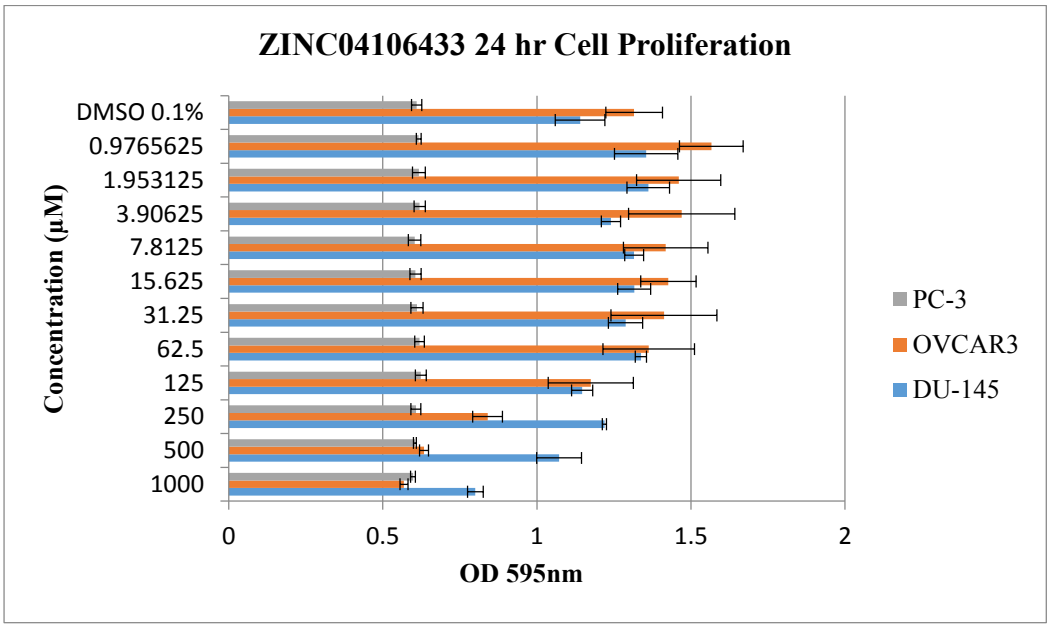

ZINC07982118 24 hr Cell Proliferation
Concentration (µM)
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
0
0.2
0.4
0.6
0.8
1
1.2
1.4
OD 595nm
PC-3
DU-145

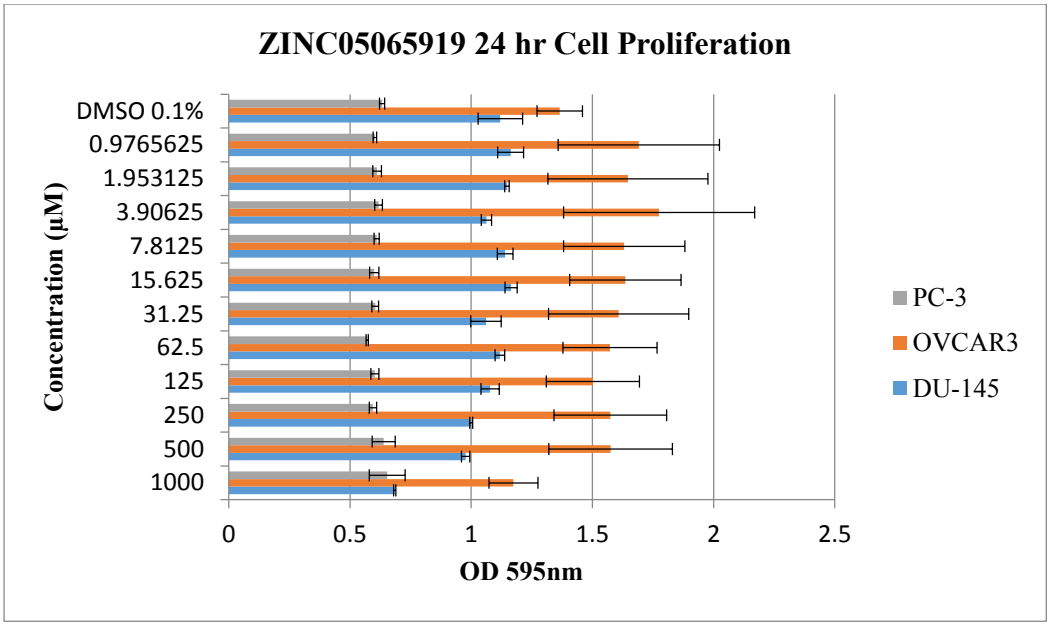

ZINC05250513 24 hr Cell Proliferation
Concentration (µM)
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
0
0.5
1
1.5
2
OD 595nm
PC-3
OVCAR3
DU-145

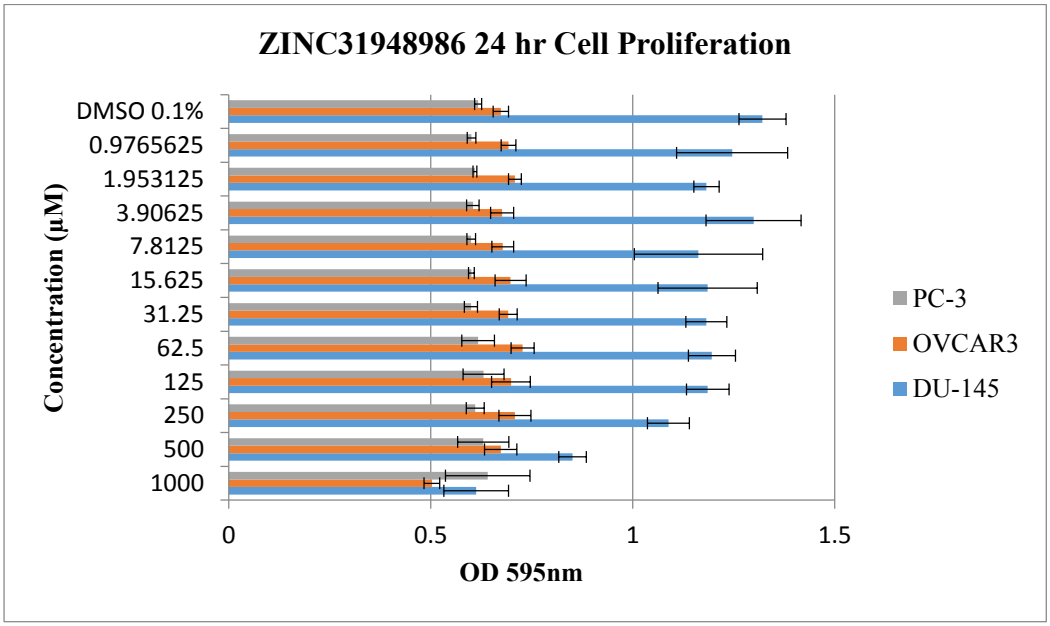

ZINC32123960 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.5
1
1.5
2
2.5
OD 595nm
PC-3
OVCAR3
DU-145

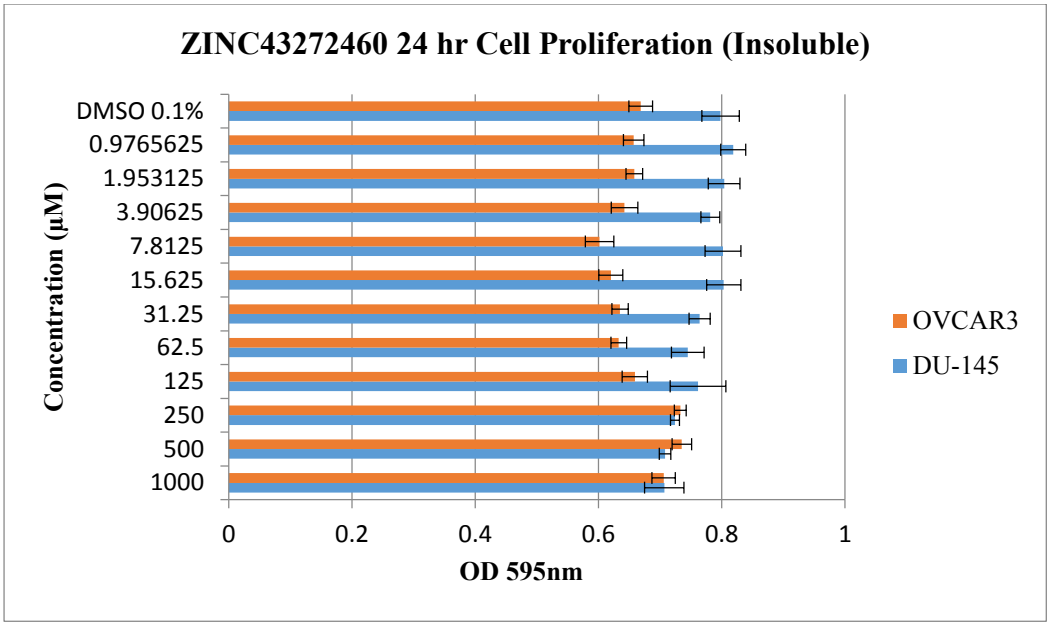

ZINC23146805 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.5
1
1.5
2
OD 595nm
PC-3
DU-145

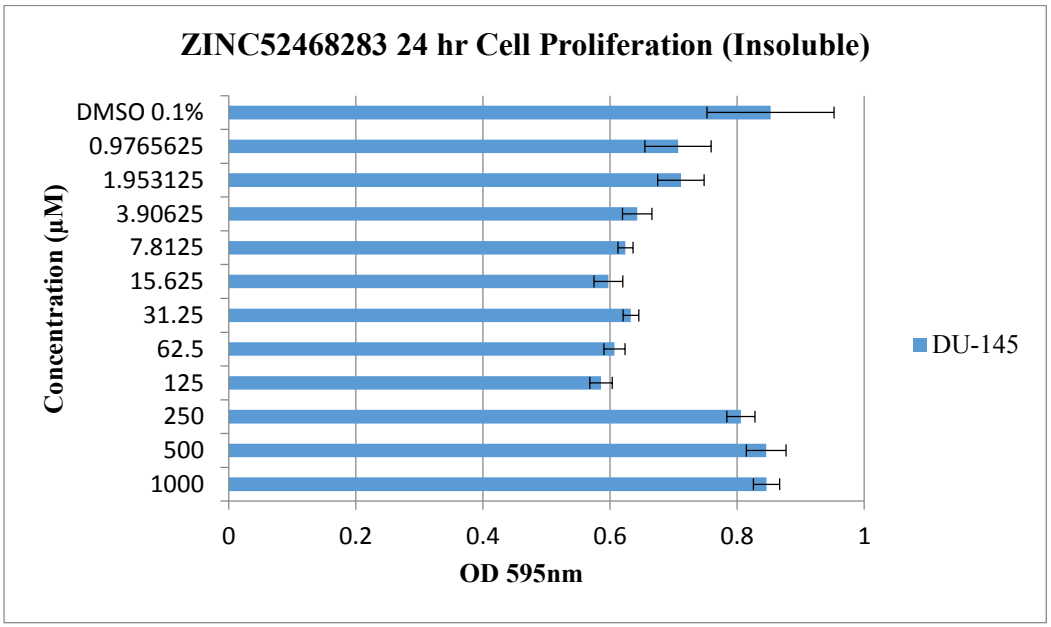

ZINC07982128 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
OD 595nm
PC-3
OVCAR3
DU-145

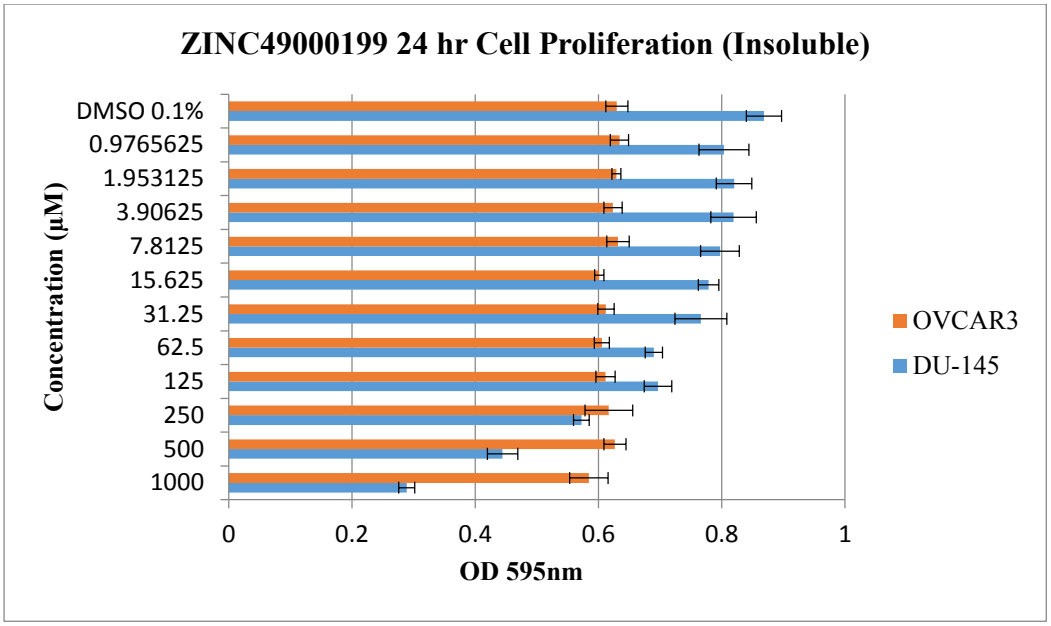

ZINC04106433 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.5
1
1.5
2
OD 595nm
PC-3
OVCAR3
DU-145

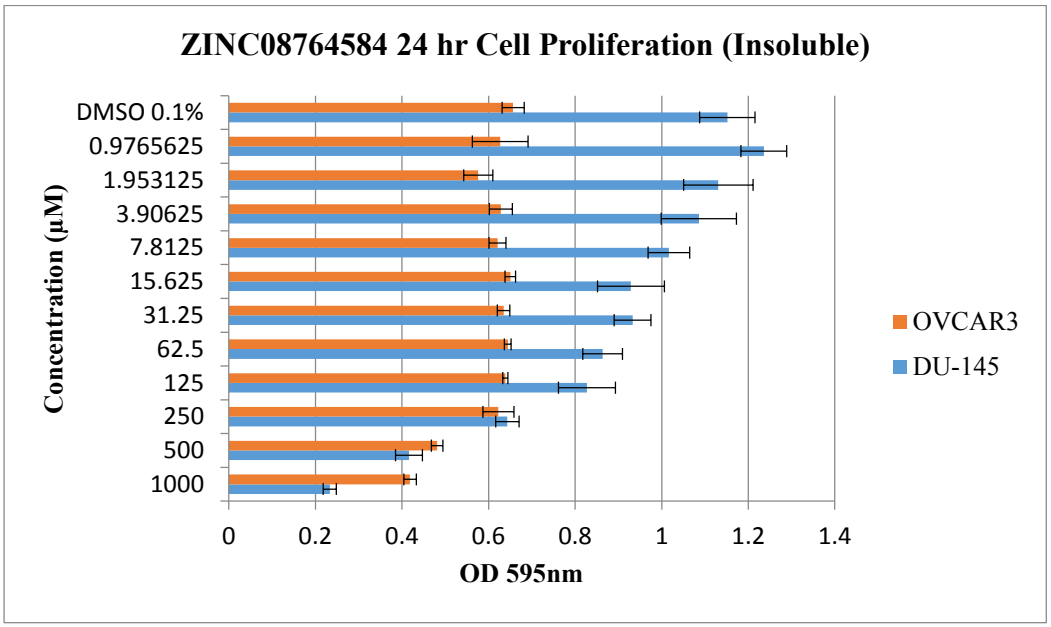

ZINC05065919 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.5
1
1.5
2
2.5
OD 595nm
PC-3
OVCAR3
DU-145


ZINC31948986 24 hr Cell Proliferation
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.5
1
1.5
OD 595nm
PC-3
OVCAR3
DU-145


ZINC43272460 24 hr Cell Proliferation (Insoluble)
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.2
0.4
0.6
0.8
1
OD 595nm
OVCAR3
DU-145

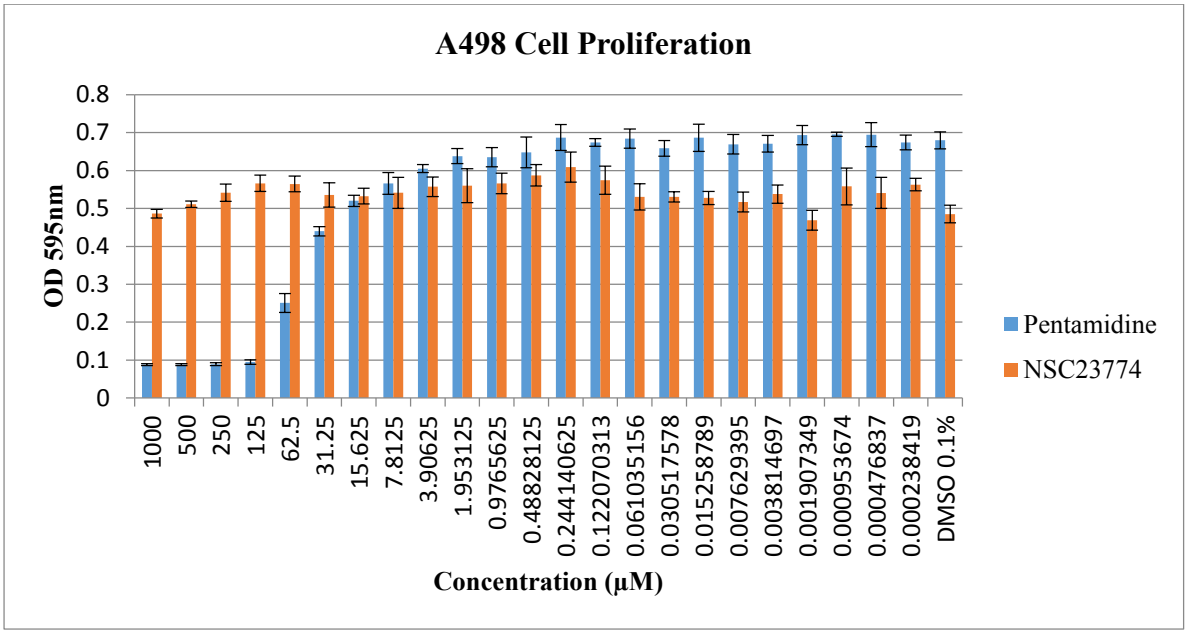

ZINC52468283 24 hr Cell Proliferation (Insoluble)
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
Concentration (µM)
0
0.2
0.4
0.6
0.8
1
OD 595nm
DU-145

