# IDENTIFICATION OF ALLOSTERIC AKT INHIBITORS BY COMPUTATIONAL METHODS

by

Oya Gürsoy Yılmaz

B.S., Chemical Engineering, Izmir Institute of Technology, 2011

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering
Boğaziçi University

2013

# ACKNOWLEDGEMENTS

This thesis study would not have been possible without the guidance and the help of a couple of individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, I want to express my sincere gratitudes to my supervisor and co-supervisor; Prof. Kutlu Ülgen and Asst. Prof. Elif Özkırımlı, due to their helps and generosities in sharing their experiences with me through these two years we worked together.

I would like to thank to Seval Aladag for her technical support and guidance during my study. Grateful thanks go to Gizem Özbüyükkaya, who became a real friend than a co-worker and never denied her boundless help to me whenever I needed. I would also like to thank to my group members Begüm Alaybeyoğlu, Utku Deniz and Deniz Menekşedağ for their moral support and optimistic attitude during long hours I spent in the laboratory. My warmest appreciations are for my dear friends İrem, Gizem, Cihan, Begüm, Erge, Deniz, İdil, Onur and Yeşim, without whom Bogazici would not mean that much for me.

I would heartily like to thank my family for their patience and steadfast encouragement to complete this study. Next to the last, I appreciate them; my lovely four footed babies Kiraz and Şans, with all my heart for being my inspiration and vitality in the completion of this research work. Lastly, my one and only, Melih Can Yılmaz; I know that we would never be what we are today without being such unique soul mates for each other.

# ABSTRACT


## IDENTIFICATION OF ALLOSTERIC AKT INHIBITORS BY COMPUTATIONAL METHODS

AKT kinases are known to be major contributors to the development of many human cancers according to preclinical data and studies of tumor specimens. Design and development of small molecule inhibitors targeting AKT and its signaling pathway is therefore an area of intense research in the field of cancer therapeutics. To date, most of the small molecule inhibitors target the kinase domain of AKT. However, due to the high degree of homology in the ATP-binding pocket among different serine/threonine kinases, achieving selectivity for these inhibitors remains a major problem. The evolutionarily less conserved PH domain and the interface it makes with the kinase domain is an attractive alternative to develop more selective inhibitors. The aim of this thesis study is to carry out in silico screening to identify potential allosteric inhibitors expected to bind the cavity between kinase and PH domains of Akt1. To this end, both structure-based and ligand-based pharmacophore models were developed. Structure-based pharmacophore models were constructed using important structural features of Akt1. The 3D structural information of previously identified allosteric Akt inhibitors was used to develop a ligand-based pharmacophore model. Database screening was performed based on 3D similarity to the selected pharmacophore hypotheses. Glide was used to predict binding modes of the ligands. Top scoring hits were further analyzed considering 2D similarity between the compounds, interactions with Akt1, fitness to pharmacophore models, ADME and druglikeness criteria. Derivatives of 3-methyl-xanthine, quinoline-4-carboxamide and 2-[4-(cyclohexa-1,3-dien-1-yl)-1H-pyrazol-3-yl]phenol were proposed as potential leads for allosteric inhibiton of Akt1.

# ÖZET

## ALLOSTERİK AKT İNHBİTÖRLERİNİN HESAPSAL YÖNTEMLER İLE BELİRLENMESİ

Klinik öncesi bilgilere ve tumor çeşitlerine gore, AKT kinazların birçok kanser türünün gelişiminde etkisinin olduğu bilinmektedir. Bu kanser türlerinin tedavisinde yol alabilmek için, AKT ve bulunduğu sinyal sistemini hedef alan küçük moleküllü AKT inhibitörlerinin üzerine çalışmalar yürütülmelidir. Şu ana kadar yapılan girişimlerin çoğunluğu AKT'nin kinaz bölgesini hedef alan inhibitörleri geliştirme üzerine yapılmıştır. Fakat, farklı serin/tironin kinazların ATP bağlanma ceplerindeki yüksek derecede benzerlikten ötürü, bu inhibitörler için seçicilik önemli bir problem teşkil etmektedir. AKT'nin PH bölgesindeki dizilim benzerliğinin kinaz bölgesindekine göre belirgin şekilde az olmasından ötürü, bu bölge ile etkileşen daha seçici inhibitörlerin geliştirilmesi önemlidir. Bu çalışmanın amacı sanal metodlar kullanarak AKT'nin PH ve kinaz bölgelerinin arasında kalan boşluğa bağlanması hedeflenen potansiyel allosterik AKT inhibitörleri belirlemektir. Bu amaç için, hem mevcut ligand yapılarını hem de protein yapısını baz alan farmakofor modeller oluşturulmuştur. Protein yapısını baz alan farmakofor modeller, AKT'nin allosterik bağlanma bölgesinin önemli özelliklerinin belirlenmesi ve kullanılması ile oluşturulmuştur. Literatürde önceden belirlenmiş ligandların üç boyutlu yapısal bilgileri kullanılarak ligandları baz alan bir farmakofor model oluşturulmuştur. Oluşturulan farmakofor modellere üç boyutlu benzerliğe göre veritabanı taraması yapılmıştır. Taramadan geçen bileşiklerin Glide yazılımı kullanılarak (doklama) bağlanma bölgeleri tahmin edilmiştir. Doklamadan en yüksek skoru elde eden bileşikler, birbirleri arasında iki boyutlu benzerlik, AKT1 ile yaptıkları etkileşimler, seçilen farmakor modellere uygunluk, ADME ve ilaç benzerliği kriterlerine göre detaylı olarak incelenmiştir. 3-metil-ksantin, kinolin-4-karboksamit ve 2-[4-(siklohekza-1,3-dien-1-il)-1H-pirazol-3-il]fenol türevleri AKT1'in allosterik inhibisyonu için potansiyel öncü yapılar olarak önerilmektedir.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1. INTRODUCTION

The AGC kinase group is one of the most evolutionarily conserved groups of human kinome [1]. Each family of AGC kinases is a potential drug target due to its physiological role related to many of human diseases. Akt (also named as protein kinase B, PKB) is a phospholipid binding-serine/threonine kinase, of AGC superfamily and is a key mediator of the PI3K cell survival signaling pathway [2]. Activation through deregulation, amplification and rearrangement of the phosphatidylinositol 3-kinase (PI3K)/ Akt pathway occur frequently in many human cancers and is associated with tumor growth, increased metastasis, and resistance to therapy [3]. Human Akt has three isoforms; Akt1/PKBα, Akt2/PKBβ and Akt3/PKBγ which share more than 85% similarity [4].

Many cellular processes such as cell survival and insulin signaling are affected by Akt [5, 6]. Various cancer types involving leukemia, breast, pancreatic, ovarian and prostate are caused by activation of Akt due to genetic mutations of its own and/or its interrelated upstream and downstream proteins (PTEN, PI3K) [3]. A large number of small molecule inhibitors targeting Akt are being developed for suppressing its activation and over-expression in various cancer types, [7].

Structural studies on Akt isoforms show that, Akt consist of three conserved domains; an N-terminal pleckstrin homology (PH) domain (Akt1 residues 1-133), a central kinase catalytic (CAT) domain (Akt residues 149-408) and a C-terminal extension (EXT) domain (Akt1 residues 449-480) [8, 9]. Among the Akt isoforms, the PH domains are 80% identical, CAT domain is 90% identical and C-terminal extension (EXT) is 70% identical [10]. Types of small molecule inhibitors reported for Akt differ in the target sites on the protein [11]. The inhibitors have been found to bind to ATP binding pocket in the catalytic domain [12-14], to the PH domain [15, 16], or to the region between PH and kinase domains [2, 9, 17].

The ATP-site targeting compounds represent the mostly studied Akt inhibitors both preclinically and in literature [18]. While the high degree of homology (over 50% sequence

identity) between the kinase catalytic domains of serine/threonine kinases makes selectivity a major issue [19, 20]; their N-terminal pleckstrin homology (PH) domains are divergent (about 30% identity), which allows development of more selective allosteric inhibitors that target the PH domain or the region between PH and kinase domains [21, 22].

In this thesis study, novel allosteric Akt1 inhibitors targeting the cavity between kinase and PH domains are identified by a virtual screening campaign combining database filtering with ligand and structure based pharmacophore models and docking. The proposed scaffolds are determined by 2D similarity analysis of docking hits and considering interactions with Akt1, fitness to pharmacophore hypothesis, drug-likeness and ADME properties. The content is as follows: Background information about the physiological role, structure, inhibition and previous work on Akt are included in Theory section. The tools used for all virtual screening approaches throughout the study are explained in Methods section. Results and Discussion section covers all computational output including figures and tables with detailed commentaries. The final section of Conclusion summarizes the workflow in general with specification of final findings.

# 2. THEORY

Recent progress in cancer therapies has altered the methods of discovering novel drug candidates. The traditional, non-target specific and relatively cytotoxic therapeutics, some of which are in pre-clinical and some are in clinical stage presently have been replaced by more specifically designed drugs. The key points for identifying prospective therapies are finding proteins that are mutated or causing mutations in cancer cells, determining the effects of these mutations in the signaling pathway and ultimately targeting these proteins or their signaling pathways [23].

## 2.1 Human Kinome

In 2002, Manning *et al*. defined protein kinase complement of the human genome as ‘human kinome’ and identified a total of 518 protein kinase genes [24]. Protein kinases are one of the largest families of genes in eukaryotes that have been studied extensively. Major groups of human kinome are CAMK, CMGC, RGC, TK, TKL, STE, CKI and AGC kinases with 134 families and 201 subfamilies Figure 2.1[25].

The family classification is based primarily on sequence comparison of catalytic domains; and also considers knowledge of sequence similarity and domain structure outside of the catalytic domains and known biological functions [26]. Human kinase distribution by major groups are shown in Figure 2.2.

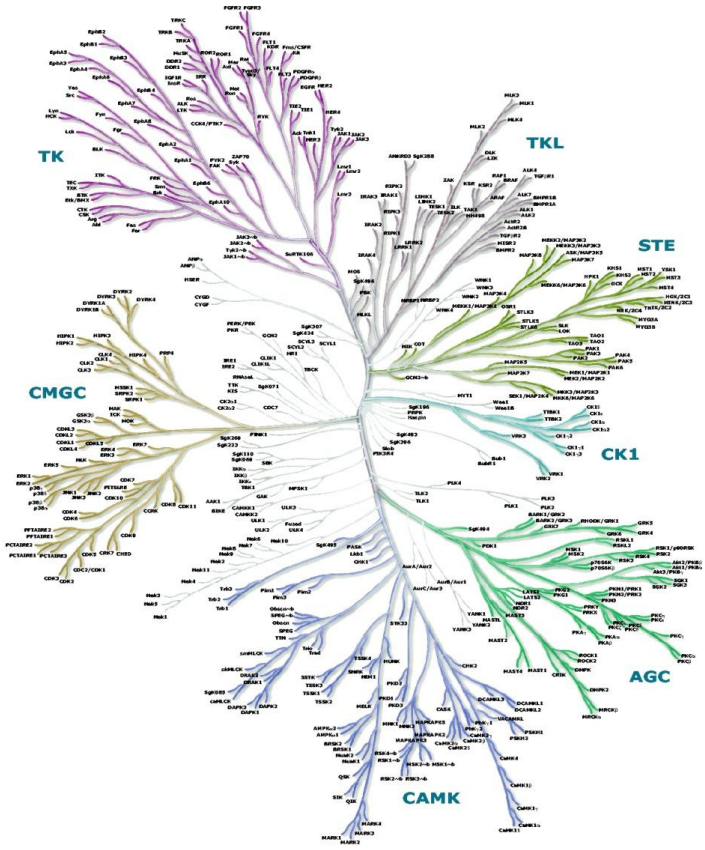



Figure 2.1. Tree representation of 7 major groups including 134 families and 201 subfamilies of human kinome (Adopted from [25]).

| Group | Families | Subfamilies | Human kinases |
|---|---|---|---|
| AGC | 14 | 21 | 63 |
| CAMK | 17 | 33 | 74 |
| CK1 | 3 | 5 | 12 |
| CMGC | 8 | 24 | 61 |
| Other | 37 | 39 | 83 |
| STE | 3 | 13 | 47 |
| Tyrosine kinase | 30 | 30 | 90 |
| Tyrosine kinase–like | 7 | 13 | 43 |
| RGC | 1 | 1 | 5 |
| Atypical-PDHK | 1 | 1 | 5 |
| Atypical-Alpha | 1 | 2 | 6 |
| Atypical-RIO | 1 | 3 | 3 |
| Atypical-A6 | 1 | 1 | 2 |
| Atypical-Other | 7 | 7 | 9 |
| Atypical-ABC1 | 1 | 1 | 5 |
| Atypical-BRD | 1 | 1 | 4 |
| Atypical-PIKK | 1 | 6 | 6 |
| Total | 134 | 201 | 518 |

Figure 2.2. Number of kinases in major human kinase groups, families and subfamilies (Adopted from [2]).

## 2.2 AGC Kinases

Among all others, AGC kinases are the third largest group with 14 families, namely, PDK1, Akt/PKB, SGK, RSK, PKA, PKG, PKC, PKN/PRK, NDR/LATS, MAST, YANK, DMPK, GRK and SGK494, including 21 subfamilies and a total of 63 kinases. The AGC kinase group comprises 12% of the human kinome and is one of the most evolutionary conserved groups. AGC kinases are represented widely within eukaryotes (i.e. all vertebrates, fungi, plants, unicellular algae, and protists) [1]. Each family of AGC kinases has a physiological role and is involved in human disease, therefore is a potential drug target.

PDK1 is a master kinase and responsible for the phosphorylation of the activation-loop site, therefore activating at least 23 other AGC kinases (i.e. Akt/PKB, SGK, PKC, S6K, RSK). PDK1 is vitally important in cancer and diabetes due to its regulatory role in PI3K/AKT pathway for AKT and AKT-dependent cellular functions. It is studied widely as being a prospective target for cancer treatment [27].

Akt/PKB kinases have major role in PI3K pathway and effect many cellular processes such as cell survival (Akt1) and insulin signaling (Akt2). Akt amplification leads to cancer, diabetes, cardiovascular and neurological diseases. As a result of overexpression of all three isoforms of Akt (Akt1, Akt2, Akt3) in various cancer types, a large number of small molecule inhibitors targeting Akt are being developed for suppressing its activation [7].

SGK has three isoforms with similar biochemical properties and downstream effectors. Among the three isoforms, SGK1 is the most studied one and has a role in cell proliferation and survival, in addition to tumor growth [28] and regulation of blood pressure. SGK is also related to development and complications of diabetes (hypertension) and neurological disorders [29]; therefore diabetic-associated hypertension is aimed to be treated by use of SGK inhibitors [30].

Family RSK has three commonly studied subfamilies which are; subfamily of p90 ribosomal S6 kinases (RSK1–4), MSK and S6K. Activation of RSK subfamily results in

phosphorylation of multiple substrates related to differentiation, survival, growth and proliferation [31]. Due to RSK isoforms being key effectors of the Ras-stimulated Raf-MEK-ERK pathway, RSK inhibitors are used for prevention of cellular invasion and metastasis [32]. MSK subfamily (MSK1, MSK2) are key elements for negative feedback mechanisms required for counteracting the production of pro-inflammatory cytokines [33], therefore inhibitors targeting MSK subfamily are potential anti-inflammatory treaters [34]. Members of S6K subfamily are important for cell growth, proliferation, and cell differentiation, protein synthesis, cell cycle progression, and metabolism [35]. Since their over-expression causes some metabolic disorders, targeting S6K may be considered for the treatment of obesity and metabolic disorders [36].

PKA family has a primary effect on glucose homeostasis and triglyceride storage. Several other functions are formation of long term memory [37], cell migration, breast cancer resistance to chemotherapy and modulation of the immune response and inflammation [38].

PKG, PKC and PKN/PRK have similar biological and physiological functions and share common structural properties. PKG is the main effector kinase of the NO/cGMP signaling cascade and plays a central role in the regulation of cardiovascular and neurological functions. PKC has three subfamilies according to regulation prerequisites by secondary messengers, (namely typical, classical and novel) and play an important role in lung and other cancers. PKN/PRK are effectors of small GTPases which are involved in several processes such as cytoskeletal regulation, vesicle transport, cell adhesion, glucose transport, cell cycle and apoptosis, therefore some of PKN/PKR isoforms can be considered as target candidates for the development of anti-cancer drugs [1].

NDR and LATS kinases are responsible for a range of biological functions including cell division, apoptosis and tumor suppression [39]. Both LAT1-LATS2 and NDR1-NDR2 kinases are known to be participated in Hippo signaling pathway which regulates cell proliferation and apoptosis to control organ size [40].

Biological functions of SGK494, MAST and YANK families of kinases are mostly unknown. However, several studies suggest that MAST kinases have relation with breast

cancer and regulate mitosis [41]; YANK kinases are related with celiac disease, Ellis-van Crevel syndrome, coronary artery disease and so on [1].

DMPK family has four subfamilies studied widely; DMPK, ROCK, MRCK, CRIK. A mutation in gene responsible for coding DMPK (dystrophica myotonin protein kinase gene) results in a multisystem disorder (myotonic dystrophy type 1, DM1) which affects skeletal and smooth muscle, eye, heart, endocrine system and central nervous system. Therefore, some symptoms of DM1 are expected to be reversed by specific DMPK activators [42]. Both ROCK and MRCK subfamilies are downstream effectors of small GTPases of the Rho family and involved in regulation of some cellular functions like invasion, adhesion, migration, proliferation and gene expression [43]. CRIK has the primary role of abscission regulation during cytokinesis in mammalian cells. Several studies show that CRIK inhibitors are developed for treatment of hepatocellular carcinoma (HCC) and HIV infection [44].

Members of GRK family are responsible for control of GPCR signaling by Ser/Thr phosphorylation. Over/under GRK activity results in a variety of human disorders (heart failure, Parkinson's disease, mood disorders). GRK2 is considered as a drug target for the treatment of cardiovascular diseasemand GRK6 for treatment of Parkinson's disease [45].

## 2.3 AKT (Protein Kinase B)

AKT (also named as protein kinase B, PKB), a phospholipid binding-serine/threonine kinase, belongs to AGC superfamily of human kinome and is a key mediator of the PI3K cell survival signaling pathway [2, 46]. It was first identified by Stephen Staal in 1987 as viral oncogene [47]. Studies reveal that various cancer types involving leukemia, brain, breast, lung and colon are caused by activation of AKT due to genetic mutations of its own and/or its interrelated upstream and downstream proteins (PTEN, PI3K) [3].

Human AKT has three isoforms namely Akt1/PKBα, Akt2/PKBβ and Akt3/PKBγ. The three isoforms of human AKT share more than 85% similarity [4]. The sequence alignment of Akt1, Akt2 and Akt3 are shown in Figure 2.3.

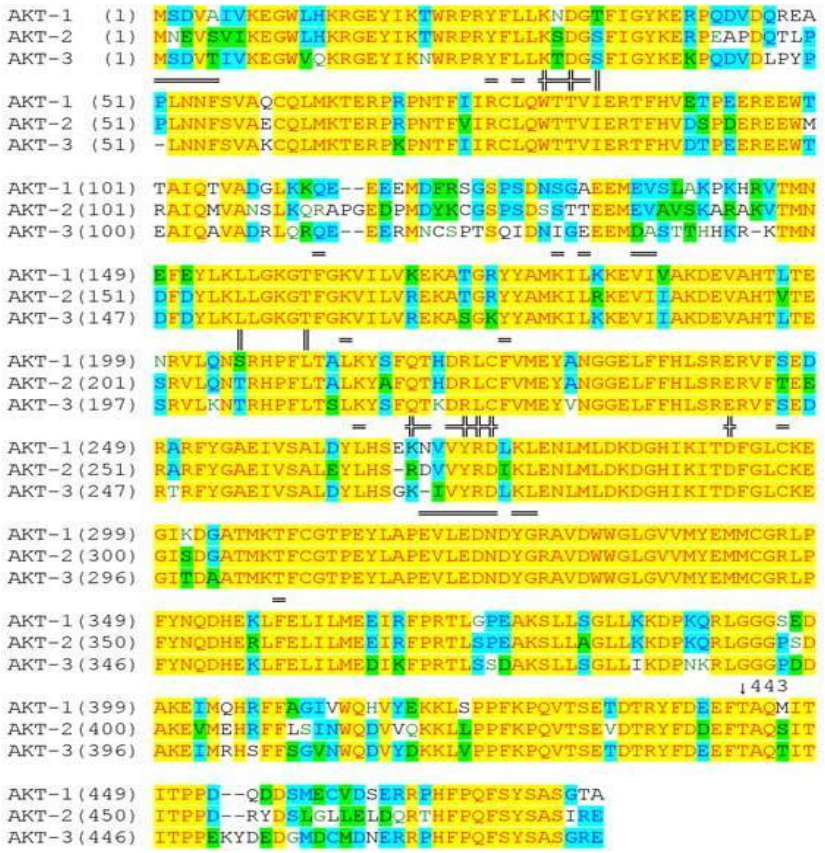
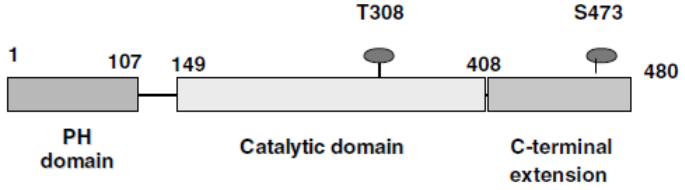



Figure 2.3. Sequence alignment of human Akt1, Akt2 and Akt3. Residues are colored if common in; all isoforms in yellow, two isoforms in cyan (Adopted from [17]).

AKT consist of three conserved domains, an N-terminal pleckstrin homology (PH) domain (residues 1-133), a central kinase catalytic (CAT) domain (149-408) and a C-terminal extension (EXT) domain (residues 408-480) Figure 2.4 [48].




| *PAIR* | *PH* | *LINK* | *CAT* | *EXT* |
|---|---|---|---|---|
| AKT1/AKT2 | 80 | 46 | 90 | 66 |
| AKT1/AKT3 | 84 | 40 | 88 | 76 |
| AKT2/AKT3 | 76 | 17 | 87 | 70 |

Figure 2.4. Residue numbers and pairwise identity (%) for three domains of AKT1 (Adopted from [10]).

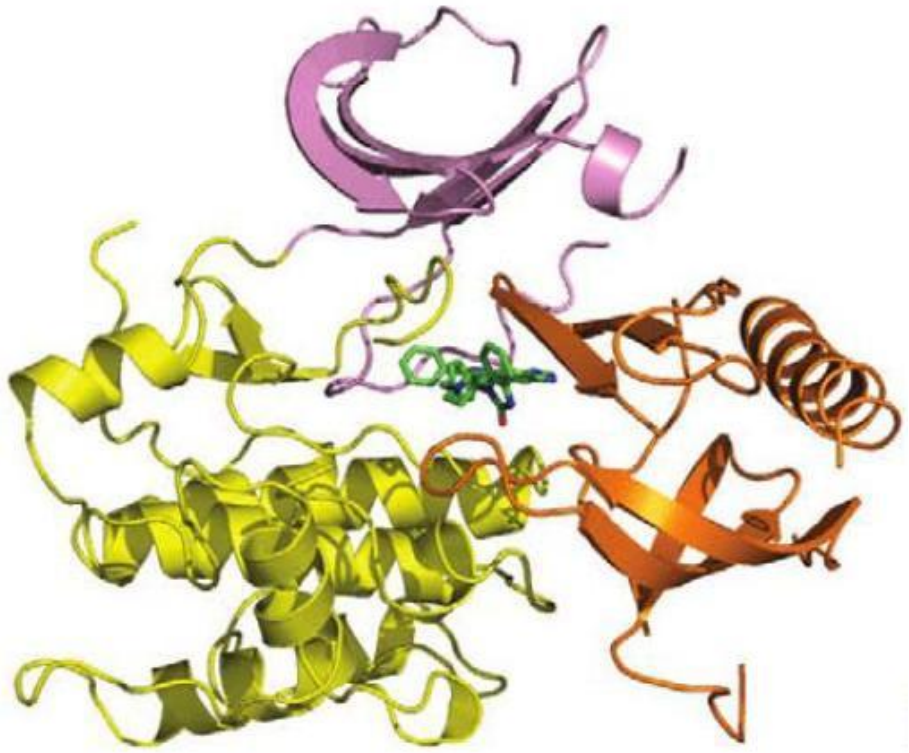


Figure 2.5. Cartoon representation of AKT1 with allosteric Inhibitor VIII (green); PH domain (orange), N-lobe (pink) and C-lobe (yellow) of the kinase domain (Adopted from [17]).

For all isoforms, PH domain and kinase domain are separated by a 39-amino acid hinge region [4]. For the kinase domains, binding pocket residues are the same [49]. Among the AKT isoforms, the PH domains are nearly 80% identical, CAT domain is nearly 90% identical and C-terminal extension (EXT) is nearly 70% identical [10]. Cartoon representation of PH and kinase domains is shown in Figure 2.5.

Each AKT isoform is associated with different types of cancers. For instance, AKT-1 is amplified in leukemia and breast cancer, AKT2 is amplified in pancreatic, breast and ovarian tumors. AKT-3 is over expressed in hormone insensitive breast and prostate cancers [50]. The downstream effectors of AKT and related diseases [5] are shown in Figure 2.6.

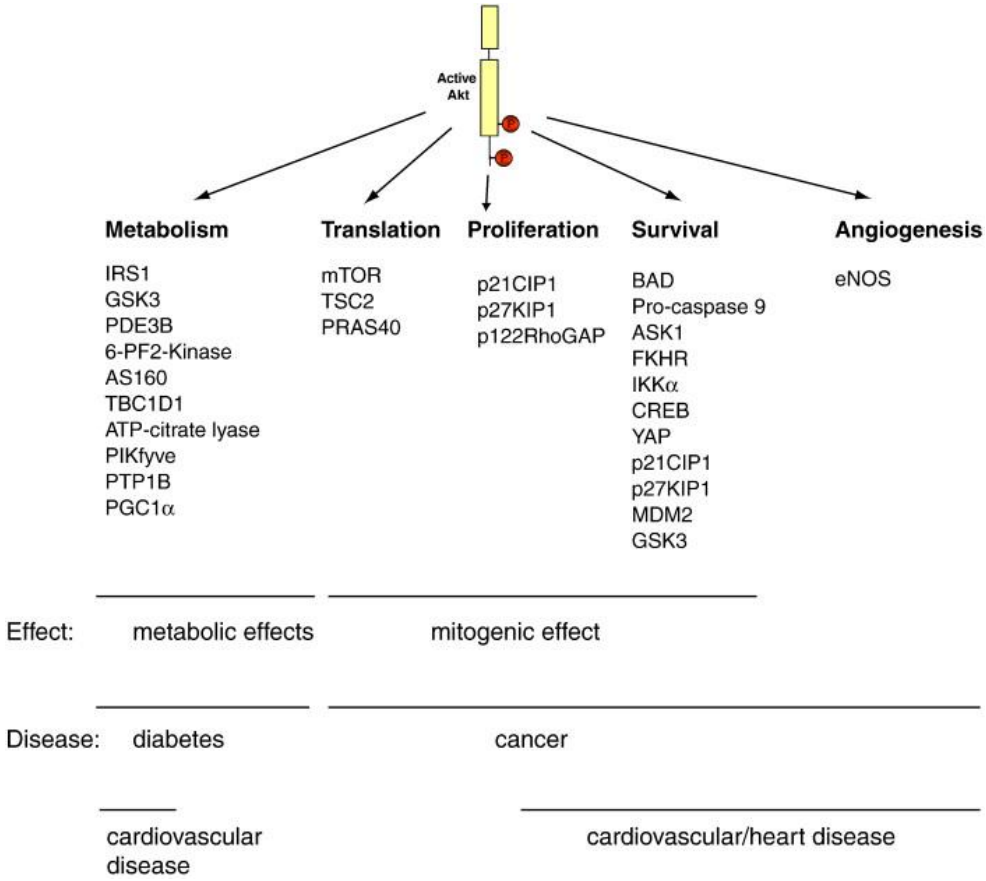



Figure 2.6. The downstream targets of AKT and their related diseases (Adopted from [7]).

Since AKT has a vital role for treatment of various cancer types related with (PI3K)/Akt pathway and its components, the recent progress in treatment of these cancer types is heading for identification of novel small molecule inhibitors targeting the AKT signaling pathway [51]. Types of small molecule inhibitors reported for AKT up to date differ in the target sites on the protein [11]. One of the three target sites for inhibition is ATP-binding pocket in catalytic kinase domain. Two other sites are called *allosteric* sites [10], either target directly the PH domain or the region between PH and kinase domains. The ATP-site targeting molecules represents the most studied and preclinically advanced AKT inhibitors described until now [52]. Although the focus has been on discovering potential ATP-competitive inhibitors targeting the kinase domain which resulted in crystallization of many ligands with their target protein (as potential drug candidates for further studies) available in Protein Data Bank (PDB), selectivity remained as a major issue since the kinase domains of different serine/threonine kinases are highly conserved with over 50% sequence identity [19, 20]. However, their N-terminal pleckstrin homology (PH) domains are divergent (less than 30% identity), that allows development of more

selective *allosteric* inhibitors which target the PH domain or the region between PH and kinase domains [21, 22].

## 2.4 PI3K/AKT Signaling Pathway

The phosphatidylinositol 3-kinase (PI3K)/ Akt pathway is one of these pathways whose activation through mutation, amplification and rearrangement occurs frequently in many human cancers and is associated with tumor growth, increased metastasis, and resistance to therapy [3]. Therefore, mutations of its components make either itself or its mediators appealing as a therapeutic target for cancer researchers [53]. An activating mutation in the pleckstrin homology domain of Akt1 (E17K) that results in growth factor independent membrane translocation of Akt and increased Akt phosphorylation levels was identified in melanoma, breast, colorectal and ovarian cancers [54-58]. The PI3K/AKT signaling pathway is represented in Figure 2.7.

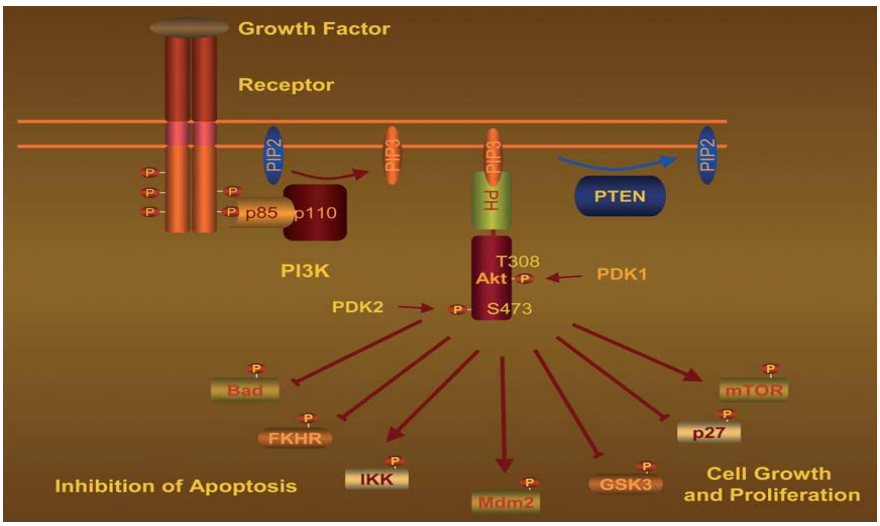



Figure 2.7. Schematic representation of PI3K/Akt pathway (Adopted from [59]).