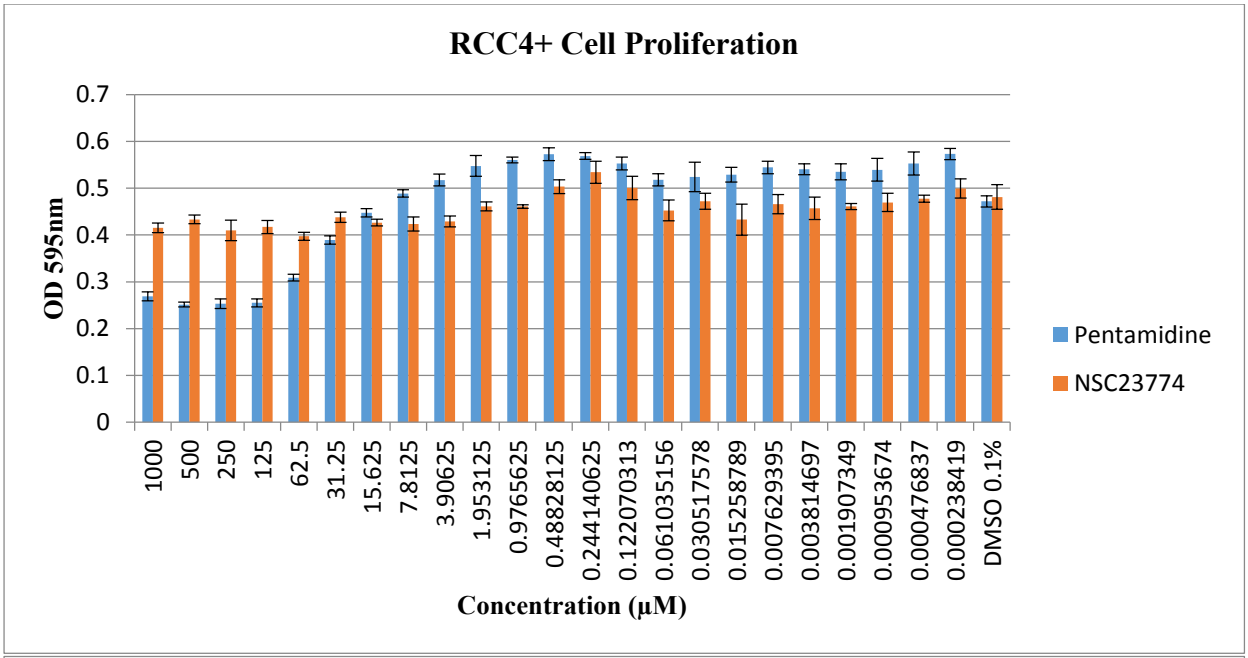

ZINC49000199 24 hr Cell Proliferation (Insoluble)
Concentration (µM)
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
0
0.2
0.4
0.6
0.8
1
OD 595nm
OVCAR3
DU-145

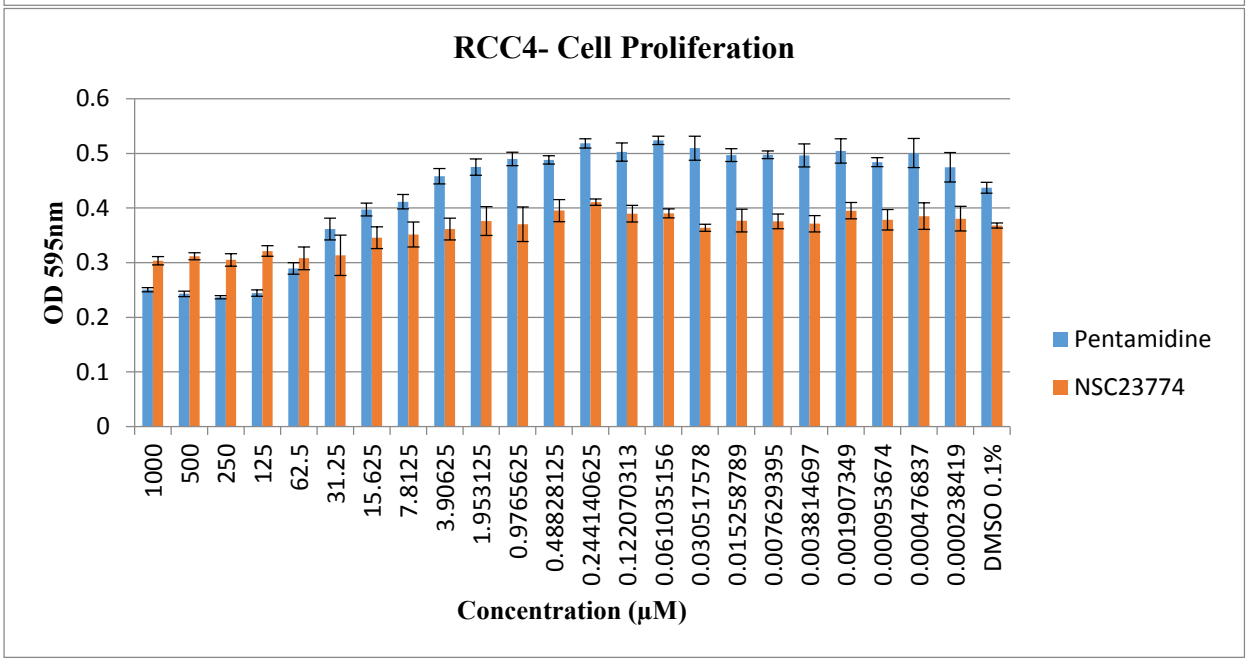

ZINC08764584 24 hr Cell Proliferation (Insoluble)
Concentration (µM)
DMSO 0.1%
0.9765625
1.953125
3.90625
7.8125
15.625
31.25
62.5
125
250
500
1000
0
0.2
0.4
0.6
0.8
1
1.2
1.4
OD 595nm
OVCAR3
DU-145

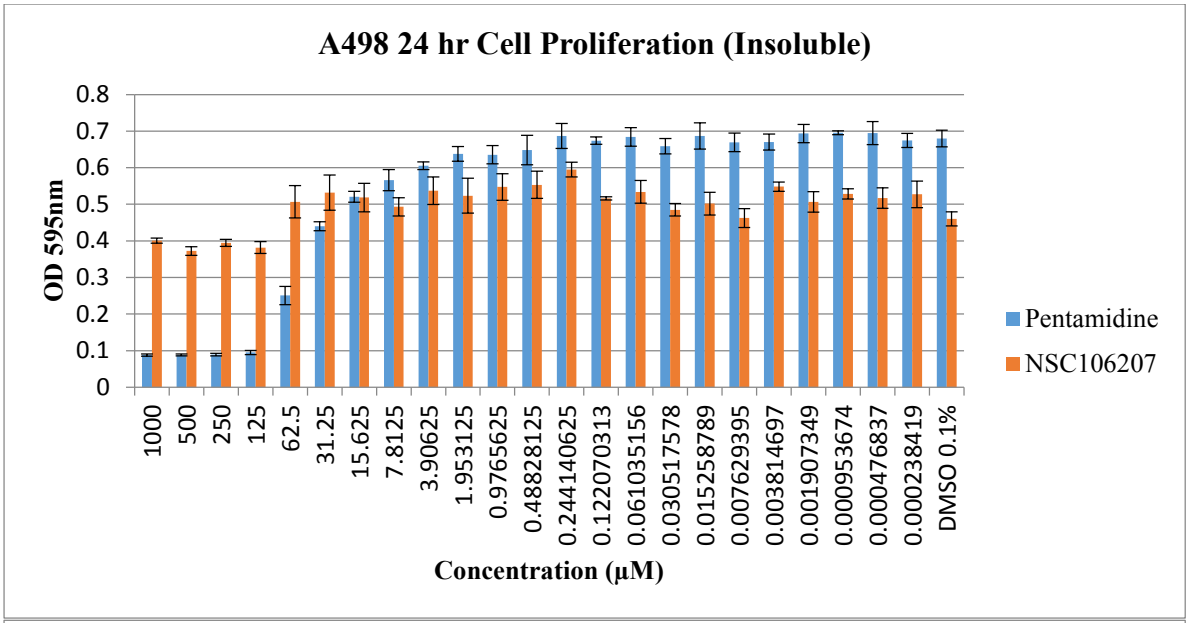
### 6.3.4 Appendix 3.4: Biological Results of NCI Compounds identified from Pentamidine Substructure Search

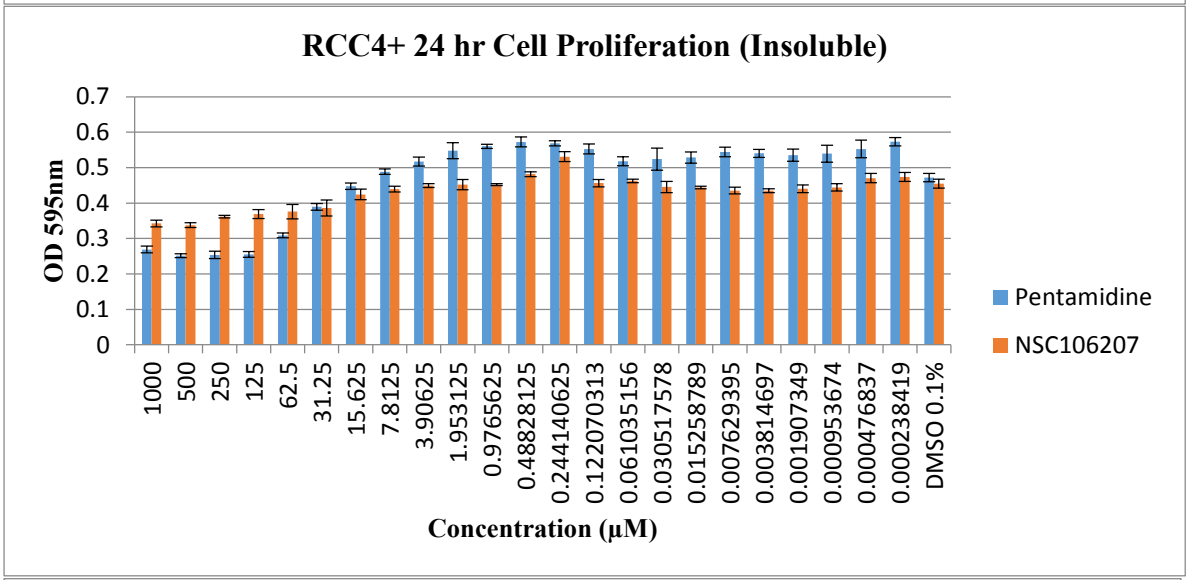
The MTT proliferation graph of NCI compounds identified from the sub-structure search performed for Pentamidine in Chapter 3. The graphs below include the proliferation MTT assay, 24 hour post-treatment in A498, RCC4+ and RCC4- cancer cell line. DMSO 0.1% was used as vehicle control. Data shown represents the mean ± S.D. of four independent experiments performed in duplicate. OD = absorbance.

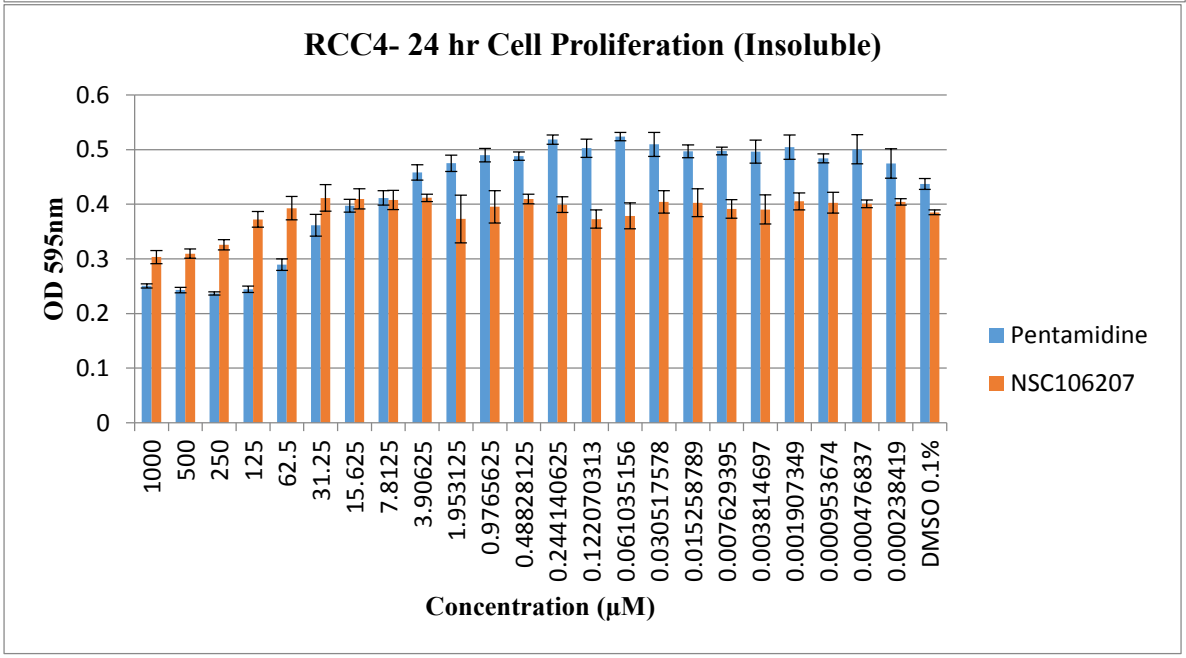


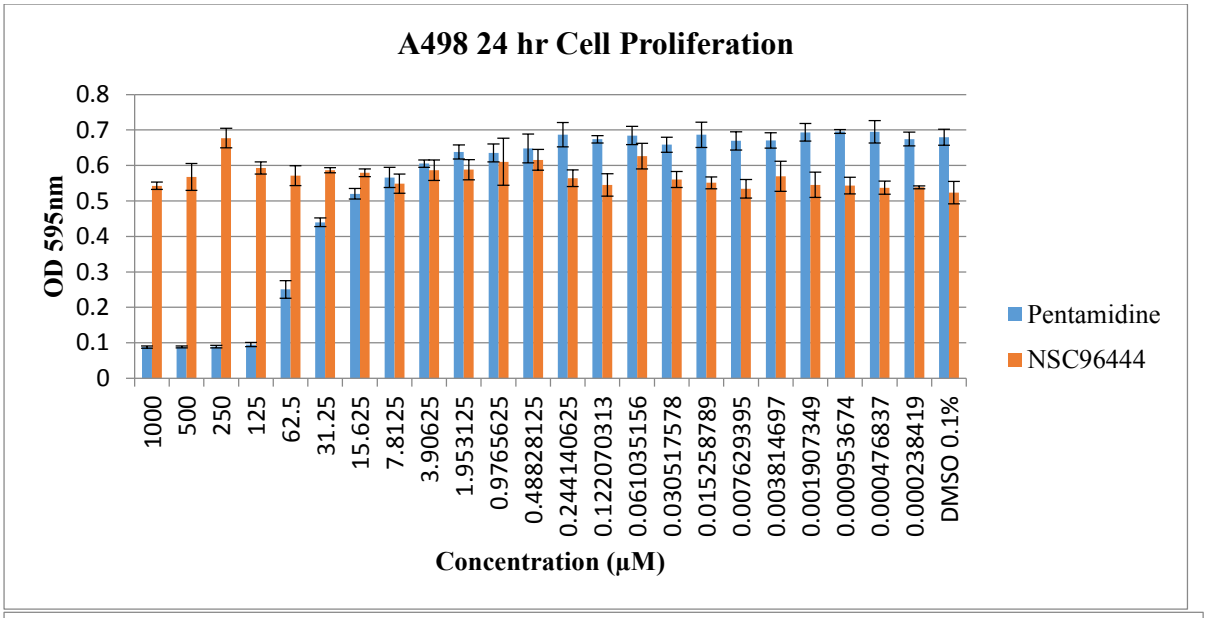

RCC4+ Cell Proliferation
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC23774

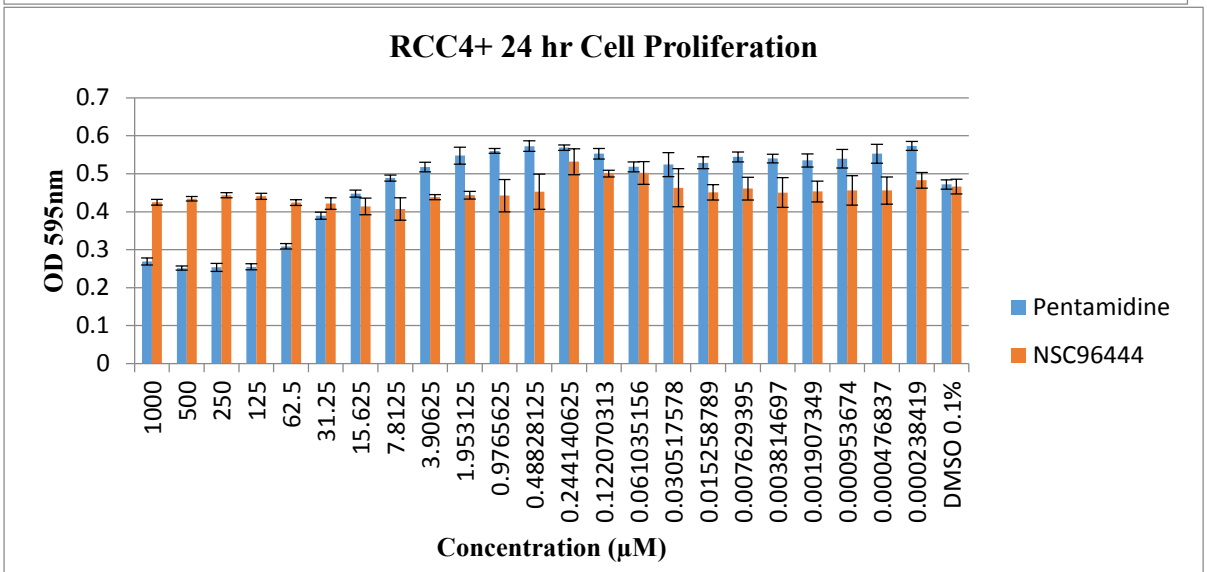

RCC4- Cell Proliferation
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC23774

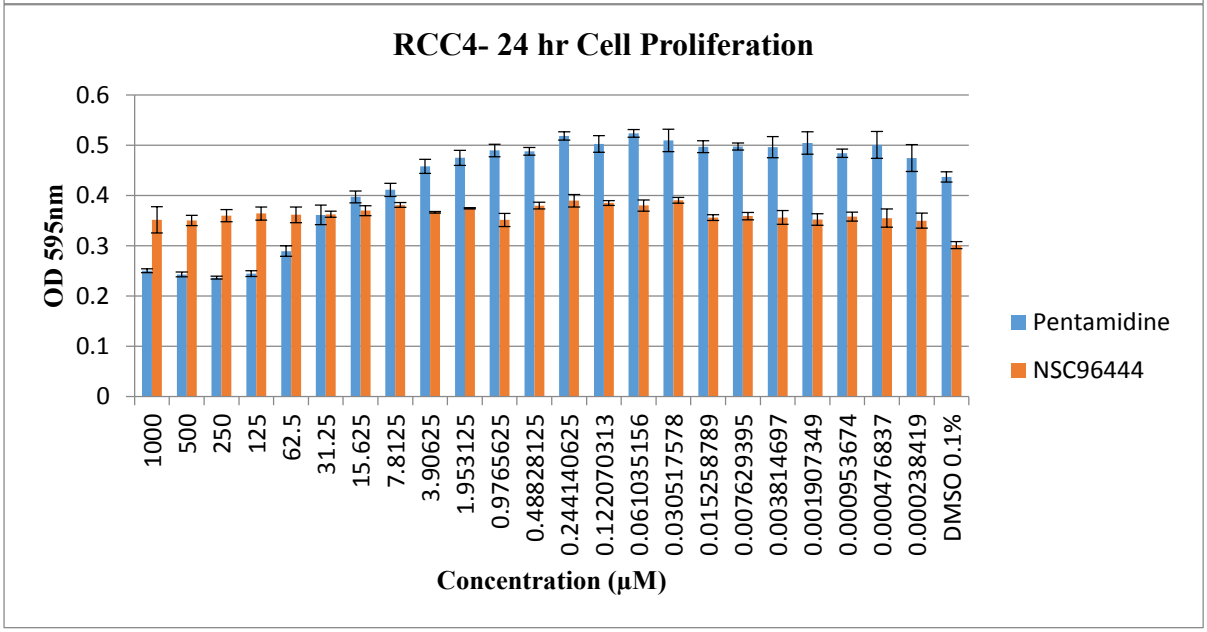

A498 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC106207

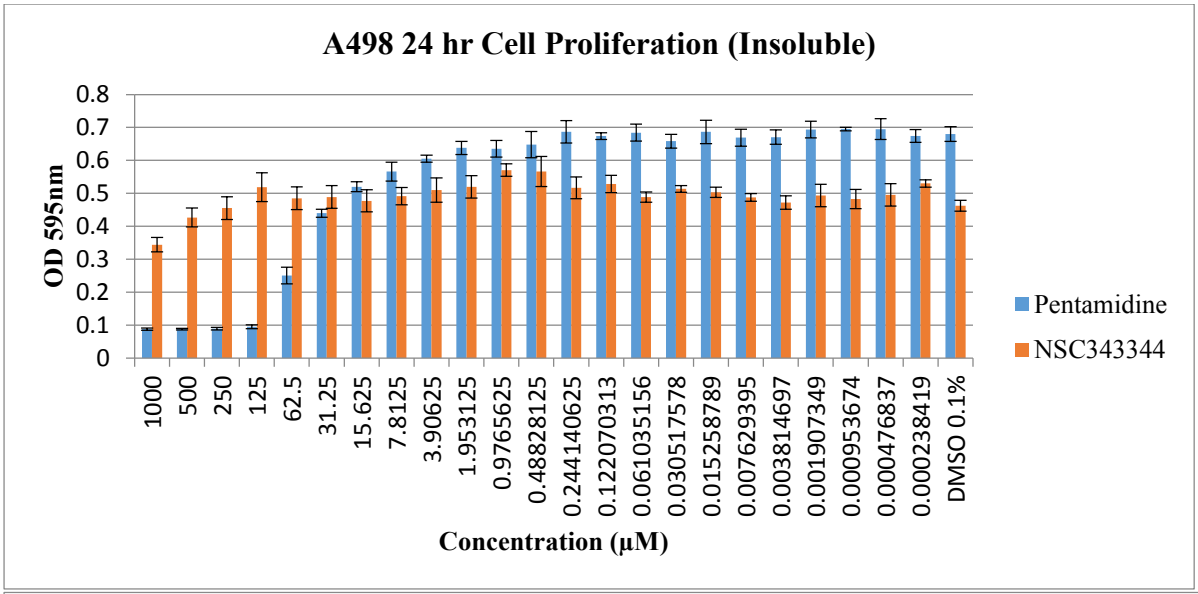

RCC4+ 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC106207

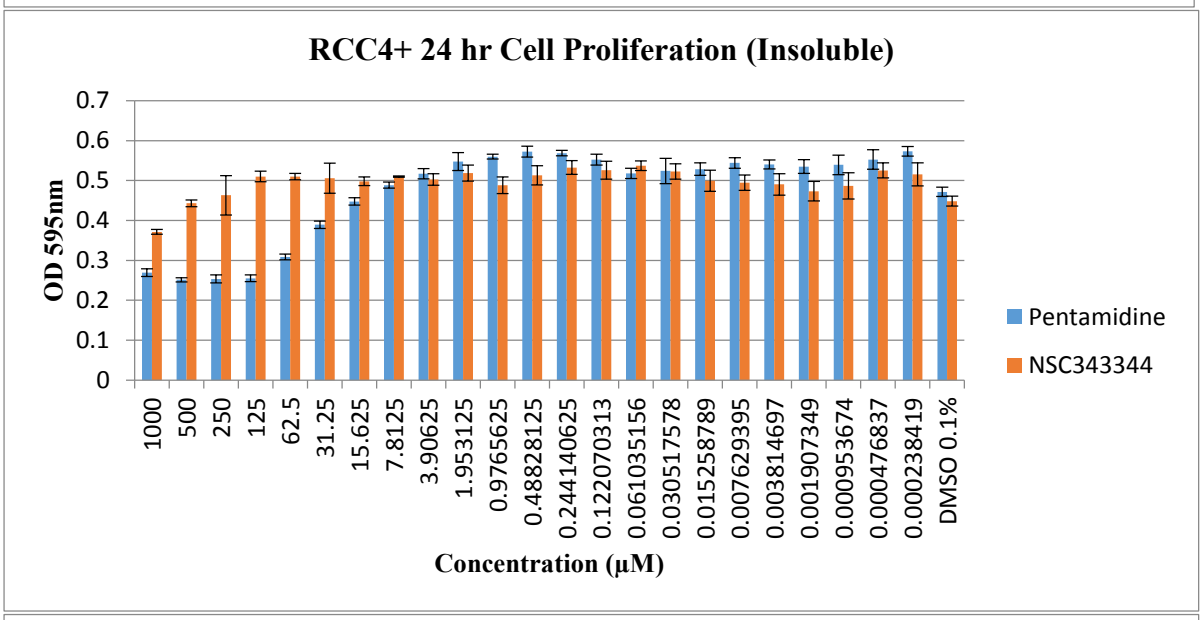

RCC4- 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC106207

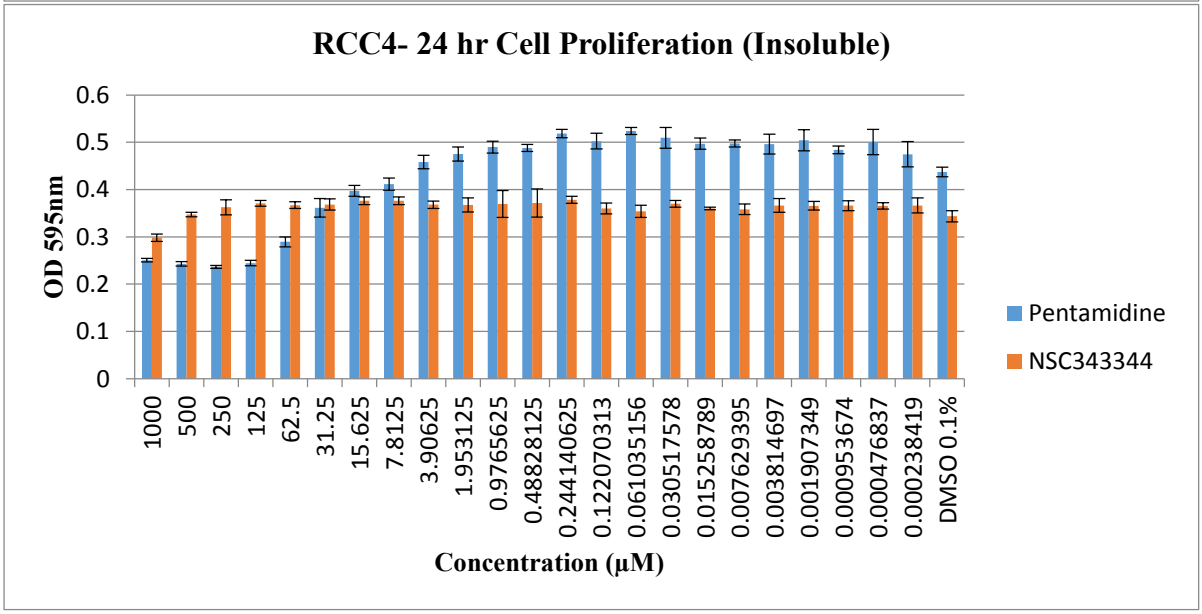

A498 24 hr Cell Proliferation
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC96444

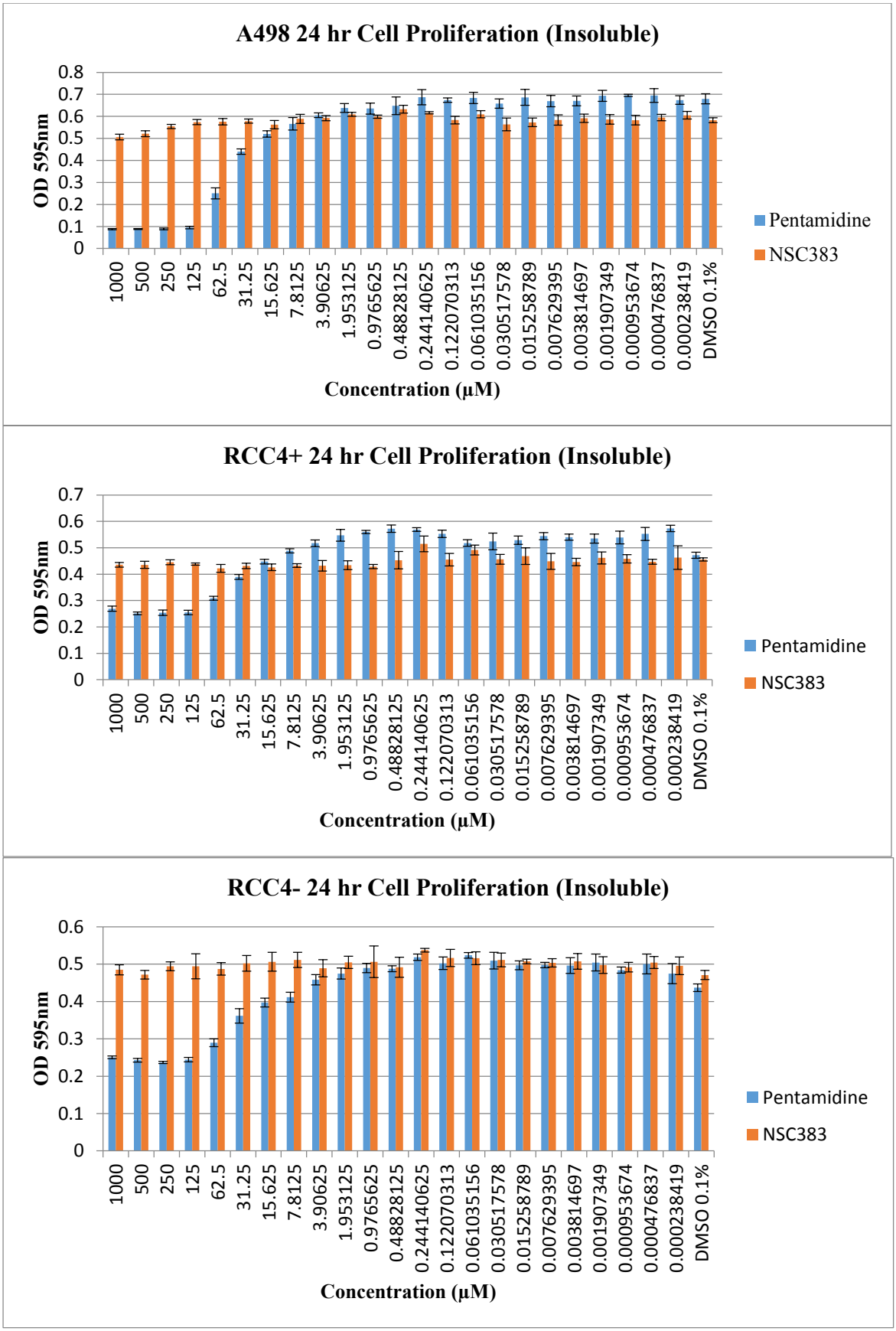

RCC4+ 24 hr Cell Proliferation
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC96444

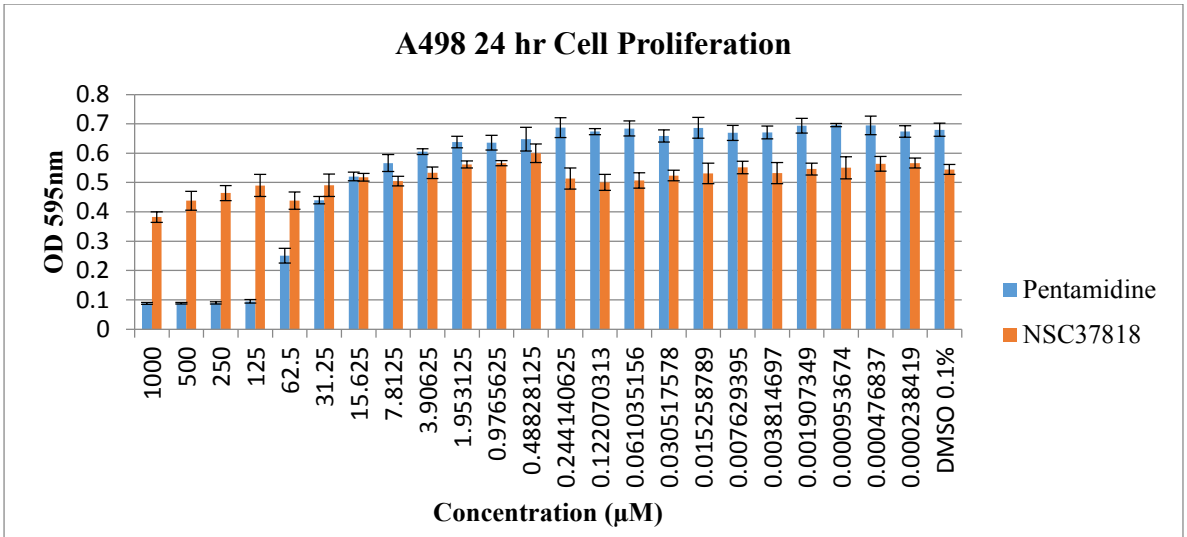

RCC4- 24 hr Cell Proliferation
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC96444