Phosphatidylinositol-3 kinases (PI3K) are members of a major lipid kinase family and phosphorylate inositol ring in inositol phospholipids. PI3Ks have three classes according to their structural characteristics and substrate specificity [60]. Class I enzymes, which are heterodimers with a catalytic subunit (p110) and an adaptor/regulatory subunit (p85), and activated directly by cell surface receptors, are the most commonly studied one among the three classes. Class I PI3Ks are also divided into class IA enzymes and class IB

enzymes. Class IA enzymes are activated by receptors with protein tyrosine kinase activity (Receptor Protein Tyrosine Kinase, RPTK), GPCRs and certain oncogenes (small G protein RAS). Class IB enzymes are regulated by GPCRs [61].

Growth factor stimulation was followed by RPTK activation which results in the pairing of PI3K (p85) with the receptor and results in allosteric activation of the catalytic subunit (p110) of PI3K. PI3K activation leads to conversion of phosphatidylinositol-4,5-bisphosphate (PI-4,5-$P_2$, PIP2) to phosphatidylinositol-3,4,5-trisphosphate (PI-3,4,5-$P_3$, PIP3) by catalytic subunit p110 at the membrane. Binding to two distinct protein-lipid binding domains, namely, the FYVE [62] and pleckstrin homology (PH) domain [63] are required to observe the polyphosphoinositide effects on cells. PH domains are found in many proteins, including phosphoinositide-dependent kinase 1 (PDK1) and Akt/PKB [64]. Since PIP3 and PIP2 are restricted to the plasma membrane, Akt and PDK1 also translocate to the plasma membrane. The colocalization of activated Akt allows Akt to be phosphorylated by PDK1 on threonine 308 which leads to partial activation. Phosphorylation of serine 473 by mTOR protein kinase results in full activation of Akt [3]. Akt activation intercedes the inhibition and activation of several targets, leading to cellular survival, growth and proliferation [59].

## 2.5 Allosteric Mechanism of PKB/AKT Inhibition

Translocation of PKB to the plasma membrane due to the interaction of its PH domain with the lipid products of the PI3K (PIP2 and PIP3) is an important step for its activation [17]. Thr 308 in the activation loop of PKBα /Akt1 kinase domain and Ser 473 in the hydrophobic motif of the C-terminal part (Thr 309 and Ser 474 in PKBβ/Akt 2) are two regulatory sites necessary for AKT activation. The phosphorylation of both of these regulatory sites results in full activation of AKT [2].

The interaction of PH and kinase domains maintains the inactive conformation of AKT, called “PH-in” conformation in basal conditions. Prevention of Thr 308 phosphorylation is sustained by the “PH-in” conformation. Upon stimulation, Thr 308 phosphorylation is allowed by the conformation change to “PH-out” by interaction of PH-domain with phosphoinositides at the plasma membrane [65].

This interaction between two domains of AKT represents a novel mechanism for allosteric control. Molecular dynamics enabled the observation of a PH-induced cavity formation in the AKT kinase domain. Through this PH-induced cavity, while AKT remains in inactive form, Trp 80 residue is accessible from outside. A specific AKT inhibitor, namely AKT Inhibitor VIII, was docked into PH induced cavity of AKT1 and interacted with Trp80; maintained AKT in inactive form and prevented the phosphorylation of Thr 308 and Ser 473 and thereby blocking the full AKT activation [9].

Inactive conformer of AKT can also be maintained by the interaction of PH domain with the hydrophobic motif (HM) [66]. Trp 80 residue of the PH domain interacts with HM and secures the protein in the inactive conformation. The phosphorylation of first Ser 473 followed by Thr 308 upon conformation change would trigger the full activation of PKB [2].

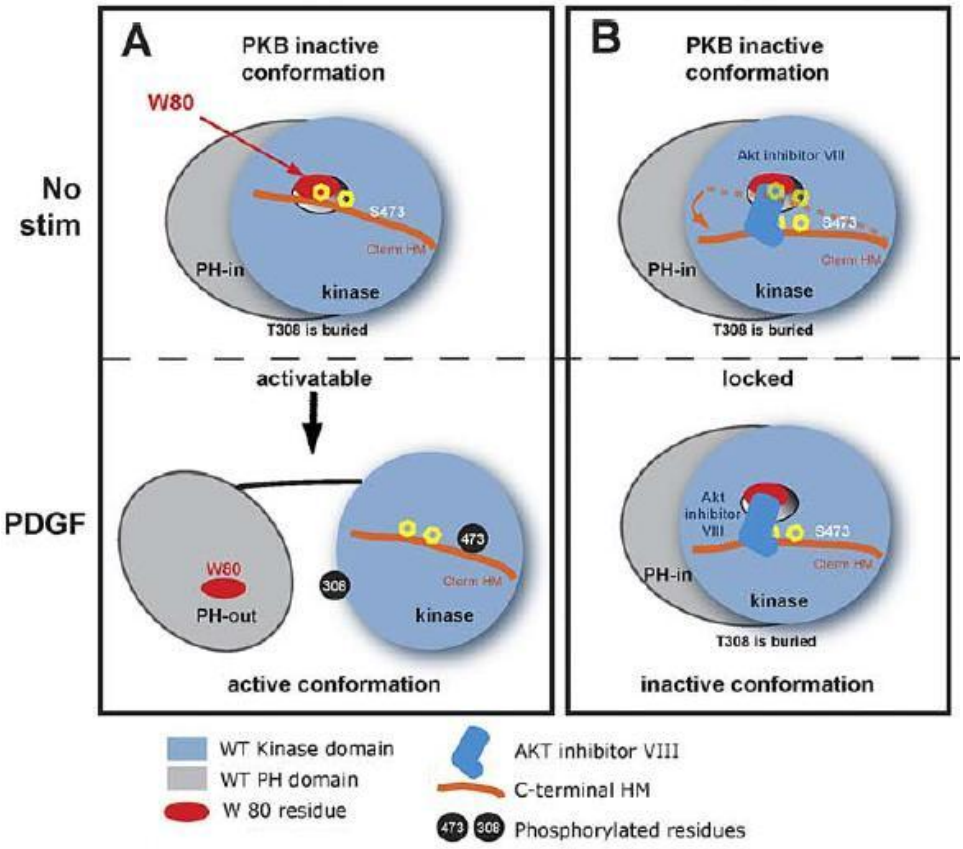



Figure 2.8. Schematic representation of active and inactive conformations of AKT1 and allosteric mechanism of AKT1 inhibition by Inhibitor VIII (Adopted from [9]).

In Figure 2.8 (A), PKB is in its inactive conformation (PH-in) before PDGF stimulation (upper panel). The PH-induced cavity created from the interaction of the PH domain and the kinase domain maintains regulatory residues Thr308 and Ser473 unphosphorylated by keeping PH domain residue Trp 80 (red oval) in interaction with the Phe469 of C terminal HM (orange). PDGF stimulation (lower panel) causes separation of PH and kinase domains (PH-out conformation), allowing the phosphorylation of the regulatory residues Thr308 and Ser 473. Figure 2.8 (B) displays allosteric AKT1 Inhibitor VIII bound to the Trp 80 in inactive conformation of AKT1 (upper panel). The position of C-terminal HM without inhibitor (dotted orange line) and with inhibitor (solid orange line) are represented. In the lower panel, AKT is locked in the PH-in conformation due to the binding of AKT inhibitor VIII. Phosphorylation of Thr308 upon PDGF stimulation is restricted by prevention of the conformational change. The phosphorylation of Ser 473 is also prevented due to the binding of AKT inhibitor VIII in the same region as HM.

## 2.6 Previous Studies on Allosteric Inhibiton of AKT1

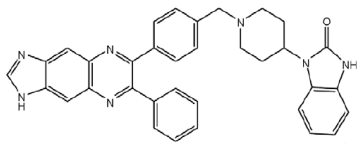
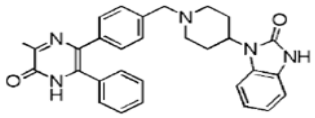
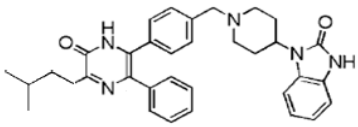
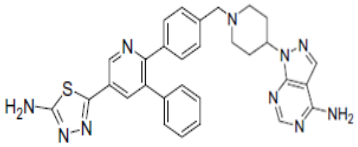
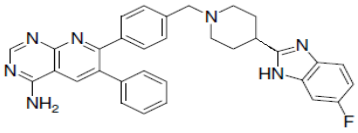
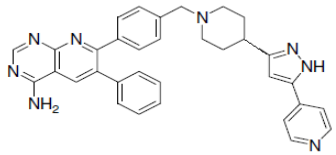
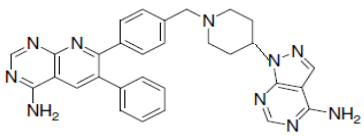
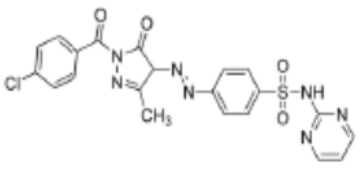
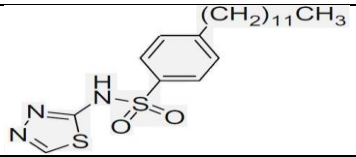
Some of the studies focusing on developing allosteric inhibitors for AKT by experimental or in silico techniques conducted by different research groups are summarized below in Table 2.1.

Lindsley *et al.* (2005) [4] employed an iterative analog library synthesis approach to develope two novel series of selective, allosteric Akt kinase inhibitors capable of inhibiting the three Akt isozymes based on either a 2,3-diphenylquinoxaline core or a 5,6-diphenylpyrazin-2(1H)-one core derived from a high throughput screening campaign. Their HTS leads displayed no inhibition in assays with Akt mutants lacking the PH domain; suggesting inhibition of Akt through an allosteric binding site which was not competitive with ATP binding site.

Cuny *et al.* (2009) [67] used in silico approaches to develop novel inhibitors targeting the pleckstrin homology domain of Akt1. They performed molecular docking, in silico ADME studies and QSAR study to guide chemical design and lead optimization. GOLD docking and scoring was used to analyze the interactions between the inhibitors and the Akt1 PH domain. Based on the QSAR study and metabolism predictions, their optimized

inhibitor was experimentally validated with significant in vitro and in vivo antitumor activity.

Table 2.1. Experimentally identified allosteric AKT1 inhibitors.

| **Assigned name** | **Isozyme selecti-vity** | **Chemical structure** | **AKT1 $IC_{50}$** | **AKT2 $IC_{50}$** | **Experi m./Co mput.** | **Ref.** |
|---|---|---|---|---|---|---|
| **INH. VIII** | AKT1 |  | 58 nM | 210 nM | E | [17] |
| Comp.1 | AKT1 |  | 760 nM | 24000 nM | C&E | [4] |
| Comp2 | AKT1 |  | 212 nM | 325 nM | C&E | [4] |
| Comp3 | AKT2 |  | 126 nM | 22 nM | E | [16] |
| Comp4 | AKT1 |  | 18 (+/- 3) nM | 239 nM | E | [68] |
| Comp5 | AKT1 |  | 21 (+/- 6) nM | 85 nM | E | [68] |
| Comp6 | AKT1 |  | 14 (+/- 0.1) nM | 99 nM | E | [68] |
| Comp7 | AKT1 |  | 13 µM | - | C&E | [48] |
| Comp8 | AKT1 |  | 6.3 µM | - | C | [67] |

Mahadevan *et al.* (2008) [48] applied in silico library screening and molecular docking to identify a novel class of non-lipid based compounds which bind selectively to the PH domain of Akt1. Some compounds showed PH domain binding selectivity for Akt against IRS1 and PDK1 and induced inhibiton of Akt in cells in addition to inhibition of cancer cell proliferation. However, their lead compound was not reported to achieve blood concentrations required to inhibit AKT in cells which was attributed to rapid metabolism and elimination.

In the study of Calleja *et al.* (2009) [9], using a multidisciplinary approach including molecular modelling, classical biochemical assays, and Förster resonance energy transfer (FRET) and by determining the spatial organization of different regions of the protein in inactive Akt, they discovered a cavity at the interface of kinase and PH domains of the inactive form. They suggested a new mechanism of allosteric control of Akt by the interaction of the PH domain with kinase domain that had not been observed so far. In another study of Calleja *et al.* [2], the novel mechanism of action of a highly specific protein kinase B inhibitor, AKT inhibitor VIII was elucidated, suggesting that simultaneous binding of the inhibitor to both kinase and PH domains through the cavity ‘‘locks’’ protein kinase B in an inactive conformation by interacting with Trp 80 and prevents the phosphorylation of Thr 308 and Ser 473 resulting in blocking the full Akt1 activation.

Another study was conducted by Wu *et al.* (2010) [17], proposing a 2.7Å resolution co-crystal structure of human Akt1 containing both the PH and kinase domains with a selective allosteric inhibitor; Akt Inhibitor VIII, bound in the interface (PDB code: 3O96). The structure reveals the interactions between the PH and kinase domains, as well as the critical amino residues that mediate binding of the inhibitor to Akt1.

Ashwell *et al*. (2012) [69] employed biochemical and biophysical screening strategy for identification and optimization of small molecule Akt1 inhibitors which target the same allosteric site with Inhibitor VIII. They crystallized the unphosphorylated Akt1 structure at 2.25Å (pdb code: 4EJN) and proposed a series of allosteric Akt inhibitors which potently inhibit intracellular Akt activation in vitro. Results of their in vivo pharmacodynamic and pharmacokinetic studies with two structures (12e, 12j) of series indicated that their series

are effective at inhibiting the activation of Akt and an additional downstream effector (p70S6) following oral dosing in mice.

## 2.7 Computer Aided Drug Design by Virtual Screening

Computer aided drug design (CADD) is performed by molecular modeling software and applications to discover or study drugs candidates and related biological molecules. The goal is to reduce the overall cost and shorten the required time associated with the discovery and development of a new drug, by identifying the most promising candidates for lead optimization and experimental testing. CADD can be categorized into three major types; structure-based drug design, ligand-based drug design and molecular dynamics simulations [70, 71].

Molecular dynamics simulations allow atoms and molecules to interact for a period of time followed by calculation of the motion for every atom, can be played to examine the overall behavior and provide detailed information on the fluctuations and conformational changes of proteins [72-74]. Structure and ligand based approaches can be explored within the scope of virtual screening, which is defined as a computational technique that corresponds to quick search of large libraries of chemical structures in order to identify those structures which are most likely to bind to a target receptor, typically a protein [75, 76].

Structure-based virtual screening requires the three-dimensional knowledge of the receptor structure and achieved by identifying complementary orientations of small molecules (ligands) in the binding site of a target receptor and to evaluate the located poses with scoring functions (docking) [77, 78]. Throughout docking process, ligand flexibity is considered by using an ensemble of conformations and target protein is held fixed [79-81]. Some of the available structure based virtual screening softwares are, MOE, FlexX, FRED, GOLD and Glide.

Ligand-based virtual screening takes the knowledge of ligands that interact with target receptor into consideration. This approach either employs statistical techniques to relate structural features to biological activities in order to identify specific structural features of a candidate ligand required for interacting with its target (QSAR); or identifies

and uses information of 2D/3D similarities between a set of structurally diverse ligands that binds to the target receptor (pharmacophore modeling) [82]. A pharmacophore is defined as "an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response" by IUPAC. Pharmacophore models can be built by using ligand information or binding site information of target receptor [83, 84].

# 3. METHODS

For all the virtual screening studies, products of Schrödinger's Small Molecule Drug Discovery Suite (2011) were used with Linux operating system on a HP xw6600 Workstation. The built in programs included in Schrödinger's Suite were; Protein Preparation Wizard [85] for structure preparation of the receptor protein, LigPrep [86] for ligand preparation, ConfGen [87, 88] for generating multiple conformations of ligands, QikProp [89] for ADME predictions of drug candidates, Phase [90-92] for database preparation, e-Pharmacophores [83, 84] for structure based pharmacophore modeling and Glide [93, 94] for binding site determination of the receptor protein and ligand docking. Maestro was the unified interface for all Schrödinger software. The workflow of the methodology in this study is given in Figure 3.1.

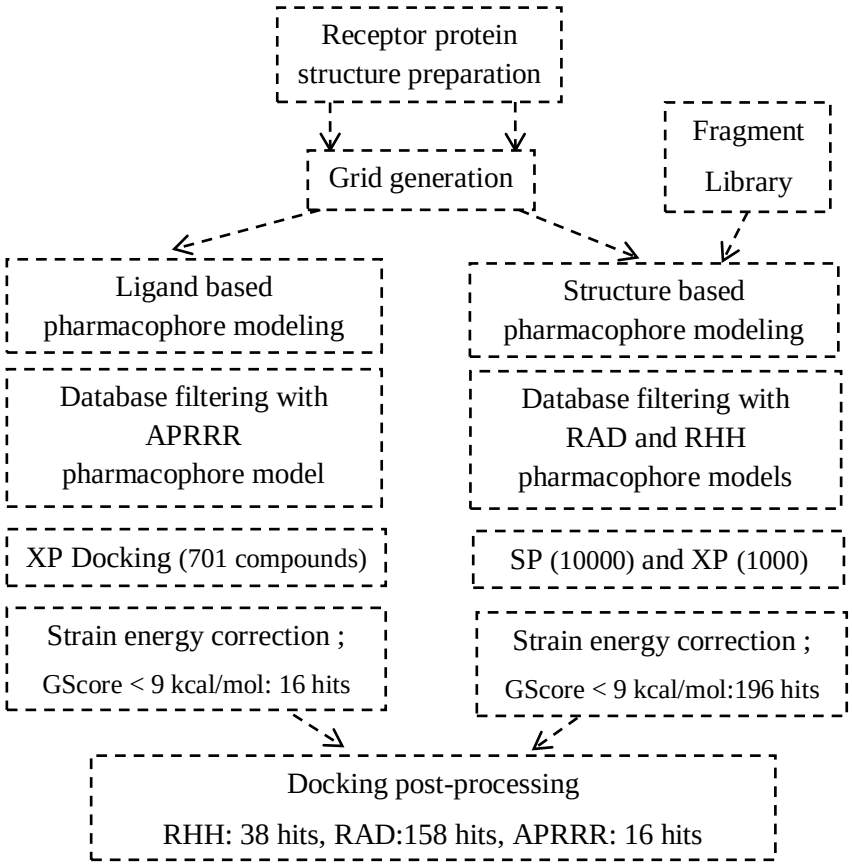



Figure 3.1. Workflow of the methodology.

## 3.1 Receptor Protein Structure Preparation

The structure file from the Protein Data Bank (PDB) was prepared for use in the molecular modeling calculations. In the PDB, there was only one AKT1 structure with the ligand bound to the allosteric region between kinase and PH domains (PDB code: 3O96). Protein Preparation Wizard (PPW) [85] of Schrödinger Suite was used to add hydrogens and to assign correct charge states. Water molecules beyond 5Å from the protein were deleted. The bond orders were assigned and missing hydrogens were added. Coordinates of missing loops (residue numbers: 45-49, 88-92, 188-199, 298-313) were predicted using Prime module. The coordinates for the loop between residues 113-145 could not be predicted since Prime is limited to predict loops with fewer than 20 residues. This loop was localized on the outer part of the protein and is not in the vicinity of the binding site therefore its presence was not considered to be crucial for docking calculations.

After fixing any structural defects, the structure was inspected and unwanted parts of the structure, such as water molecules, were deleted. There was only Chain A and native ligand (code: IQO) for the protein, and Water 455 was kept for further calculations due to its bridging function. Before optimization of hydrogen bonding network, protonation states were generated for the native ligand between pH range of 7.0±4.0 by Epik module [95]. Five different protonation states were generated and sorted by increasing state penalty, which includes a reward for the number of hydrogen bonds formed. State1 was chosen with the lowest state penalty of 0.18 kcal/mol and charge of +1. Remaining protonation states had higher state penalties of 4.68, 4.96, 5.39 and 5.59 kcal/mol.

The orientation of hydroxyl or thiol groups, the terminal amide groups in asparagine (Asn) and glutamine (Gln), and the ring of histidine (His) cannot be determined from the X-ray structure. Therefore, upon choosing the protonation state, hydrogen-bonding network was optimized by reorienting hydroxyl and thiol groups, water molecules, amide groups of asparagine and glutamine, and the imidazole ring in histidine. Optimizing the orientation of these groups is an iterative process, which passes over all the groups whose H-bonds need to be optimized for multiple times. Protonation states of histidine, aspartic acid and glutamic acid and tautomeric states of histidine were predicted at neutral pH.

As a final step, the structure is refined by minimization. Heavy atoms were restrained during minimization. For minimization, Optimized Potentials for Liquid Simulations (OPLS_2005) force field was chosen with the maximum allowed RMSD of 0.30Å for heavy atoms (default value). The protein preparation process was completed after this minimization step and the final structure was saved with (.mae) extension to be used in other Schrödinger applications. Workflow of the protein preparation process is shown in Figure 3.2.

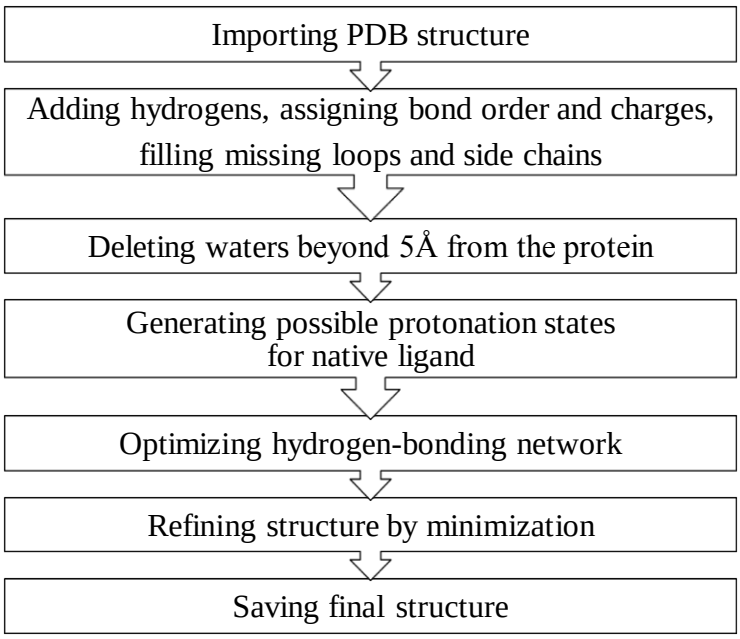



Figure 3.2. Protein Preparation Process Workflow.

## 3.2 Receptor Grid Generation

Binding site of protein is defined by a grid via several different sets of fields that provide progressively more accurate scoring of the ligand poses. Each field on the grid is identical in size and represents the shape and properties of the active site. Receptor Grid Generation panel in Glide [93] software of Maestro interface was used for grid generation of receptor protein. The protonated, optimized and minimized output structure obtained from Protein Preparation Wizard was used as input structure in this stage. Four tabs exist in Receptor Grid Generation panel which are receptor, site, constraints and rotatable groups.

In the Receptor tab, ‘pick to identify ligand’ option was selected prior to picking the ligand in the workspace. In this way, the ligand was excluded from grid generation since everything not defined as the ligand is treated as part of the receptor by Glide. Receptor is not allowed to be flexible; however, van der Waals radii of non polar atoms, which are represented as Van der Waals spheres, were scaled in order to decrease penalties for close contacts and earn a little flexibility. Since the binding site is a cavity which is buried between two domains of the protein, default scaling value of 1.0 (no scaling) was reduced to 0.8 with 0.25 partial charge cutoff to relax the binding site.

In the Site tab, position of the scoring grids and the way they were prepared from the structure in the workspace were determined. Glide uses two cubic boxes to define where the energetic calculations should be done. The two boxes share a common center. The ‘outer’ or ‘enclosing’ box is the larger one that contains all atoms of the binding ligand. The ‘inner’ or ‘ligand diameter midpoint box’ is the other one, in which the acceptable positions for the ligand center during the site point search should lie. In a rigid docking run, ligand center is defined as the midpoint of the line drawn between the two most widely separated atoms. The outer box should be large enough to contain all atoms of a possible ligand when the ligand center is placed at an edge or vertex of the inner box. Too small box sizes may restrict the number of possible ligands to be bound whereas too large boxes need more space both on disk and in memory for the scoring grids and results in requirement for longer computational time. Glide allows users to define the box sizes manually, however since our receptor contains a native ligand in the workspace, ligands at similar size to workspace ligand were chosen to be docked. The center of the outer box (20Å) was chosen as the centroid of the workspace ligand and dimensions of the inner box was left as default (10Å).

In the Constraints tab, four types of constraints can be set to be applied in docking calculations; positional, H-bond, metal and hydrophobic. Constraints that are intended to be applied in docking should be defined during receptor grid generation step. No constraint was set since no important receptor-ligand information that could be set as a constraint among these four types was attained from literature.

In the Rotatable Groups tab, hydroxyl groups of residues such as Ser, Thr and Tyr can be allowed to adopt different orientations when different ligands are docked in order to produce the most favorable interaction. Two rotatable groups were detected on Tyr 263 and rotation was not allowed since Tyr 263 was not located in a close region to active site.

### 3.3 Generation of 3D Database

Preparation of the database to be screened was performed in the 'Generate Phase Database' panel which is an application of Phase software. ZINC, which is a free database provided by the Shoichet Laboratory in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF) containing over 21 million purchasable compounds for virtual screening in ready-to-dock 3D formats, was chosen as the source of ligands [96]. Among the libraries containing purchasable and drug-like compounds provided by 236 vendors available in ZINC database, 38 largest compound libraries with a total of 499,026 compounds were obtained in SDF format to be given as input. The names of vendors and the number of unique, drug-like molecules are presented in Table 3.1. The access date for all vendors is 05/2012.

Compounds supplied by vendors were prepared in a few steps. Ionization and tautomer states of molecules were generated using Epik [95] utility at a pH range of 7±2. Highly energetic states were removed. Stereochemical information was obtained from the 3D geometry of the structures. Starting chiralities were selected by LigPrep [86], based on the chemistry of naturally occurring steroids, fused rings and peptides. Up to 32 stereoisomers were generated and then filtered to retain those with the lowest energies. Conformers were generated by ConfGen [87] utility with a maximum conformer number of 100 per compound and 10 per rotatable bond. Energy minimization was carried out in 50 steps for the conformers using the OPLS_2005 force field and a distance dependent dielectric constant of 4 was left as default.

Although compounds which were classified as 'drug-like' were obtained from ZINC, filtering was applied based on ADME (absorption, distribution, metabolism, excretion) properties and Lipinski's Rule of Five. ADME properties were predicted by QikProp module [89] and include; SASA (total solvent accessible surface area in $Å^2$),

FOSA (hydrophobic component of the SASA, associated with a saturated carbon and attached hydrogen), FISA (hydrophilic component of the SASA, associated with the solvent accessible surface area on N, O, and H on heteroatoms), PISA (π component of the SASA, π is referred to a carbon and attached hydrogen), WPSA (weakly polar component of the SASA, takes the polar components as electronegative atoms such as nitrogen and oxygen; and halogens, phosphorus and sulphur atoms), PSA (van der Waals surface area of polar nitrogen and oxygen atoms), percentage of human oral absorption (HOA%). Lipinski's Rule of Five is used to evaluate drug-likeness or determine if a chemical compound with a certain pharmacological or biological activity has properties that would make it a likely orally active drug in humans. According to Rule of Five, an orally active drug has at most one violation of the following criteria [97] :

- At most 5 hydrogen bond donors (nitrogen or oxygen atoms with one or more hydrogen atoms)
- At most 10 hydrogen bond acceptors (nitrogen or oxygen atoms)
- A molecular mass less than 500 daltons
- An octanol-water partition coefficient (log P) not greater than 5

Structures which did not satisfy this rule were not added. It should be noted that, to pre-filter with Rule of five, QikProp should be run to predict the necessary properties. Additionally, structures that had reactive groups were removed by ligfilter script. Pre-defined set of reactive groups are; Acyl halides, Phosphinyl halides, Phosphines, Sulfonyl halides, Phosphonyl halides, Alkyl sulfonates, Sulfinyl halides, Alkali metals, Epoxides, Sulfenyl halides, Alkaline-earth metals, Azides, Alkyl halides without fluorine, Lanthanide series metals, Diazonium compounds, Anhydrides, Actinide series metals, Isonitriles, Perhalomethylketones, Transition metals, Halopyrimidines, Aldehydes, Other metals, 1,2-Dicarbonyls, Formates, Toxic nonmetals, Michael acceptors, Peroxides, Noble gases, Beta-heterosubstituted carbonyls, R-S-O-R, Carbodiimides, Diazo compounds, Isothiocyanates, Silyl enol ethers, R-N-S-R, Isocyanates, Nitroalkanes, Disulfides.

Finally, 499,026 compounds resulted in ~26 million conformations after filtration. Database was saved as Phase database and ready to be used with extension 'phasedb'.

Table 3.1. Names of vendors and number of unique drug-like compounds provided.

| Vendor name | Number of unique, drug-like compounds | Vendor name | Number of unique, drug-like compounds |
|---|---|---|---|
| Acros Organics | 3023 | Matrix Scientific | 1419 |
| Aldrich CPR | 29902 | Maybridge | 3557 |
| Amadis Chemical | 4836 | Molport BB | 4162 |
| American Custom Chemicals Corp. | 19284 | NCI Plated 2007 | 21235 |
| Anward | 3067 | Oakwood Chemical | 6024 |
| Apollo Scientific | 4291 | Otava | 11580 |
| Asinex | 16992 | PBMR Labs | 21679 |
| AsisChem | 16814 | Pharmeks | 24183 |
| BioSynth | 2240 | Princeton BioMolecular Research | 15233 |
| ChemBridge | 9051 | Ryan Scientific BB | 6300 |
| ChemDiv | 43579 | Specs | 15224 |
| Chemical Block | 9765 | Synton-Lab | 2645 |
| Enamine | 5436 | TCM Database @ Taiwan | 17693 |
| Florida Heterocyclic Compounds | 3467 | TimTec | 10686 |
| Fluorochem | 5303 | Toslab | 4935 |
| Frontier Scientific Services | 4173 | Tractus | 4632 |
| IBScreen | 13943 | UORSY | 26441 |
| Innovapharm | 20332 | Vitas-M | 63822 |
| Life Chemicals | 6772 | Zelinsky Institute | 15306 |
| Total number of entries = 499,026 | | | |

### 3.4 Structure Based Pharmacophore Modeling (e-Pharmacophores)

For allosteric inhibiton of AKT, only a single crystallized protein-ligand complex existed in PDB at the time this study initiated. Due to lack of a set of ligands ready to be used for ligand-based pharmacophore modeling in PDB, initially, structure based pharmacophore modeling based on allosteric binding site of AKT receptor was applied by using E-pharmacophores script [83, 84] which is supported by Phase and Glide softwares.

Initially, the fragment library obtained from Schrödinger [93] was docked to grid generated protein structure. This fragment library consists of 441 unique small fragments (1-7 ionization/tautomer variants; 6-37 atoms; MW range 32-226) derived from molecules in the medicinal chemistry literature. The set includes a total of 667 fragments with accessible low energy ionization and tautomeric states and metal and state penalties for each compound from Epik. The fragments were docked using extra precision mode of Glide without constraints and results were saved on Glide XP descriptor file. Glide XP descriptor file, which carries the information about fragment binding modes to receptor and energetic terms, was then given as input for E-pharm script.

E-pharm script used the information stored in descriptor file to sum the energetic contributions that construct GlideScore of all fragments on each atom center. The summed values allowed detection of the possible pharmacophore sites on atoms. These sites are ranked based on their energies and best ranked set of sites was chosen to develop pharmacophore hypothesis. There are six features checked in Phase which are defined as a ‘site’ ; hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively charged group (N), positively charged group (P), and aromatic ring (R). Phase allows to construct a hypothesis with pharmacophoric sites changing between three and seven. Two hypotheses were constructed; RHH and RAD. Finally, ZINC database was filtered separately using RHH and RAD hypotheses to reduce the number of structures to be docked in the next stage. The number of structures in database were reduced to 10,000 after filtering with both hypotheses.