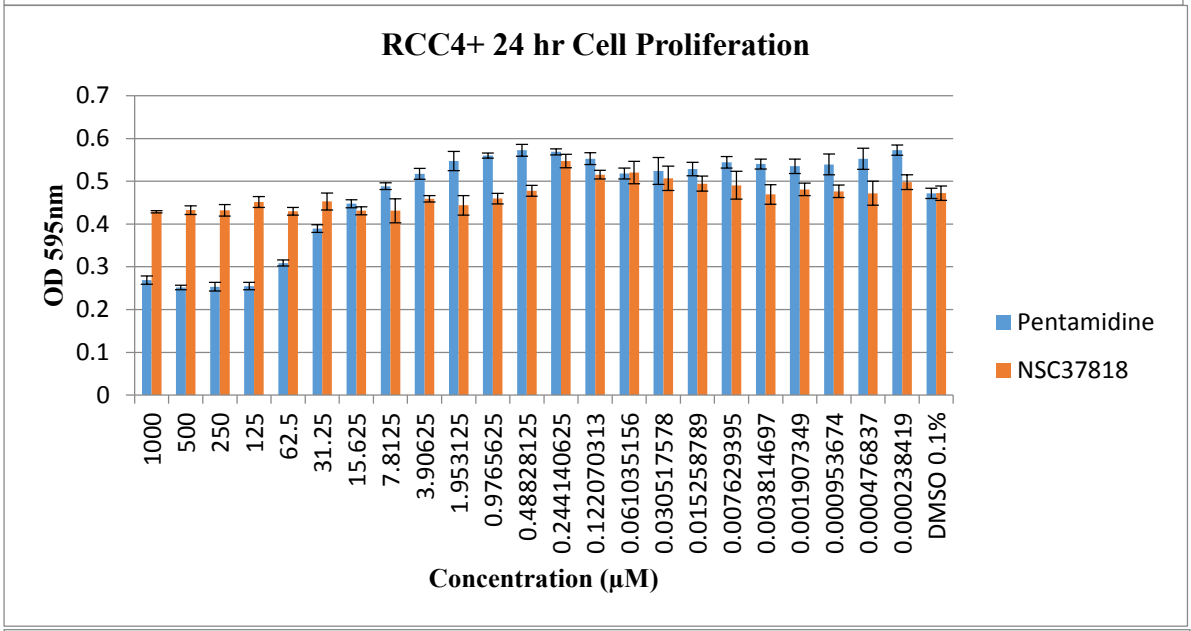

A498 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC343344

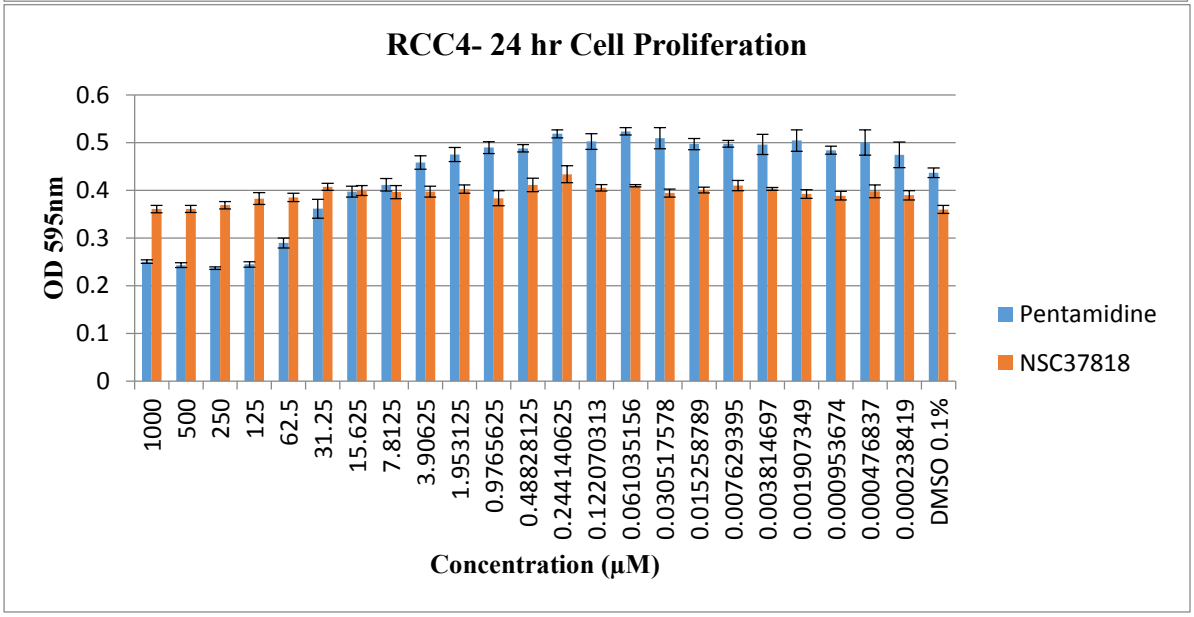

RCC4+ 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC343344

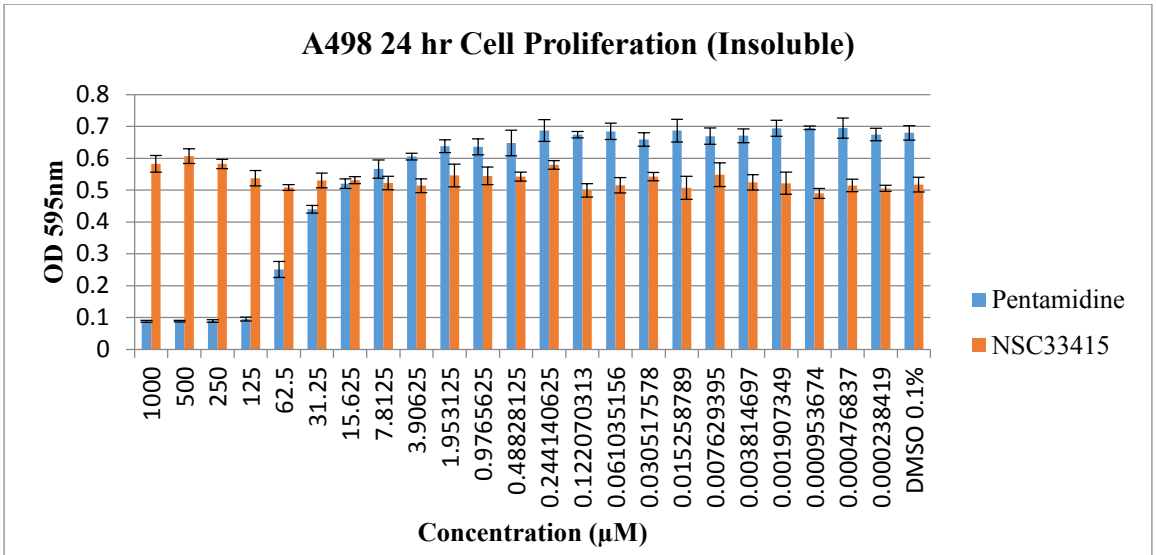

RCC4- 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC343344

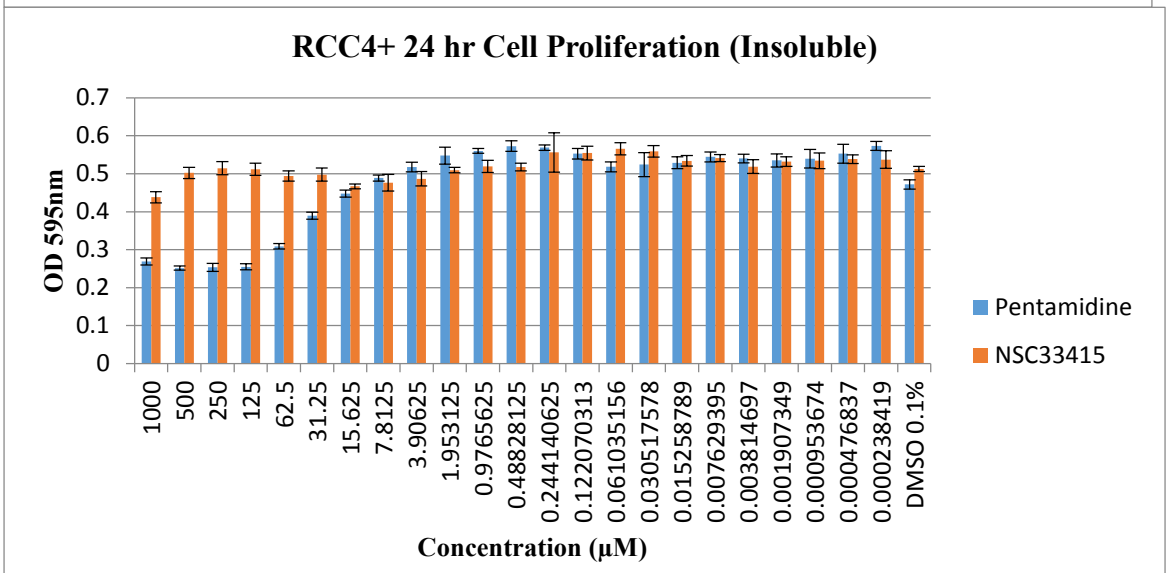

A498 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC383
RCC4+ 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC383
RCC4- 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC383

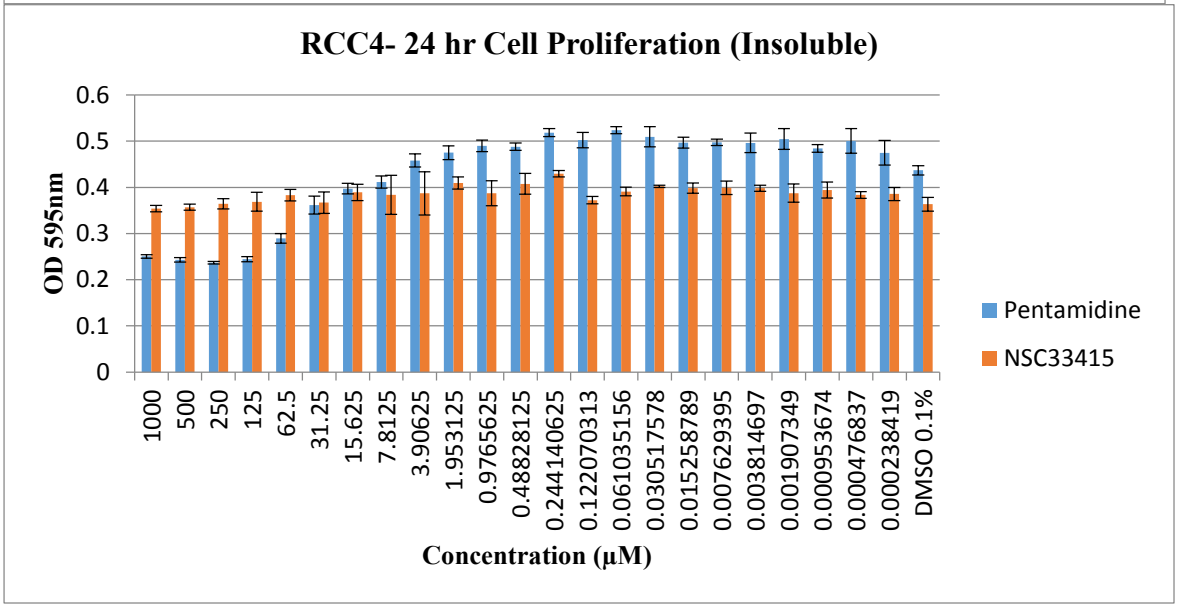

A498 24 hr Cell Proliferation
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC37818

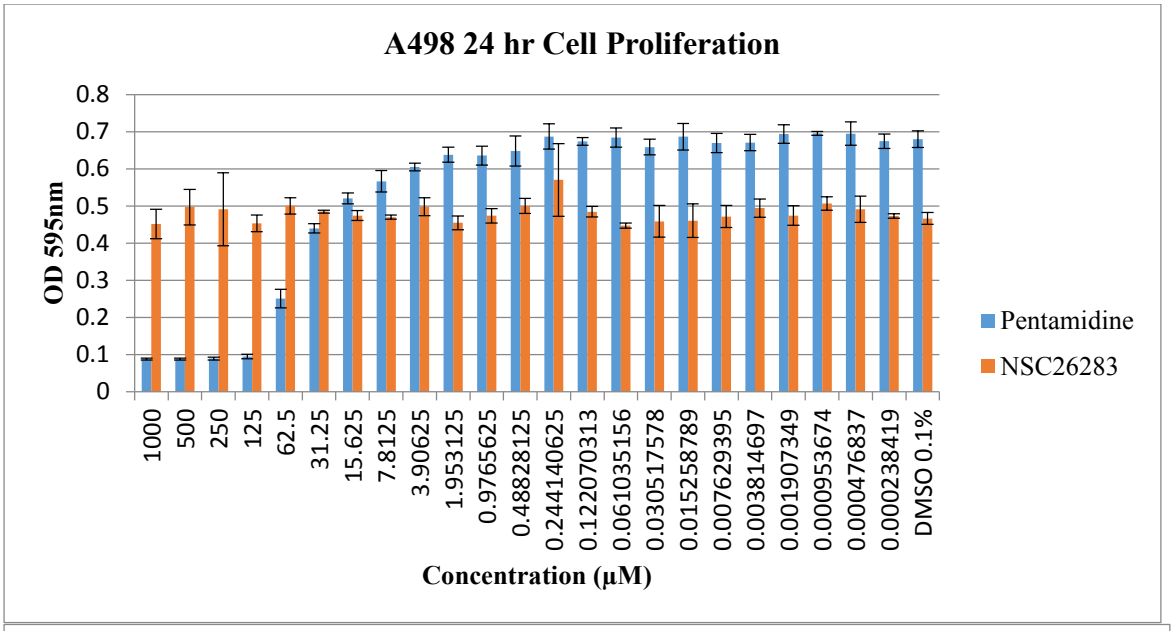

RCC4+ 24 hr Cell Proliferation
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC37818

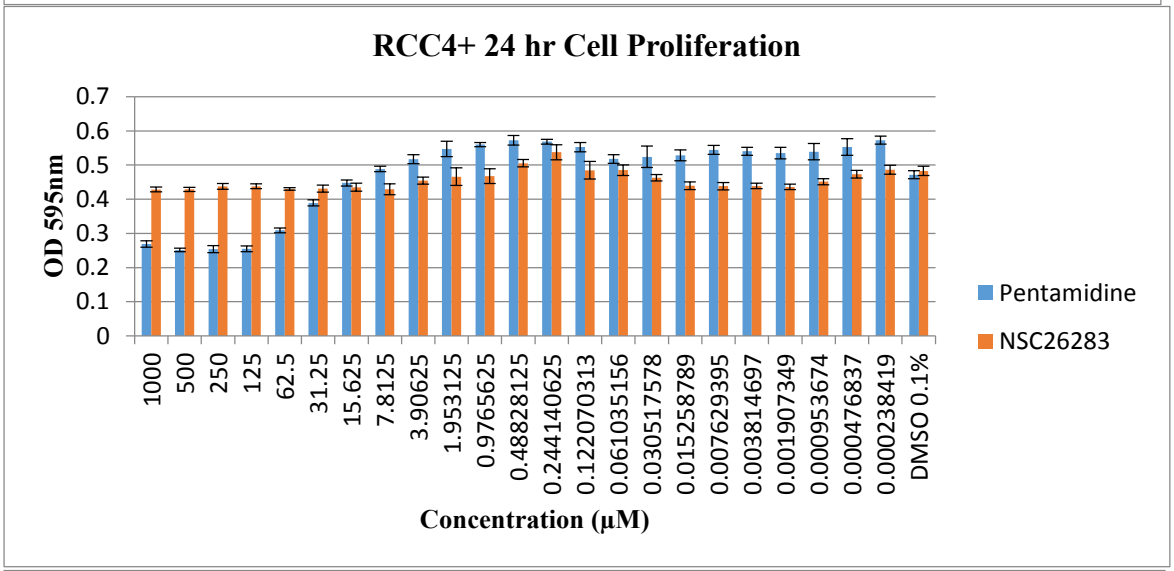

RCC4- 24 hr Cell Proliferation
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC37818