## 3.5 Ligand Based Pharmacophore Modeling

To construct a ligand based pharmacophore hypothesis, seven small molecule allosteric Akt1 inhibitors that have been obtained from several studies in literature [4, 16, 17, 98] were used. The information of these ligands was used to identify a set of pharmacophore site points (features) and model hypothesis. Develop Pharmacophore Model workflow of Phase module, which has four steps of Prepare Ligands, Create Sites, Find Common Pharmacophores and Score Hypothesis, was used.

In 'prepare ligands' step, 2D structures of the ligands were drawn manually, followed by conversion to low-energy 3D structures with LigPrep module [86], where hydrogen atoms were added, protonation was done to produce ionization states at neutral pH, maximum 32 stereoisomers were generated for each ligand, and an energy minimization was performed to generate maximum 100 conformers for each ligand by ConfGen [87].

The pharmacophore sites for all ligands were created in the second step. As mentioned previously, the six pharmacophore features provided by Phase to define sites were; hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively charged group (N), positively charged group (P) and aromatic ring (R). Hydrophobic sites included rings, halogenic moieties, isopropyl groups or t-butyl groups and chains having at most 4 carbons. Aromatic ring sites were localized at the center of each ring. Negatively and positively charged group sites were located on the charged atom of the corresponding group.

Next, common pharmacophores were identified among pre-created sites. Common pharmacophores are defined as the pharmacophores that have the same feature with the same spatial arrangement. Identical sites were grouped, and if a group had at least one pharmacophore from each ligand in the set, it was taken as a common pharmacophore. Phase constructs a variant list for the common pharmacophores, which each 'variant' includes a number of feature types. The number of the pharmacophore features in a variant can be minimum three and maximum seven. In this study, minimum four and maximum seven site points were selected to be included in a variant and all seven ligands were set to

match each site point. With these criteria, twelve variants were identified in the variant list. These twelve variants allowed detection of 413 hypotheses with seven common pharmacophore models including five pharmacophore sites (AADRR, PRRRR, APRRR, DRRRR, AAADR, AAPRR, ADRRR). The number of hypotheses contributed to each common pharmacophore model varied between 1 and 129. The model which had the highest number of hypotheses contribution, APRRR, was chosen to be scored in next step.

In 'Scoring Hypotheses' step, common pharmacophore models were scored to find optimum alignments to active set of inhibitors. Assessment of the models was based on survival score; which is a combination of (default weighting factor of 1.0) vector score, site score and volume score. Vector score is the average cosine of the angles formed by corresponding pairs of vector features (acceptors, donors, and aromatic rings) in the aligned structures. Site score is the root-mean-squared deviation in the site-point positions. Volume score is based on the overlap of van der Waals spheres of the non-hydrogen atoms in each pair. After scoring, the APRRR hypothesis with the third highest survival score was selected since the first two hypotheses exhibited low fitness to native ligand (Inhibitor VIII).

## 3.6 Docking and Scoring

Glide (Grid-based Ligand Docking with Energetics) software [93, 94] was used for docking. Favorable interactions between the ligands in the pre-filtered database and receptor protein, AKT1 were searched using rigid docking. Flexibility of ligands was provided by generating conformations for each input ligand. The ligand-receptor interactions of poses are assessed by a series of hierarchical filters by Glide. The filtering begins with investigation of the spatial fit of the ligand pose to pre-defined binding site generated in the grid generation stage. The final step of the algorithm is composed of energy-minimization of the poses using OPLS-AA nonbonded ligand-receptor interaction energy followed by scoring. For scoring, Glide's own empirical scoring function GlideScore, which is based on ChemScore [99] was used.

### 3.6.1 Glide Docking Hierarchy

For each ligand conformation, initially, ‘site points’, which are on equally spaced (2Å) grid, are selected in the binding site. This selection is done by grouping both receptor surface to site point and ligand surface to ligand center distances separately with a width of 1Å, followed by comparing distance intervals from the site point to the receptor surface and from the ligand center to the ligand surface. If the match is acceptable, Glide places the ligand center at the site point, otherwise discards the site point. The second hierarchical step is investigation of location of atoms that stays within an acceptable distance from the line between the most widely separated atoms, which is called as ‘ligand diameter’. The orientation is discarded if many steric clashes exist with the receptor. Next, atoms, which are capable of making hydrogen bonds with the receptor are scored, and if this score is acceptable, all interactions are scored. This step is called ‘greedy scoring’ since atoms are not only scored considering their actual x,y,z coordinates relative to receptor; they are also forced to move ±1Å in all coordinates in order to achieve a better score. Schrödinger’s discretized version of the ChemScore empirical scoring function, namely GlideScore, was used in all scoring steps. The third hierarchical step consists of energy minimization with rigid-body translations and rotations. Energy minimization is performed on the pre-computed OPLS-AA van der Waals and electrostatic grids for the receptor. Top ranked poses are then sampled which improves the energy score by inspection of torsional motion about the ligand core and end-group rotatable bonds. In the final hierarchical step, energetically minimized ligand poses are re-scored with GlideScore which takes ChemScore as basis and adds a steric clash term, rewards and penalties such as buried polar terms to penalize electrostatic mismatches, amide twist penalties, hydrophobic enclosure terms, and excluded volume penalties. Energetic contributors of GlideScore with their weighting factors are given in Equation 3.1 and descriptions of the contributors are represented in Table 3.2.

$$GScore = 0.05 \times vdW + 0.15 \times Coul + Lipo + Hbond + Metal + Rewards + RotB + Site$$

Equation 3.1

Table 3.2. GlideScore components [93].

| Component | Description |
|---|---|
| vdW | Van der Waals energy. This term is calculated with reduced net ionic charges on groups with formal charges, such as metals, carboxylates, and guanidiniums. |
| Coul | Coulomb energy. This term is calculated with reduced net ionic charges on groups with formal charges, such as metals, carboxylates, and guanidiniums. |
| Lipo | Lipophilic term derived from hydrophobic grid potential. Rewards favorable hydrophobic interactions. |
| Hbond | Hydrogen-bonding term. This term is separated into differently weighted components that depend on whether the donor and acceptor are neutral, one is neutral and the other is charged, or both are charged. |
| Metal | Metal-binding term. Only the interactions with anionic or highly polar acceptor atoms are included. If the net metal charge in the apo protein is positive, the preference for anionic or polar ligands is included; if the net charge is zero, the preference is suppressed. |
| Rewards | Rewards and penalties for various features, such as buried polar groups, hydrophobic enclosure, correlated hydrogen bonds, amide twists, and so on. This category covers all terms other than those explicitly mentioned. |
| RotB | Penalty for freezing rotatable bonds. |
| Site | Polar interactions in the active site. Polar but non-hydrogen-bonding atoms in a hydrophobic region are rewarded. |

### 3.6.2 Standard Precision and Extra Precision Modes for Glide Docking

Glide provides three different docking options; high-throughput virtual screening (HTVS), standard precision (SP) and extra precision (XP). Difference of these modes are their scoring functions and sampling algorithms. HTVS mode scans very large number of ligands rapidly with a highly restricted conformational sampling than other two modes and advanced settings are not available. SP docking scans ligands of unknown quality in large numbers with moderate rapidity. XP docking is the most discriminating mode and includes

additional terms over the SP scoring function, therefore requires considerably more CPU time.

In XP docking, optimization of the parameters for solvation and sampling are more detailed compared to SP. Penalties are assigned to structures in which one or more groups are inadequately solvated. Although descriptors are also used in Glide SP scoring for solvation, as a result of improved sampling in XP docking, higher penalties are assigned to serious violations of physical principles. XP GlideScore specifically rewards occupancy of well-defined hydrophobic pockets by hydrophobic ligand groups. Hydrophobic reward terms are also employed in empirical scoring functions such as ChemScore and the SP version of GlideScore in the form of lipophilic-lipophilic pair terms. However, simple pair terms underestimate hydrophobic effects and the hydrophobic term in GlideScore XP was developed to offset this underestimation. XP GlideScore also includes improvements to the scoring of hydrogen bonds as well as detection of buried polar groups, and detection of pi-cation and pi-pi stacking interactions.

### 3.6.3 Virtual Screening Workflow

To run an entire sequence of jobs for docking, Virtual Screening Workflow (VSW) was used. Using this workflow, pre-filtered ZINC database was docked to grid-generated AKT1 structure with with SP docking followed by XP docking. If SP docking and XP docking are run separately, a minimization should be performed on ligands to relieve their strains between two docking modes. An advantage of VSW is that, the ligands are passed from one stage to the next one without any intervention. Another advantage is that, if a subjob fails, the master job attempts to rerun the job a few times before quitting. If the master job fails due to a system or network failure, it picks up the calculation from the latest point for which it has results and can restart from them.

Preparation of database ligands and filtering according to Lipinski's rule were already done previously, therefore these options were skipped. For receptor to be docked, pre-generated grid file and for ligand source, ZINC database filtered with structure-based pharmacophore model were chosen. No constraints were applied. Epik state penalties for ligands were incorporated to the docking score. Docking and scoring of atoms having more

than 300 atoms and 50 rotatable bonds were not allowed. To decrease the penalties for close contacts, vdW radius of nonpolar atoms, which have less than 0.15 partial elementary charges ($1.602\times10^{-19}$ coulombs), were scaled by 0.8. For both SP and XP, flexible docking method was chosen with penalties on non-planar conformations for amide bonds and post-docking minimization was performed. Additionally, for both docking modes, ten poses per compound state, at maximum, were generated and only the best scoring state was kept in the final hit list. After SP docking, top 10% of the docked ligands according to SP GScore was allowed to pass to XP docking stage. For XP docking, XP Descriptor information was saved and all docked ligands (100%) were chosen to be kept in the final hit list. Completing all the selection of options and parameters, job was started. Output files were created in subdirectories of the working directory. The path to the output files was given at the end of the *log file*. As output, a *pose viewer* file (_pv.mae) was created, containing the receptor AKT1 and the final hits of docked ligands.

### 3.7 Post-Processing of Docking by Strain Rescoring

Glide docking fits the ligands into a fixed receptor structure with scaled vdW radii, thus some ligand strain is allowed in order to compensate for the lack of receptor flexibility. After docking, the scaling was removed and Strain-Rescore script was applied to output pose viewer file to identify ligands with strain. Ligands with more than 4 kcal/mol energy difference (default value) between the docked and free conformations received penalties and quarter of this energy difference was added to the GlideScore by MacroModel utility [100]. Strain-corrected GlideScore values were used for further evaluation of the hits.

### 3.8 Similarity Analysis

Investigation of similarity between docked ligands to AKT1 structure was conducted to determine a diverse set of chemical scaffolds as potential inhibitors using ChemMine Tools [101] which is a free online service for analyzing and clustering small molecules by structural similarities or physicochemical properties.

Two dimensional similarity of ligands was identified by using atom pairs as structural descriptors. ‘Atom pairs’ is a structural descriptor type, defined by the shortest paths among the non-hydrogen atoms in a molecule. Each path is described by the types of atoms in a pair, the length of their shortest bond path, the number of their pi electrons and the non-hydrogen atoms bonded to them. The number of atom pairs describing a molecule grows with its number of atoms. In order to use atom pairs for similarity comparisons, enumeration of their common and unique atom pairs should be followed by employing these numbers to compute a similarity coefficient such as Tanimoto coefficient in this tool.

The Tanimoto Coefficient (TC) has several different definitions. In this tool, it is defined as ‘c/(a+b+c)’, namely the proportion of the features shared among two compounds divided by their union. The variable c is the number of features common in both compounds, while a and b are the number of features that are unique in one or the other compound, respectively. The Tanimoto coefficient has a range from 0 to 1 with higher values indicating greater similarity than lower ones. However, Tanimoto coefficient of 1 does not mean that two compounds are identical in three dimensional space; it indicates that two compounds have identical structural descriptors.

Hierarchical clustering requires a distance matrix of all-against-all compound distances that is generated by subtracting the similarity measure from one (1 - TC). SMILES codes of ligands were input and the resulting distance matrix was then passed to the clustering program automatically. The program hierarchically joins the most to least similar compounds using single linkage. A tree file was generated to visualize two-dimensional similarity of ligands and identify common scaffolds.

# 4. RESULTS AND DISCUSSION

One of the major contributors for the development of human cancers such as leukemia or breast cancer is AKT. Within the scope of this study, novel ligands for allosteric inhibition of AKT1 activity were identified by virtual screening methods to take a step in the treatment of these several cancer types. Database screening was performed via structure based pharmacophore building and docking, followed by further analyses of hits considering predicted ADME properties and key ligand-protein interactions to propose a diverse set of chemical scaffolds as allosteric inhibitors for AKT1.

## 4.1 Receptor Protein Structure Preparation and Grid Generation

The coordinates of receptor protein-ligand complex were obtained from PDB (code: 3O96). The hydrogen atom coordinates were added and bond orders were assigned. Missing loops (Residue numbers: 45-49, 88-92, 188-199, 298-313) were predicted by Prime application of Suite. The coordinates for the loop between residues Gln 113 and Val 145 could not be predicted since Prime is limited to predict loops with fewer than 20 residues. However, since this loop is localized on the outer part of the protein and is not in the vicinity of the binding site, its presence was not considered to be crucial for docking calculations. Out of 21 water molecules present in the coordinate file, water 455 was maintained for further calculations due to its bridging function (Kettle, 2012). All remaining water molecules were deleted.

Protonation states were generated at pH range of 7.0±4.0 by Epik module and hydrogen-bonding network was optimized by reorienting hydroxyl and thiol groups, water molecules, amide groups of asparagine and glutamine, and the imidazole ring in histidine. Five generated states with their state penalties (kcal/mol), formal charges (Q) and possible intermolecular H-bond counts are given in Table 4.1. $S_1$ protonation state was selected due to its low state penalty.

Table 4.1. Ligand states generated by Protein Preparation Wizard.

| **PDB Code-Ligand** | **$S_1$** | **$S_2$** | **$S_3$** | **$S_4$** | **$S_5$** |
|---|---|---|---|---|---|
| 3O96 – IQO | 0.18 kcal/mol Q=+1 H bond count=0 | 4.68 kcal/mol Q=+2 H bond count=2 | 4.96 kcal/mol Q=0 H bond count=0 | 5.39 kcal/mol Q=0 H bond count=0 | 5.59 kcal/mol Q=0 H bond count=0 |

Finally, minimization was done with restraints on the heavy atoms by Impref utility of Impact application using OPLS_2005 force field with the maximum allowed RMSD of 0.30 for heavy atoms. Figure 4.1 shows AKT1 structure before and after protein preparation.

Protein preparation was followed by grid generation, in which the binding site was defined by a grid consisting of several different sets of fields. These fields were equal in size and provided more accurate scoring of the ligand poses. The center of the cube of 20Å size was chosen as the centroid of the native ligand (PDB code: IQO) of the complex (PDB code: 3O96). For the inner box, which defines the boundaries for the center of a ligand to be docked, center was the same with outer box with a size of 10Å. Rotation of hydroxyl groups on serines and tyrosines were not allowed and no H-bond constraint was applied since there was not any information about rotation or a specific H-bond interaction with the receptor for a candidate ligand. The van der Waals scaling value of nonpolar atoms was reduced to 0.8 from default value of 1.0 (no scaling) with 0.25 partial charge cutoff. The vdW scaling 0.8 mimicked the presence of a larger binding site allowing larger ligands to be docked. Outer and inner boxes defining the binding site for the receptor are represented in Figure 4.2

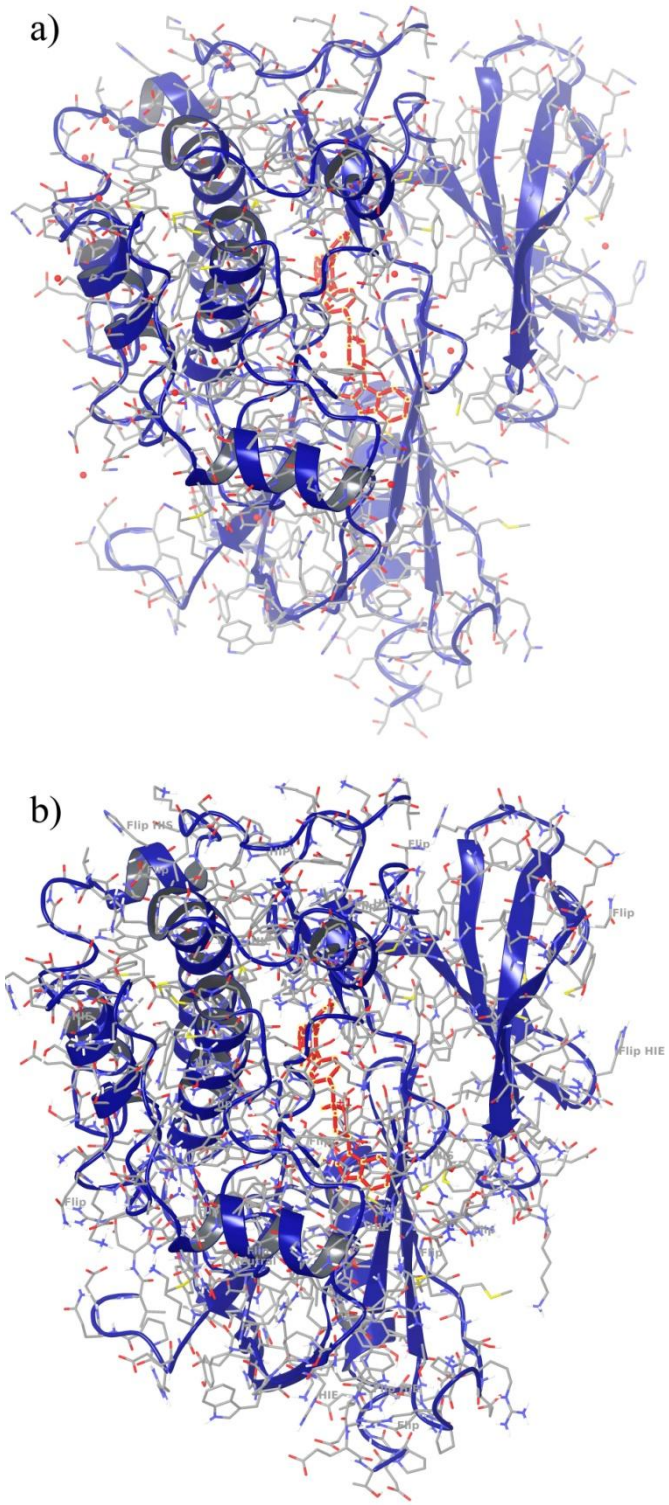



Figure 4.1. Structure of AKT1 (a) before preparation (b) after preparation. Backbone is shown in cartoon and side chains are in licorice representation. Coloring: Secondary Structure (atom based), native ligand (IQO) (red tube), water molecules (red circles).

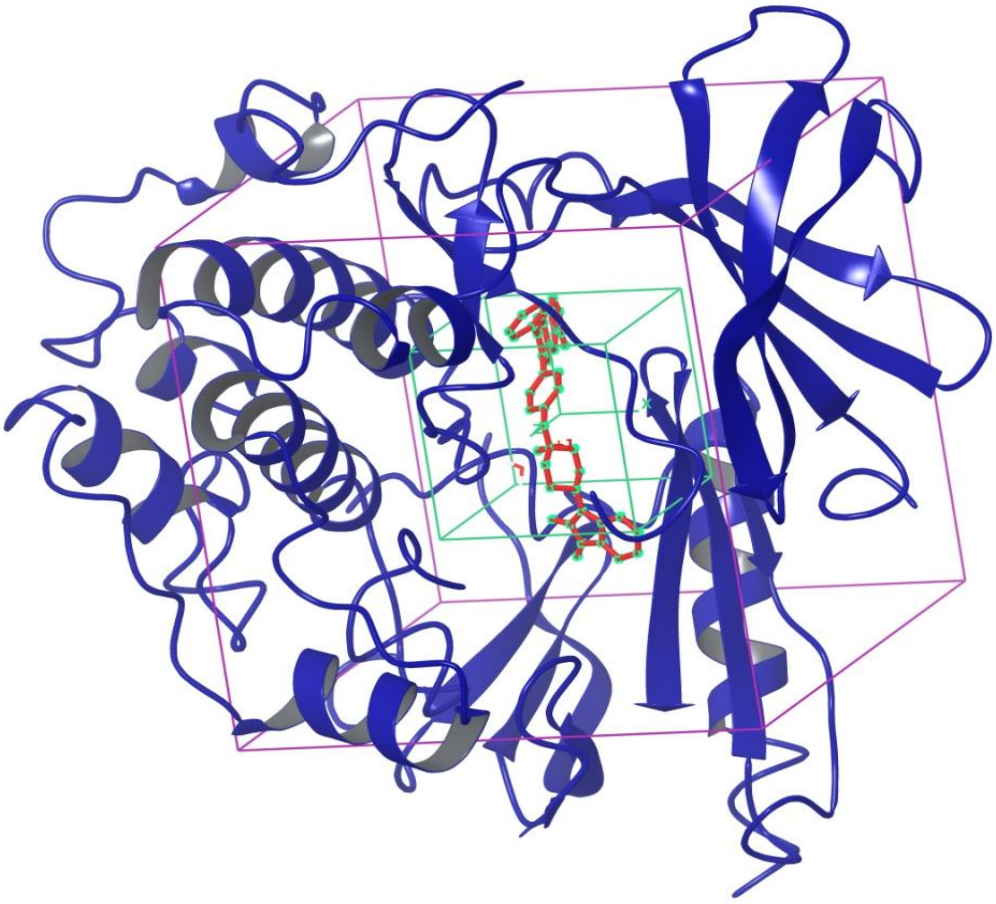


Figure 4.2. Outer (purple) and inner (green) boxes for AKT1 (blue). Native ligand is in red. Center of native ligand is the center for both boxes.

## 4.2 Allosteric Binding Site Analysis of AKT1 in Complex with Inhibitor VIII

The only full length AKT1-allosteric inhibitor complex structure that was available in PDB when this study initiated was AKT1- Inhibitor VIII complex (PDB code: 3O96). The interactions between Inhibitor VIII and the binding site were investigated in detail.

The co-crystal structure of AKT1 contains residues 2-429 in 3O96 complex. This region comprises the kinase domain and PH domain. Inhibitor VIII binds to AKT1 in an allosteric binding site located at the interface of PH domain and N- and C- lobes of the kinase domain. Inhibitor VIII and the interacting residues in the allosteric binding site are shown in Figure 4.3.

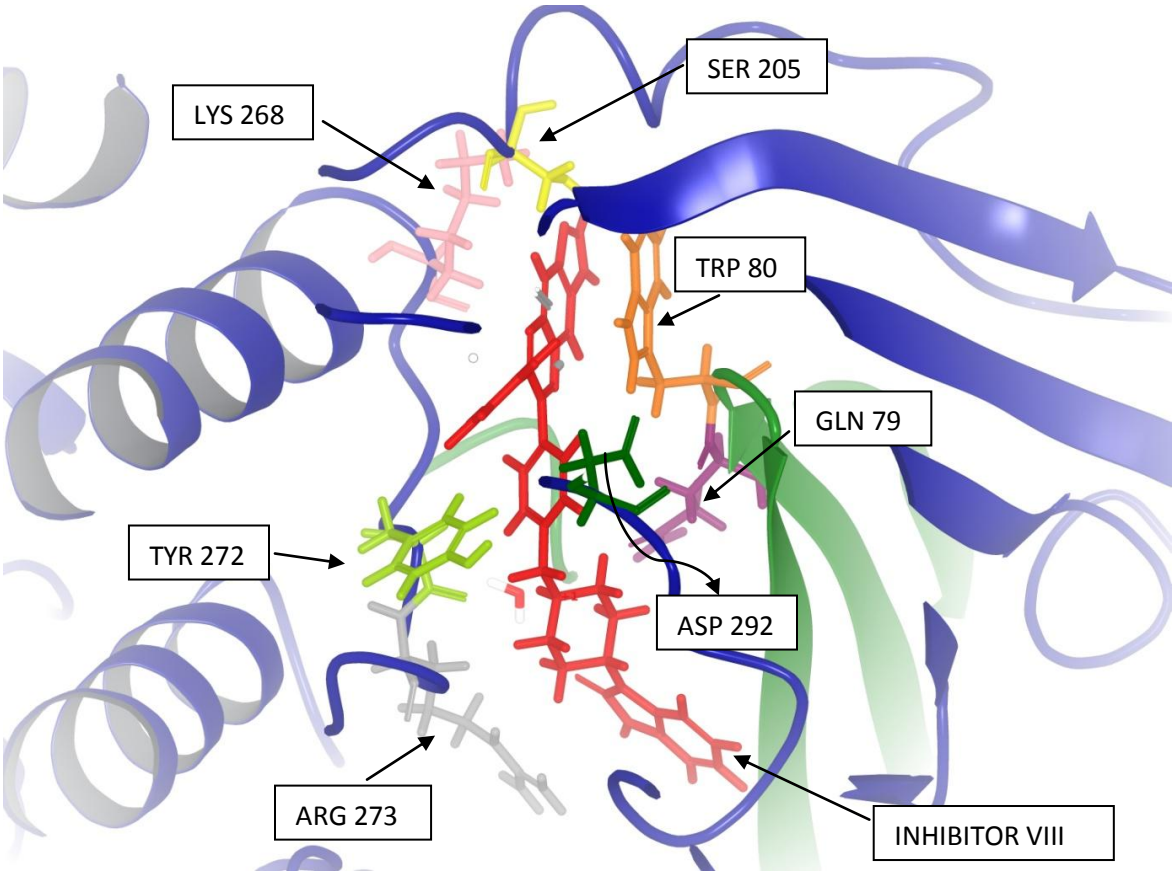



Figure 4.3. Analysis of Inibitor VIII in allosteric binding site of AKT1 (Coloring: PH domain in green ribbon, kinase domain in blue ribbon, Inhibitor VIII in red tube).

Residues of Gln 79, Trp 80, Ser 205, Lys 268, Tyr 272 and Asp 292 were in contact with Inhibitor VIII. The interaction diagram obtained from Maestro and the interaction diagram obtained from the PDB site are represented in Figure 4.4.

Although they were mostly similar, there were a few differences between the maps obtained from Maestro after protein preparation and from the PDB. The π-π stacking interaction between Trp 80 and the tricyclic substructure of Inhibitor VIII was identified in both maps. The π-π interaction with Tyr 272 present in the PDB map (Fig.xx.b) is replaced by a π-cation interaction with Arg 273 (Fig.xx.a), because interacting atom of Tyr 272 is 4Å away in the prepared structure file while interacting atoms of Arg 273 and Tyr 272 are both 3Å away in the initial structure. Another difference is the absence of H-bonds between Ser 205 and Lys 268 and the ligand in the prepared structure. The interacting atoms of these residues move by 0.4Å (Ser205) and 0.5Å (Lys268) upon minimization.

These differences in the structure before and after minimization suggest that these residues may be more dynamic.

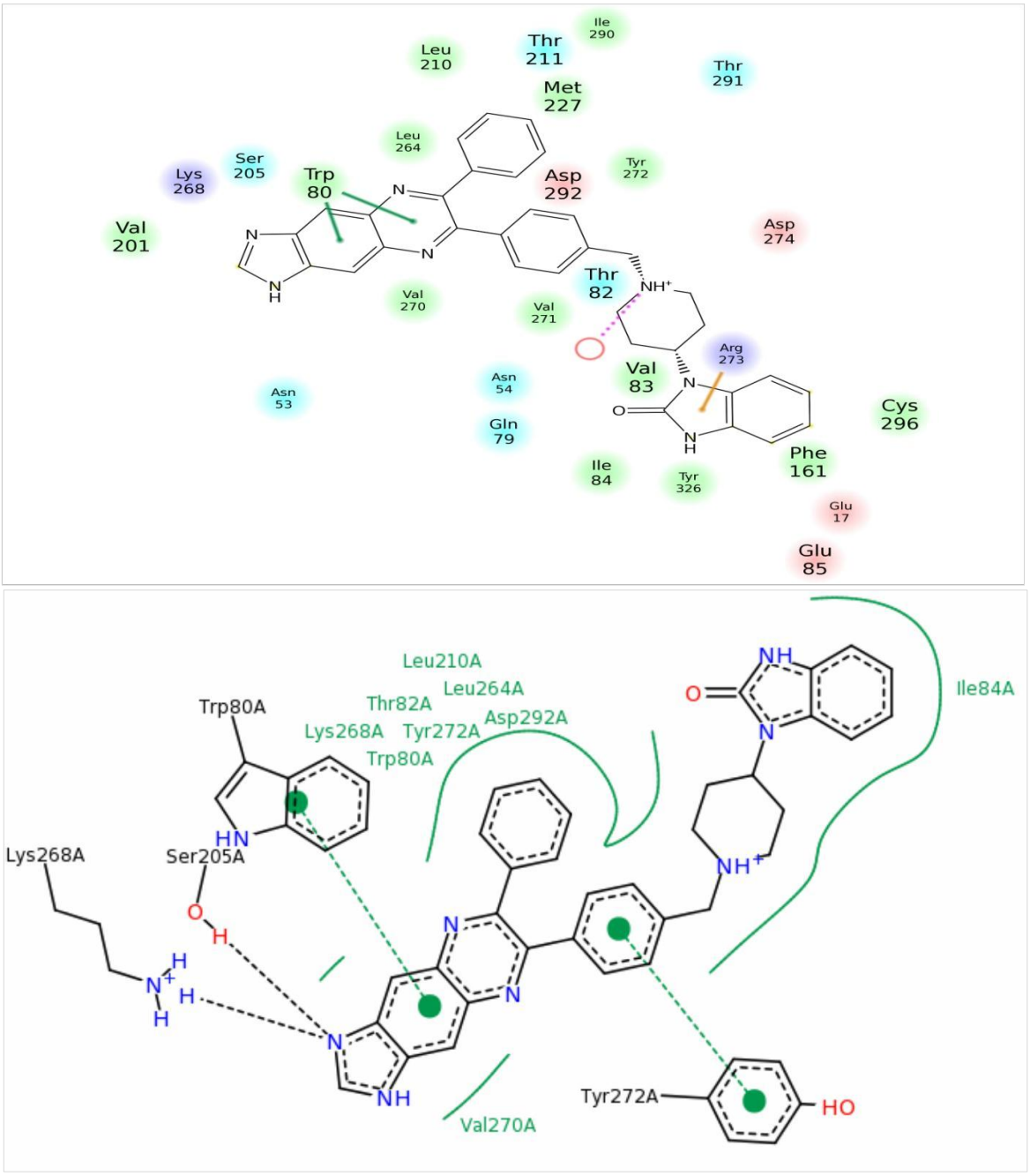



Figure 4.4. Interaction diagrams for Inhibitor VIII and AKT1 obtained from (a) Maestro (b) PDB.

## 4.3 Analysis of Sequence Similarity between Akt and other PH domain including kinases

The BLAST algorithm [102] calculates similarity scores for local alignments (the most similar regions between two sequences) between the query sequence and subject sequences using specific scoring matrices and returns a table of the best matches from the database. The hit table includes several useful pieces of information, including the query coverage, E-value and max identity. Query coverage is the percent of the query sequence that overlaps the subject sequence and max. identity is the percent similarity between the query and subject sequences over the length of the coverage area. The Expect value (e-value) of the hit is defined as the number of hits one can expect to see by chance when searching a database of a particular size. It decreases exponentially as the score of the match increases. For example, an E value of 1 assigned to a hit can be interpreted as meaning that in a database of the current size one might expect to see 1 match with a similar score simply by chance. The lower the E-value, or the closer it is to zero, the more 'significant' the match is.

The BLAST algorithm was used to detect sequence similarity of Akt1 (1-480) of homo sapiens with the proteins in homo sapiens organism. Protein-protein blast was performed by excluding models and using NCBI Reference Sequence Database, which is an integrated and non-redundant set of reference sequences including genomic, transcript, and protein. A total of 110 hits matched with varying max. identity percentages between 32% and 51% (Figure 4.5).

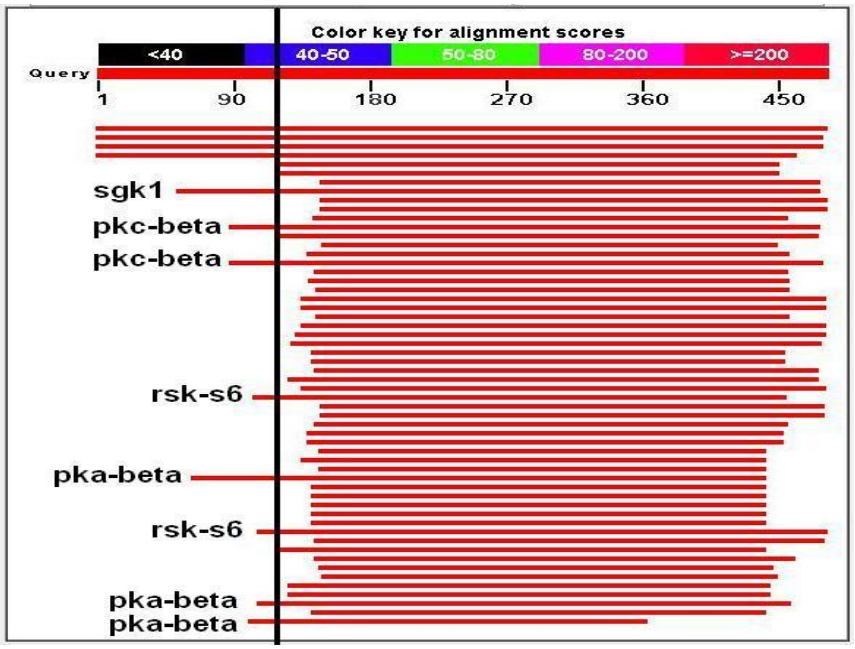



Figure 4.5. Sequence similarity analysis for Akt1 using NCBI-BLAST. Black solid line indicates PH domain region (res: 1-113) of Akt1 approximately.

The identities of PH domain of Akt1 (1-113) were investigated and four classes of kinases (PKA-β, PKC-β, SGK1, RSK-S6) showed partial sequence similarity to PH domain of Akt1 (Table 4.2). Instead of PH domain in Akt1, SGK1 has a PX domain (res: 41-72) which is similar to PH domain for phosphoinositide binding function. PKC-β does not have a PH domain similar region but the similar region with PH domain (~80-120) corresponds to cysteine-rich zinc binding domain of PKC-β. The similar region of RSK-S6 (~90-120) with PH domain of Akt1 is a part of catalytic kinase domain of ribosomal protein S6 kinase. The similar region of PKA-β is also a part of catalytic kinase domain of PKA-β. SGK1, PKC-β and RSK-S6 exhibited higher than 75% sequence overlaps with Akt1, however the max sequence identities were about 40% and the highest contribution to 40% was from residues with residue numbers greater than 113 (non-PH domain region of Akt). To sum up, all four kinases that showed sequence similarity with PH domain of Akt1 indicated low e-values and low percentages of maximum identity for a maximum 87% and minimum 54% query coverages, which makes Akt1 PH domain targeting or interacting compounds highly selective between all serine/threonine kinases.

Table 4.2. Similarity descriptors of Akt1 obtained by NCBI-BLAST.

| AKT1 accession number: AAL55732 / refseq_protein / exclude models / organism: homo sapiens / protein-protein blast | | | | |
|---|---|---|---|---|
| **Description** | **Query cover** | **E value** | **Max. identity** | **Accession number** |
| Serine/threonine- protein kinase Sgk1 | 87% | 2e-107 | 43% | NP_001137150.1 |
| protein kinase C beta type | 81% | 9e-99 | 42% | NP_002729.2 |
| Ribosomal protein S6 kinase | 77% | 2e-83 | 40% | NP_055311.1 |
| cAMP-dependent protein kinase beta | 54% | 4e-66 | 44% | NP_997461.1 |

## 4.4 Validation of Glide Docking

Prior to docking, docking protocol was validated by re-docking Inhibitor VIII (co-crystal ligand of 3O96 complex in PDB) to AKT1 using Glide XP mode. After protein preparation step, Inhibitor VIII coordinates were saved separately. Grid box for AKT1 binding site was generated and Inhibitor VIII was docked to AKT1 without any constraints. The XP GScore was -10.94 kcal/mol, which represents a good binding affinity according to score scale of Glide docking protocol. RMSD of all heavy atoms between docked and co-crystal forms of Inhibitor VIII was calculated as 0.38Å. RMSD values lower than 2.0-2.5Å are referred as acceptable in literature (Pierce, 2008) (Hevener, 2009), therefore 0.38Å is a well enough value. The superimposition of docked and co-crystal forms of Inhibitor VIII is shown in Figure 4.6.

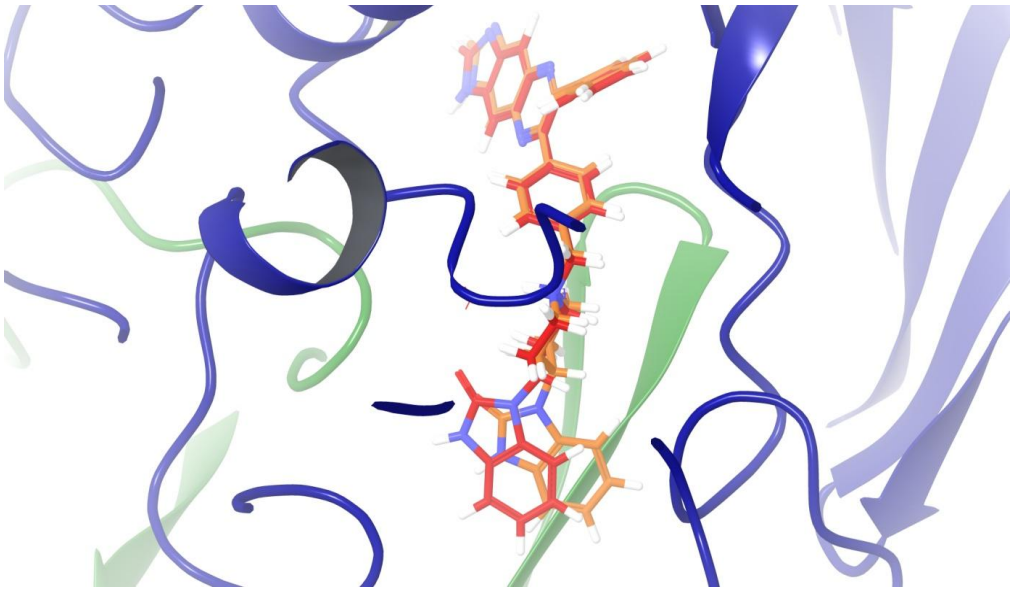


Figure 4.6. Superimposition of docked (red) and co-crystal (orange) forms of Inhibitor VIII located in allosteric binding site of AKT1 (PH domain green, kinase domain dark blue). Atom coloring for ligands: N atoms (light blue), H atoms (white), O atoms (red).

In addition to RMSD calculation, interaction diagram of docked form was analysed to check the key ligand-protein interactions. Interaction diagram for re-docked form of Inhibitor VIII is represented in Figure 4.7.

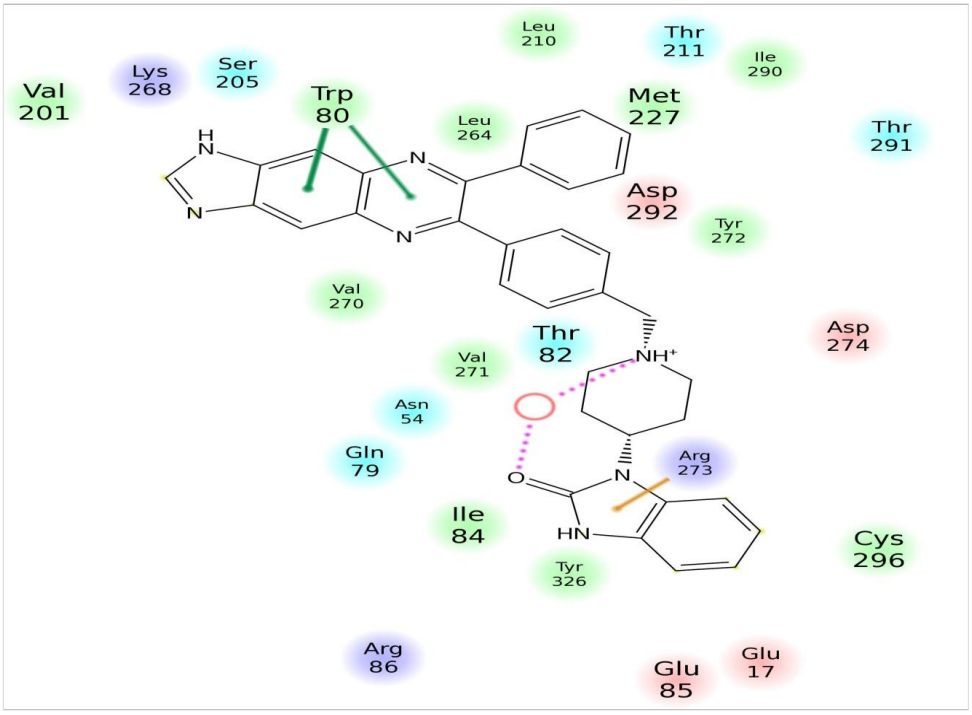



Figure 4.7. Interaction diagram for re-docked form of Inhibitor VIII and AKT1.

The double π-π stacking interaction with Trp 80 and the π-cation interaction with Arg 273 were both identified after docking. Water 455 was observed to interact with oxygen and make bridging interaction with primary amine. The low RMSD value and identification of the same key interactions with pre-docked form of Inhibitor VIII and AKT1 complex validated the docking protocol prior to further docking experiments.

## 4.5 Docking, Binding Mode and Interaction Analyses of Experimentally Identified Inhibitors

The allosteric AKT1 inhibitors which were identified by experimental studies in literature were investigated and a total of eight potent inhibitors were obtained from five different studies [4], [16], [98], [48], [67]. Six of these inhibitors had IC50 values in nanomolar range and two of them had IC50 values in micromolar range. Five inhibitors showed better inhibitory activity for AKT1 compared to AKT2. List of these inhibitors are given in Table 2.1.

Similarities of two dimensional structures were investigated using hierarchical clustering based on Tanimoto coefficient. Inhibitor VIII exhibited highest similarity with Compound 1 and Compound 2, with a Tanimoto coefficient value of 0.6, where the identical compounds give a Tanimoto coefficient of 1.0. The tree file and heatmap of eight experimentally identified inhibitors and Inhibitor VIII are given in Figure 4.8.

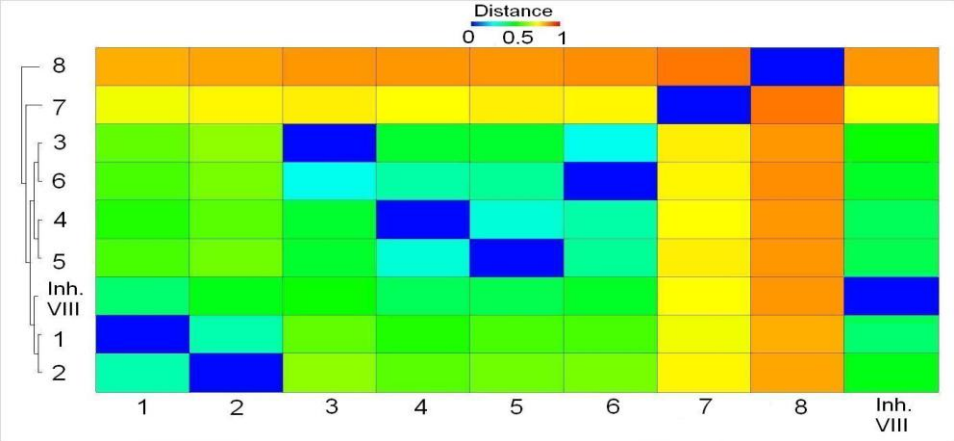



Figure 4.8. Heatmap for Inhibitor VIII and experimentally identified ligands.

The heatmap assigns a blue color code for identical ligands wheras the ligands which show no two dimensional similarity share a red colored cell. The cell color changes from red to blue as the similarity between two ligands increase.

Two dimensional chemical structures of the eight experimentally identified inhibitors and Inhibitor VIII were drawn by a chemical structure drawing tool (e-molecules.org). Structures were trasformed into 3D by LigPrep [86] with varying ionization and tautomerization states at neutral pH and 25 stereoisomers were created for all nine inhibitors. For each inhibitor, energetically minimized conformers were generated by ConfGen [87] using a RMSD value of 1Å to detect redundant conformers, and conformers whose energy was more than 25kcal/mol (104.67 kJ/mol) higher than the lowest energy conformer were eliminated. Finally, 391 conformers were obtained for nine inhibitors (including Inhibitor VIII) and were docked to grid-generated AKT1 structure using Glide XP mode without applying any constraints.

For the nine inhibitors included in docking, the conformer for each ligand which showed the highest binding affinity (based on GScore value) for AKT1 were examined in detail. Glide XP GScores for the highest ranked conformer of each ligand, its rank among all docked conformers of nine ligands and the key interactions they make with AKT1 are shown in Table 4.3.

The docking results showed that, the top scoring conformation of the experimentally identified ligands could bind to the allosteric site with good affinity except for Compound8. Six of the eight ligands made the π-π stacking interaction with Trp 80 residue of AKT1 necessary for the allosteric mechanism of inhibiton (Wu, Voegtli, 2010). Compound8 did not make any of the common interactions with AKT1 since it has a bulky dodecane tail which probably inhibited its fitting in the cavity between two domains of the protein (Figure 4.9, gray). Alignment of all experimentally identified ligands with Inhibitor VIII is represented in Figure 4.9.

Table 4.3. Docking scores, ranks and interaction analyses for experimentally identified allosteric AKT1 inhibitors.

| **Assigned name** | **Glide XP GScore (kcal/mol)** | **Rank among all docked conformers (391)** | **Interactions with AKT1** |
|---|---|---|---|
| INHIBITOR VIII | -10.94 | 1 | **Trp 80 (π-π)**, Trp 80 (π-π), Arg 273 (π-cat), Water 455 (H-bond) |
| Compound1 | -10.0 | 18 | **Trp 80 (π-π)**, Arg 273 (π-cat), Arg 273 (π-cat), Water 455 (H-bond) |
| Compound2 | -9.95 | 20 | **Trp 80 (π-π)**, Arg 273 (π-cat), Water 455 (H-bond) |
| Compound3 | -10.14 | 12 | Arg 273 (π-cat), Water 455 (H-bond), Glu 85 (H-bond) |
| Compound4 | -10.10 | 13 | **Trp 80** (π-π), Trp 80 (π-π), Trp 80 (H-bond) |
| Compound5 | -10.27 | 8 | **Trp 80** (π-π), Arg 273 (π-cat), Water 455 (H-bond), Tyr 326 (π-π) |
| Compound6 | -10.79 | 5 | **Trp 80** (π-π), Arg 273 (π-cat), Water 455 (H-bond), Glu 85 (H-bond) |
| Compound7 | -9.91 | 21 | **Trp 80** (π-π), Asn 53 (H-bond) |
| Compound8 | -7.75 | 349 | Thr 211 (H-bond) |

These ligands were proposed to be allosteric inhibitors by experimental studies but no structural information is available for their binding modes therefore, it is possible that these ligands bind elsewhere on the structure. Fortunately, seven ligands were docked with good scores and six ligands made the compulsory π-π stacking interaction with Trp 80, suggesting that at least six ligands bind in the designated allosteric site.

The experimentally identified ligands were examined further to define the interacting sites on ligands and check if the same structural parts make the same interactions with AKT1. The common substructures of ligands are shown in Table 4.4 with proper nomenclature.

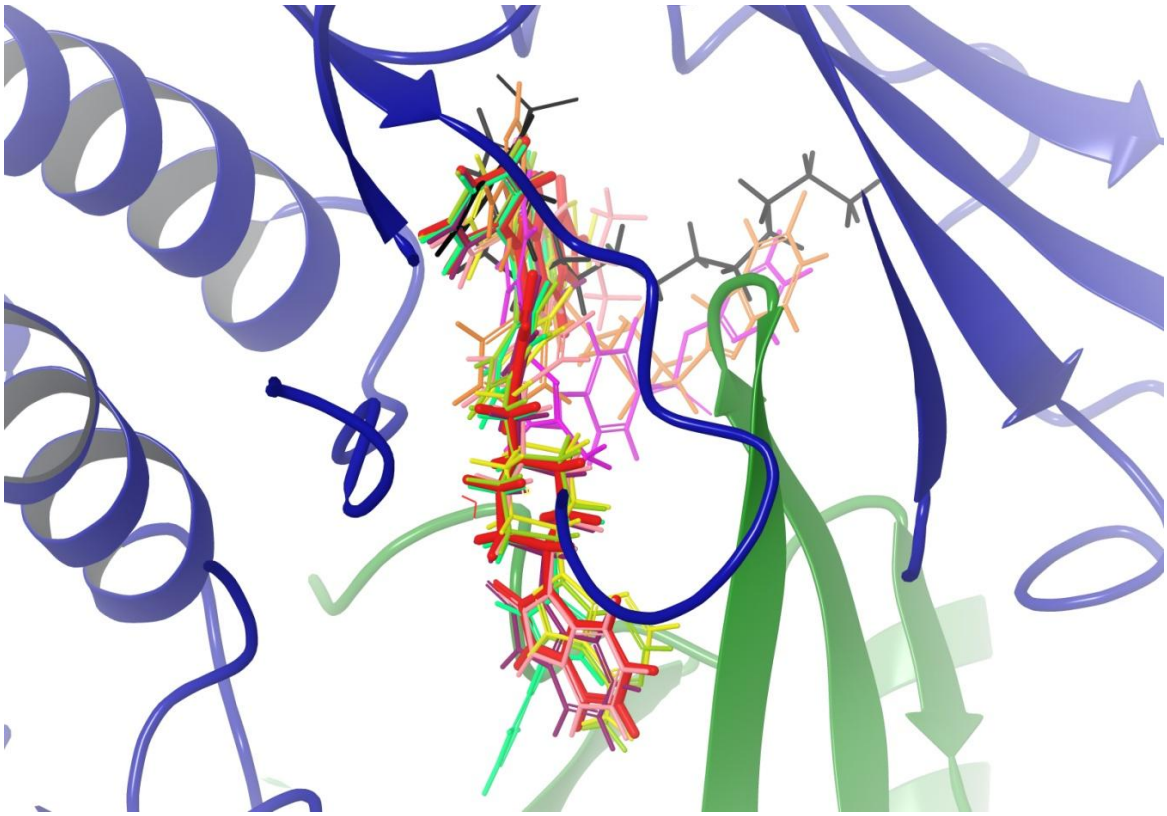
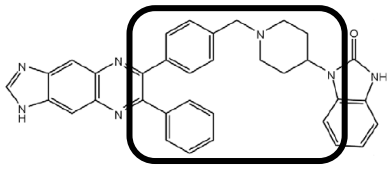
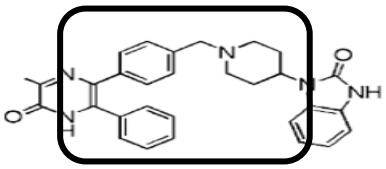
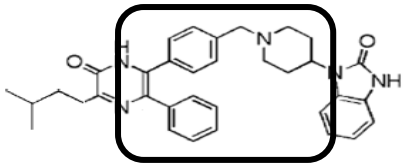
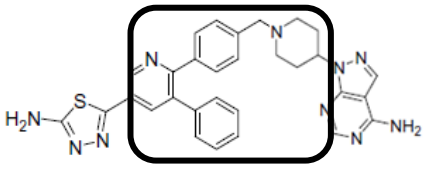
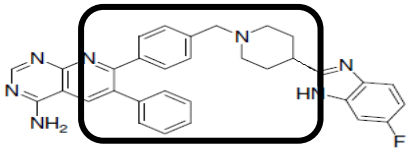
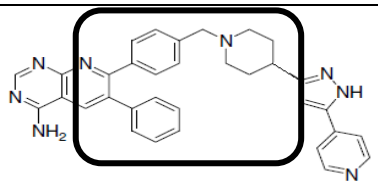
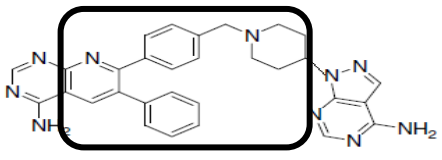


Figure 4.9. Alignment of Inhibitor VIII and 8 experimentally identified ligands in allosteric binding site (Coloring: PH domain in green ribbon, kinase domain in blue ribbon, Inhibitor VIII in bold red, other ligands in varying colors).

When the interaction maps were obtained, Inhibitor VIII, compound1 and compound2 were observed to make the π-π stacking interaction with Trp 80 through their pyrazine ring. The same interaction was observed between pyridine ring and Trp80 for compounds 4,5 and 6. The π-cation interaction with Arg 273 was observed between (1,3dihydro-2H-benzimidazol-2-one) sub-structure of compounds 1,2 and Inhibitor VIII.

Table 4.4. Common sub-structures of experimentally identified allosteric AKT1 inhibitors shown in black square.

| **Assigned name** | **Chemical structure** | **Interactions with AKT1** | **Common ligand sub-structures** |
|---|---|---|---|
| INHIBITOR VIII |  | **Trp 80 (π-π)**, Trp 80 (π-π), Arg 273 (π-cat), Water 455 (H-bond) | 2-phenyl-3-[4-(piperidin-1-ylmethyl)phenyl] **pyrazine** |
| Compound1 |  | **Trp 80 (π-π)**, Arg 273 (π-cat), Arg 273 (π-cat), Water 455 (H-bond) | 6-phenyl-5-[4-(piperidin-1-ylmethyl)phenyl]-1,2-dihydro**pyrazine** |
| Compound2 |  | **Trp 80 (π-π)**, Arg 273 (π-cat), Water 455 (H-bond) | |
| Compound3 |  | Arg 273 (π-cat), Water 455 (H-bond), Glu 85 (H-bond) | 3-phenyl-2-[4-(piperidin-1-ylmethyl)phenyl] **pyridine** |
| Compound4 |  | **Trp 80** (π-π), Trp 80 (π-π), Trp 80 (H-bond) | |
| Compound5 |  | **Trp 80** (π-π), Arg 273 (π-cat), Water 455 (H-bond), Tyr 326 (π-π) | |
| Compound6 |  | **Trp 80** (π-π), Arg 273 (π-cat), Water 455 (H-bond), Glu 85 (H-bond) | |

## 4.6 Structure Based Pharmacophore Modeling with E-Pharmacophores

Pharmacophore modeling can be done in two ways; ligand based or structure based. Ligand based pharmacophore modeling requires an initial set of ligands to identify their common sites as pharmacophores, whereas structure based pharmacophore modeling uses the structural information of the binding site. E-pharmacophores script was used to apply structural pharmacophore modeling. Running this script requires initial docking of Schrödinger fragment library (Glide) to allosteric binding site, to obtain descriptive information collected from binding modes of fragments to receptor and energetic terms, followed by determination of possible pharmacophoric sites and finally hypothesis construction using selected sites.

Schrödinger Fragment Library (Glide) consisting of 441 unique (1-7 ionization/tautomer variants; 6-37 atoms; MW range 32-226) and giving a total of 667 small fragments with accessible low energy ionization and tautomeric states were docked using Glide XP mode without applying any constraints, and the information was stored in the Glide XP descriptor file. Binding modes of all fragments in allosteric binding site are shown in Figure 4.10.

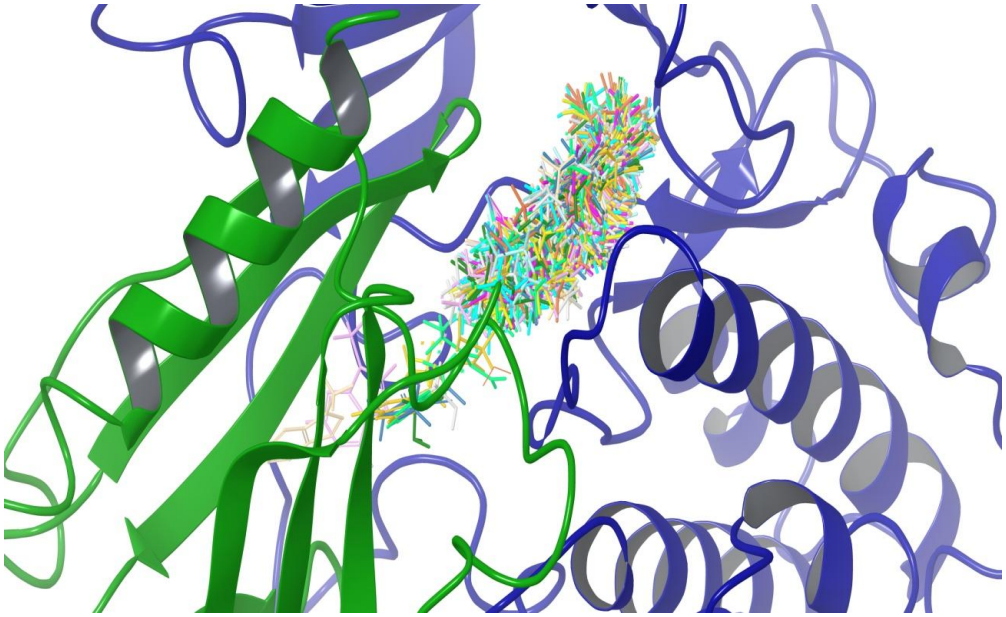


Figure 4.10. Binding modes of all fragments on AKT1 allosteric site generated by Glide XP docking (Protein structure represented in ribbons and fragments in wires. Coloring: PH domain in green, kinase domain in blue).

The pose viewer file obtained from Glide XP docking was the input for E-pharmacophores script. The GlideScore of all fragments were composed of their energetic contributions on atom centers. Possible pharmacophore sites were determined by E-pharm script and ranked based on their energies. Seven important sites were identified E-pharm script and their ranks, types, scores and x,y,z coordinates are listed in Table 4.5. Number codes in feature labels were assigned randomly by the E-pharmacophores script. Locations of pharmacophore sites in binding site are also shown Figure 4.11.

Table 4.5. Scores and coordinates of seven pharmacophore features by E-pharm.

| Rank | Feature label | Score | X | Y | Z | type |
|---|---|---|---|---|---|---|
| 1 | R2246 | -8.4 | 6.8315 | -1.704 | 10.3065 | R |
| 2 | H1462 | -5.34 | 5.9594 | -3.426 | 9.4635 | H |
| 3 | A575 | -4.0 | 11.389 | -5.9983 | 11.7999 | A |
| 4 | H1384 | -3.39 | 7.7423 | -1.3332 | 12.9548 | H |
| 5 | D937 | -2.29 | 6.6722 | 2.3768 | 9.6869 | D |
| 6 | D1305 | -2.20 | 13.7871 | -3.6748 | 5.5687 | D |
| 7 | A598 | -2.14 | 9.8486 | 0.4118 | 11.4 | A |

The minimum number of sites to construct a hypothesis is three and the maximum is seven. Hypothesis composed of three sites were selected to allow sampling of a wider range of compounds because as the number of sites increases, the pharmacophore model becomes more restrictive. Several criteria were taken into account while choosing the sites for hypothesis modeling;

- The sites with higher energy contribution were induced if they were less than 10Å away from each other.
- To provide inhibiton by a ligand bounded to allosteric site, it was known from literature [17] that a pi-stacking interaction between the ligand and Trp80 of AKT1 is a must. Therefore, the only ring feature detected among seven sites, R2246 was included in all hypotheses. The alignment of all pharmacophore site points on Inhibitor VIII is shown is Figure 4.12.

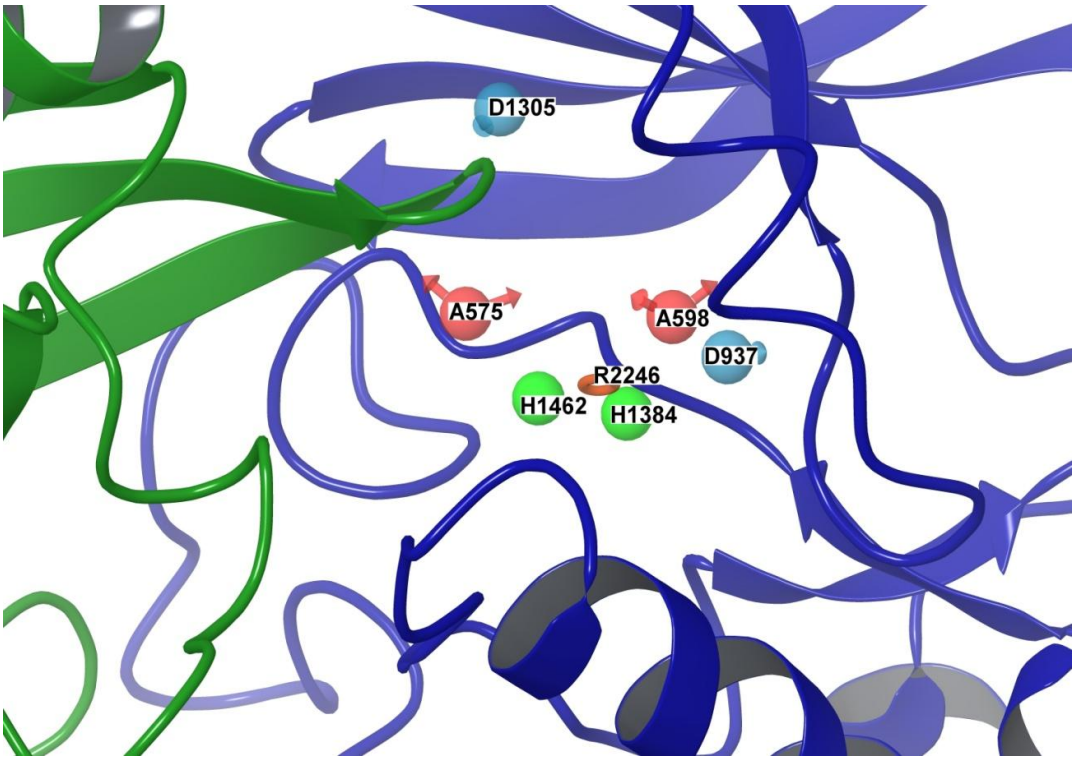



Figure 4.11. Location of seven detected pharmacophore sites in binding site of AKT1 (Coloring: PH domain in green and kinase domain in blue).