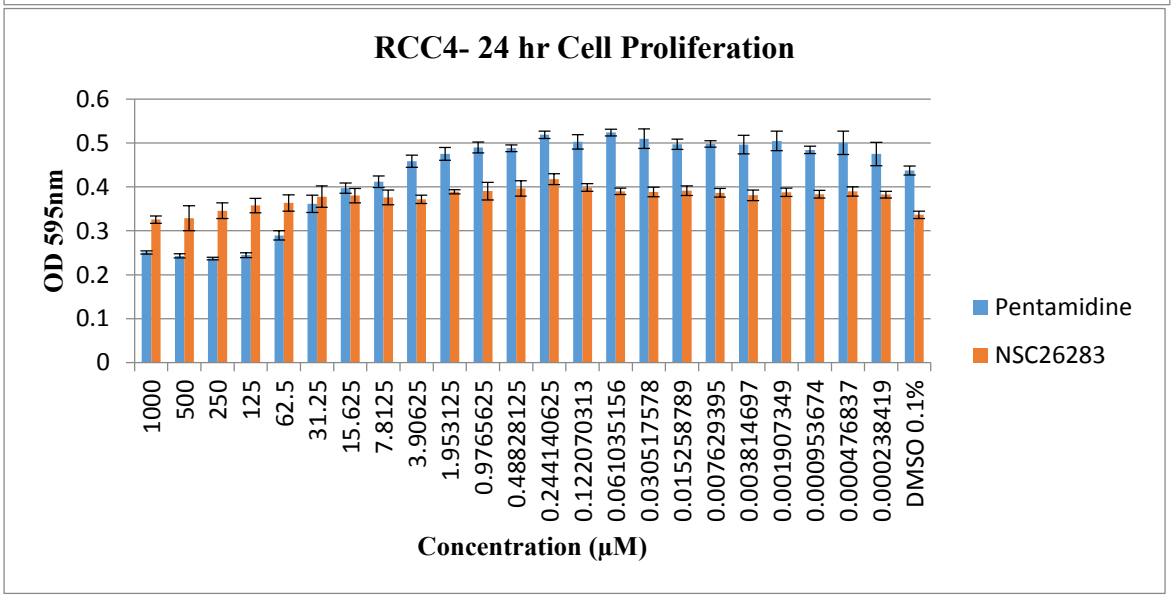

A498 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC33415

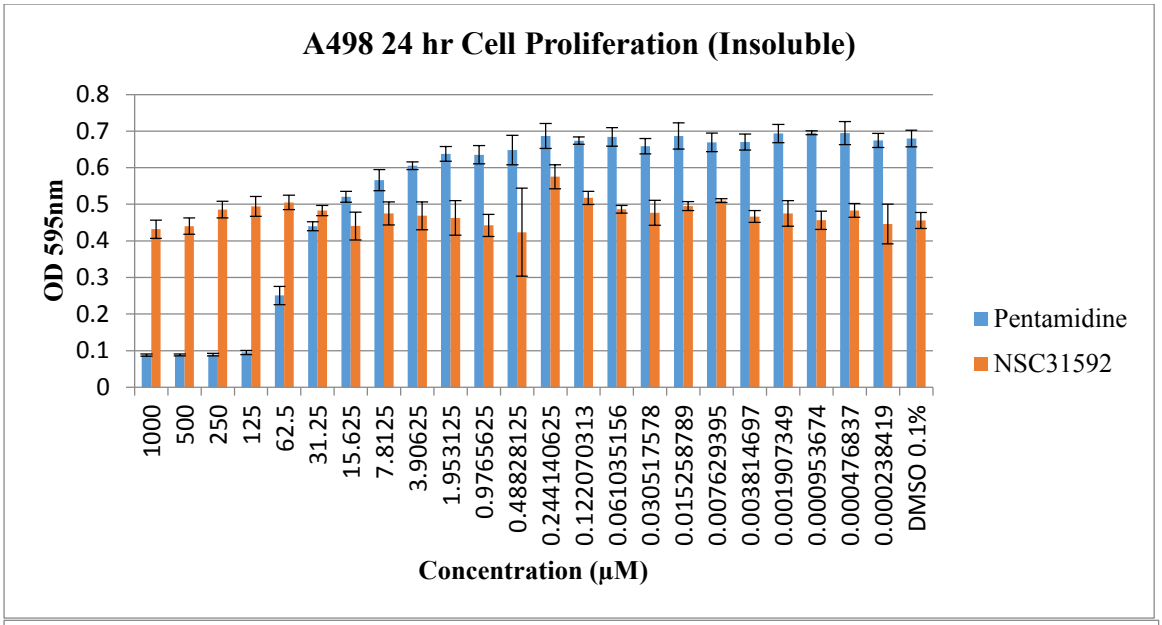

RCC4+ 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC33415

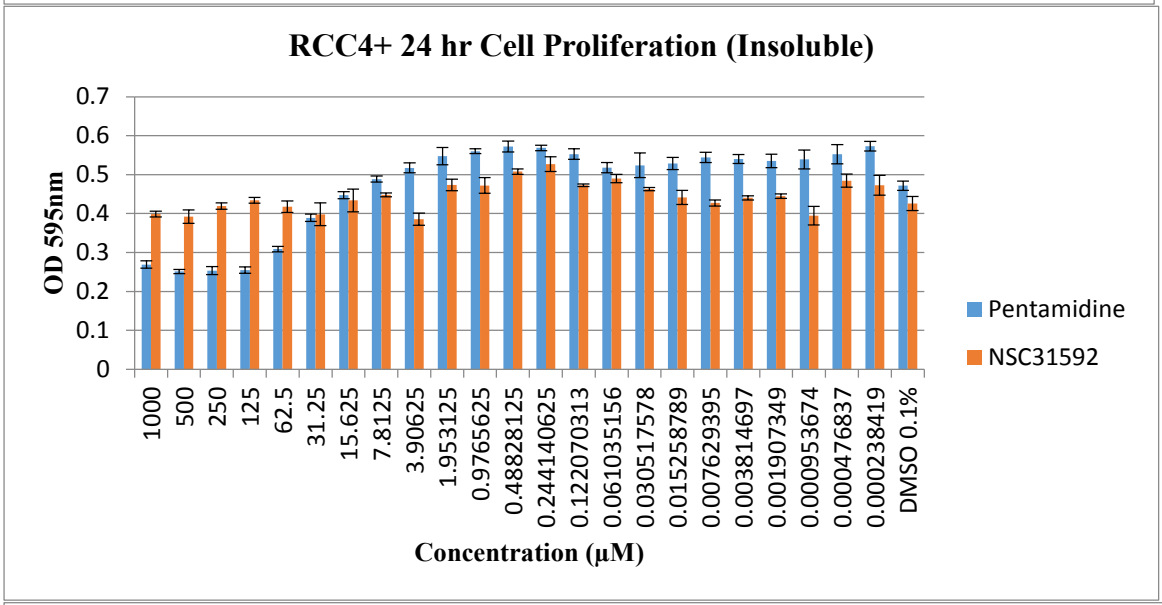

RCC4- 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC33415

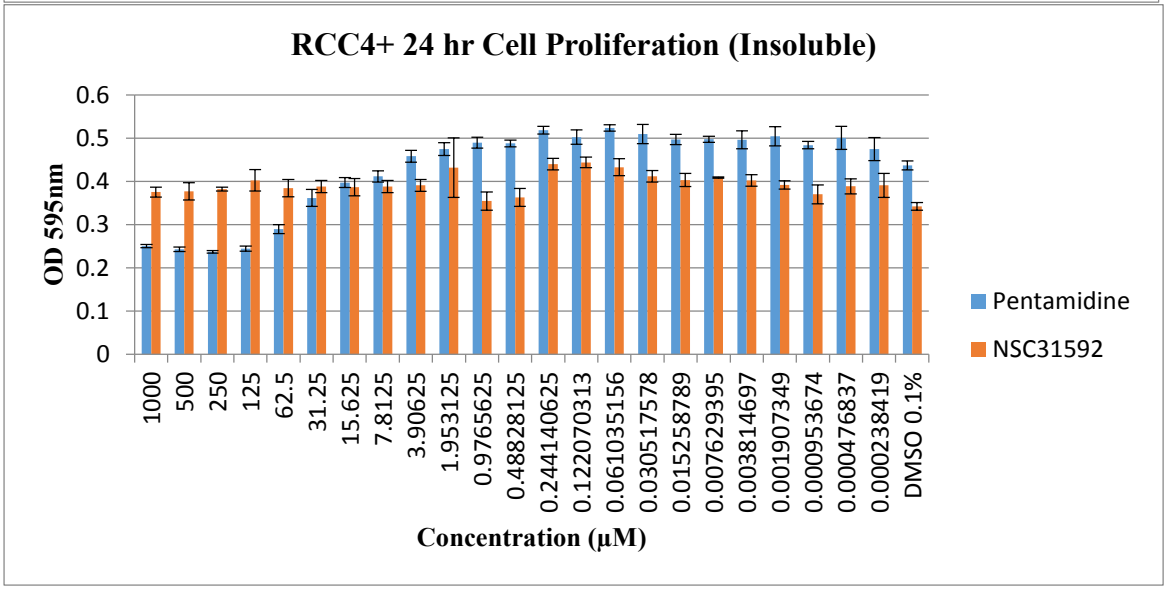

A498 24 hr Cell Proliferation
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC26283

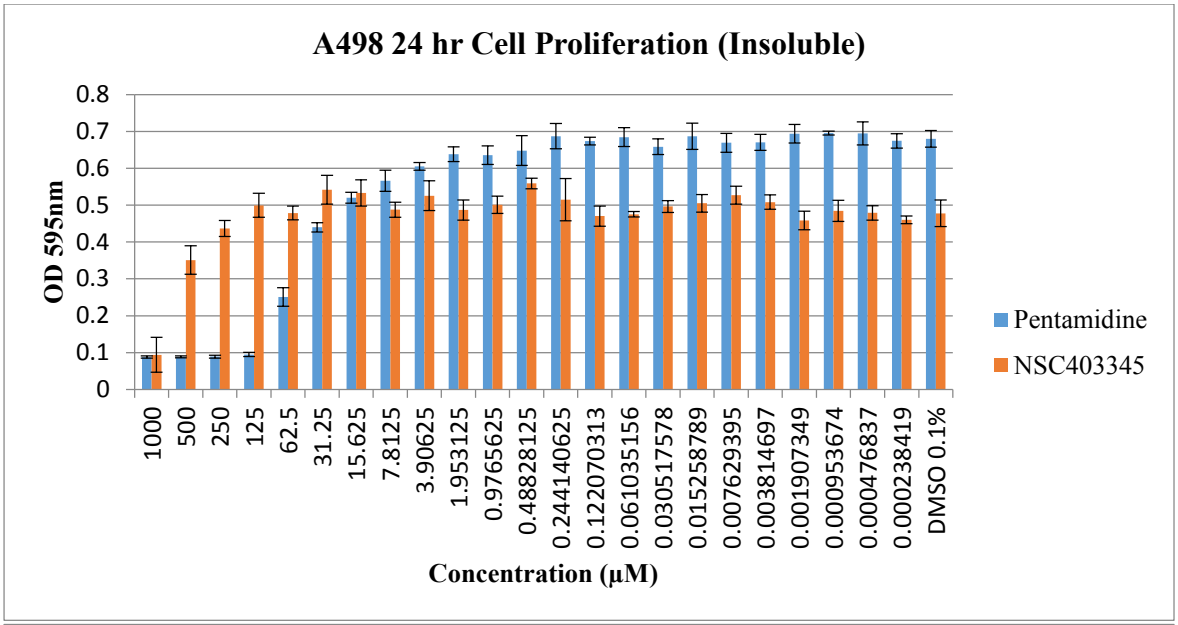

RCC4+ 24 hr Cell Proliferation
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC26283

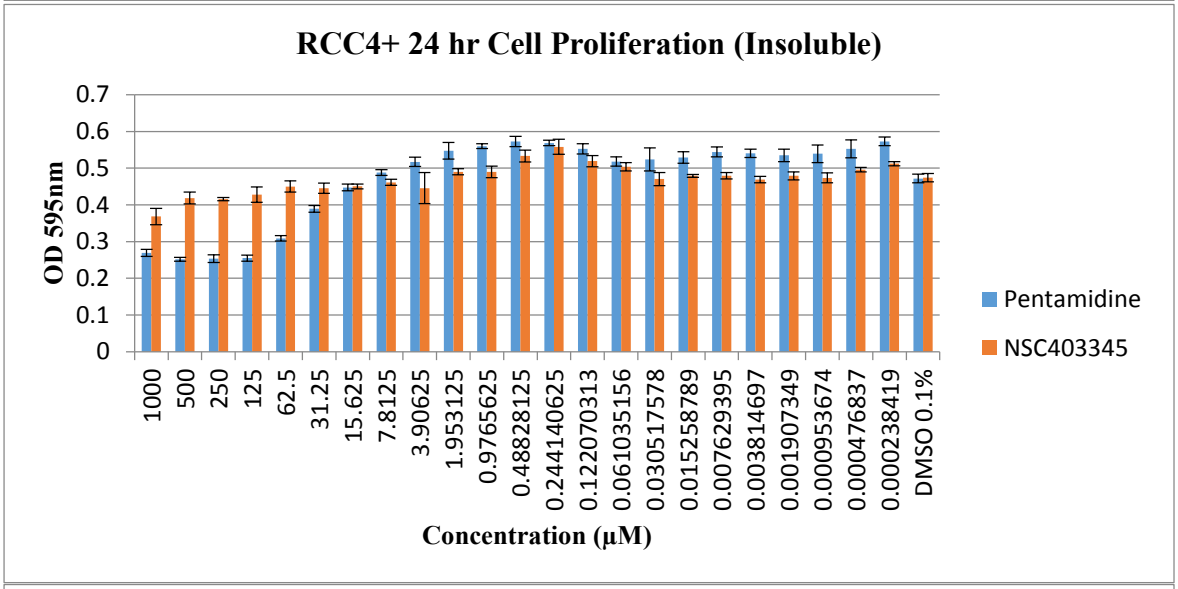

RCC4- 24 hr Cell Proliferation
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC26283

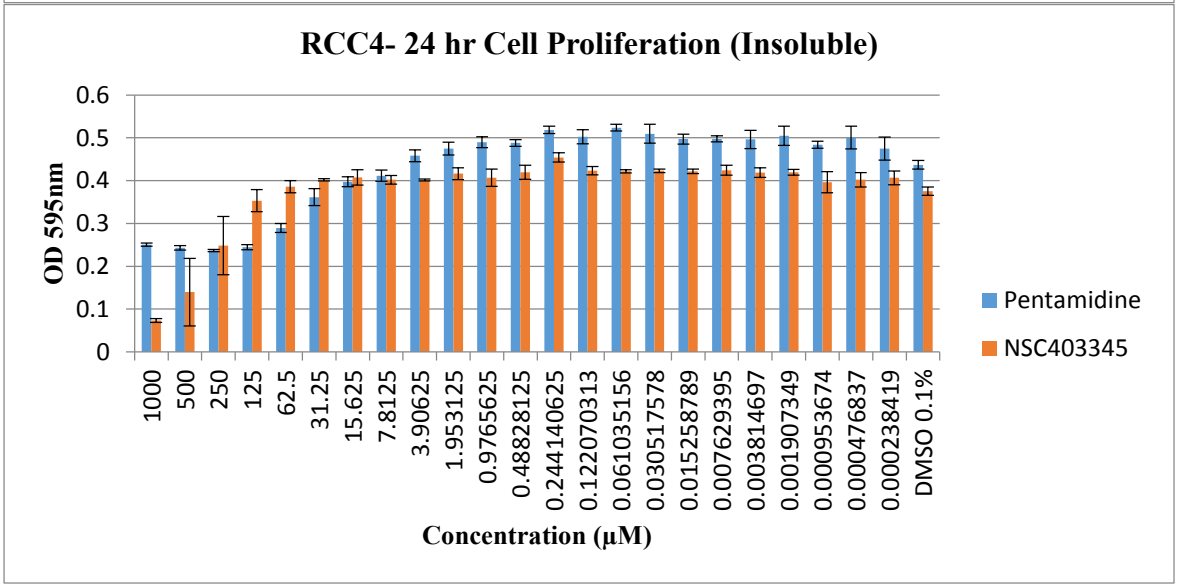

A498 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC31592

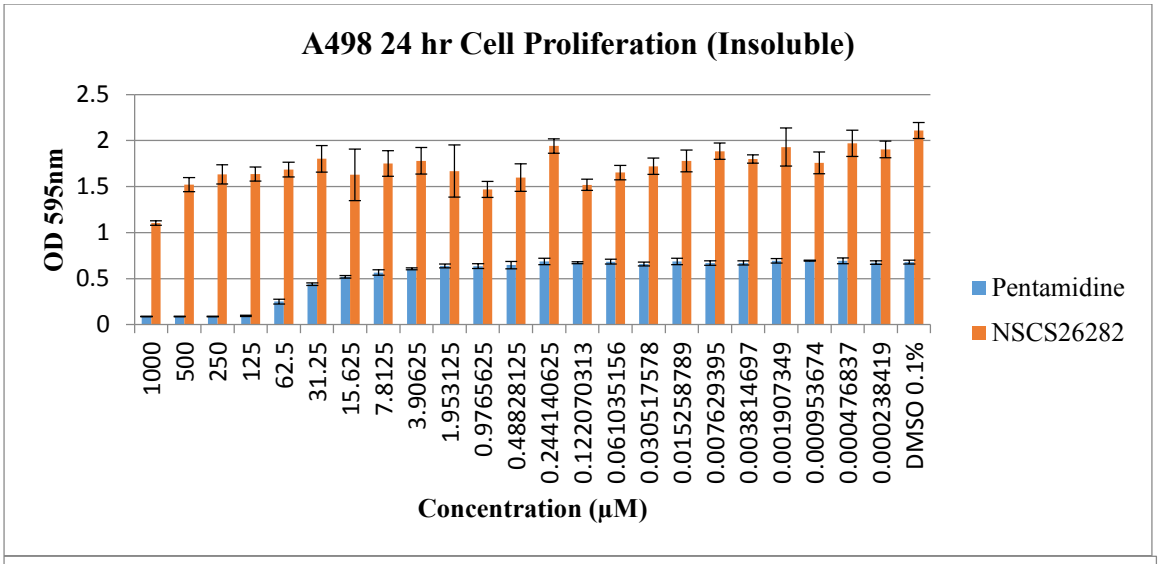

RCC4+ 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC31592

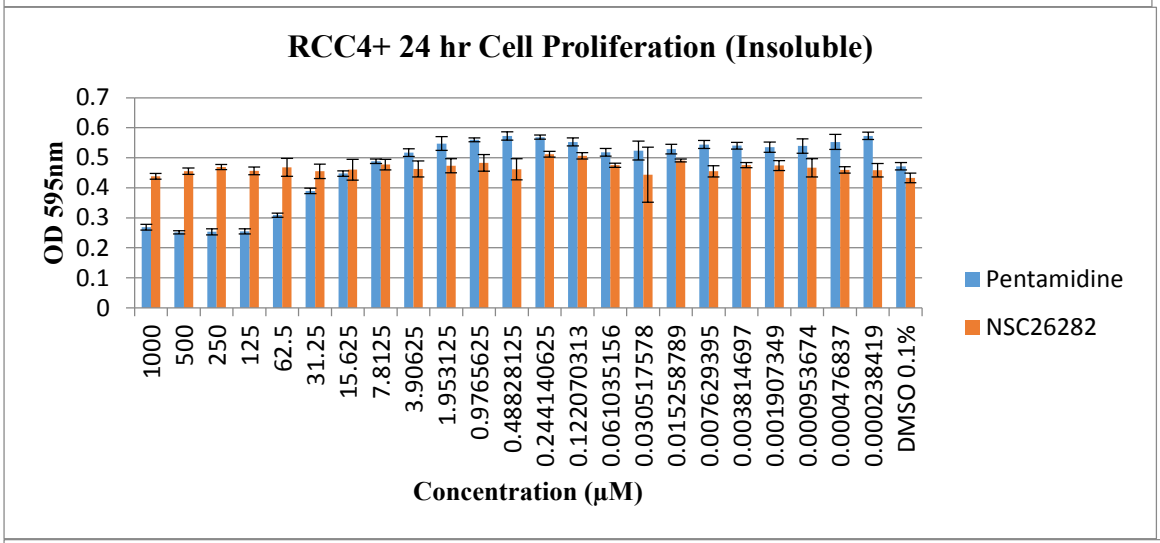

RCC4+ 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC31592

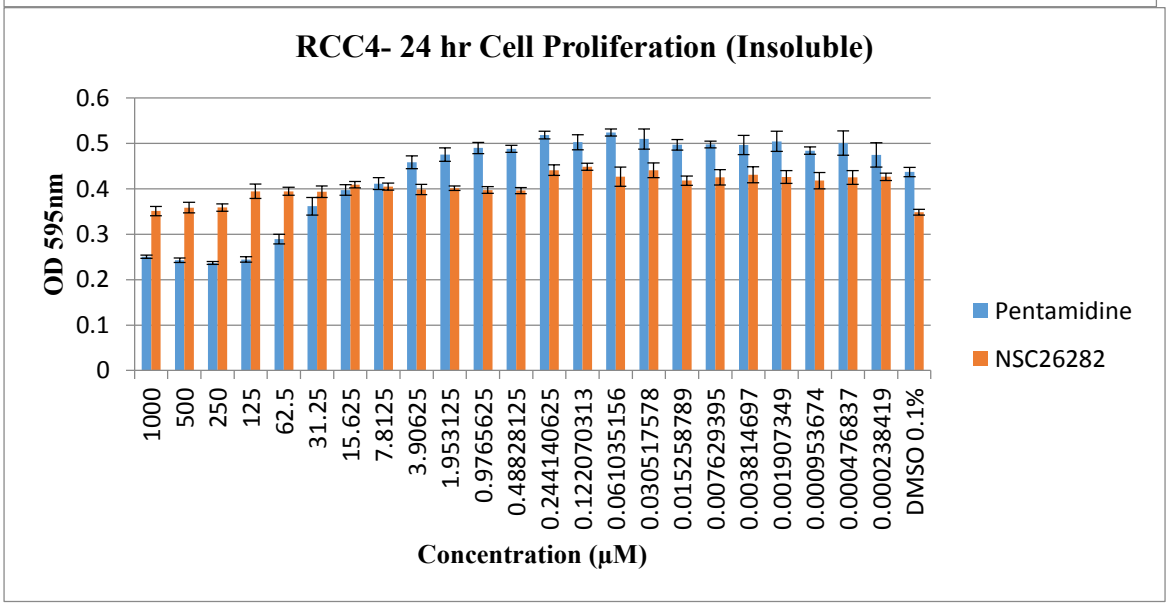

A498 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC403345


RCC4+ 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC403345


RCC4- 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC403345

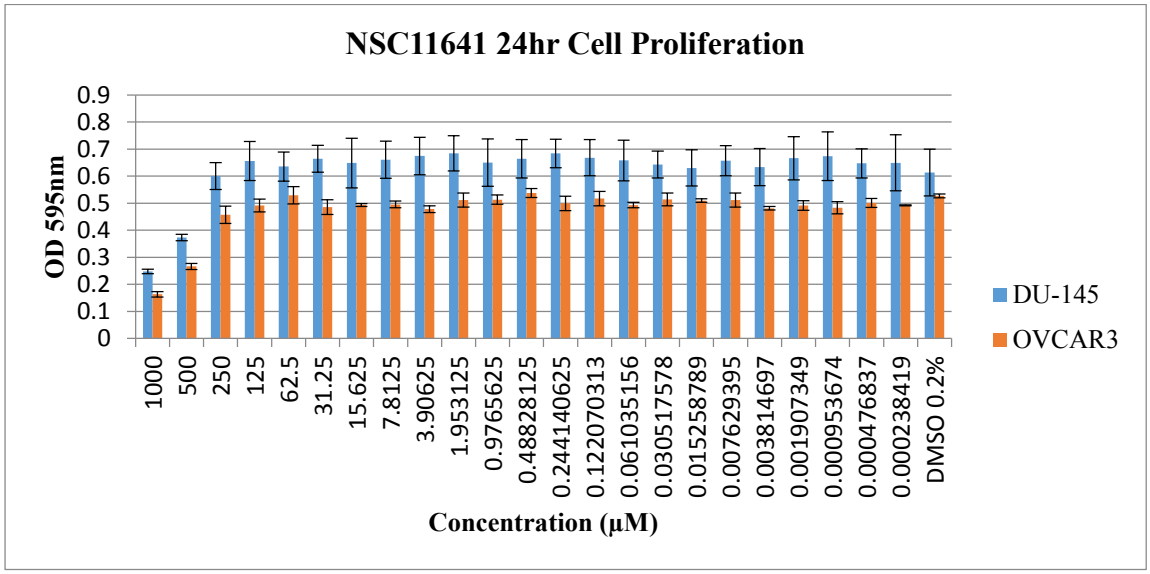

A498 24 hr Cell Proliferation (Insoluble)
OD 595nm
2.5
2
1.5
1
0.5
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSCS26282