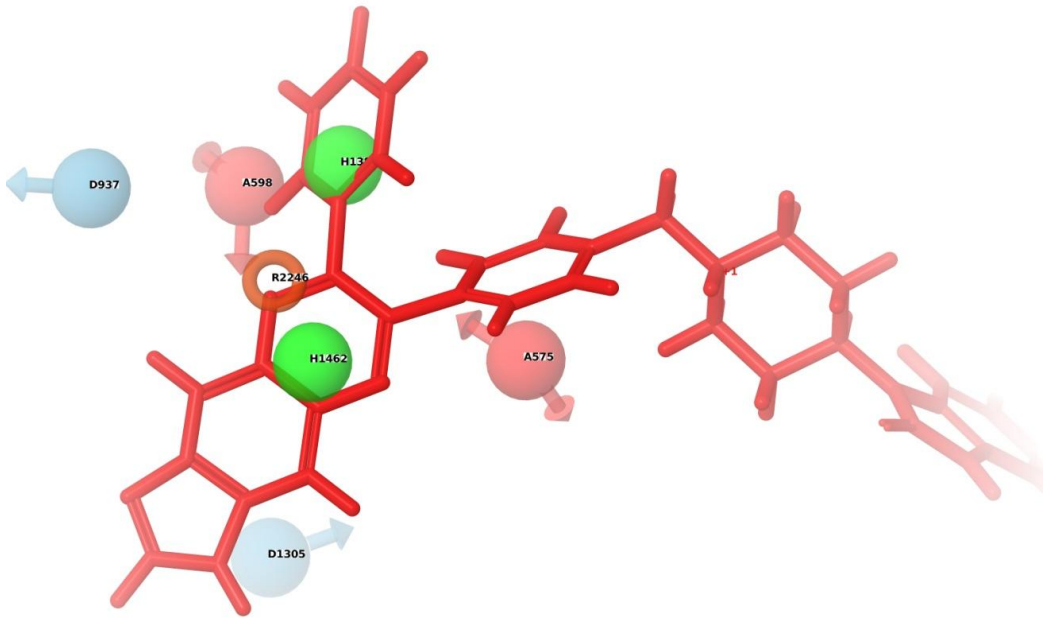



Figure 4.12. Seven pharmacophore sites detected by E-pharm superimposed to Inhibitor VIII (Coloring: Ring in orange, Hydrophobic sites in green balls, Hydrogen Bond Acceptor sites in red balls, Hydrogen Bond Donor sites in blue balls, Inhibitor VIII in red tube).

As a result, one hypothesis was created using R2246, H1462, H1384 (RHH hypothesis) and the other was created using R2246, A575, D937 (RAD hypothesis). All site to site distances are listed in Table 4.6 and inter-site distances for both RHH and RAD hypotheses are shown in Figure 4.13. Note that RHH would select for ligands that would have similar hydrophobic character with the Inhibitor VIII triple ring substructure, while RAD hypothesis has a different chemical scaffold from Inhibitor VIII with D937 4.1Å away from it.

Table 4.6. Site to site distances of seven detected pharmacophore sites.

| Site1 | Site2 | Distance (Å) |
|---|---|---|
| A575 | A598 | 6.605 |
| A575 | D937 | 9.842 |
| A575 | D1305 | 7.069 |
| A575 | H1384 | 6.033 |
| A575 | H1462 | 6.446 |
| A575 | R2246 | 6.438 |
| A598 | D937 | 4.109 |
| A598 | D1305 | 8.137 |
| A598 | H1384 | 3.146 |
| A598 | H1462 | 5.797 |
| A598 | R2246 | 3.844 |
| D937 | D1305 | 10.208 |
| D937 | H1384 | 5.058 |
| D937 | H1462 | 5.851 |
| D937 | R2246 | 4.131 |
| D1305 | H1384 | 9.827 |
| D1305 | H1462 | 8.747 |
| D1305 | R2246 | 8.644 |
| H1384 | H1462 | 4.444 |
| H1384 | R2246 | 2.825 |
| H1462 | R2246 | 2.106 |

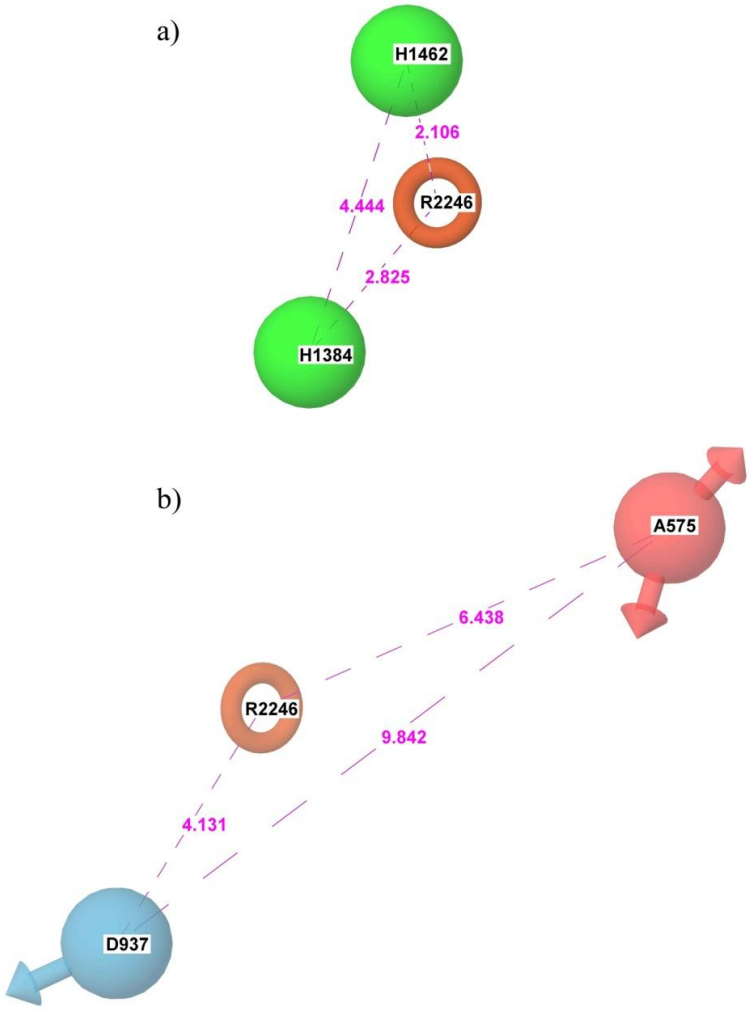



Figure 4.13. Inter-site distances for (a) RHH hypothesis (b) RAD hypothesis (Coloring: Ring in orange, Hydrophobic sites in green, Hydrogen Bond Acceptor sites in red and Hydrogen Bond Donor sites in blue).

## 4.7 Ligand-based pharmacophore modeling

In order to consider protein specificity and isozyme selectivity, several small molecule allosteric inhibitors have been discovered for Akt1. The information of these ligands was utilized to determine a set of common features, namely a pharmacophore model, which the potential inhibitors were expected to match. The number of features to be included in the pharmacophore model was selected as five because it was aimed to use the existing ligand information as much as possible without being too restrictive. Higher

feature numbers could be too restrictive while less would be lack of pharmacophore information resulting in non-specific ligand selection.

Alignment of compounds 1 to 6 (Table 4.2) and Inhibitor VIII provided determination of twelve variants. These twelve variants allowed detection of 413 hypotheses with seven common pharmacophore models including five pharmacophore sites (AADRR, PRRRR, APRRR, DRRRR, AAADR, AAPRR, ADRRR). The number of hypotheses contributed to each common pharmacophore model varied between 1 and 129. The decision of pharmacophore model to be selected was based on the criteria; (a) each pharmacophore site should match all seven ligands (b) the model should exhibit high fitness to Inhibitor VIII (c) the ring structure which interacts with Trp 80 on Akt1 should be represented in the model (d) the model should have high number of hypotheses contribution among 413 detected hypotheses. The pharmacophore model APRRR was built satisfying the stated criteria (Figure 4.14). The model exhibited full fitness to Compound6 (score: 3.00/3.00) and the second best fitting compound was Inhibitor VIII (score: 2.71/3.00). The tricyclic feature set forms a hydrophobic moiety in the allosteric binding site, which holds for most of the type II kinase inhibitors [18]. Feature R14 was the representative of the ring structure making π-π stacking interaction with Trp 80 and feature P12 stands for the primary amine ($NH^+$) interacting with Water 455. Filtering the database with APRRR hypothesis based on 3D similarity yielded 701 matching compounds with fitness scores higher than zero.

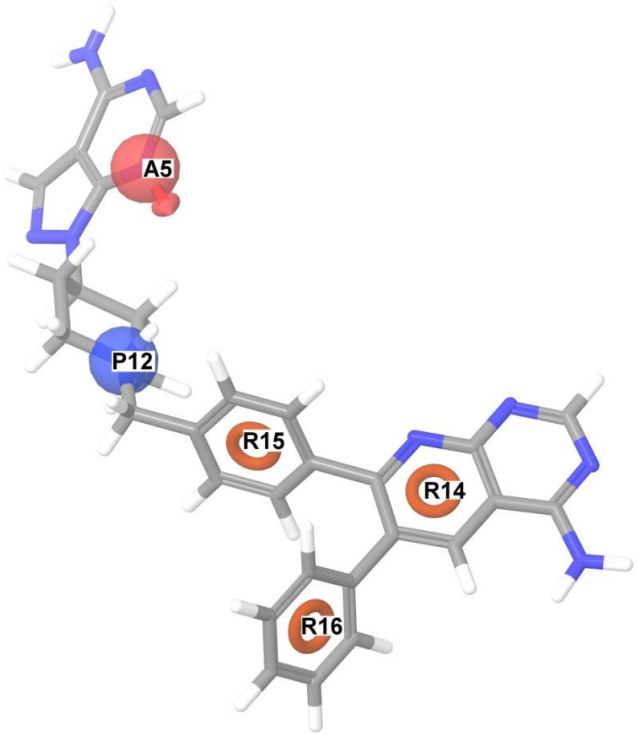



Figure 4.14. Ligand-based pharmacophore model APRRR.

## 4.8 Docking with Glide

The virtual docking experiments were performed using Glide application of Schrödinger (ref). Two docking modes were used which were standard precision (SP) and extra precision (XP). For scoring the binding affinity of the poses, Glide's own scoring function GlideScore (GScore) was employed. The XP mode differs from SP mode by addition of some extra terms to scoring function and consequently consumes more CPU time. Therefore, filtering of ZINC database with the structure based pharmacophore models RHH and RAD was followed by two-step docking protocol. Two different docking runs were performed for each hypothesis separately. After filtering by using both RHH and RAD hypothesis, the initial ~26 million conformations of 499,026 unique compounds provided by ZINC was reduced to 10,000 compounds for each SP docking run. Docking with SP mode and scoring with SP scoring function was followed by re-docking top 1000 molecules (10%) using XP docking mode and scoring with XP scoring function. Virtual Screening Workflow was utilized to provide automatic passing of top scored 10%

molecules from SP docking to XP docking. 701 compounds filtered with APRRR hypothesis were docked using only XP mode due to small number of ligands.

### 4.8.1 Virtual Screening Workflow for compunds selected with pharmacophore models RHH and RAD

ZINC database was prepared and filtered with RHH and RAD pharmacophore hypothesis separately. A total of 10,000 compounds were ready to be docked into grid generated structure of AKT1. The binding site of ligands was defined in grid generation step as 20Å cubic box surrounding Inhibitor VIII and has the same center with Inhibitor VIII. The center of each ligand to be docked was set to be in an inner cubic box of 10Å sharing the same center with the outer box. The filtered ZINC database and grid generated receptor file was given as input to Virtual Screening Workflow (VSW) application for flexible docking of ligands while the receptor was kept rigid. No hydrogen bond or positional constraints were defined in grid-generation step. Epik state penalties for ligands were incorporated to the docking score. Docking and scoring of molecules with more than 300 atoms and 50 rotatable bonds were not allowed. To soften the potential nonpolar atoms of the ligand, vdW radius of these atoms which have less than 0.15 partial charges were scaled by 0.8. For both docking modes, maximum ten poses per compound state was generated and only best scoring state was kept in the final hit list.

Successive docking of 10,000 compounds with SP and XP modes resulted in 997 docked ligands for the RHH hypothesis and 1000 docked ligands for the RAD hypothesis. The hits obtained by RAD hypothesis showed better GScore values (-12.2 to -8.3 kcal/mol) compared to the hits obtained by RHH hypothesis (-11.4 to -2.3 kcal/mol). Figure 4.15 shows the scatter plot of XP GScore values for RHH hits and RAD hits, respectively.

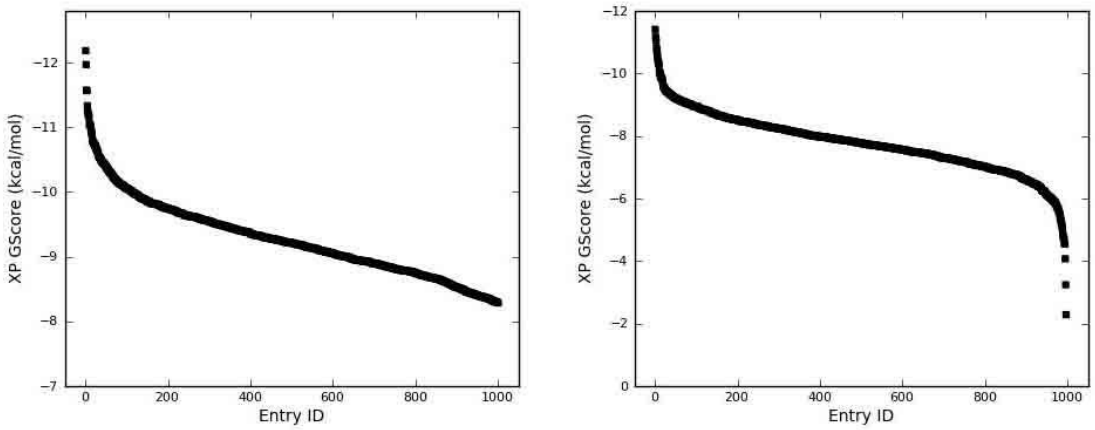



Figure 4.15. Scatter plot of XP GScore values for ZINC compounds filtered with RHH hypothesis (Total: 997 compounds) (left), RAD hypothesis (Total: 1000 compounds) (right) and docked to AKT1.

According to Figure 4.15, for compounds selected by the RHH hypothesis, 997 docked ligands with Glide XP mode showed a sharp decrease in their GScore values from -11.4 kcal/mol to almost -9 kcal/mol. After -9 kcal/mol, the decrease in GScore values was significantly smaller. Therefore, cutoff GScore value was chosen as -9.00 kcal/mol for docking hits of RHH hypothesis. In order to consider equality for binding affinity on evaluations of the hits obtained by starting with two different hypotheses of RHH and RAD, the same cutoff value of -9.00 kcal/mol was also chosen for docking hits of RAD hypothesis. As a result, 88 hits were obtained from RHH hypothesis and 630 hits were obtained from RAD hypothesis followed by strain energy correction.

### 4.8.2 Glide XP Docking of compunds selected with pharmacophore model APRRR

Pre-filtering ZINC database with ligand-based APRRR hypothesis matched only 701 ligands and docking to grid generated Akt1 structure performed in only XP mode (with the same docking parameters stated in the previous section) due to survival of a small number of ligands. The GScore values of 701 compounds varied between -10.7 to -0.5 kcal/mol. In order to consider equality for binding affinity on evaluations of the hits obtained by starting with APRRR and hits obtained by starting two structure-based hypotheses of RHH and RAD, the same GScore cutoff value of -9.00 kcal/mol was also applied for docking hits of APRRR hypothesis. 64 hits were obtained with a better GScore

value than -9.00 kcal/mol. Figure 4.16 shows the scatter plot of XP GScore values for APRRR hits.

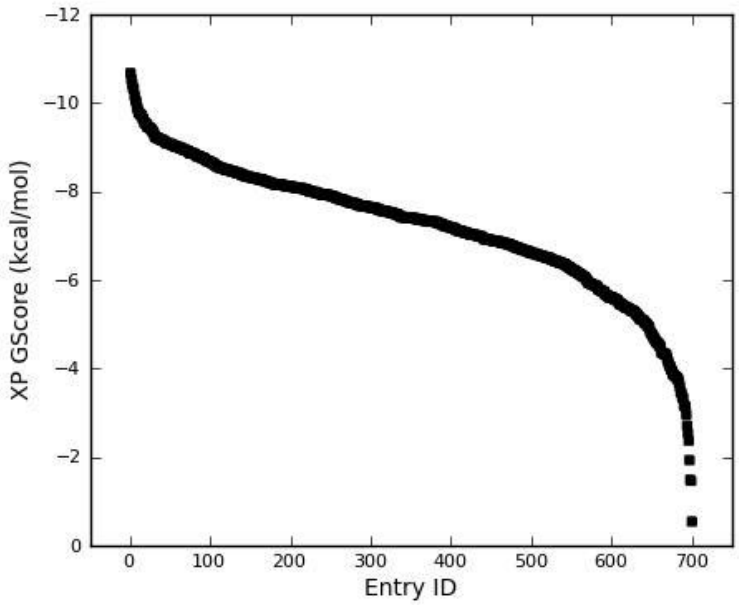



Figure 4.16. Scatter plot of XP GScore values for ZINC compounds filtered with APRRR hypothesis (Total: 701 compounds) and docked to AKT1.

## 4.9 Strain Energy Correction of Docked Ligands

Strain energy correction was applied to docked ligands by using Strain-Rescore (Glide) script to identify ligands with too much strain. Ligands with more than 4 kcal/mol energy difference (default value) between the docked and free forms received penalties and quarter of the energy difference (so called strain penalty) was added to the GScore. Strain-corrected GScore values were used for further evaluations of the hits. Eventually, with the cutoff value of -9 kcal/mol, 38 compounds based on the RHH hypothesis, 158 compounds based on the RAD hypothesis and 16 compounds based on the APRRR hypothesis were selected and further analyzed.

## 4.10 Similarity and Interaction Analyses for Selected Hits

The selected hits were first evaluated based on chemical similarity in order to determine the types of scaffolds selected by pharmacophore filtering and docking. To this

end, ChemMine Tools were used to identify two dimensional similarities between all docked and filtered ligands to AKT1 structure; 38 compounds from hits of RHH hypothesis and 158 compounds from hits of RAD hypothesis by using unique SMILES codes for each compound. Tanimoto Coefficient (TC) (chemmine) was used to measure similarity and has a range from 0 to 1 with higher values indicating greater similarity. Hierarchical clustering was performed to join the most to least similar compounds using single linkage and create a tree file to visualize the similarity and identify common scaffolds between a total of 196 compounds from hits of structure based pharmacophore models RHH and RAD. For 16 hits of APRRR hypothesis, clustering was not applied due to small number of surviving compounds. The numbers in labels of compounds in the tree (i.e.RAD9999) were assigned by Glide during docking.

The similarity trees included some branches with many members, indicating enrichment of these similar compounds in the top compounds after docking. There were five significant clusters in similarity tree of 196 compounds, namely; RAD-family1 (27), RAD-family2 (20), RAD-family3 (12), RAD-family4 (22) and RHH-family5 (13). Each cluster and all hits of APRRR hypothesis were discussed individually considering the compulsory π-π stacking interaction with Akt1, ADME criteria and drug-likeness, for the compounds included.

### 4.10.1 Assessment of RAD-family1

The similarity tree of the first cluster including 27 compounds obtained from ZINC database filtered with RAD hypothesis is represented in Figure 4.17.

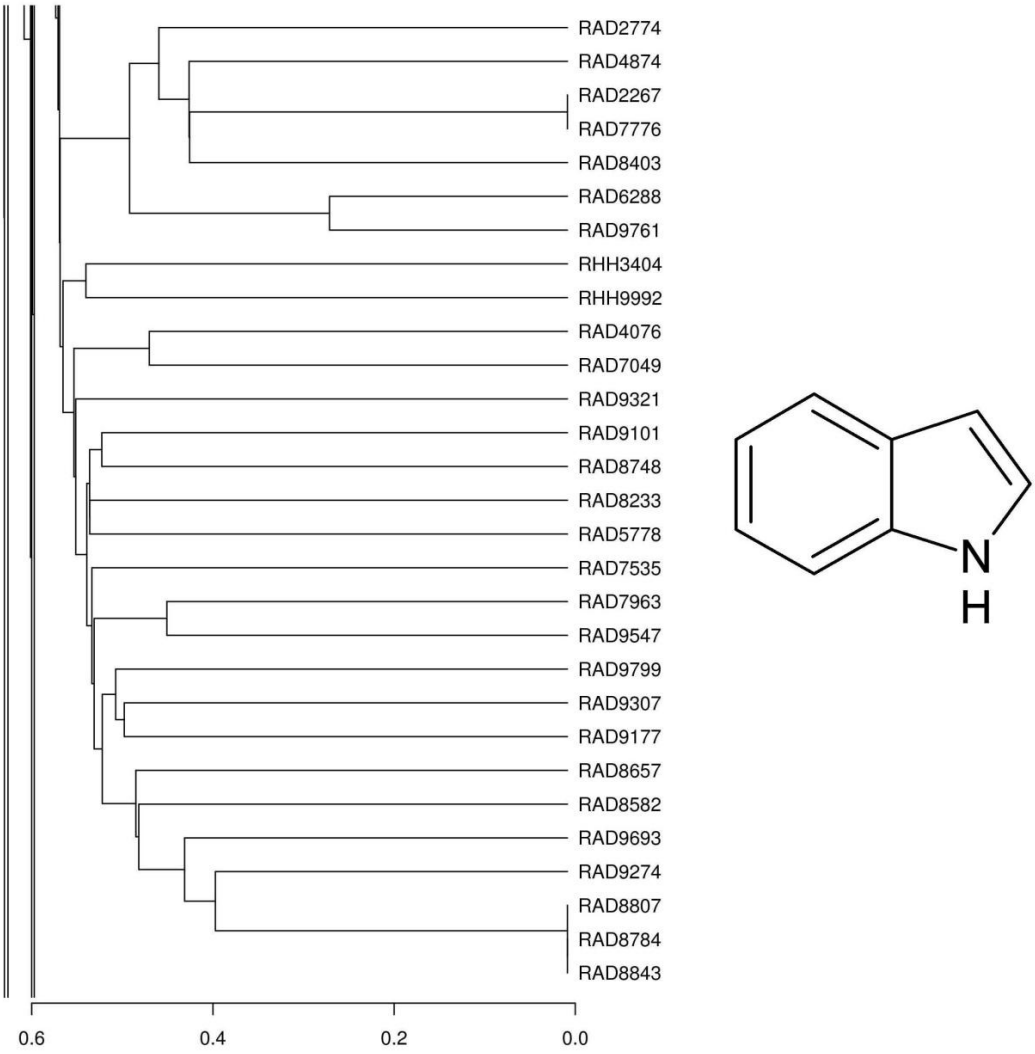



Figure 4.17. Similarity tree for RAD-family1 cluster based on Tanimoto Coefficient (left), similar sub-structure 1H-indole for all compounds in RAD-family1 cluster (right).

The Tanimoto similarity coefficient for the farthest compounds in RAD-family1 cluster was calculated as ~0.45. The GlideScore values for the 27 compounds varied between -11.1 kcal/mol and -9.07 kcal/mol. This cluster contains compounds that share the 1*H*-indole scaffold. The interaction diagrams were generated for all of the compounds to identify the ligand-protein interactions and check for the compulsory π-π stacking interaction of a ring structure of ligand and Trp 80 of AKT1. The interactions of the 27 compounds in RAD-family1 cluster with the protein are summarized in Table 4.7.

Table 4.7. Interactions with AKT1 for 27 compounds in RAD-family1 cluster.

| | TRP 80 | THR 211 | LYS 268 | ASN 53 | SER 205 | ASP 292 | TYR 272 |
|---|---|---|---|---|---|---|---|
| RAD 2267 | | H (b) | | | | | |
| RAD 2774 | | H (b) | | | | | |
| RAD 4076 | | H (b) | | | | | |
| RAD 4874 | pi, H (b) | | H (b), H (s) | H (s) | | | |
| RAD 5778 | | H (b) | | | | | |
| RAD 6288 | | H (b) | | | | | |
| RAD 7049 | | H (b) | | | | | |
| RAD 7535 | | H (b) | | H (b), H (s) | | | |
| RAD 7776 | H (s) | | | | H (b) | | |
| RAD 7963 | pi | | | | H (b) | | |
| RAD 8233 | | H (b) | | | | | |
| RAD 8403 | | | | | H (b) | H (s) | |
| RAD 8582 | | H (b) | | | | | |
| RAD 8657 | | H (b) | | | | | |
| RAD 8748 | | H (b) | | | | | |
| RAD 8784 | | H (b) | | H (s) | | | |
| RAD 8807 | pi | | | | H (b) | | H (b) |
| RAD 8843 | | H (b) | | | | | |
| RAD 9101 | | H (b) | | | | | |
| RAD 9177 | | H (b) | | | | | |
| RAD 9274 | | H (b) | | | | | |
| RAD 9307 | pi, pi, H (b) | | | | | | |
| RAD 9321 | | H (b) | | H (s) | | | |
| RAD 9547 | | H (b) | | | | | |
| RAD 9693 | | H (b) | | | | | |
| RAD 9761 | | | | | H (b) | | |
| RAD 9799 | | H (b) | | H (b) | | | |

Only 4 of the 27 compounds make the π-π stacking interaction with Trp 80 residue of AKT1. 20 compounds make hydrogen bond with backbone of polar residue Thr 211. Since the compulsory π-π stacking interaction with Trp 80 was observed for just a few compounds in this cluster, it was decided to discard this family and carry on evaluations with other clusters.

### 4.10.2 Assessment of RAD-family2

The similarity tree of the second cluster including 20 compounds obtained from ZINC database filtered with RAD hypothesis is represented in Figure 4.18.

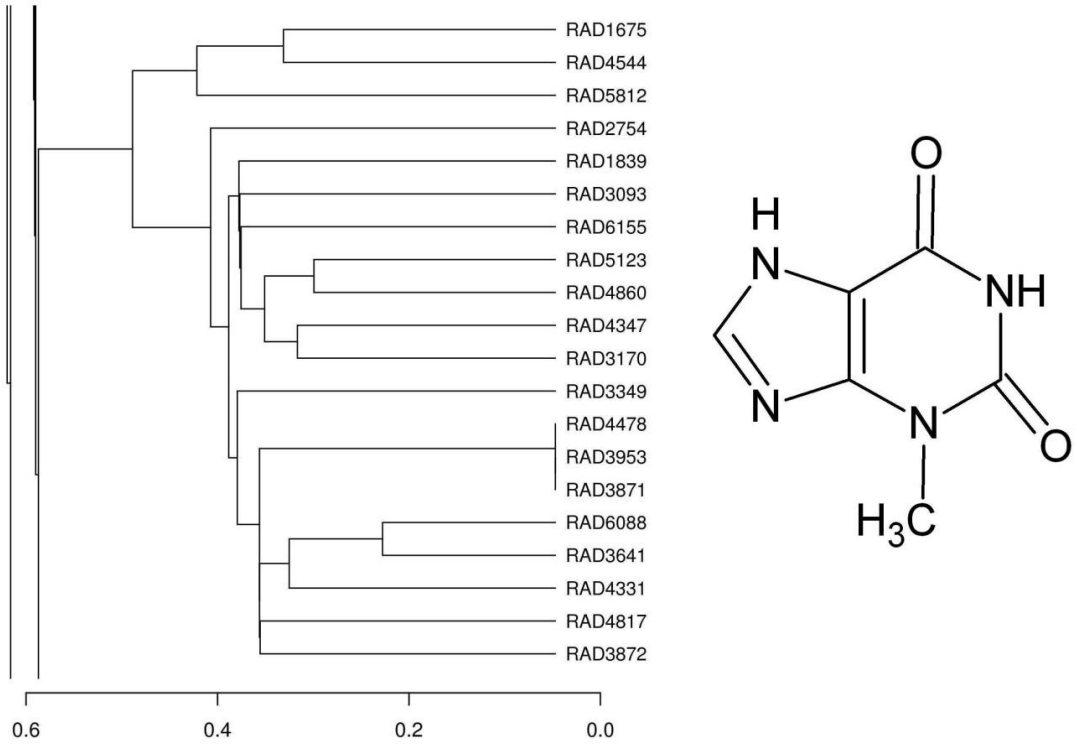



Figure 4.18. Similarity tree for RAD-family2 cluster based on Tanimoto Coefficient (left), similar sub-structure 3-methyl-xanthine for all compounds in RAD-family2 cluster (right).

The Tanimoto similarity coefficient for the farthest compounds in RAD-family2 cluster was calculated as ~0.55. The GlideScore values for the 20 compounds varied between -10.2 kcal/mol and -9.0 kcal/mol. This cluster contains compounds that share the 3-methyl-xanthine scaffold. The interaction diagrams were generated for all of the

compounds to identify the ligand-protein interactions and check for the compulsory π-π stacking interaction of a ring structure of ligand and Trp 80 of AKT1. The interactions of the 20 compounds in RAD-family2 cluster with the protein are summarized in Table 4.8.

Table 4.8. Interactions with AKT1 for 20 compounds in RAD-family2 cluster. Discarded compounds are shown in bold.

| | TRP 80 | THR 211 | SER 205 | ASP 292 | TYR 272 | THR 81 | GLN 79 |
|---|---|---|---|---|---|---|---|
| RAD 1675 | pi | | H (b) | | | | |
| RAD 1839 | pi | | H (b) | | | | |
| RAD 2754 | pi | | H (b) | | | | |
| RAD 3093 | pi | | H (b) | | | | |
| RAD 3170 | pi | | H (b) | H (s) | | | |
| RAD 3349 | pi | | H (b) | | | | |
| RAD 3641 | pi | | H (b) | | | | |
| RAD 3871 | pi | | H (b) | H (s) | | | |
| RAD 3872 | pi | | H (b) | H (s) | | H (b) | |
| RAD 3953 | pi | | H (b) | H (s) | | | |
| **RAD 4331** | | | H (b) | | | | |
| RAD 4347 | pi | | H (b) | | | | |
| RAD 4478 | pi | | H (b) | H (s) | | | |
| RAD 4544 | pi | | | | pi | | H (b) |
| RAD 4817 | pi | | H (b) | H (s) | | | |
| RAD 4860 | pi | | H (b) | | | | |
| **RAD 5123** | | H (b) | | | | | |
| RAD 5812 | pi | | | | pi | | H (b) |
| RAD 6088 | pi | | H (b) | | | | |
| RAD 6155 | pi | | H (b) | | | | |

18 of the 20 compounds make the π-π stacking interaction with Trp 80 residue of AKT1. Additionally, 6 of these 18 compounds make hydrogen bond with side chain of negatively charged residue Asp 292 and 16 of these 18 compounds make hydrogen bond with backbone of polar residue Ser 205. Ser 205 makes the only direct hydrogen bond with

Inhibitor VIII in the crystal structure (Wu, 2010). Since the compulsory π-π stacking interaction with Trp 80 was observed for almost all compounds in this cluster, all but two members of this family were kept for further analyses (the discarded compounds that do not make interaction with Trp80 are shown in bold in Table 4.8.

### 4.10.3 Assessment of RAD-family3

The similarity tree of the third cluster including 12 compounds obtained from ZINC database filtered with RAD hypothesis is given in Figure 4.19.

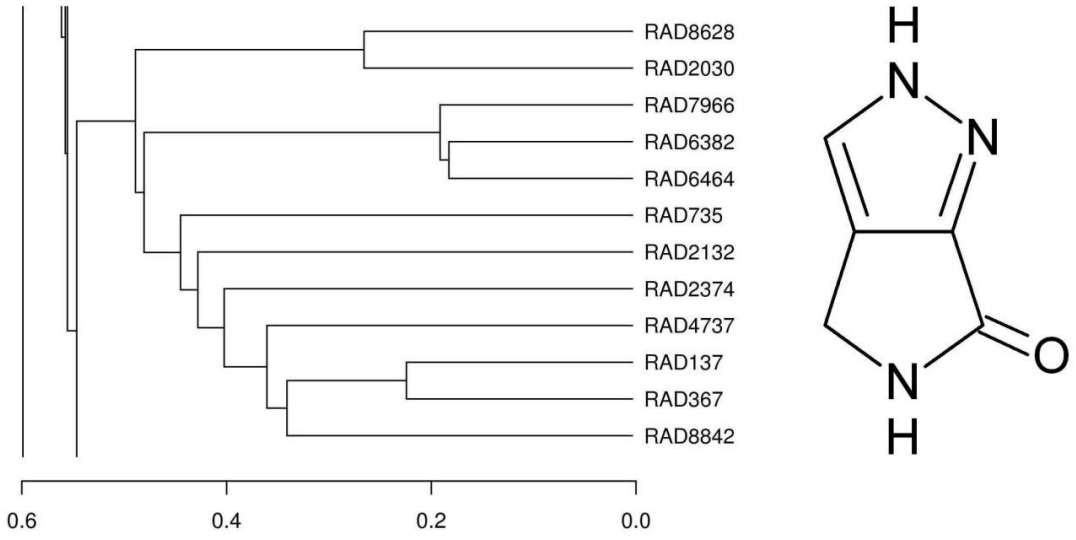



Figure 4.19. Similarity tree for RAD-family3 cluster based on Tanimoto Coefficient (left), similar sub-structure 4,5-dihydropyrrolo[3,4-c]pyrazol-6(2H)-one for nine compounds in RAD-family3 cluster (right).