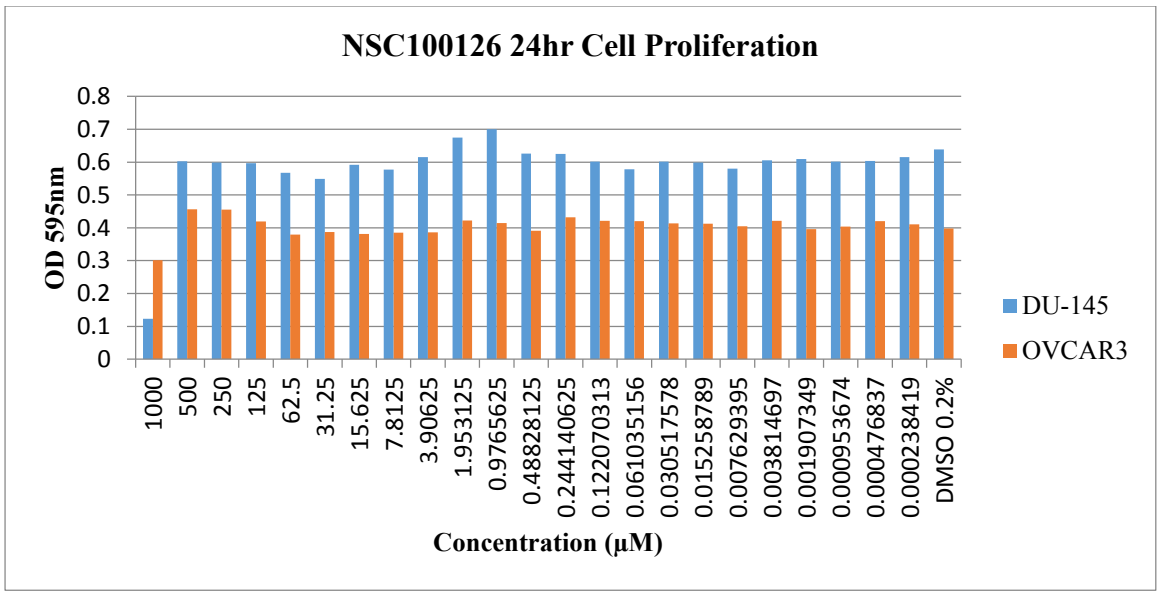

RCC4+ 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC26282

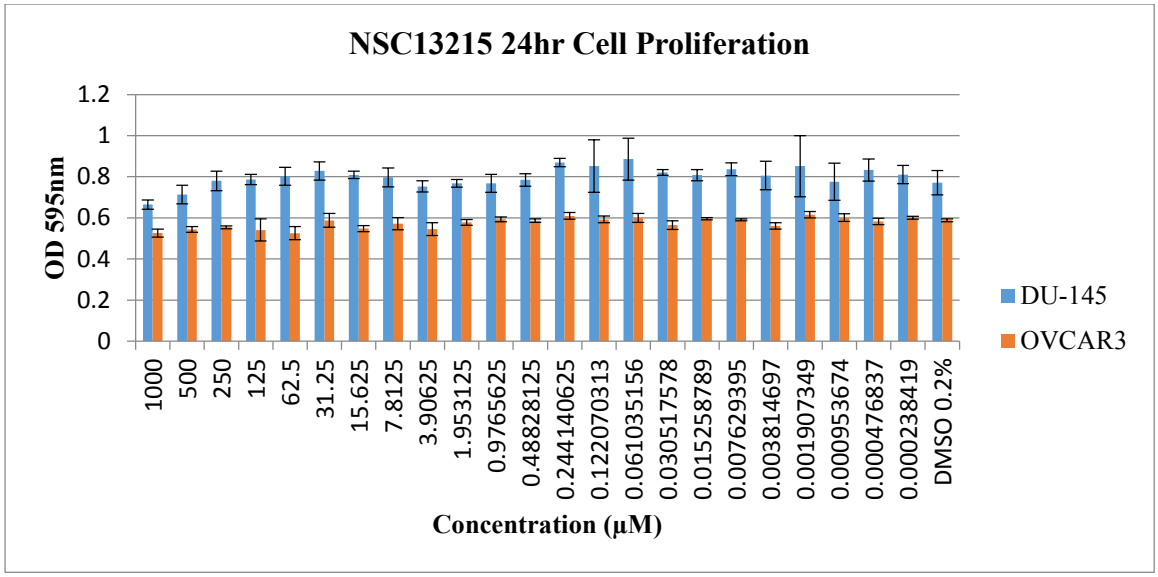

RCC4- 24 hr Cell Proliferation (Insoluble)
OD 595nm
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.1%
Concentration (µM)
Pentamidine
NSC26282

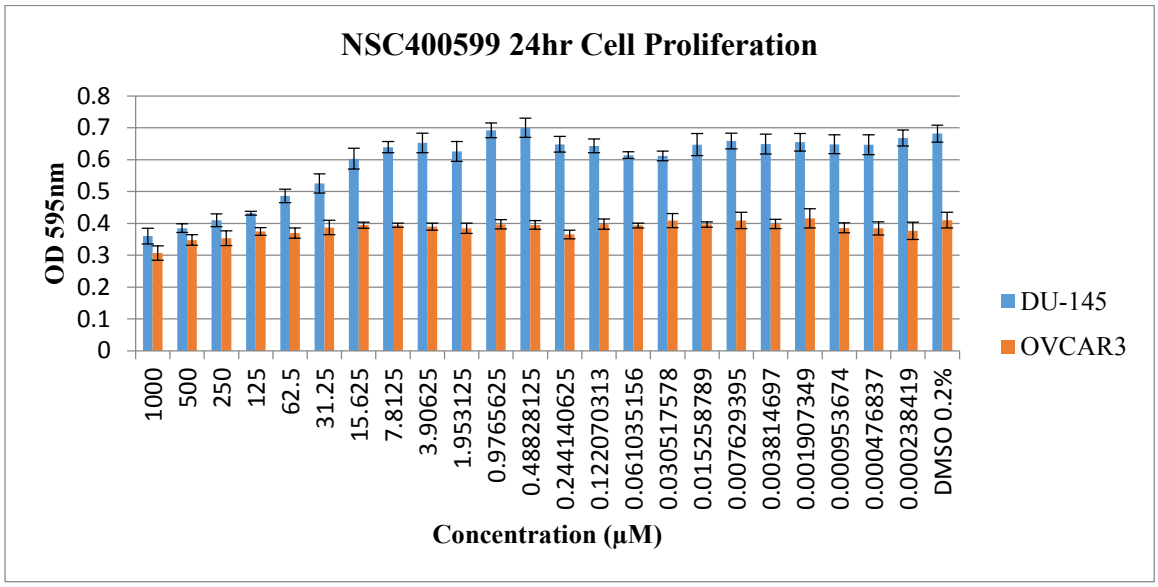
### 6.3.5 Appendix 3.5: Biological Results of NCI Compounds identified from QNZ LBVS

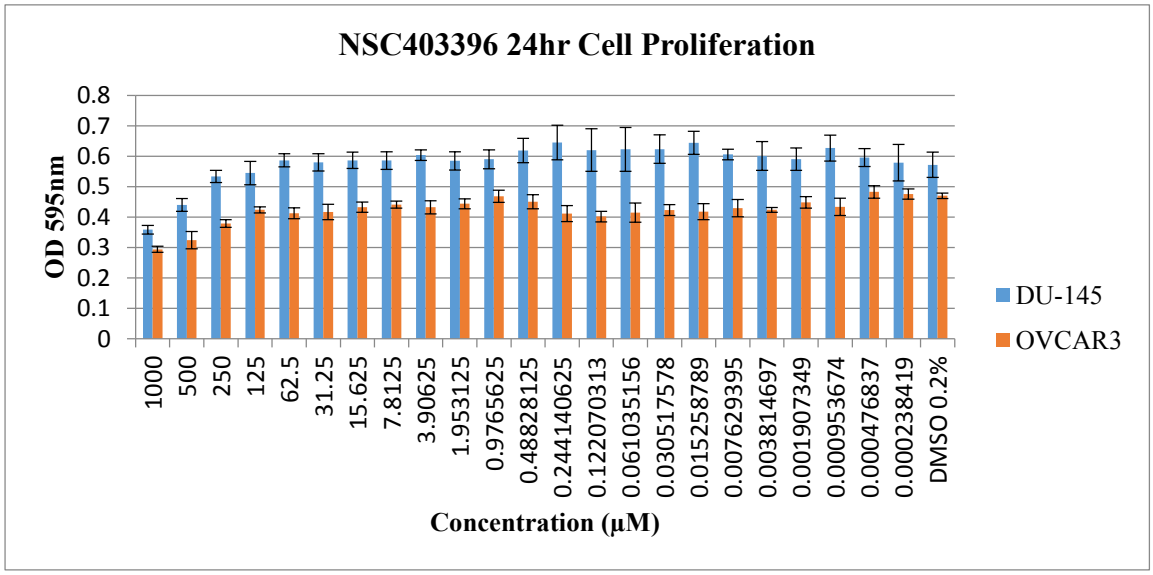
The MTT proliferation graph of NCI compounds identified from the LBVS performed for QNZ in Chapter 3 are depicted below. The graphs include the proliferation MTT assay, 24 hour post-treatment in DU-145, PC-3 and OVCAR3 cancer cell line. DMSO 0.2% was used as vehicle control. Data shown represents mean ± S.D. of four independent experiments performed in duplicate. OD = absorbance.

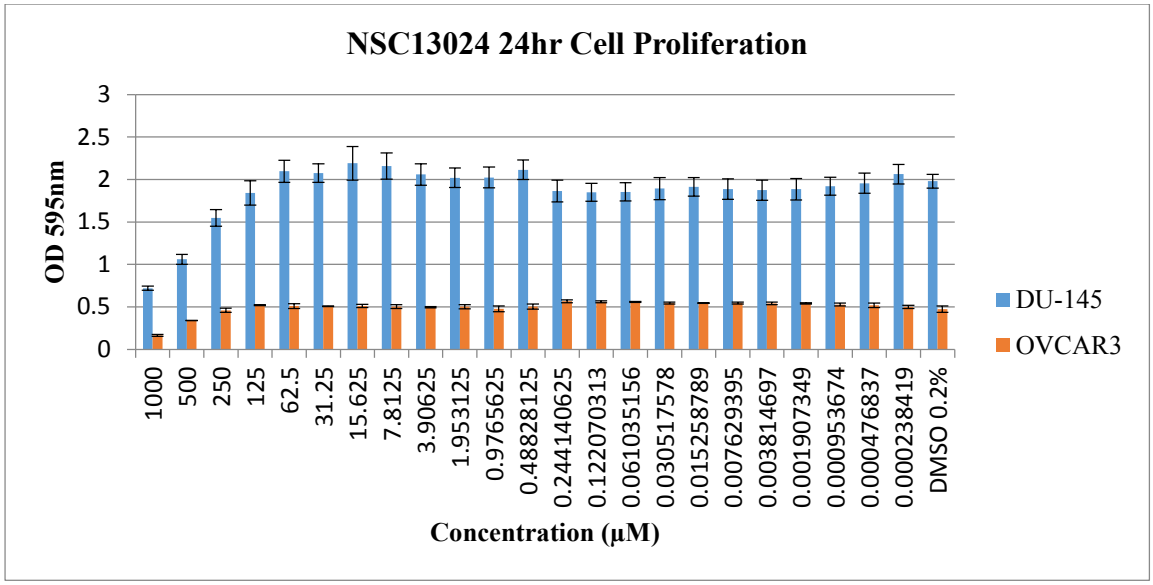


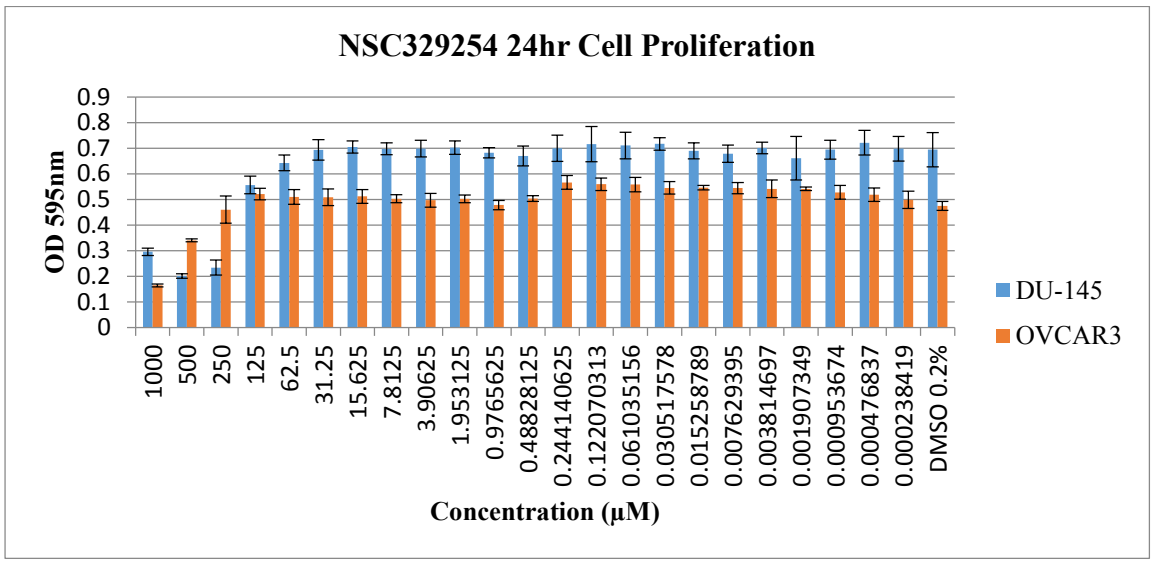



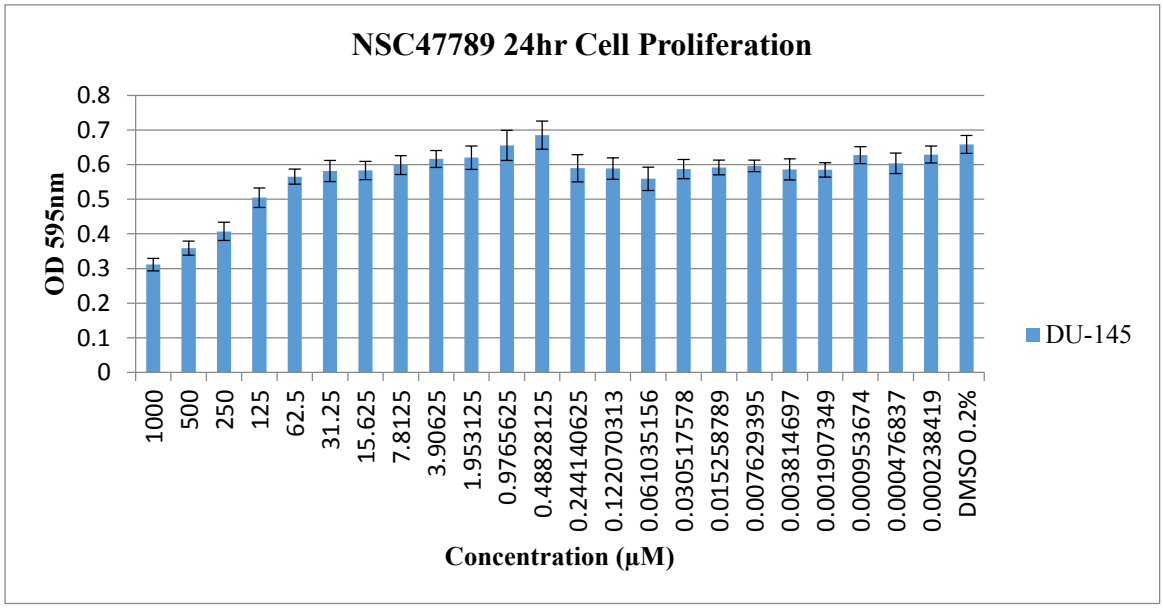

NSC13215 24hr Cell Proliferation
OD 595nm
1.2
1
0.8
0.6
0.4
0.2
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3

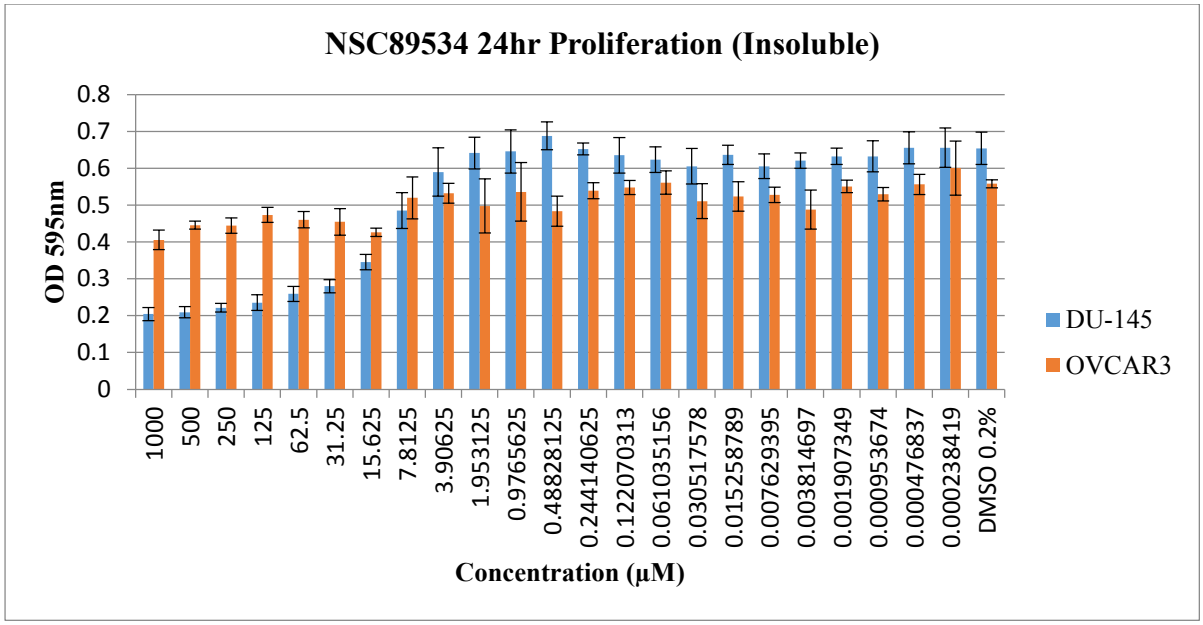

NSC400599 24hr Cell Proliferation
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3

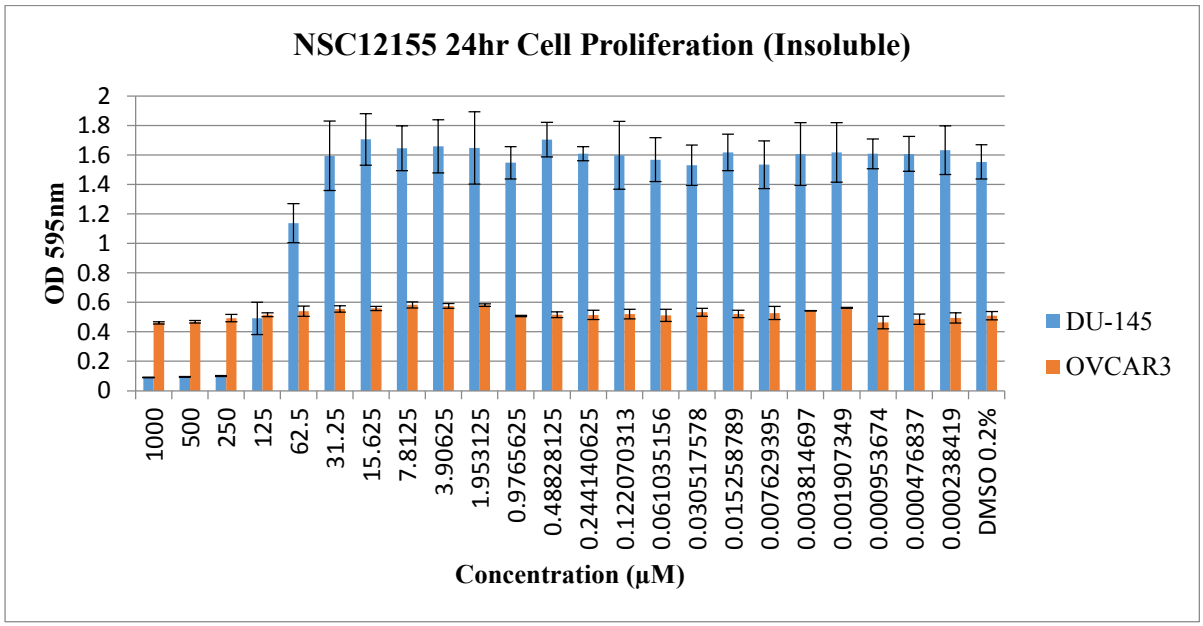

NSC403396 24hr Cell Proliferation
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3

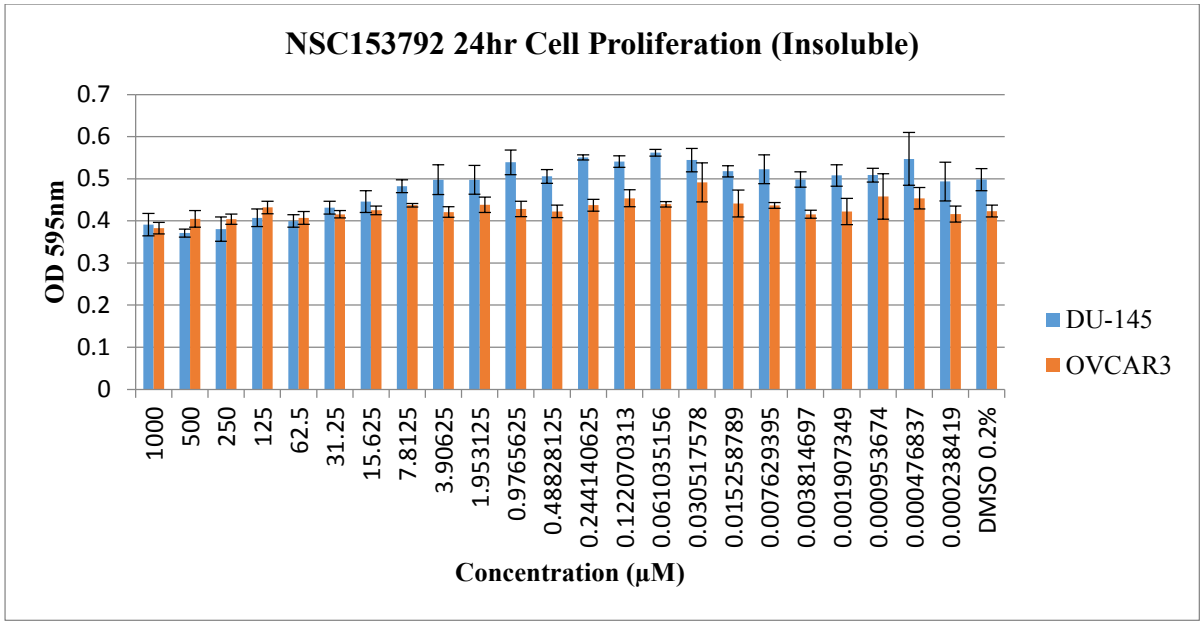

NSC13024 24hr Cell Proliferation
OD 595nm
3
2.5
2
1.5
1
0.5
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3

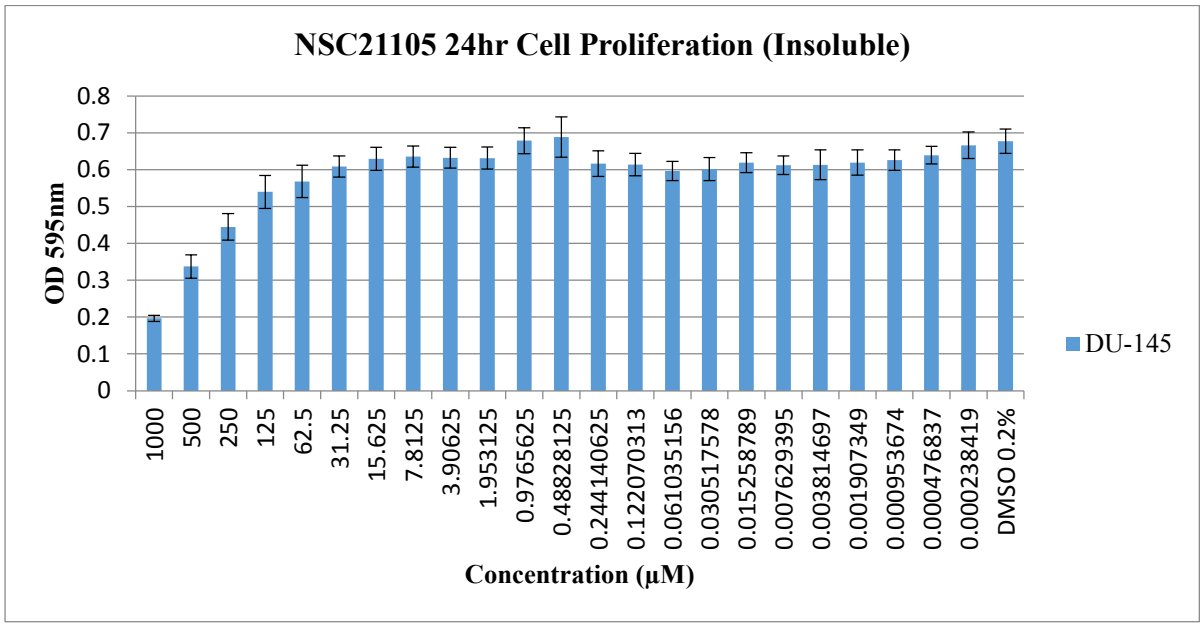

NSC329254 24hr Cell Proliferation
OD 595nm
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3

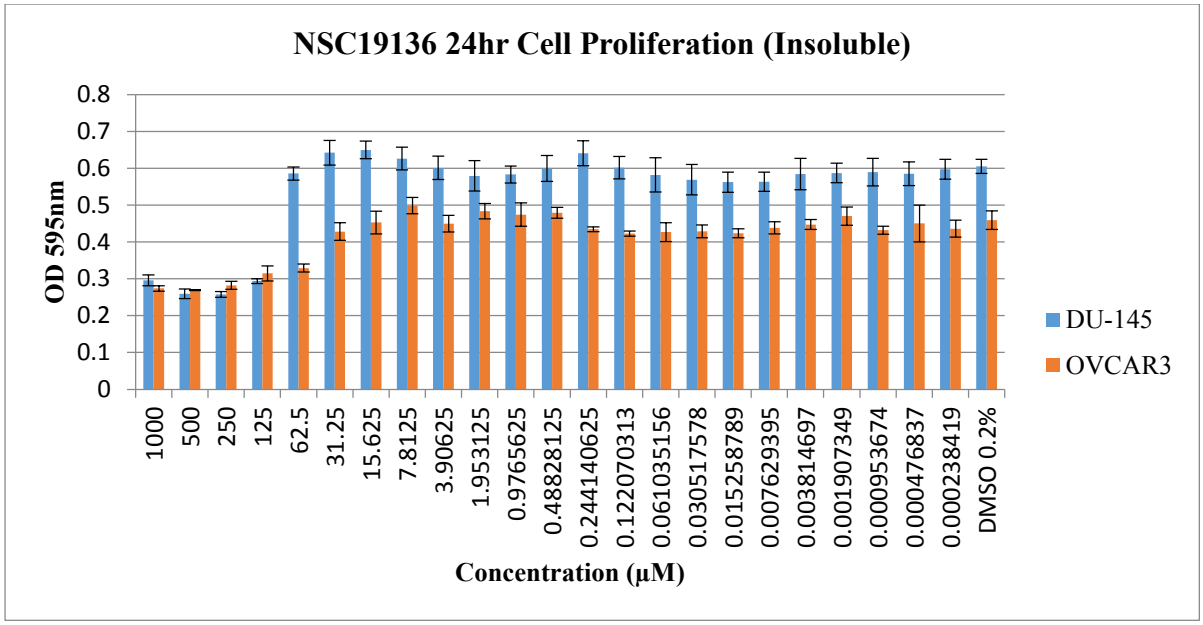

NSC47789 24hr Cell Proliferation
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145

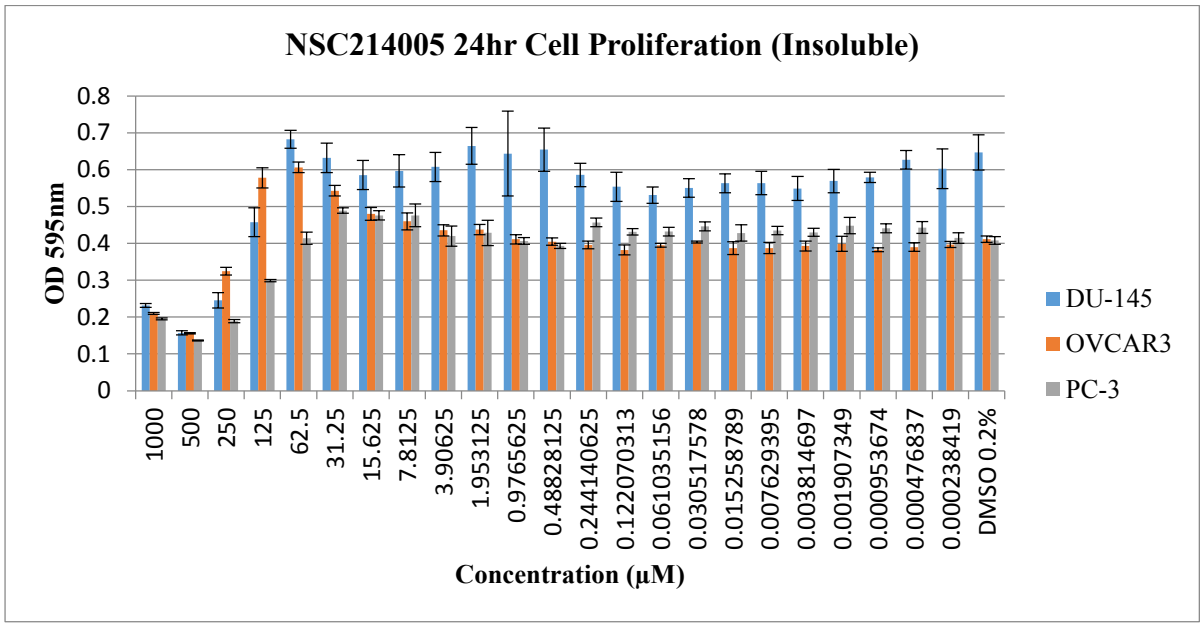

NSC89534 24hr Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3

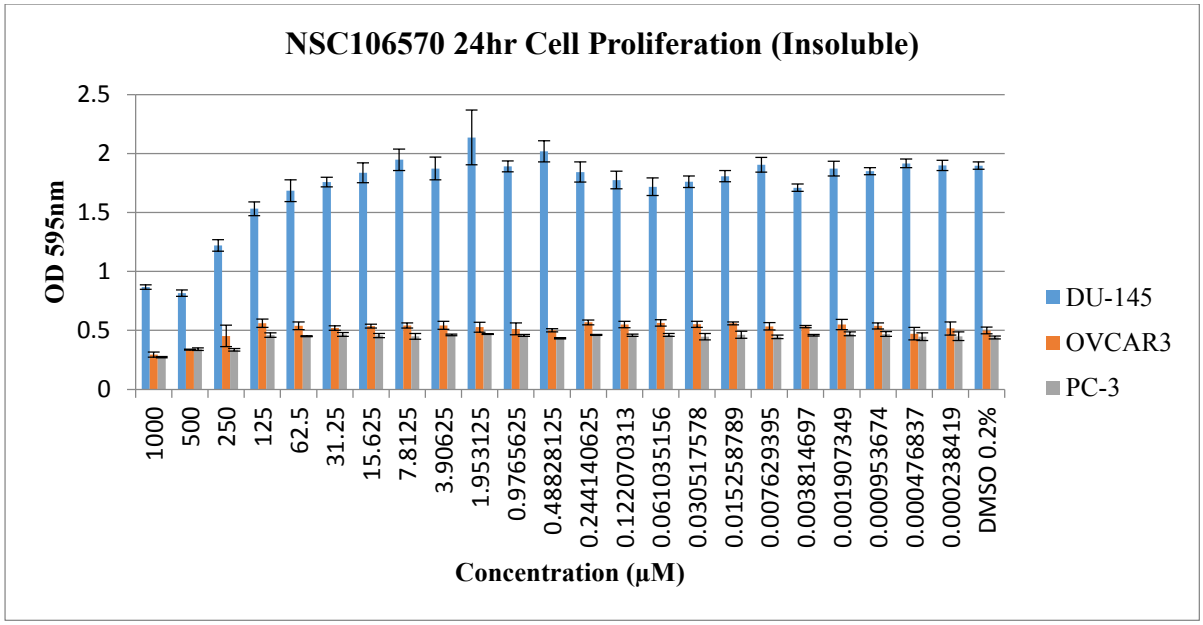

NSC12155 24hr Cell Proliferation (Insoluble)
OD 595nm
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3

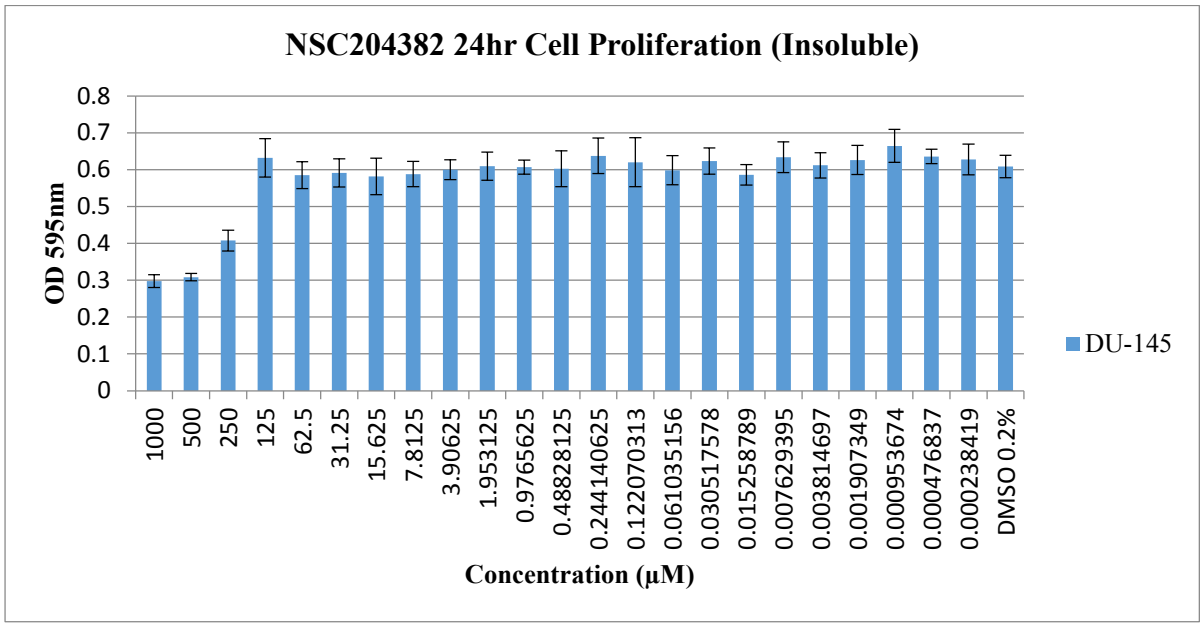

NSC153792 24hr Cell Proliferation (Insoluble)
OD 595nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3