The highest Tanimoto similarity coefficient between members of the RAD-family3 cluster was ~0.55. The GlideScore values for the 12 compounds varied between -10.3 kcal/mol and -9.0 kcal/mol. This cluster contains nine compounds (except RAD6464, RAD6382 and RAD7966) that share the 4,5-dihydropyrrolo[3,4-c]pyrazol-6(2*H*)-one scaffold. The interaction diagrams were generated for all of the compounds to identify the ligand-protein interactions and check for the compulsory π-π stacking interaction of a ring structure of ligand and Trp 80 of AKT1. The interactions of the 12 compounds in RAD-family3 cluster with the protein are summarized in Table 4.9.

Table 4.9. Interactions with AKT1 for 12 compounds in RAD-family3 cluster.

| | TRP 80 | THR 211 | LYS 268 | ASN 53 | GLN 79 | LEU 78 | Water 455 | THR 291 | VAL 271 |
|---|---|---|---|---|---|---|---|---|---|
| RAD 137 | pi | | | | H (s) | | | | |
| RAD 367 | | | | | | | H (s) | | |
| RAD 735 | pi | H (b) | | | | | H (s) | | |
| RAD 2030 | H (b) | | | H (s) | H (s) | H (b) | | | |
| RAD 2132 | pi | H (b) | | | | | | | |
| RAD 2374 | pi | | H (s) | | H (s) | | | H (b) | |
| RAD 4737 | | H (b) | | | H (s) | | | | |
| RAD 6382 | | | | | | | H (s) | | |
| RAD 6464 | | | | | | | H (s) | | |
| RAD 7966 | | | | | | | H (s) | | |
| RAD 8628 | | | | H (b) | H (s) | | | | |
| RAD 8842 | H (s) | H (b) | | | | | | | H (b) |

Only 4 of 12 compounds make the π-π stacking interaction with Trp 80 residue of AKT1. Most common hydrogen bonds were made between ligands and backbone of Thr 211, side chain of Gln 79 residues and conserved water 455, however none of these interactions were observed for more than 5 compounds. Since the compulsory π-π stacking interaction with Trp 80 was observed for just a few compounds in this cluster, it was decided to discard this family and carry on evaluations with other clusters.

### 4.10.4 Assessment of RAD-family4

The similarity tree of the fourth cluster including 22 compounds obtained from ZINC database filtered with RAD hypothesis is represented in Figure 4.20.

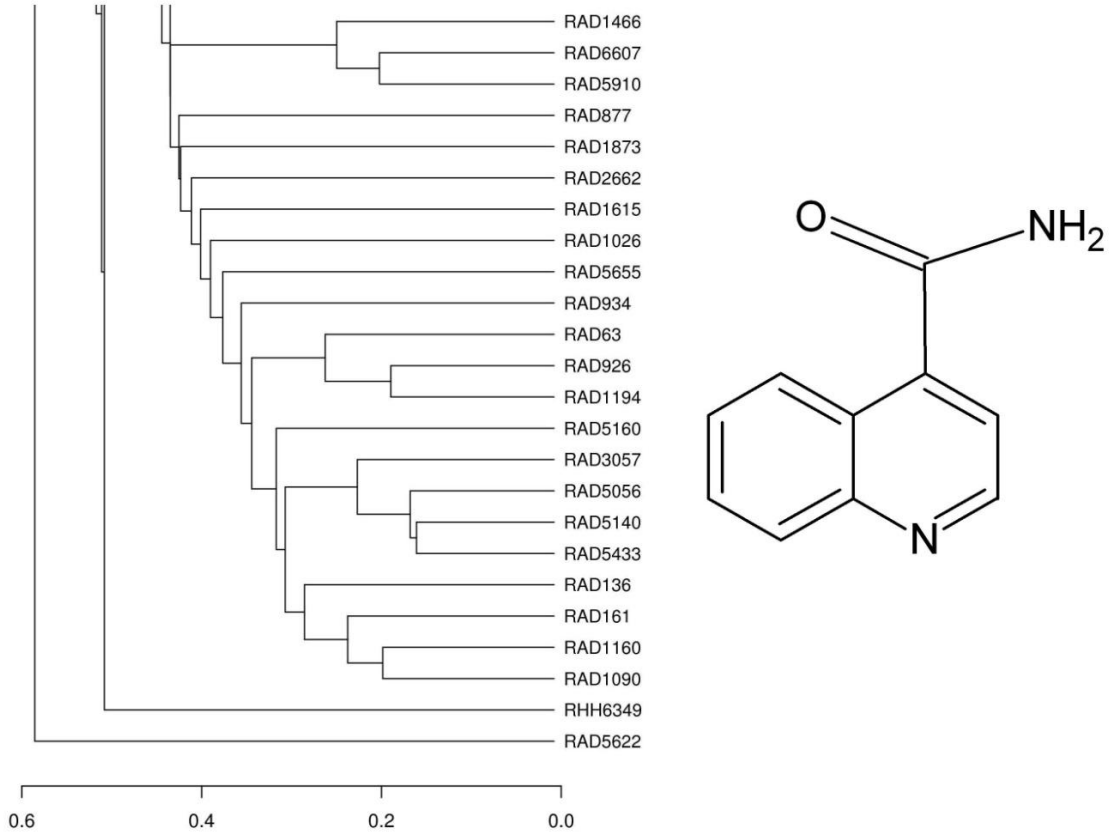



Figure 4.20. Similarity tree for RAD-family4 cluster based on Tanimoto Coefficient (left), similar sub-structure quinoline-4-carboxamide for all compounds in RAD-family4 cluster (right).

The Tanimoto similarity coefficient for the farthest compounds in RAD-family4 cluster was calculated as ~0.55. The GlideScore values for the 22 compounds varied between -10.6 kcal/mol and -9.0 kcal/mol. This cluster contains compounds that share the quinoline-4-carboxamide scaffold. The interaction diagrams were generated for all of the compounds to identify the ligand-protein interactions and check for the compulsory π-π stacking interaction of a ring structure of ligand and Trp 80 of AKT1. The interactions of the 22 compounds in RAD-family4 cluster with the protein are summarized in Table 4.10.

Table 4.10. Interactions with AKT1 for 22 compounds in RAD-family4 cluster.

| | TRP 80 | THR 211 | LYS 268 | ASN 53 | ASP 292 | TYR 272 | GLN 79 | Water 455 | GLY 294 | THR 82 |
|---|---|---|---|---|---|---|---|---|---|---|
| RAD 63 | pi | | | | | pi | | | | |
| RAD 136 | pi | | | | | | H (s) | H (s) | | |
| RAD 161 | pi | | H (s) | | | | H (s) | H (s) | | |
| RAD 877 | pi | | | | | | H (s) | H (s) | | |
| RAD 926 | pi, pi | | | H (s) | | pi | H (s) | H (s) | | |
| RAD 934 | pi | | | | | pi | | | H (b) | |
| RAD 1026 | pi | | | | | H (b) | | | | |
| RAD 1090 | pi | | H (s) | | | | H (s) | H (s) | | |
| RAD 1160 | pi, pi | | H (s) | | | pi | H (s) | H (s) | | |
| RAD 1194 | pi | | | | | pi | | H (s) | | |
| RAD 1466 | pi | | | | H (s) | | | | | |
| RAD 1615 | pi | | | | | | | | | |
| RAD 1873 | pi | | | | H (s) | | | | | H (s) |
| RAD 2662 | pi, pi | | | H (s) | | pi, H (s) | | | | |
| RAD 3057 | pi | | | | | pi | H (s) | H (s) | | |
| RAD 5056 | pi | | | | | | H (s) | H (s) | | |
| RAD 5140 | pi | | | | | | H (s) | H (s) | | |
| RAD 5160 | pi | | | | | | H (s) | H (s) | | |
| RAD 5433 | pi | | | | | pi | H (s) | H (s) | | |
| RAD 5655 | pi | H (b) | | | | | | | | |
| RAD 5910 | pi | | | | | | | H (s) | | |
| RAD 6607 | pi | | | | H (s) | | | | | |

All compounds in RAD-family4 were observed to make the π-π stacking interaction with Trp 80 and 3 compounds made double π-π stacking interaction. Additionally, 11 compounds make hydrogen bond with side-chain of polar residue Gln 79 and 13 compounds make hydrogen bond with conserved water 455. Interestingly, all compounds interacting with Gln 79 are also observed to make hydrogen bond with Water 455. The maximum number of ligand-protein interactions was at most four for a single compound. Since the compulsory π-π stacking interaction with Trp 80 was observed for all compounds

in this cluster, it was decided to keep this family for further evaluations like considering ADME criteria and drug-likeness.

### 4.10.5 Assessment of RHH-family5

The similarity tree of the fifth cluster including 13 compounds obtained from ZINC database filtered with RHH hypothesis is represented in Figure 4.21.

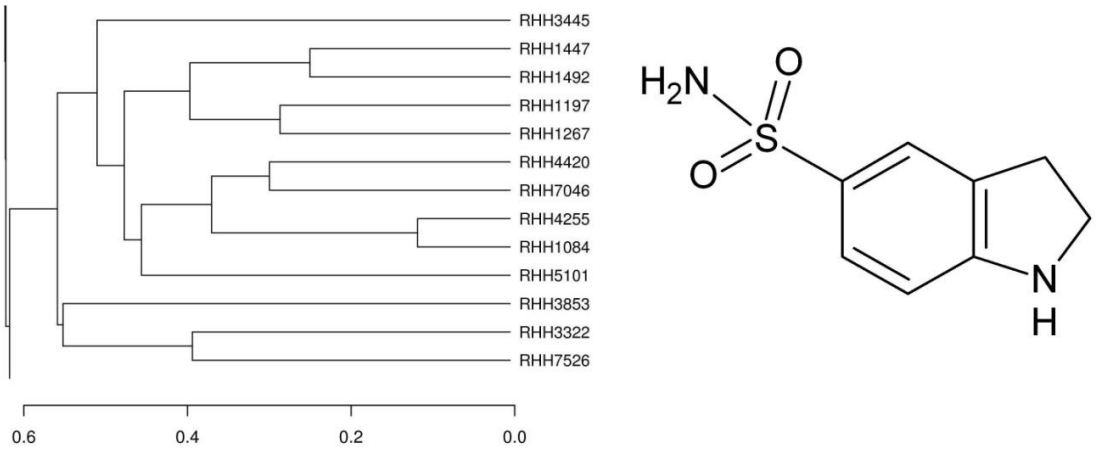



Figure 4.21. Similarity tree for RHH-family5 cluster based on Tanimoto Coefficient (left), similar sub-structure 2,3-dihydro-1H-indole-5-sulfonamide for all compounds in RHH-family5 cluster (right).

The Tanimoto similarity coefficient for the farthest compounds in RHH-family5 cluster was calculated as ~0.45. The only cluster detected from docking hits of ZINC database filtered with RHH hypothesis is RHH-family5. This cluster contains compounds that share the 2,3-dihydro-1*H*-indole-5-sulfonamide scaffold. The GlideScore values for the 13 compounds varied between -9.56 kcal/mol and -9.0 kcal/mol. The interaction diagrams were generated for all of the compounds to identify the ligand-protein interactions and check for the compulsory π-π stacking interaction of a ring structure of ligand and Trp 80 of AKT1. The interactions of the 13 compounds in RHH-family5 cluster with the protein are summarized in Table 4.11.

Table 4.11. Interactions with AKT1 for 13 compounds in RHH-family5 cluster. Discarded compounds are shown in bold.

| | TRP 80 | ASN 53 | SER 216 |
|---|---|---|---|
| **RHH 1084** | H (b) | | H (s) |
| **RHH 1197** | H (b) | | |
| RHH 1267 | pi, H (b) | | |
| RHH1447 | pi, pi, H (b) | | |
| RHH 1492 | pi, pi, H (b) | | |
| RHH 3322 | pi, H (b) | | |
| **RHH 3445** | H (b) | | |
| RHH 3853 | pi | H (s) | |
| RHH 4255 | pi, H (b) | | |
| **RHH 4420** | H (b) | | |
| **RHH 5101** | H (b) | | |
| RHH 7046 | pi, H (b) | H (s) | |
| RHH 7526 | pi, pi | | |

8 of 13 compounds in RHH-family5 are observed to make the π-π stacking interaction and 11 compounds make hydrogen bond with Trp80. Additionally, 2 hydrogen bonds are made with side chain of polar residue Asn 53 and a single hydrogen bond is made with polar residue Ser 216. Since the compulsory π-π stacking interaction with Trp 80 was observed for a high percentage of compounds in this cluster, all but four members of this family were kept for further analyses (the discarded compounds that do not make interaction with Trp80 are shown in bold in Table 4.11).

### 4.10.6 Assessment of hits of APRRR hypothesis

The interactions of the 16 hits of APRRR hypothesis with Akt1 are summarized in Table 4.12.

.

Table 4.12. Interactions with AKT1 for 16 hits of APRRR hypothesis. Discarded compounds are shown in bold.

| | Trp 80 | Tyr 272 | Thr 211 | Ala 58 | Leu 78 | Asn 53 | Ser 205 | Leu 268 | Cys 296 | Wat 455 | Asp 292 | Phe 293 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ZINC 36733118 | pi | | | | | | | | | | | |
| ZINC 09060134 | pi | pi | | H (b) | H (b) | | | | | | | |
| ZINC 09410451 | pi | pi | H (b) | H (b) | H (b) | | | | | | | |
| ZINC 09410511 | pi, pi | pi | H (b) | | | H (b) | H (b) | | | | | |
| ZINC 09410906 | pi | pi | | H (b) | H (b) | | | | | | | |
| ZINC 13624933 | pi, H (b) | | | | H (b) | | | | | | | |
| ZINC 11866403 | pi, H (b) | | | | H (b) | | | | | | | |
| ZINC 19800260 | pi, pi | | | | | H(s) | | | | | | |
| ZINC 32623328 | pi | | | | | | | H (s) | | | | |
| ZINC 12027062 | pi | | | | | | | | | | | |
| ZINC 12278696 | pi, pi, H(b) | | | | | | | | | | | |
| **ZINC 12021111** | H (s) | pi | H (b) | | | | | | | | | |
| **ZINC 18046815** | | | H (b) | | | H (s) | | | | | | |
| **ZINC 20393001** | | | | | | | | | H (b) | H (s) | | |
| **ZINC 09318983** | | | | | | | | | | | | H(b) |
| **ZINC 02639096** | | | | | | | | | | | H (s) | |

11 of 16 hits of APRRR hypothesis are observed to make the π-π stacking interaction and four compounds make hydrogen bond with Trp80. Additionally, five hydrogen bonds are made with backbone of Leu 78 and a π-π stacking interaction was observed between five compounds and Tyr 272. Eleven hits were kept for further analyses (the discarded compounds that do not make interaction with Trp80 are shown in bold in Table 4.12).

## 4.11 Evaluation of Pharmacokinetic Properties for Selected Similar Families

Clustering of similar compounds which have better GScore values than -9.0 kcal/mol resulted in detection of five clusters. Four clusters included compounds based on the RAD pharmacophore hypothesis and there was a single cluster for compounds based on the RHH pharmacophore hypothesis, followed by docking. The clusters were further examined for the presence of the compulsory π-π stacking interaction with Trp 80. As a result, RAD-family1 and RAD-family3 were eliminated. ADME and drug-likeness of compounds in RAD-family2, RAD-family, RHH-family5 and APRRR-1 were evaluated.

ADME (absorption, distribution, metabolism, excretion) predictions were performed by QikProp program of Maestro interface. QikProp also provides acceptable ranges for comparing the predicted properties of compounds with those of 95% of known drugs. SASA and its contributors FOSA, FISA, PISA and WPSA, PSA and percentage of Human Oral Absorption were considered as key descriptors for ADME criteria. The ranges of these descriptors were given in corresponding property tables. The compounds that were out of these ranges were eliminated from further analyses.

Each compound has a unique title assigned by Maestro, and a ZINC code assigned by the database. Number of stars (#stars) shows the number of descriptor values that fall outside the 95% range of similar values for known drugs. SASA is defined as the total solvent accessible surface area ($Å^2$). FOSA is the hydrophobic component of the SASA, which is associated with a saturated carbon and attached hydrogen. FISA is the hydrophilic component of the SASA, which is associated with the solvent accessible surface area on N, O, and H on heteroatoms. PISA is the π component of the SASA and π is referred to a carbon and attached hydrogen. WPSA is the weakly polar component of the SASA, which takes the polar components as electronegative atoms such as nitrogen and oxygen, halogens, phosphorus and sulphur atoms. PSA is the van der Waals surface area of polar nitrogen and oxygen atoms. Percentage of human oral absorption (HOA%) is accepted as good when greater than 80% and poor when less than 25%. For each compound, the properties described above were represented in Table 4.14 for RAD-family2, Table 4.16 for RAD-family4 and Table 4.18 for RHH-family5. The discarded compounds which were out of given ranges were indicated by strikethrough.

In addition to ADME properties, drug-likeness properties were also estimated by QikProp. The drug-likeness of a compound was assessed according to Lipinski's Rule of Five [97], which considers molecular weight (<500 Da), number of hydrogen bond acceptors (≤10) and donors (≤5), and octanol-water partition coefficient (≤5). Octanol-water partition coefficient is associated with lipophilicity; permeation through biological membranes becomes easier as lipophilicity becomes higher up to a limit value however higher lipophilicity results in poorer solubility in water. The non-integer values of hydrogen bond donors or acceptors are due to evaluation of these values from varying conformations. The acceptable upper and/or lower limits of drug-likeness properties of compounds are given in corresponding tables for ease of comparison with those of 95% of known drugs.

### 4.11.1 Predicted Drug-Likeness and ADME Properties for RAD-family2

Initially, drug-likeness was investigated according to Lipinski's Rule of Five. Predicted properties are given in Table 4.13 for RAD-family2.

Table 4.13. Drug-likeness properties for RAD-family2 considering Lipinski's Rule of Five.

| Title | ZINC Code | MW | donorHB | accptHB | QPlogPo/w | RuleOfFive |
|---|---|---|---|---|---|---|
| | | 130.0<br>725.0 | 0.0<br>6.0 | 2.0<br>20.0 | -2.0<br>6.5 | #violations |
| 3871 | ZINC00839381 | 406.46 | 3 | 9.15 | 2.01 | 0 |
| 3872 | ZINC02373214 | 426.87 | 3 | 9.15 | 2.08 | 0 |
| 5812 | ZINC00626346 | 471.52 | 3 | 7.95 | 4.16 | 0 |
| 4817 | ZINC04102500 | 436.90 | 3 | 9.95 | 1.92 | 0 |
| 4860 | ZINC19904676 | 458.56 | 3 | 9.95 | 3.00 | 0 |
| 3170 | ZINC13810798 | 388.43 | 5 | 8.95 | 0.88 | 0 |
| 4544 | ZINC13512705 | 482.93 | 3 | 9.45 | 3.63 | 0 |
| 1675 | ZINC05295022 | 462.51 | 3 | 9.95 | 3.04 | 0 |
| 2754 | ZINC00625908 | 447.92 | 3 | 7.95 | 3.63 | 0 |
| 3953 | ZINC00839381 | 406.46 | 3 | 9.15 | 2.03 | 0 |
| 4347 | ZINC04454246 | 399.45 | 2 | 7.95 | 2.86 | 0 |
| 3349 | ZINC04386258 | 401.46 | 3 | 7.95 | 3.03 | 0 |
| 1839 | ZINC00822339 | 417.46 | 2 | 8.7 | 2.93 | 0 |
| 3093 | ZINC05360472 | 442.52 | 2 | 9.95 | 2.33 | 0 |
| 6155 | ZINC00662936 | 449.51 | 2 | 7.95 | 3.72 | 0 |
| 3641 | ZINC00850903 | 404.48 | 2 | 7.45 | 3.15 | 0 |
| 4478 | ZINC00839381 | 406.46 | 3 | 9.15 | 2.02 | 0 |
| 6088 | ZINC00633233 | 390.46 | 2 | 7.45 | 2.97 | 0 |

All compounds in RAD-family2 have predicted drug-likeness properties in the same range with those of 95% of known drugs and also no violations to Rule of Five exist.

Table 4.14. Predicted ADME properties for RAD-family2. Eliminated molecules are indicated by strikethrough and out-of-range property shown in bold.

| Title | ZINC code | #stars | SASA | FOSA | FISA | PISA | WPSA | PSA | HOA% |
|---|---|---|---|---|---|---|---|---|---|
| | | | 300.0<br>1000.0 | 0.0<br>750.0 | 7.0<br>330.0 | 0.0<br>450.0 | 0.0<br>175.0 | 7.0<br>200.0 | <25% low<br>>80% high |
| 3871 | ZINC00839381 | 0 | 697.35 | 267.91 | 183.85 | 227.56 | 18.03 | 133.60 | 79.04 |
| 3872 | ZINC02373214 | 0 | 691.11 | 221.87 | 196.52 | 179.77 | 92.94 | 135.22 | 77.26 |
| ~~**5812**~~ | ~~**ZINC00626346**~~ | ~~0~~ | ~~791.28~~ | ~~141.97~~ | ~~153.13~~ | ~~496.18~~ | ~~**0.00**~~ | ~~121.50~~ | ~~96.83~~ |
| 4817 | ZINC04102500 | 0 | 731.72 | 314.07 | 166.72 | 179.26 | 71.67 | 128.57 | 70.64 |
| ~~**4860**~~ | ~~**ZINC19904676**~~ | ~~0~~ | ~~851.74~~ | ~~553.77~~ | ~~144.74~~ | ~~153.23~~ | ~~**0.00**~~ | ~~127.30~~ | ~~80.65~~ |
| ~~**3170**~~ | ~~**ZINC13810798**~~ | ~~0~~ | ~~684.70~~ | ~~311.90~~ | ~~206.60~~ | ~~166.19~~ | ~~**0.00**~~ | ~~150.27~~ | ~~57.78~~ |
| 4544 | ZINC13512705 | 0 | 802.37 | 196.21 | 157.09 | 377.37 | 71.69 | 132.04 | 93.06 |
| ~~**1675**~~ | ~~**ZINC05295022**~~ | ~~0~~ | ~~766.12~~ | ~~281.48~~ | ~~154.43~~ | ~~330.22~~ | ~~**0.00**~~ | ~~127.31~~ | ~~90.03~~ |
| 2754 | ZINC00625908 | 0 | 745.23 | 358.77 | 134.55 | 180.25 | 71.67 | 121.97 | 96.88 |
| 3953 | ZINC00839381 | 0 | 696.42 | 267.95 | 182.15 | 227.49 | 18.83 | 133.64 | 79.42 |
| ~~**4347**~~ | ~~**ZINC04454246**~~ | ~~0~~ | ~~695.49~~ | ~~386.07~~ | ~~139.34~~ | ~~170.09~~ | ~~**0.00**~~ | ~~112.89~~ | ~~91.55~~ |
| ~~**3349**~~ | ~~**ZINC04386258**~~ | ~~0~~ | ~~721.16~~ | ~~417.13~~ | ~~134.59~~ | ~~169.44~~ | ~~**0.00**~~ | ~~119.88~~ | ~~93.36~~ |
| ~~**1839**~~ | ~~**ZINC00822339**~~ | ~~0~~ | ~~705.88~~ | ~~416.42~~ | ~~127.32~~ | ~~162.14~~ | ~~**0.00**~~ | ~~120.91~~ | ~~94.00~~ |
| ~~**3093**~~ | ~~**ZINC05360472**~~ | ~~0~~ | ~~761.98~~ | ~~460.04~~ | ~~146.80~~ | ~~155.15~~ | ~~**0.00**~~ | ~~117.64~~ | ~~76.39~~ |
| ~~**6155**~~ | ~~**ZINC00662936**~~ | ~~0~~ | ~~741.39~~ | ~~309.43~~ | ~~140.12~~ | ~~291.84~~ | ~~**0.00**~~ | ~~110.83~~ | ~~96.46~~ |
| 3641 | ZINC00850903 | 0 | 678.66 | 356.37 | 140.42 | 170.33 | 11.54 | 109.95 | 93.10 |
| 4478 | ZINC00839381 | 0 | 698.02 | 269.26 | 182.61 | 227.98 | 18.16 | 132.95 | 79.32 |
| 6088 | ZINC00633233 | 0 | 682.65 | 343.97 | 137.71 | 179.55 | 21.42 | 111.05 | 92.46 |

The molecules with titles 5812, 4860, 3170, 1675, 4347, 3349, 1839, 3093 and 6155 were eliminated from RAD-family2 due to being out of range for some ADME properties. For the remaining 9 compounds, fitness to RAD pharmacophore hypothesis was considered as a last elimination criteria before selecting final hits to be proposed.

### 4.11.2 Predicted Drug-likeness and ADME Properties for RAD-family4

Predicted drug-likeness properties according to Lipinski's Rule of Five are given in Table 4.15 for RAD-family4.

Table 4.15. Drug-likeness properties for RAD-family4 considering Lipinski's Rule of Five.

| Title | ZINC Code | MW | donorHB | accptHB | QPlogPo/w | RuleOfFive |
|---|---|---|---|---|---|---|
| | | 130.0<br>725.0 | 0.0<br>6.0 | 2.0<br>20.0 | -2.0<br>6.5 | #violations |
| 926 | ZINC02103221 | 425.49 | 2 | 5.75 | 4.88 | 0 |
| 161 | ZINC16392469 | 485.56 | 2 | 6.5 | 4.81 | 0 |
| 1466 | ZINC13321559 | 441.49 | 3 | 5.25 | 4.83 | 0 |
| 1873 | ZINC16392736 | 475.52 | 3 | 9.5 | 2.60 | 0 |
| 5655 | ZINC05936944 | 385.42 | 2 | 6.5 | 3.81 | 0 |
| 5433 | ZINC13321319 | 428.51 | 2 | 6 | 4.38 | 0 |
| 1160 | ZINC02753923 | 485.56 | 2 | 6.5 | 4.89 | 0 |
| 877 | ZINC01073862 | 430.52 | 2 | 6.5 | 4.15 | 0 |
| 5140 | ZINC02867568 | 413.49 | 2 | 5 | 4.59 | 0 |
| 5910 | ZINC13322330 | 398.42 | 3 | 6 | 3.58 | 0 |
| 5056 | ZINC08917432 | 431.48 | 2 | 5 | 4.90 | 0 |
| 2662 | ZINC20158772 | 466.50 | 2 | 7 | 4.87 | 0 |
| 1194 | ZINC02867745 | 409.49 | 2 | 5 | 4.98 | 0 |
| 5160 | ZINC01073706 | 479.00 | 3 | 6.5 | 4.72 | 0 |
| 136 | ZINC00656127 | 445.49 | 2 | 6.5 | 4.06 | 0 |
| 1090 | ZINC01073894 | 459.52 | 2 | 6.5 | 4.46 | 0 |
| 6607 | ZINC13321539 | 411.46 | 3 | 4.5 | 4.81 | 0 |
| 3057 | ZINC13321889 | 442.53 | 2 | 6.5 | 4.42 | 0 |
| 1026 | ZINC01068312 | 445.49 | 2 | 7 | 3.89 | 0 |
| 63 | ZINC02818676 | 411.42 | 2 | 6.5 | 3.54 | 0 |
| 934 | ZINC02742477 | 426.47 | 1 | 7 | 4.70 | 0 |
| 1615 | ZINC01049548 | 433.48 | 3 | 8.75 | 2.85 | 0 |

All compounds in RAD-family4 have predicted drug-likeness properties in the same range with those of 95% of known drugs and also no violations to Rule of Five exist. The predicted ADME properties are shown in Table 4.16.

Table 4.16. Predicted ADME properties for RAD-family4. Eliminated molecules are indicated by strikethrough and out-of-range property shown in bold.

| Title | ZINC Code | #stars | SASA | FOSA | FISA | PISA | WPSA | PSA | HOA% |
|---|---|---|---|---|---|---|---|---|---|
| | | | 300.0<br>1000.0 | 0.0<br>750.0 | 7.0<br>330.0 | 0.0<br>450.0 | 0.0<br>175.0 | 7.0<br>200.0 | <25% low<br>>80% high |
| ~~**926**~~ | ~~**ZINC02103221**~~ | ~~0~~ | ~~763.01~~ | ~~173.73~~ | ~~124.91~~ | ~~**464.38**~~ | ~~**0.00**~~ | ~~100.27~~ | ~~100.00~~ |
| 161 | ZINC16392469 | 0 | 753.98 | 334.14 | 113.05 | 265.99 | 40.80 | 108.22 | 100.00 |
| ~~**1466**~~ | ~~**ZINC13321559**~~ | ~~0~~ | ~~788.01~~ | ~~227.79~~ | ~~164.55~~ | ~~395.67~~ | ~~**0.00**~~ | ~~107.16~~ | ~~100.00~~ |
| 1873 | ZINC16392736 | 0 | 726.33 | 183.82 | 190.74 | 349.85 | 1.92 | 128.79 | 81.32 |
| ~~**5655**~~ | ~~**ZINC05936944**~~ | ~~0~~ | ~~665.64~~ | ~~169.35~~ | ~~76.17~~ | ~~420.12~~ | ~~**0.00**~~ | ~~77.16~~ | ~~100.00~~ |
| 5433 | ZINC13321319 | 0 | 720.11 | 179.14 | 129.83 | 369.89 | 41.24 | 99.46 | 100.00 |
| 1160 | ZINC02753923 | 0 | 745.39 | 278.51 | 107.73 | 318.99 | 40.17 | 107.84 | 100.00 |
| 877 | ZINC01073862 | 0 | 748.59 | 311.65 | 148.72 | 256.33 | 31.90 | 103.05 | 100.00 |
| 5140 | ZINC02867568 | 0 | 705.96 | 150.31 | 126.87 | 390.28 | 38.50 | 90.86 | 100.00 |
| ~~**5910**~~ | ~~**ZINC13322330**~~ | ~~0~~ | ~~714.73~~ | ~~63.68~~ | ~~183.66~~ | ~~**467.39**~~ | ~~**0.00**~~ | ~~112.40~~ | ~~88.27~~ |
| 5056 | ZINC08917432 | 0 | 712.30 | 150.16 | 119.03 | 351.90 | 91.21 | 90.65 | 100.00 |
| ~~**2662**~~ | ~~**ZINC20158772**~~ | ~~0~~ | ~~801.66~~ | ~~228.39~~ | ~~116.39~~ | ~~**456.88**~~ | ~~**0.00**~~ | ~~112.46~~ | ~~100.00~~ |
| ~~**1194**~~ | ~~**ZINC02867745**~~ | ~~0~~ | ~~739.55~~ | ~~163.52~~ | ~~126.56~~ | ~~**449.47**~~ | ~~**0.00**~~ | ~~92.55~~ | ~~100.00~~ |
| 5160 | ZINC01073706 | 0 | 761.06 | 267.57 | 135.55 | 247.66 | 110.28 | 102.58 | 100.00 |
| 136 | ZINC00656127 | 0 | 712.19 | 220.49 | 126.12 | 333.86 | 31.72 | 110.66 | 100.00 |
| 1090 | ZINC01073894 | 0 | 743.82 | 238.73 | 129.75 | 340.55 | 34.79 | 110.53 | 100.00 |
| ~~**6607**~~ | ~~**ZINC13321539**~~ | ~~0~~ | ~~772.01~~ | ~~154.28~~ | ~~161.30~~ | ~~**456.44**~~ | ~~**0.00**~~ | ~~99.01~~ | ~~100.00~~ |
| 3057 | ZINC13321889 | 0 | 735.51 | 207.46 | 136.94 | 350.89 | 40.22 | 101.85 | 100.00 |
| 1026 | ZINC01068312 | 0 | 744.46 | 146.84 | 179.28 | 383.61 | 34.72 | 123.91 | 90.83 |
| ~~**63**~~ | ~~**ZINC02818676**~~ | ~~0~~ | ~~666.46~~ | ~~72.11~~ | ~~136.26~~ | ~~**458.09**~~ | ~~**0.00**~~ | ~~111.03~~ | ~~96.07~~ |
| ~~**934**~~ | ~~**ZINC02742477**~~ | ~~0~~ | ~~773.99~~ | ~~258.27~~ | ~~97.56~~ | ~~418.17~~ | ~~**0.00**~~ | ~~85.33~~ | ~~100.00~~ |
| ~~**1615**~~ | ~~**ZINC01049548**~~ | ~~0~~ | ~~773.30~~ | ~~92.71~~ | ~~199.84~~ | ~~**478.82**~~ | ~~1.94~~ | ~~116.83~~ | ~~81.25~~ |

The ADME properties of the molecules with titles 926, 1466, 5655, 5910, 2662, 1194, 6607, 63, 934 and 1615 were out of range. For the remaining 12 compounds, fitness to RAD pharmacophore hypothesis was considered as a last elimination criteria before selecting final hits to be proposed.