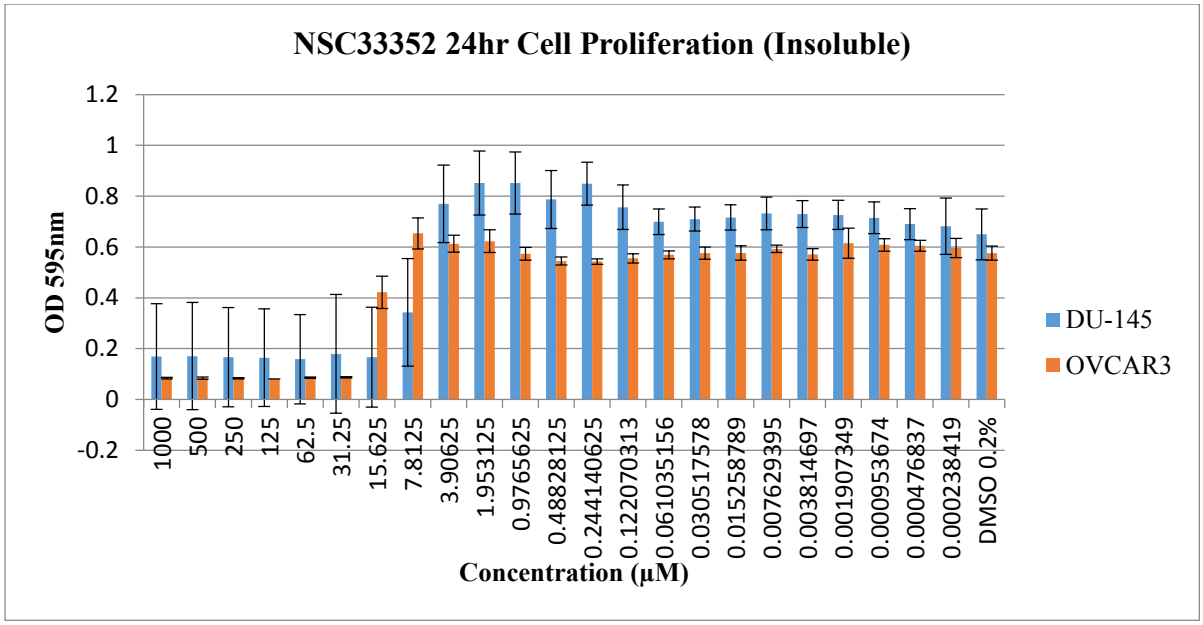

NSC21105 24hr Cell Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145

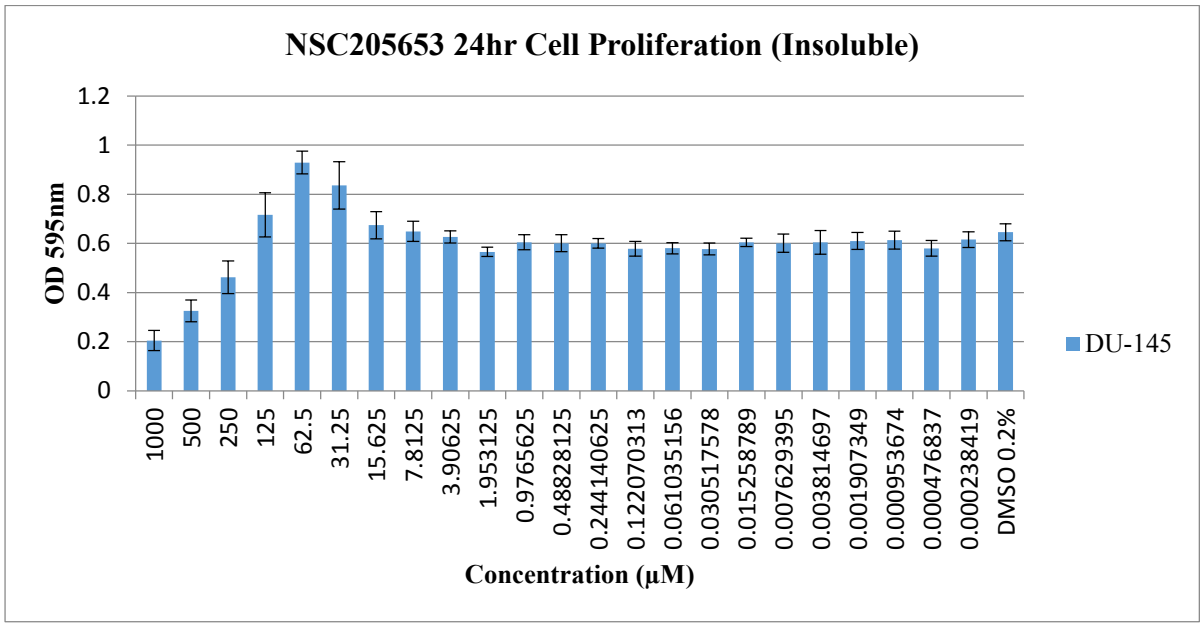

NSC19136 24hr Cell Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3

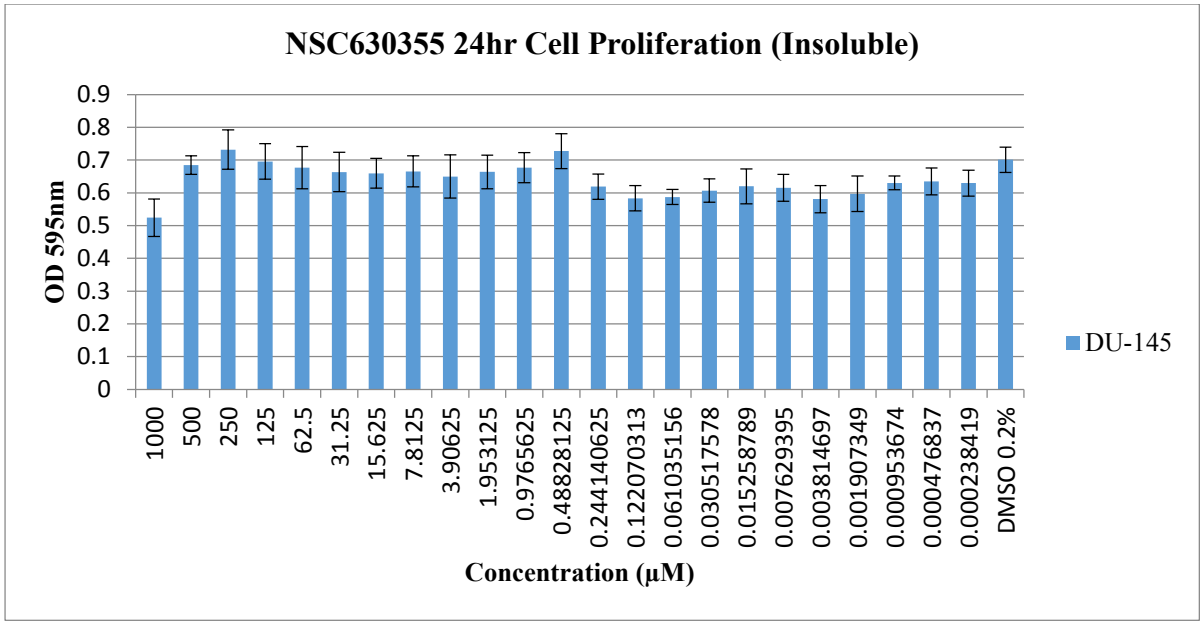

NSC214005 24hr Cell Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3
PC-3

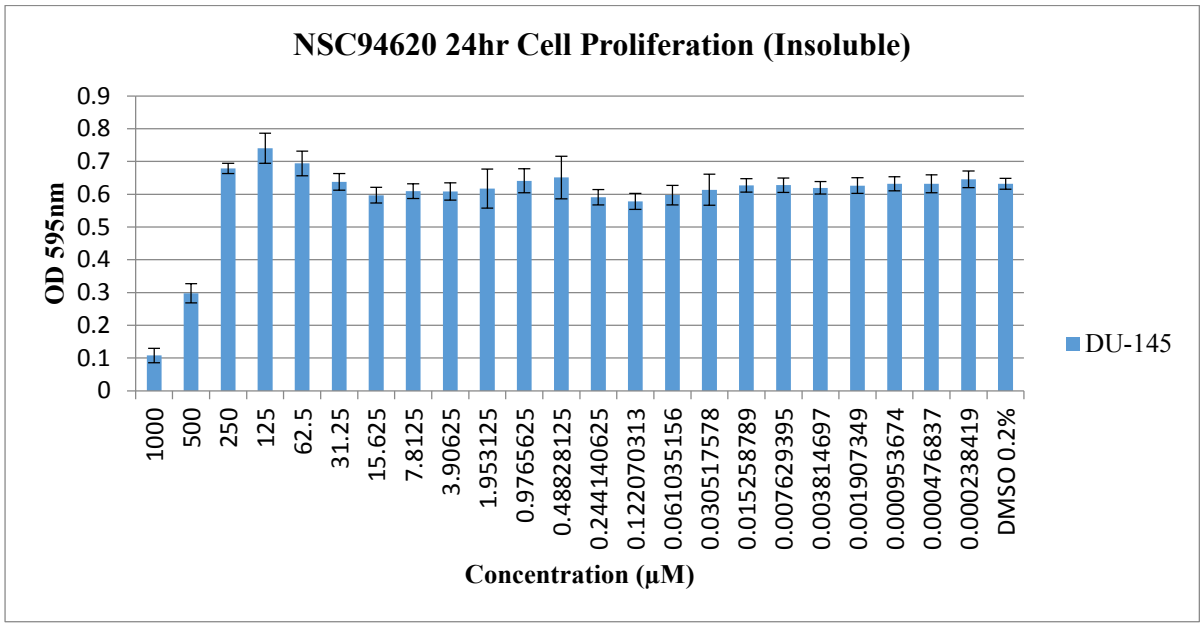

NSC106570 24hr Cell Proliferation (Insoluble)
OD 595nm
2.5
2
1.5
1
0.5
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3
PC-3

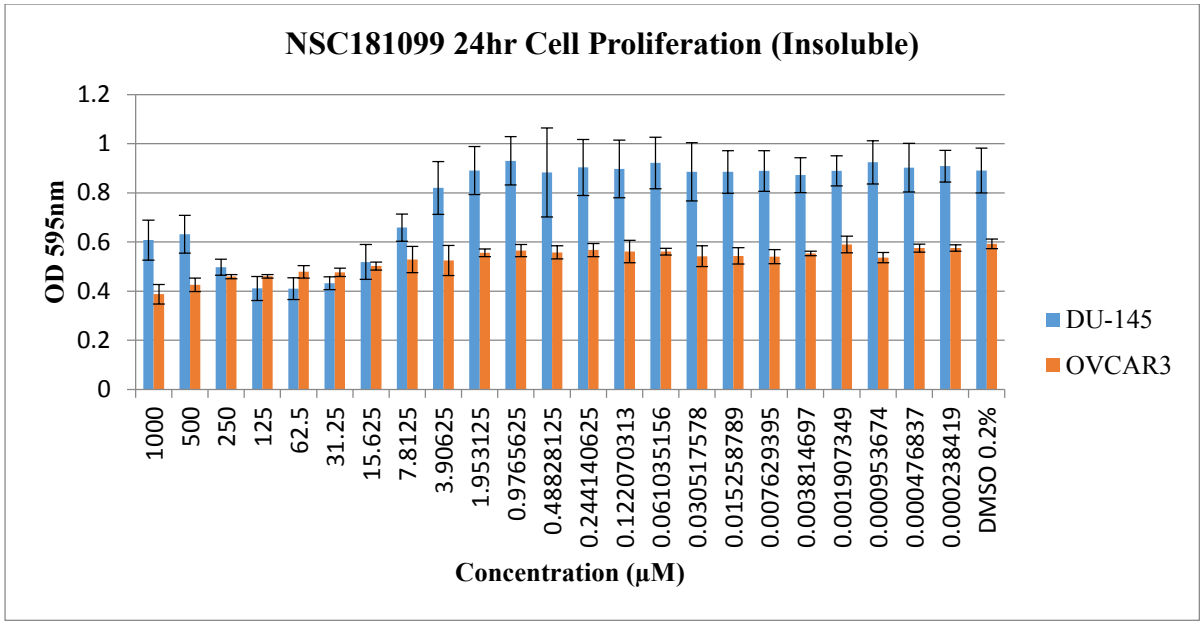

NSC204382 24hr Cell Proliferation (Insoluble)
OD 595nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145

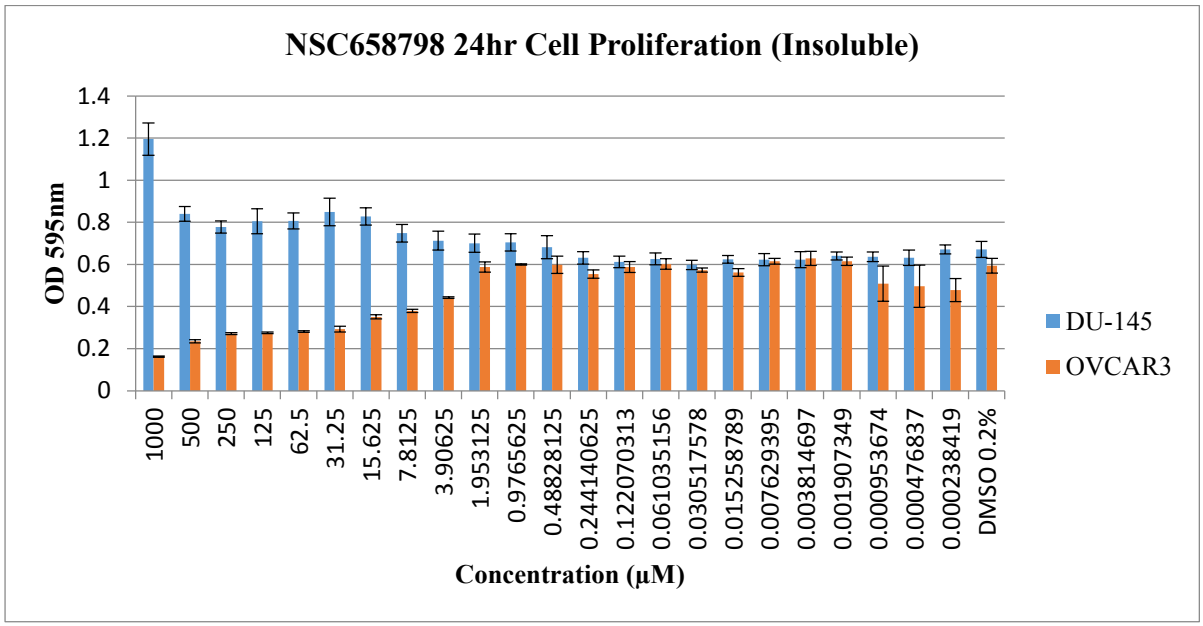

NSC33352 24hr Cell Proliferation (Insoluble)
OD 595nm
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3

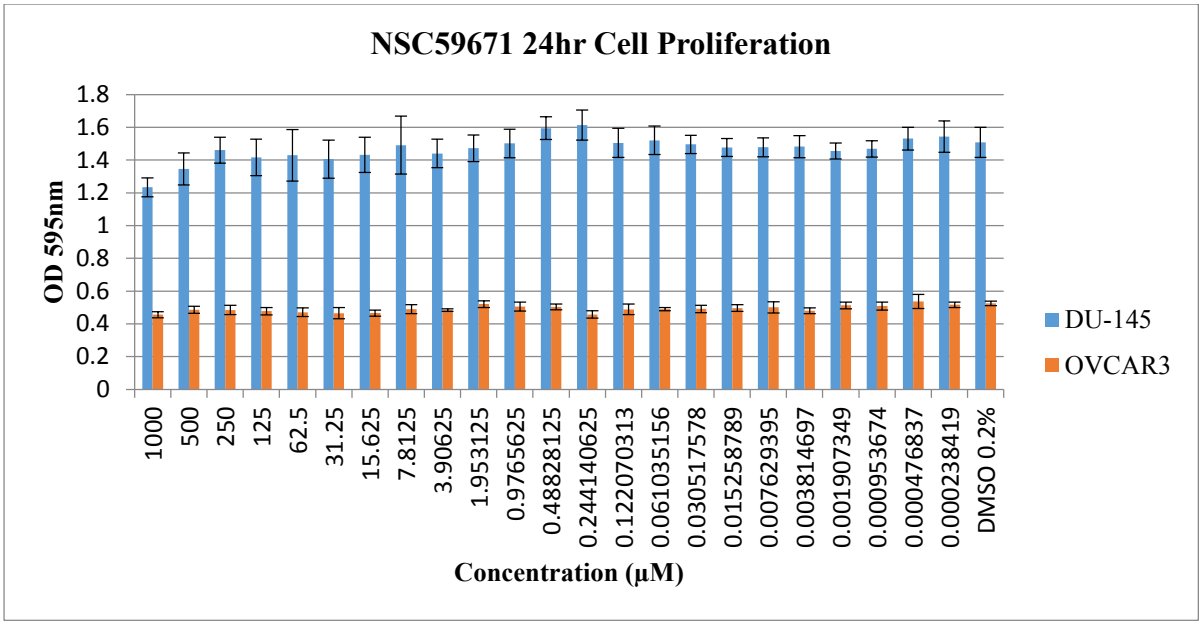

NSC205653 24hr Cell Proliferation (Insoluble)
OD 595nm
1.2
1
0.8
0.6
0.4
0.2
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145


NSC630355 24hr Cell Proliferation (Insoluble)
OD 595nm
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145


NSC94620 24hr Cell Proliferation (Insoluble)
OD 595nm
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145


NSC181099 24hr Cell Proliferation (Insoluble)
OD 595nm
1.2
1
0.8
0.6
0.4
0.2
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3


NSC658798 24hr Cell Proliferation (Insoluble)
OD 595nm
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3


NSC59671 24hr Cell Proliferation
OD 595nm
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1000
500
250
125
62.5
31.25
15.625
7.8125
3.90625
1.953125
0.9765625
0.48828125
0.244140625
0.122070313
0.061035156
0.030517578
0.015258789
0.007629395
0.003814697
0.001907349
0.000953674
0.000476837
0.000238419
DMSO 0.2%
Concentration (µM)
DU-145
OVCAR3