### 4.11.3 Predicted ADME Properties for RHH-family5

Predicted drug-likeness properties according to Lipinski's Rule of Five are given in Table 4.17 for RHH-family5.

Table 4.17. Drug-likeness properties for RHH-family5 considering Lipinski's Rule of Five.

| Title | ZINC Code | mol MW | donorHB | accptHB | QPlogPo/w | RuleOfFive |
|---|---|---|---|---|---|---|
| | | 130.0<br>725.0 | 0.0<br>6.0 | 2.0<br>20.0 | -2.0<br>6.5 | #violations |
| 3322 | ZINC09397086 | 461.59 | 1 | 10 | 2.51 | 0 |
| 1197 | ZINC09429279 | 477.64 | 1.25 | 9.25 | 3.59 | 0 |
| 1447 | ZINC09429089 | 463.98 | 1.25 | 9.25 | 3.49 | 0 |
| 4255 | ZINC20529671 | 461.96 | 1 | 9 | 3.79 | 0 |
| 7046 | ZINC20529899 | 445.53 | 0 | 10 | 2.09 | 0 |
| 1267 | ZINC09429204 | 457.59 | 1.25 | 9.25 | 3.41 | 0 |
| 5101 | ZINC09396806 | 441.54 | 1 | 10 | 2.62 | 0 |
| 7526 | ZINC09397114 | 473.56 | 1 | 10 | 3.37 | 0 |
| 1084 | ZINC20529871 | 445.51 | 1 | 9 | 3.37 | 0 |
| 1492 | ZINC09429139 | 457.59 | 1.25 | 9.25 | 3.57 | 0 |
| 3853 | ZINC09117054 | 455.57 | 0 | 9.25 | 3.11 | 0 |
| 3445 | ZINC06866689 | 420.52 | 0 | 9.5 | 2.56 | 0 |
| 4420 | ZINC20529887 | 441.54 | 1 | 9 | 3.05 | 0 |

All compounds in RHH-family5 have predicted drug-likeness properties in the same range with those of 95% of known drugs and also no violations to Rule of Five exist.

Table 4.18. Predicted ADME properties for RHH-family5. Eliminated molecules are indicated by strikethrough and out-of-range property shown in bold.

| Title | ZINC Code | #stars | SASA | FOSA | FISA | PISA | WPSA | PSA | HOA% |
|---|---|---|---|---|---|---|---|---|---|
| | | | 300.0<br>1000.0 | 0.0<br>750.0 | 7.0<br>330.0 | 0.0<br>450.0 | 0.0<br>175.0 | 7.0<br>200.0 | <25% low<br>>80% high |
| 3322 | ZINC09397086 | 0 | 772.00 | 385.45 | 152.56 | 185.74 | 48.24 | 103.95 | 83.83 |
| 1447 | ZINC09429089 | 0 | 743.78 | 326.45 | 138.00 | 206.91 | 72.43 | 109.02 | 95.49 |
| ~~**4255**~~ | ~~**ZINC20529671**~~ | ~~0~~ | ~~836.29~~ | ~~356.63~~ | ~~178.92~~ | ~~227.17~~ | ~~73.56~~ | ~~113.16~~ | ~~90.28~~ |
| 7046 | ZINC20529899 | 0 | 786.45 | 420.34 | 155.26 | 208.94 | 1.91 | 112.70 | 80.48 |
| ~~**1267**~~ | ~~**ZINC09429204**~~ | ~~0~~ | ~~819.39~~ | ~~422.04~~ | ~~129.62~~ | ~~267.72~~ | ~~**0.00**~~ | ~~108.61~~ | ~~94.22~~ |
| 7526 | ZINC09397114 | 0 | 759.19 | 395.38 | 131.96 | 183.15 | 48.70 | 100.84 | 95.81 |
| ~~**1492**~~ | ~~**ZINC09429139**~~ | ~~0~~ | ~~766.22~~ | ~~441.79~~ | ~~133.42~~ | ~~190.12~~ | ~~**0.89**~~ | ~~105.36~~ | ~~96.74~~ |
| ~~**3853**~~ | ~~**ZINC09117054**~~ | ~~0~~ | ~~718.92~~ | ~~436.39~~ | ~~107.47~~ | ~~174.78~~ | ~~**0.02**~~ | ~~77.79~~ | ~~100.00~~ |

The ADME properties of the molecules with titles 4255, 1267, 1492 and 3853 were out of range. For the remaining 4 compounds, fitness to RHH pharmacophore hypothesis was considered as a last elimination criteria before selecting final hits to be proposed.

### 4.11.4 Predicted Drug-Likeness and ADME Properties for APRRR hits

Initially, drug-likeness was investigated according to Lipinski's Rule of Five. Predicted properties are given in Table 4.19 for APRRR hits.

All compounds of APRRR hypothesis have predicted drug-likeness properties in the same range with those of 95% of known drugs and also no violations to Rule of Five exist. Predicted ADME properties for hits of APRRR hypothesis are shown in Table 4.20.

The molecules with titles ZINC12278696, ZINC09410906, ZINC09060134, ZINC12027062, ZINC3673311 and ZINC32623328 were eliminated due to being out of range for WPSA. For the remaining 5 compounds, fitness to APRRR pharmacophore hypothesis was considered as a last elimination criteria before selecting final hits to be proposed.

Table 4.19. Drug-likeness properties for hits of APRRR hypothesis considering Lipinski's Rule of Five.

| Title | MW | donorHB | accptHB | QPlogPo/w | RuleOfFive |
|---|---|---|---|---|---|
| ZINC36733118 | 420.507 | 1 | 7.7 | 3.346 | 0 |
| ZINC09060134 | 450.536 | 4 | 7.75 | 2.241 | 0 |
| ZINC09410451 | 454.955 | 4 | 7 | 2.614 | 0 |
| ZINC09410511 | 440.928 | 4 | 7.5 | 2.131 | 0 |
| ZINC09410906 | 480.563 | 4 | 8.5 | 2.407 | 0 |
| ZINC13624933 | 485.376 | 2 | 7.75 | 2.932 | 0 |
| ZINC11866403 | 440.925 | 2 | 7.75 | 2.857 | 0 |
| ZINC19800260 | 486.588 | 1 | 7.7 | 4.717 | 0 |
| ZINC32623328 | 435.522 | 1 | 8.45 | 3.874 | 0 |
| ZINC12027062 | 462.544 | 0 | 8.45 | 3.106 | 0 |
| ZINC12278696 | 483.609 | 0 | 8.25 | 4.834 | 0 |

Table 4.20. Predicted ADME properties for hits of APRRR hypothesis. Eliminated molecules are indicated by strikethrough and out-of-range property shown in bold.

| Title | #stars | SASA | FOSA | FISA | PISA | WPSA | PSA | HOA % |
|---|---|---|---|---|---|---|---|---|
| | | 300.0<br>1000.0 | 0.0<br>750.0 | 7.0<br>330.0 | 0.0<br>450.0 | 0.0<br>175.0 | 7.0<br>200.0 | <25% low<br>>80% high |
| ZINC09410511 | 0.00 | 755.46 | 221.15 | 201.99 | 260.63 | 71.69 | 113.66 | 61.32 |
| ZINC09410451 | 0.00 | 781.24 | 302.69 | 199.22 | 207.65 | 71.69 | 113.12 | 63.98 |
| ~~**ZINC12278696**~~ | ~~0.00~~ | ~~736.50~~ | ~~260.09~~ | ~~48.28~~ | ~~428.13~~ | ~~**0.00**~~ | ~~53.14~~ | ~~91.66~~ |
| ~~**ZINC09410906**~~ | ~~0.00~~ | ~~837.03~~ | ~~478.45~~ | ~~199.24~~ | ~~159.34~~ | ~~**0.00**~~ | ~~126.24~~ | ~~62.76~~ |
| ~~**ZINC09060134**~~ | ~~0.00~~ | ~~794.14~~ | ~~395.42~~ | ~~199.21~~ | ~~199.51~~ | ~~**0.00**~~ | ~~121.39~~ | ~~61.80~~ |
| ZINC11866403 | 0.00 | 767.24 | 299.17 | 143.76 | 252.68 | 71.63 | 96.82 | 74.81 |
| ~~**ZINC12027062**~~ | ~~0.00~~ | ~~621.88~~ | ~~316.20~~ | ~~41.22~~ | ~~264.46~~ | ~~**0.00**~~ | ~~64.28~~ | ~~100.00~~ |
| ZINC13624933 | 0.00 | 772.29 | 299.17 | 143.77 | 251.98 | 77.38 | 96.82 | 75.25 |
| ZINC19800260 | 0.00 | 747.66 | 331.00 | 11.61 | 363.82 | 41.24 | 49.78 | 100.00 |
| ~~**ZINC36733118**~~ | ~~0.00~~ | ~~670.56~~ | ~~312.03~~ | ~~87.41~~ | ~~271.12~~ | ~~**0.00**~~ | ~~79.39~~ | ~~92.43~~ |
| ~~**ZINC32623328**~~ | ~~0.00~~ | ~~802.77~~ | ~~425.63~~ | ~~89.29~~ | ~~287.85~~ | ~~**0.00**~~ | ~~85.07~~ | ~~95.20~~ |

## 4.12 Consideration of Fitness to Pharmacophore Hypothesis

The compounds were further examined based on their fitness to chosen pharmacophore hypothesis. Each compound has a ‘fitness score’ which measures its proximity to matching pharmacophore site points and how well the matching vector features such as acceptors, donors and aromatic rings overlay with the hypothesis. Fitness score is a weighted linear combination of site, vector and volume scores and ranges from -1.0 to 3.0. The site score is the root-mean-squared deviation (RMSD) in the site point positions and has a range of 0.0 to 1.0. The vector score is the average cosine of the angles formed by corresponding pairs of vector features (acceptors, donors, aromatic rings) in the aligned structures and has a range of -1.0 to 1.0. The volume score is based on the overlap of van der Waals models of the non-hydrogen atoms in each pair of structures and has a range of 0.0 to 1.0.

As mentioned above, the fitness score ranges from -1.0 to 3.0, a fitness score of -1.0 can be thought as 0% proximity to chosen pharmacophore hypothesis whereas 3.0 corresponds to exact match (100% proximity) of all site points to compound features. The fitness scores versus strain corrected GScores for the docked compounds with better GScores than -9.0 kcal/mol are shown in Figure 4.22 for hits obtained by filtering ZINC with RAD hypothesis, Figure 4.23 for hits obtained by filtering ZINC with RHH hypothesis and Figure 4.24 for hits obtained by filtering ZINC with APRRR hypothesis.

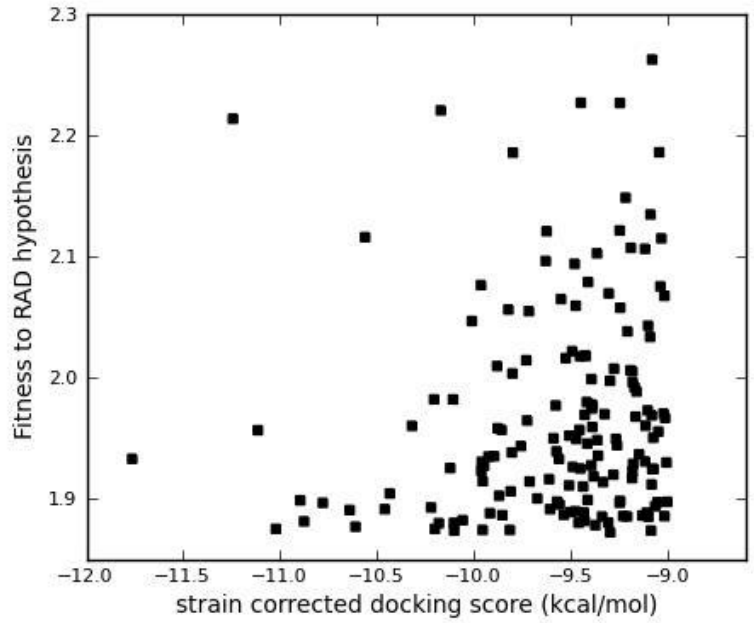



Figure 4.22. Scatter plot of Fitness scores vs. GScores for ZINC compounds filtered with RAD hypothesis and docked to AKT1 (GScore ≤ -9.0 kcal/mol. Total: 158 compounds).

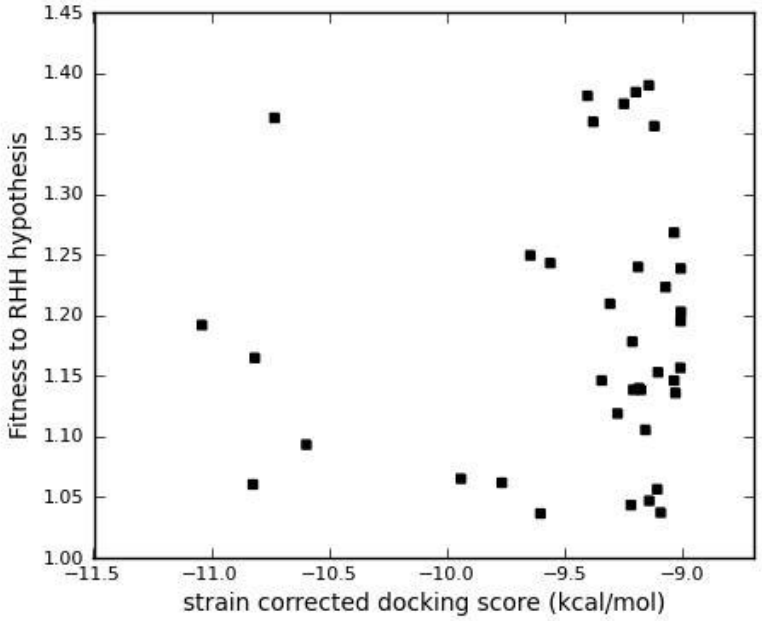



Figure 4.23. Scatter plot of Fitness scores vs. GScores for ZINC compounds filtered with RHH hypothesis and docked to AKT1. (GScore ≤ -9.0 kcal/mol. Total: 38 compounds).

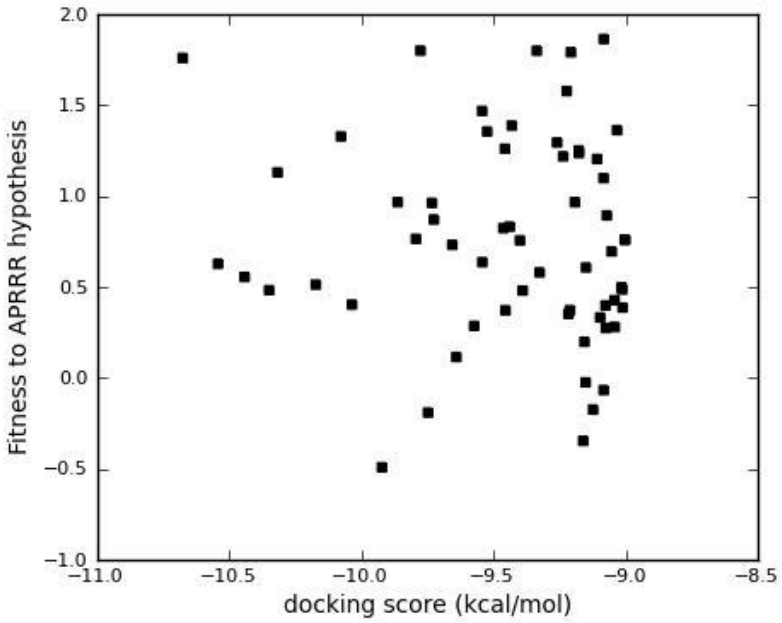



Figure 4.24. Scatter plot of Fitness scores vs. GScores for ZINC compounds filtered with APRRR hypothesis and docked to AKT1. (GScore ≤ -9.0 kcal/mol. Total: 64 compounds).

The elimination of database compounds initiated with structure based and ligand based pharmacophore building and filtering; therefore, for the final compounds, relevancy with the pharmacophore hypothesis was recognized as an important parameter. Accordingly, it was decided to keep the compounds with a fitness score better than 1.8 for structure based pharmacophore models RHH and RAD; which is equivalent to 70% and higher proximity to chosen pharmacophore hypothesis. The threshold for fitness score was decreased to 1.0 for APRRR hypothesis since more site points were included in this hypothesis which makes the hypothesis harder to match. The alignment, vector, volume and fitness scores for compounds in RAD-family2, RAD-family4, RHH-family5 and hits of APRRR hypothesis were given in Table 4.21, Table 4.22, Table 4.23 and Table 4.24, respectively. The eliminated compounds are shown in bold and indicated with a strikethrough.

Table 4.21. Fitness Score and its contributors for compounds in RAD-family2.

| Title | ZINC Code | Align Score | Vector Score | Volume Score | Fitness |
|---|---|---|---|---|---|
| 3871 | ZINC00839381 | 0.35 | 0.95 | 0.32 | 1.98 |
| 3872 | ZINC02373214 | 0.35 | 0.95 | 0.32 | 1.98 |
| 4817 | ZINC04102500 | 0.36 | 0.95 | 0.31 | 1.96 |
| 4544 | ZINC13512705 | 0.34 | 0.92 | 0.33 | 1.96 |
| 2754 | ZINC00625908 | 0.27 | 0.93 | 0.31 | 2.02 |
| 3953 | ZINC00839381 | 0.35 | 0.95 | 0.32 | 1.98 |
| 3641 | ZINC00850903 | 0.34 | 0.95 | 0.32 | 1.99 |
| 4478 | ZINC00839381 | 0.25 | 0.85 | 0.32 | 1.97 |
| 6088 | ZINC00633233 | 0.41 | 0.95 | 0.32 | 1.93 |

Table 4.22. Fitness Score and its contributors for compounds in RAD-family4.

| Title | ZINC Code | Align Score | Vector Score | Volume Score | Fitness |
|---|---|---|---|---|---|
| 161 | ZINC16392469 | 0.15 | 0.90 | 0.45 | 2.22 |
| 1873 | ZINC16392736 | 0.13 | 0.85 | 0.31 | 2.06 |
| 5433 | ZINC13321319 | 0.41 | 0.97 | 0.31 | 1.94 |
| 1160 | ZINC02753923 | 0.07 | 0.78 | 0.37 | 2.10 |
| 877 | ZINC01073862 | 0.28 | 0.99 | 0.36 | 2.12 |
| 5140 | ZINC02867568 | 0.52 | 0.96 | 0.42 | 1.95 |
| 5056 | ZINC08917432 | 0.41 | 0.99 | 0.30 | 1.95 |
| 5160 | ZINC01073706 | 0.32 | 0.92 | 0.30 | 1.95 |
| 136 | ZINC00656127 | 0.12 | 0.98 | 0.35 | 2.23 |
| 1090 | ZINC01073894 | 0.22 | 0.98 | 0.30 | 2.10 |
| 3057 | ZINC13321889 | 0.32 | 0.94 | 0.33 | 2.01 |
| 1026 | ZINC01068312 | 0.19 | 0.93 | 0.34 | 2.11 |

Table 4.23. Fitness Score and its contributors for compounds in RHH-family5. Eliminated molecules are indicated by strikethrough.

| Title | ZINC Code | Align Score | Vector Score | Volume Score | Fitness |
|---|---|---|---|---|---|
| ~~**3322**~~ | ~~**ZINC09397086**~~ | ~~0.86~~ | ~~0.65~~ | ~~0.30~~ | ~~**1.24**~~ |
| ~~**1447**~~ | ~~**ZINC09429089**~~ | ~~0.84~~ | ~~0.75~~ | ~~0.31~~ | ~~**1.36**~~ |
| ~~**7046**~~ | ~~**ZINC20529899**~~ | ~~0.82~~ | ~~0.53~~ | ~~0.27~~ | ~~**1.12**~~ |
| ~~**7526**~~ | ~~**ZINC09397114**~~ | ~~0.80~~ | ~~0.51~~ | ~~0.26~~ | ~~**1.11**~~ |

Table 4.24. Fitness Score and its contributors for compounds in APRRR-1 cluster. Eliminated molecules are indicated by strikethrough.

| Title | Align Score | Vector Score | Volume Score | Fitness |
|---|---|---|---|---|
| ZINC09410511 | 0.98 | 0.73 | 0.42 | 1.33 |
| ZINC09410451 | 0.87 | 0.96 | 0.56 | 1.80 |
| ~~**ZINC11866403**~~ | ~~2.46~~ | ~~0.47~~ | ~~0.24~~ | ~~**0.34**~~ |
| ~~**ZINC13624933**~~ | ~~2.46~~ | ~~0.63~~ | ~~0.24~~ | ~~**0.17**~~ |
| ~~**ZINC19800260**~~ | ~~1.64~~ | ~~0.42~~ | ~~0.30~~ | ~~**0.77**~~ |

All the compounds in RAD-family2 and RAD-family4 showed higher than 73% relevancy to RAD hypothesis. However, all compounds in RHH-family5 and 3 hits of APRRR hypothesis had significantly low fitness scores than the defined threshold values of 1.8 and 1.0, respectively, hence these compounds were decided to be discarded.

### 4.13 Binding Affinities and Interactions of Survival Compounds

In each cluster, a number of compounds were filtered considering the existence of compulsory π-π stacking interaction with Trp 80, ADME properties and fitness to pharmacophore hypothesis, as described in previous sections. The remaining 9 compounds in RAD-family2, 12 compounds in RAD-family4 and 2 hits of APRRR hypothesis were subjected to a final assessment based on the strain energies. The interactions between Akt1 and the surviving compounds were examined in detail.

Glide XP GScore, strain corrected GScore and its main contributors evdW, coul, lipo, Hbond and rotB; which were defined in methods section previously, are listed in Table 4.25 for compounds in RAD-family2, compounds in RAD-family4 and hits of APRRR hypothesis. As mentioned previously, ligands with more than 4 kcal/mol energy difference between the docked and free forms received penalties, and quarter of the strain energy (namely strain penalty) was added to the GScore. Since the binding site is buried between two domains of AKT1, it was expected to observe some strain energy for the docked compounds. For binding sites that are located near receptor surfaces, it is recommended to eliminate the ligands that receive strain penalties; whereas a user defined strain energy threshold can be determined for ligands that bind to buried binding sites. In this manner, it was decided to allow strain energy up to 5kcal/mol for the compounds. Quarter of the energy difference was added to XP GScore, and compounds that gained more than 5 kcal/mol strain energy were eliminated. This was the final elimination for the docked compounds. Eliminated compounds are shown in bold and indicated by strikethrough in Table 4.25.

The interaction maps of the surviving 7 compounds that belong to RAD-family2, 11 compounds that belong to RAD-family4 and 2 hits of APRRR hypothesis were obtained to check the compulsory π-π stacking interaction with Trp 80. The protein-ligand interactions for compounds in RAD-family2, RAD-family4 and APRRR-1 clusters were summarized in Table 4.26, Table 4.27 and Table 4.28.

Table 4.25. GlideScore, strain corrected GlideScore and GlideScore contributors for RAD-family2 and RAD-family4.

| Title | ZINC Code | XP GScore | bound (kcal/mol) | free (kcal/mol) | strain (kcal/mol) | strain penalty | strain corr. Gscore | glide evdw | glide ecoul | glide lipo | glide Hbond | glide rotB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **RAD-FAMILY2** | | | | | | | | | | | | |
| 3871 | ZINC00839381 | -10.21 | 34.32 | 30.87 | 3.46 | 0.00 | -10.21 | -46.39 | -10.28 | -5.06 | -1.98 | 0.36 |
| 3872 | ZINC02373214 | -10.23 | 35.38 | 30.91 | 4.46 | 0.12 | -10.11 | -47.58 | -10.31 | -4.66 | -2.41 | 0.33 |
| ~~**4817**~~ | ~~**ZINC04102500**~~ | ~~-10.55~~ | ~~37.69~~ | ~~31.07~~ | ~~**6.61**~~ | ~~0.65~~ | ~~-9.88~~ | ~~-50.21~~ | ~~-6.14~~ | ~~-5.67~~ | ~~-2.08~~ | ~~0.31~~ |
| 4544 | ZINC13512705 | -9.74 | 48.40 | 44.36 | 4.04 | 0.01 | -9.73 | -56.10 | -4.21 | -6.60 | -0.96 | 0.22 |
| 2754 | ZINC00625908 | -9.45 | 35.84 | 32.67 | 3.17 | 0.00 | -9.43 | -50.62 | -1.62 | -6.21 | -1.18 | 0.26 |
| ~~**3953**~~ | ~~**ZINC00839381**~~ | ~~-9.86~~ | ~~34.90~~ | ~~29.15~~ | ~~**5.75**~~ | ~~0.44~~ | ~~-9.42~~ | ~~-47.30~~ | ~~-8.64~~ | ~~-4.88~~ | ~~-1.97~~ | ~~0.36~~ |
| 3641 | ZINC00850903 | -9.41 | 33.75 | 28.77 | 4.98 | 0.25 | -9.16 | -48.29 | -2.61 | -5.46 | -1.18 | 0.31 |
| 4478 | ZINC00839381 | -9.05 | 32.98 | 28.85 | 4.13 | 0.03 | -9.02 | -47.32 | -4.61 | -5.24 | -1.48 | 0.36 |
| 6088 | ZINC00633233 | -9.01 | 32.02 | 31.46 | 0.56 | 0.00 | -9.01 | -46.69 | -3.08 | -5.10 | -1.18 | 0.33 |
| **RAD-FAMILY4** | | | | | | | | | | | | |
| 161 | ZINC16392469 | -10.17 | 48.82 | 46.83 | 1.99 | 0.000 | -10.17 | -56.20 | -8.05 | -6.82 | -1.19 | 0.15 |
| 1873 | ZINC16392736 | -9.83 | 47.51 | 43.50 | 4.01 | 0.002 | -9.82 | -43.97 | -12.01 | -6.10 | -0.98 | 0.19 |
| 5433 | ZINC13321319 | -9.76 | 39.11 | 36.33 | 2.78 | 0.000 | -9.76 | -53.52 | -7.72 | -6.38 | -1.05 | 0.19 |
| 1160 | ZINC02753923 | -9.63 | 51.76 | 48.61 | 3.15 | 0.000 | -9.63 | -47.96 | -7.69 | -6.11 | -1.33 | 0.15 |
| 877 | ZINC01073862 | -9.63 | 33.42 | 32.69 | 0.72 | 0.000 | -9.63 | -55.45 | -5.10 | -6.61 | -0.98 | 0.18 |
| 5140 | ZINC02867568 | -9.59 | 51.53 | 48.65 | 2.88 | 0.000 | -9.59 | -51.54 | -7.04 | -6.36 | -1.08 | 0.20 |
| 5056 | ZINC08917432 | -9.51 | 54.62 | 52.95 | 1.67 | 0.000 | -9.51 | -55.36 | -5.50 | -6.34 | -1.00 | 0.18 |
| 5160 | ZINC01073706 | -9.85 | 39.00 | 37.24 | 1.76 | 0.000 | -9.47 | -61.53 | -6.52 | -7.01 | -1.10 | 0.15 |
| 136 | ZINC00656127 | -9.45 | 44.24 | 42.69 | 1.55 | 0.000 | -9.45 | -54.87 | -8.03 | -6.39 | -1.33 | 0.17 |
| ~~**1090**~~ | ~~**ZINC01073894**~~ | ~~-9.99~~ | ~~44.56~~ | ~~38.09~~ | ~~**6.47**~~ | ~~0.618~~ | ~~-9.37~~ | ~~-56.11~~ | ~~-6.91~~ | ~~-6.89~~ | ~~-0.82~~ | ~~0.20~~ |
| 3057 | ZINC13321889 | -9.28 | 38.63 | 34.75 | 3.89 | 0.000 | -9.28 | -48.32 | -7.50 | -6.14 | -1.04 | 0.17 |
| 1026 | ZINC01068312 | -9.19 | 41.02 | 38.23 | 2.78 | 0.000 | -9.19 | -56.97 | -5.41 | -6.15 | -0.70 | 0.22 |
| **APRRR-1 Cluster** | | | | | | | | | | | | |
| - | ZINC09410511 | -10.08 | 38.87 | 35.06 | 3.81 | 0.00 | -10.08 | -48.92 | -5.74 | -6.26 | -1.47 | 0.26 |
| - | ZINC09410451 | -9.78 | 35.44 | 33.56 | 1.88 | 0.00 | -9.78 | -52.98 | -6.98 | -7.28 | -1.67 | 0.25 |

Table 4.26. Interactions with AKT1 for 8 final compounds in RAD-family2 cluster.

| | ZINC Code | TRP 80 | SER 205 | ASP 292 | TYR 272 | THR 81 | GLN 79 |
|---|---|---|---|---|---|---|---|
| RAD 2754 | ZINC00625908 | pi | H (b) | | | | |
| RAD 3641 | ZINC00850903 | pi | H (b) | | | | |
| RAD 3871 | ZINC00839381 | pi | H (b) | H (s) | | | |
| RAD 3872 | ZINC02373214 | pi | H (b) | H (s) | | H (b) | |
| RAD 4478 | ZINC00839381 | pi | H (b) | H (s) | | | |
| RAD 4544 | ZINC13512705 | pi | | | pi | | H (b) |
| RAD 6088 | ZINC00633233 | pi | H (b) | | | | |

Table 4.27. Interactions with AKT1 for 11 final compounds in RAD-family4 cluster.

| | ZINC Code | TRP 80 | LYS 268 | ASP 292 | TYR 272 | GLN 79 | Water 455 | THR 82 |
|---|---|---|---|---|---|---|---|---|
| RAD 136 | ZINC00656127 | pi | | | | H (s) | H (s) | |
| RAD 161 | ZINC16392469 | pi | H (s) | | | H (s) | H (s) | |
| RAD 877 | ZINC01073862 | pi | | | | H (s) | H (s) | |
| RAD 1026 | ZINC01068312 | pi | | | H (b) | | | |
| RAD 1160 | ZINC02753923 | pi, pi | H (s) | | pi | H (s) | H (s) | |
| RAD 1873 | ZINC16392736 | pi | | H (s) | | | | H (s) |
| RAD 3057 | ZINC13321889 | pi | | | pi | H (s) | H (s) | |
| RAD 5056 | ZINC08917432 | pi | | | | H (s) | H (s) | |
| RAD 5140 | ZINC02867568 | pi | | | | H (s) | H (s) | |
| RAD 5160 | ZINC01073706 | pi | | | | H (s) | H (s) | |
| RAD 5433 | ZINC13321319 | pi | | | pi | H (s) | H (s) | |

Table 4.28. Interactions with AKT1 for 2 final compounds in APRRR-1 cluster.

| | Trp 80 | Tyr 272 | Thr 211 | Ala 58 | Leu 78 | Asn 53 | Ser 205 |
|---|---|---|---|---|---|---|---|
| ZINC09410451 | pi | pi | H (b) | H (b) | H (b) | | |
| ZINC09410511 | pi, pi | pi | H (b) | | | H (b) | H (b) |

ZINC00625908, ZINC00850903, ZINC00839381, ZINC02373214, ZINC00839381 and ZINC00633233 in RAD-2 cluster made π-π stacking interaction with Trp 80 through their imidazole ring of xanthine sub-structure. The hydrogen bond

interactions for the same six compounds were observed between Ser 205 and primary nitrogen (NH) of xanthine. ZINC13512705 and ZINC00626346 showed π-π stacking interaction with Tyr 272 through their imidazole ring of xanthine sub-structure and interacted with Gln 79 through a hydrogen bond. Hydrogen bond interaction was detected between side chain of Asp 292 and ZINC00839381, ZINC02373214, ZINC00839381. The key interactions observed mostly within RAD-2 cluster were determined as π-π stacking interaction with Trp 80 and hydrogen bonding interaction with Ser 205.

All members of RAD-family4 cluster made π-π stacking interaction with Trp 80 through their quinoline sub-structure. Except for ZINC01068312 and ZINC16392736, all compounds also showed hydrogen bonding interaction with side chain of Gln 79 and conserved water 455. ZINC02753923, ZINC13321889, ZINC13321319 made π-π stacking interaction and ZINC01068312 made hydrogen bond with Tyr 272. ZINC16392469 and ZINC02753923 showed hydrogen bonding interaction with side chain of Lys 268. The key interactions observed mostly within RAD-4 cluster were determined as π-π stacking interaction with Trp 80 and hydrogen bonding interaction with Ser 205.

Both two hits of APRRR hypothesis showed π-π stacking interaction with both Trp 80 and Tyr 272. ZINC09410451 made hydrogen bonds with backbone of Ala 58, backbone of Leu 78 and backbone of Thr 211. ZINC09410511 made three hydrogen bonds with backbones of Asn 53, Ser 205 and Thr 211. A generalization was not made for key interactions in this cluster due to small number of compounds included.

### 4.14 Docking to Kinase Domain of AKT1

The compounds that were docked to grid generated AKT1 receptor, scored with GlideScore and gained GScores -9.0 kcal/mol and better, namely 38 compounds from hits of RHH hypothesis and 158 compounds from hits of RAD hypothesis were re-docked to a PH-domain excluded AKT1 structure. The aim was getting an idea on binding modes and affinities of the pre-docked ligands in the absence of PH domain of AKT1.

An AKT1 structure that contains only kinase domain was selected as the receptor (PDB code: 3QKL). To define the binding site, kinase domain of AKT1 (144-480, PDB code: 3QKL) was aligned to full length AKT1 (2-429, pdb code: 3O96). For grid generation, the outer (20Å) and inner (10Å) boxes were selected in same sizes and coordinates, with their center as the centroid of the native ligand (Inhibitor VIII, pdb code: IQO) of the complex (PDB code: 3O96). Totally 196 compounds were docked to kinase domain without applying any constraints in Glide XP mode. Figure 4.25 shows the scatter plot of XP GScore values for 196 compounds of RHH hits and RAD hits docked to allosteric binding part of kinase domain. The GScore values indicated that, the compounds that show very good binding affinity (with GScore values between -11.0 kcal/mol and -9.0 kcal/mol) when docked to AKT1 in the presence of PH domain, exhibited poor binding affinity (with GScore values between -5.5 kcal/mol and 1.9 kcal/mol) in the absence of PH domain.

In addition to poor binding affinity scores, the binding modes and binding site interactions with AKT1 of 196 ligands changed. The interaction diagrams of arbitrarily selected 10 compounds from RAD-family2 and RAD-family4 were obtained to check the protein-ligand interactions. It was seen that the interactions with the AKT1 kinase domain residues in the presence of PH domain (i.e. Ser205, Lys268, Tyr 272, Asp 292) could not be observed for any of the ligands which suggests that the binding modes of these compounds were completely different. The interactions are listed in Table 4.29.

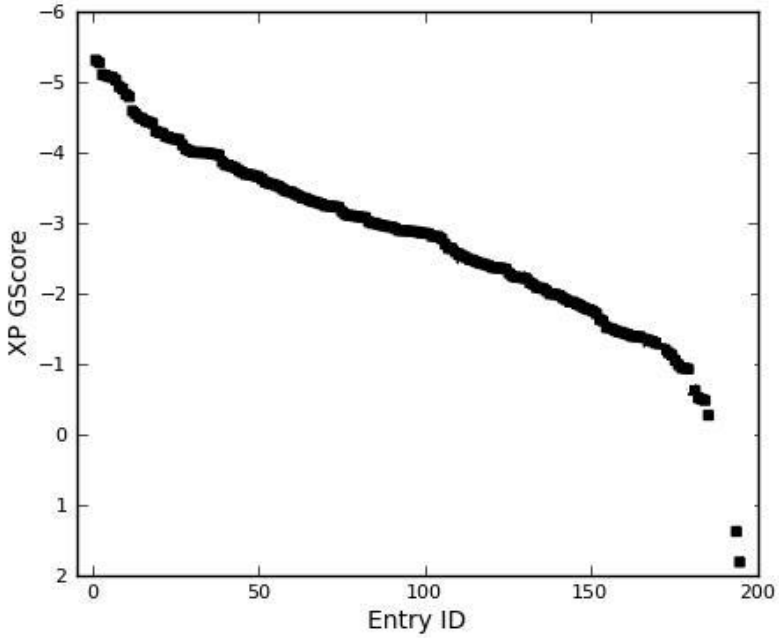



Figure 4.25. Scatter plot of XP GScore values for hits of RHH and RAD hypothesis (Total: 196 compounds) docked to allosteric binding part of kinase domain (PDB code: 3QKL).

Table 4.29. Interactions with AKT1 for 10 selected compounds from RAD-family2 and RAD-family4 clusters in the absence of PH domain.

| | TYR 229 | LYS 284 | GLU 432 | ARG 174 | LYS 214 | VAL 201 | GLU 298 | TYR 175 | LYS 289 | LYS 419 | GLU 228 | PHE 293 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RAD 136 | H (b) | pi cat | | | | | | | | | | |
| RAD 1466 | H (b) | | H (s) | pi cat | pi cat | | | | | | | |
| RAD 2754 | | | | | | H (b) | | | | | | |
| RAD 3641 | | | | | | | H (s) | | | | | |
| RAD 3872 | | pi cat | H (s) | | | | | H (b) | H (s) | H (s) | H (s) | |
| RAD 4544 | H (b) | | H (s) | | | | | H (b) | | | | |
| RAD 161 | | | | | | | | | | | | H (b) |
| RAD 877 | H (b) | pi cat | | | | | | | | | H (s) | |
| RAD 1026 | | | | | | | | | | | | pi |
| RAD 1160 | | pi cat | | | | | | | | | | |

The binding modes of compounds with titles 3872 from RAD-family2 and 161 from RAD-family4 were superimposed in the presence and absence of PH domain to see

the difference in-between (Figure 4.26). These compounds were selected as representatives since they exhibited the highest strain-corrected GScore values (3872: -10.23 kcal/mol, 161: -10.17 kcal/mol) in the presence of PH domain in addition to high fitness score to RAD hypothesis (3872: 1.98, 161: 2.22).

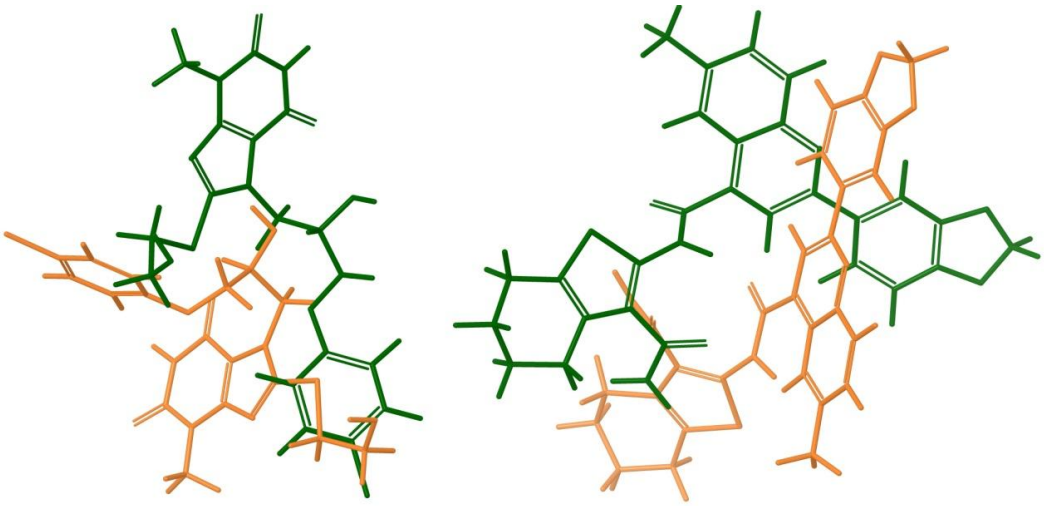


Figure 4.26. Binding modes of compounds 3872 (left) and 161 (right) in allosteric binding site of AKT1; in the absence of PH domain (orange) and in the presence of PH domain (green).

# 5. CONCLUSIONS AND RECOMMENDATIONS

## 5.1 Conclusion

In this thesis study, pharmacophore modeling and docking approaches were combined to identify a diverse set of scaffolds as potential allosteric inhibitors for Akt1. Both ligand-based and structure-based pharmacophore modeling were employed. The ligand-based pharmacophore model APRRR was based on previously identified allosteric inhibitors for Akt1 and structure-based pharmacophore models RHH and RAD used the information of allosteric binding site which was located between PH and kinase domains of Akt1. ZINC database was filtered separately based on 3D similarity to three pharmacophore hypotheses. Standard precision followed by extra precision docking were performed using Glide protocol to predict binding modes and affinities of compounds in the pre-filtered database. Docked compounds with GlideScores higher than -9.0 kcal/mol were subjected to post-docking evaluations. Hiererchical clustering was employed to determine two dimensional similarities between top scoring hits. Five clusters were identified for a total of 196 hits of structure-based pharmacophore hypotheses RHH and RAD. Only 18 compounds survived in the ligand-based pharmacophore filtering based on the APRRR hypothesis. The clusters with high population (from structure based pharmacophore filtering) or the full set of 18 compounds (from ligand based pharmacophore filtering) were further analyzed considering the existence of compulsory π-π stacking interaction with Trp 80 of Akt1, ADME and drug-likeness criteria and fitness to corresponding pharmacophore hypothesis. After postdocking processing, the compounds in the RAD-2 and RAD-4 clusters and the APRRR hypothesis remained. The key protein-ligand interactions were identified for these compounds. The common scaffolds for the surviving hits revealed three scaffolds as potential allosteric inhibitors for Akt1. Derivatives of; 3-methyl-xanthine, quinoline-4-carboxamide and 2-[4-(cyclohexa-1,3-dien-1-yl)-1H-pyrazol-3-yl]phenol were proposed for further studies as lead structures. Derivatives of methyl-xanthine were previously reported as human pancreatic lipase inhibitors [103], apnea reducers in preterm infants [103], therapeutics for severe exacerbations of chronic obstructive pulmonary disease (COPD) [104] and unique class of

drugs for the treatment of asthma [105]. Derivatives of quinoline-carboxamide moieties were reported to enhance affinity of neurokinin 1 receptor antagonists, which are used as preventors for cancer chemotherapy side-effects such as vomiting [106]. Additionally, quinoline-4-carboxamide derivatives were reported to exhibit antimicrobial effects and influenced the growth of the microorganisms [107]. As a result, through the integration of structure and ligand based pharmacophore modeling and docking against a novel allosteric site, this study identified putative allosteric inhibitors for Akt.

## 5.2 Recommendations for Future Studies

In both ligand and structure based screening, additional pharmacophore models can be developed considering different features of the ligands and protein structure to obtain a different set of inhibitors which may also show good binding affinities. Inclusion of different features in the model can enhance diversity in the results and allow comparison of a larger number of hits. Furthermore, inhibiton of Akt by the proposed molecules can be verified by experimental testing in vitro by biological assays.

For docking calculations, Glide software was used. Glide keeps the receptor rigid while allows ligand flexibility during docking. To consider receptor flexibility, induced-fit docking calculations can be performed, for a selected percentage of hits with the highest docking score.

In addition to virtual screening, Molecular Dynamics (MD) simulations can also be done to improve analysis of hits. For both apo state and ligand-bound states of Akt, performing MD simulations would give an opinion about the stability of the ligand-protein complex.

# REFERENCES


1. Arencibia, J. M., D. Pastor-Flores, A. F. Bauer, J. O. Schulze, and R. M. Biondi, "Agc Protein Kinases: From Structural Mechanism of Regulation to Allosteric Drug Development for the Treatment of Human Diseases", *Biochim Biophys Acta*, Vol. 1834 pp. 1302-1321, 2013.

2. Calleja, V., M. Laguerre, and B. Larijani, "3-D Structure and Dynamics of Protein Kinase B-New Mechanism for the Allosteric Regulation of an Agc Kinase", *J Chem Biol*, Vol. 2, No. 1, pp. 11-25, 2009.

3. Hennessy, B. T., D. L. Smith, P. T. Ram, Y. Lu, and G. B. Mills, "Exploiting the Pi3k/Akt Pathway for Cancer Drug Discovery", *Nat Rev Drug Discov*, Vol. 4, No. 12, pp. 988-1004, 2005.

4. Lindsley, C. W., Z. Zhao, W. H. Leister, R. G. Robinson, S. F. Barnett, D. Defeo-Jones, R. E. Jones, G. D. Hartman, J. R. Huff, H. E. Huber, and M. E. Duggan, "Allosteric Akt (Pkb) Inhibitors: Discovery and Sar of Isozyme Selective Inhibitors", *Bioorg Med Chem Lett*, Vol. 15, No. 3, pp. 761-4, 2005.

5. Green, C. J., O. Goransson, G. S. Kular, N. R. Leslie, A. Gray, D. R. Alessi, K. Sakamoto, and H. S. Hundal, "Use of Akt Inhibitor and a Drug-Resistant Mutant Validates a Critical Role for Protein Kinase B/Akt in the Insulin-Dependent Regulation of Glucose and System a Amino Acid Uptake", *J Biol Chem*, Vol. 283, No. 41, pp. 27653-67, 2008.

6. Garofalo, R. S., S. J. Orena, K. Rafidi, A. J. Torchia, J. L. Stock, A. L. Hildebrandt, T. Coskran, S. C. Black, D. J. Brees, J. R. Wicks, J. D. McNeish, and K. G. Coleman, "Severe Diabetes, Age-Dependent Loss of Adipose Tissue, and Mild Growth Deficiency in Mice Lacking Akt2/Pkb Beta", *J Clin Invest*, Vol. 112, No. 2, pp. 197-208, 2003.

7. Hers, I., E. E. Vincent, and J. M. Tavare, "Akt Signalling in Health and Disease", *Cell Signal*, Vol. 23, No. 10, pp. 1515-27, 2011.

8. Kumar, C. C. and V. Madison, "Drugs Targeted against Protein Kinases", *Expert Opin Emerg Drugs*, Vol. 6, No. 2, pp. 303-15, 2001.

9. Calleja, V., M. Laguerre, P. J. Parker, and B. Larijani, "Role of a Novel Ph-Kinase Domain Interface in Pkb/Akt Regulation: Structural Mechanism for Allosteric Inhibition", *PLoS Biol*, Vol. 7, No. 1, pp. e17, 2009.

10. Kumar, C. C. and V. Madison, "Akt Crystal Structure and Akt-Specific Inhibitors", *Oncogene*, Vol. 24, No. 50, pp. 7493-501, 2005.

11. Lindsley, C. W., "The Akt/Pkb Family of Protein Kinases: A Review of Small Molecule Inhibitors and Progress Towards Target Validation: A 2009 Update", *Curr Top Med Chem*, Vol. 10, No. 4, pp. 458-77, 2010.

12. Blake, J. F., R. Xu, J. R. Bencsik, D. Xiao, N. C. Kallan, S. Schlachter, I. S. Mitchell, K. L. Spencer, A. L. Banka, E. M. Wallace, S. L. Gloor, M. Martinson, R. D. Woessner, G. P. Vigers, B. J. Brandhuber, J. Liang, B. S. Safina, J. Li, B. Zhang, C. Chabot, S. Do, L. Lee, J. Oeh, D. Sampath, B. B. Lee, K. Lin, B. M. Liederer, and N. J. Skelton, "Discovery and Preclinical Pharmacology of a Selective Atp-Competitive Akt Inhibitor (Gdc-0068) for the Treatment of Human Tumors", *J Med Chem*, Vol. 55, No. 18, pp. 8110-27, 2012.

13. Freeman-Cook, K. D., C. Autry, G. Borzillo, D. Gordon, E. Barbacci-Tobin, V. Bernardo, D. Briere, T. Clark, M. Corbett, J. Jakubczak, S. Kakar, E. Knauth, B. Lippa, M. J. Luzzio, M. Mansour, G. Martinelli, M. Marx, K. Nelson, J. Pandit, F. Rajamohan, S. Robinson, C. Subramanyam, L. Wei, M. Wythes, and J. Morris, "Design of Selective, Atp-Competitive Inhibitors of Akt", *J Med Chem*, Vol. 53, No. 12, pp. 4615-22, 2010.

14. Wang, P., L. Zhang, Q. Hao, and G. Zhao, "Developments in Selective Small Molecule Atp-Targeting the Serine/Threonine Kinase Akt/Pkb", *Mini Rev Med Chem*, Vol. 11, No. 13, pp. 1093-107, 2011.

15. Moses, S. A., M. A. Ali, S. Zuohe, L. Du-Cuny, L. L. Zhou, R. Lemos, N. Ihle, A. G. Skillman, S. Zhang, E. A. Mash, G. Powis, and E. J. Meuillet, "In Vitro and in Vivo Activity of Novel Small-Molecule Inhibitors Targeting the Pleckstrin Homology Domain of Protein Kinase B/Akt", *Cancer Res*, Vol. 69, No. 12, pp. 5073-81, 2009.

16. Hartnett, J. C., S. F. Barnett, M. T. Bilodeau, D. Defeo-Jones, G. D. Hartman, H. E. Huber, R. E. Jones, A. M. Kral, R. G. Robinson, and Z. Wu, "Optimization of 2,3,5-Trisubstituted Pyridine Derivatives as Potent Allosteric Akt1 and Akt2 Inhibitors", *Bioorg Med Chem Lett*, Vol. 18, No. 6, pp. 2194-7, 2008.

17. Wu, W. I., W. C. Voegtli, H. L. Sturgis, F. P. Dizon, G. P. Vigers, and B. J. Brandhuber, "Crystal Structure of Human Akt1 with an Allosteric Inhibitor Reveals a New Mode of Kinase Inhibition", *PLoS One*, Vol. 5, No. 9, pp. e12913, 2010.

18. Garuti, L., M. Roberti, and G. Bottegoni, "Non-Atp Competitive Protein Kinase Inhibitors", *Curr Med Chem*, Vol. 17, No. 25, pp. 2804-21, 2010.

19. Rhodes, N., D. A. Heerding, D. R. Duckett, D. J. Eberwein, V. B. Knick, T. J. Lansing, R. T. McConnell, T. M. Gilmer, S. Y. Zhang, K. Robell, J. A. Kahana, R. S. Geske, E. V. Kleymenova, A. E. Choudhry, Z. Lai, J. D. Leber, E. A. Minthorn, S. L. Strum, E. R. Wood, P. S. Huang, R. A. Copeland, and R. Kumar, "Characterization of an Akt Kinase Inhibitor with Potent Pharmacodynamic and Antitumor Activity", *Cancer Res*, Vol. 68, No. 7, pp. 2366-74, 2008.

20. Luo, Y., A. R. Shoemaker, X. Liu, K. W. Woods, S. A. Thomas, R. de Jong, E. K. Han, T. Li, V. S. Stoll, J. A. Powlas, A. Oleksijew, M. J. Mitten, Y. Shi, R. Guan, T. P. McGonigal, V. Klinghofer, E. F. Johnson, J. D. Leverson, J. J. Bouska, M. Mamo, R. A. Smith, E. E. Gramling-Evans, B. A. Zinker, A. K. Mika, P. T. Nguyen, T. Oltersdorf, S. H. Rosenberg, Q. Li, and V. L. Giranda, "Potent and Selective Inhibitors of Akt Kinases Slow the Progress of Tumors in Vivo", *Mol Cancer Ther*, Vol. 4, No. 6, pp. 977-86, 2005.

21. Barnett, S. F., D. Defeo-Jones, S. Fu, P. J. Hancock, K. M. Haskell, R. E. Jones, J. A. Kahana, A. M. Kral, K. Leander, L. L. Lee, J. Malinowski, E. M. McAvoy, D. D. Nahas, R. G. Robinson, and H. E. Huber, "Identification and Characterization of Pleckstrin-Homology-Domain-Dependent and Isoenzyme-Specific Akt Inhibitors", *Biochem J*, Vol. 385, No. Pt 2, pp. 399-408, 2005.

22. Cherrin, C., K. Haskell, B. Howell, R. Jones, K. Leander, R. Robinson, A. Watkins, M. Bilodeau, J. Hoffman, P. Sanderson, G. Hartman, E. Mahan, T. Prueksaritanont, G. Jiang, Q. B. She, N. Rosen, L. Sepp-Lorenzino, D. Defeo-Jones, and H. E. Huber, "An Allosteric Akt Inhibitor Effectively Blocks Akt Signaling and Tumor Growth with Only Transient Effects on Glucose and Insulin Levels in Vivo", *Cancer Biol Ther*, Vol. 9, No. 7, pp. 493-503, 2010.

23. Yun, J., "Allosteric Akt Inhibitors as a Targeted Cancer Therapy", *Cancer Biol Ther*, Vol. 9, No. 7, pp. 504-6, 2010.

24. Manning, G., D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam, "The Protein Kinase Complement of the Human Genome", *Science*, Vol. 298, No. 5600, pp. 1912-34, 2002.

25. Workman, P., "Drugging the Cancer Kinome: Progress and Challenges in Developing Personalized Molecular Cancer Therapeutics", *Cold Spring Harb Symp Quant Biol*, Vol. 70, No. pp. 499-515, 2005.

26. Manning, G., G. D. Plowman, T. Hunter, and S. Sudarsanam, "Evolution of Protein Kinase Signaling from Yeast to Man", *Trends Biochem Sci*, Vol. 27, No. 10, pp. 514-20, 2002.

27. Bayascas, J. R., "Pdk1: The Major Transducer of Pi 3-Kinase Actions", *Curr Top Microbiol Immunol*, Vol. 346, No. pp. 9-29, 2010.

28. Bruhn, M. A., R. B. Pearson, R. D. Hannan, and K. E. Sheppard, "Second Akt: The Rise of Sgk in Cancer Signalling", *Growth Factors*, Vol. 28, No. 6, pp. 394-408, 2010.

29. Lang, F., A. Gorlach, and V. Vallon, "Targeting Sgk1 in Diabetes", *Expert Opin Ther Targets*, Vol. 13, No. 11, pp. 1303-11, 2009.

30. Ackermann, T. F., K. M. Boini, N. Beier, W. Scholz, T. Fuchss, and F. Lang, "Emd638683, a Novel Sgk Inhibitor with Antihypertensive Potency", *Cell Physiol Biochem*, Vol. 28, No. 1, pp. 137-46, 2011.

31. Romeo, Y., X. Zhang, and P. P. Roux, "Regulation and Function of the Rsk Family of Protein Kinases", *Biochem J*, Vol. 441, No. 2, pp. 553-69, 2012.

32. Romeo, Y. and P. P. Roux, "Paving the Way for Targeting Rsk in Cancer", *Expert Opin Ther Targets*, Vol. 15, No. 1, pp. 5-9, 2011.

33. Pearce, L. R., D. Komander, and D. R. Alessi, "The Nuts and Bolts of Agc Protein Kinases", *Nat Rev Mol Cell Biol*, Vol. 11, No. 1, pp. 9-22, 2010.

34. Ananieva, O., J. Darragh, C. Johansen, J. M. Carr, J. McIlrath, J. M. Park, A. Wingate, C. E. Monk, R. Toth, S. G. Santos, L. Iversen, and J. S. Arthur, "The Kinases Msk1 and Msk2 Act as Negative Regulators of Toll-Like Receptor Signaling", *Nat Immunol*, Vol. 9, No. 9, pp. 1028-36, 2008.

35. Carnevalli, L. S., K. Masuda, F. Frigerio, O. Le Bacquer, S. H. Um, V. Gandin, I. Topisirovic, N. Sonenberg, G. Thomas, and S. C. Kozma, "S6k1 Plays a Critical Role in Early Adipocyte Differentiation", *Dev Cell*, Vol. 18, No. 5, pp. 763-74, 2010.

36. Um, S. H., F. Frigerio, M. Watanabe, F. Picard, M. Joaquin, M. Sticker, S. Fumagalli, P. R. Allegrini, S. C. Kozma, J. Auwerx, and G. Thomas, "Absence of S6k1 Protects against Age- and Diet-Induced Obesity While Enhancing Insulin Sensitivity", *Nature*, Vol. 431, No. 7005, pp. 200-5, 2004.

37. Kandel, E. R., "The Molecular Biology of Memory: Camp, Pka, Cre, Creb-1, Creb-2, and Cpeb", *Mol Brain*, Vol. 5, pp. 14, 2012.

38. de Leeuw, R., K. Flach, C. Bentin Toaldo, X. Alexi, S. Canisius, J. Neefjes, R. Michalides, and W. Zwart, "Pka Phosphorylation Redirects Eralpha to Promoters of a Unique Gene Set to Induce Tamoxifen Resistance", *Oncogene*, Vol. 361, 2012.

39. Cornils, H., R. S. Kohler, A. Hergovich, and B. A. Hemmings, "Downstream of Human Ndr Kinases: Impacting on C-Myc and P21 Protein Stability to Control Cell Cycle Progression", *Cell Cycle*, Vol. 10, No. 12, pp. 1897-904, 2011.

40. Pan, D., "The Hippo Signaling Pathway in Development and Cancer", *Dev Cell*, Vol. 19, No. 4, pp. 491-505, 2010.

41. Robinson, D. R., S. Kalyana-Sundaram, Y. M. Wu, S. Shankar, X. Cao, B. Ateeq, I. A. Asangani, M. Iyer, C. A. Maher, C. S. Grasso, R. J. Lonigro, M. Quist, J. Siddiqui, R. Mehra, X. Jing, T. J. Giordano, M. S. Sabel, C. G. Kleer, N. Palanisamy, R. Natrajan, M. B. Lambros, J. S. Reis-Filho, C. Kumar-Sinha, and A. M. Chinnaiyan, "Functionally Recurrent Rearrangements of the Mast Kinase and Notch Gene Families in Breast Cancer", *Nat Med*, Vol. 17, No. 12, pp. 1646-51, 2011.

42. Harmon, E. B., M. L. Harmon, T. D. Larsen, J. Yang, J. W. Glasford, and M. B. Perryman, "Myotonic Dystrophy Protein Kinase Is Critical for Nuclear Envelope Integrity", *J Biol Chem*, Vol. 286, No. 46, pp. 40296-306, 2011.

43. Olson, M. F., "Applications for Rock Kinase Inhibition", *Curr Opin Cell Biol*, Vol. 20, No. 2, pp. 242-8, 2008.

44. Loomis, R. J., D. A. Holmes, A. Elms, P. A. Solski, C. J. Der, and L. Su, "Citron Kinase, a Rhoa Effector, Enhances Hiv-1 Virion Production by Modulating Exocytosis", *Traffic*, Vol. 7, No. 12, pp. 1643-53, 2006.

45. Gurevich, E. V., J. J. Tesmer, A. Mushegian, and V. V. Gurevich, "G Protein-Coupled Receptor Kinases: More Than Just Kinases and Not Only for Gpcrs", *Pharmacol Ther*, Vol. 133, No. 1, pp. 40-69, 2012.

46. Ruan, G. X. and A. Kazlauskas, "Focus on Molecules: Akt (Pkb)", *Exp Eye Res*, Vol. 93, No. 5, pp. 570-1, 2011.

47. Staal, S. P., "Molecular Cloning of the Akt Oncogene and Its Human Homologues Akt1 and Akt2: Amplification of Akt1 in a Primary Human Gastric Adenocarcinoma", *Proc Natl Acad Sci U S A*, Vol. 84, No. 14, pp. 5034-7, 1987.

48. Mahadevan, D., G. Powis, E. A. Mash, B. George, V. M. Gokhale, S. Zhang, K. Shakalya, L. Du-Cuny, M. Berggren, M. A. Ali, U. Jana, N. Ihle, S. Moses, C. Franklin, S. Narayan, N. Shirahatti, and E. J. Meuillet, "Discovery of a Novel Class of Akt Pleckstrin Homology Domain Inhibitors", *Mol Cancer Ther*, Vol. 7, No. 9, pp. 2621-32, 2008.

49. Medina-Franco, J. L., M. A. Giulianotti, Y. Yu, L. Shen, L. Yao, and N. Singh, "Discovery of a Novel Protein Kinase B Inhibitor by Structure-Based Virtual Screening", *Bioorg Med Chem Lett*, Vol. 19, No. 16, pp. 4634-8, 2009.

50. Lindsley, C. W., S. F. Barnett, M. E. Layton, and M. T. Bilodeau, "The Pi3k/Akt Pathway: Recent Progress in the Development of Atp-Competitive and Allosteric Akt Kinase Inhibitors", *Curr Cancer Drug Targets*, Vol. 8, No. 1, pp. 7-18, 2008.

51. LoPiccolo, J., C. A. Granville, J. J. Gills, and P. A. Dennis, "Targeting Akt in Cancer Therapy", *Anticancer Drugs*, Vol. 18, No. 8, pp. 861-74, 2007.

52. Hernandez-Campos, A., I. Velazquez-Martinez, R. Castillo, F. Lopez-Vallejo, P. Jia, Y. Yu, M. A. Giulianotti, and J. L. Medina-Franco, "Docking of Protein Kinase B Inhibitors: Implications in the Structure-Based Optimization of a Novel Scaffold", *Chem Biol Drug Des*, Vol. 76, No. 3, pp. 269-76, 2010.

53. Liu, P., H. Cheng, T. M. Roberts, and J. J. Zhao, "Targeting the Phosphoinositide 3-Kinase Pathway in Cancer", *Nat Rev Drug Discov*, Vol. 8, No. 8, pp. 627-44, 2009.

54. Stemke-Hale, K., A. M. Gonzalez-Angulo, A. Lluch, R. M. Neve, W. L. Kuo, M. Davies, M. Carey, Z. Hu, Y. Guan, A. Sahin, W. F. Symmans, L. Pusztai, L. K. Nolden, H. Horlings, K. Berns, M. C. Hung, M. J. van de Vijver, V. Valero, J. W. Gray, R. Bernards, G. B. Mills, and B. T. Hennessy, "An Integrative Genomic and Proteomic Analysis of Pik3ca, Pten, and Akt Mutations in Breast Cancer", *Cancer Res*, Vol. 68, No. 15, pp. 6084-91, 2008.

55. Zhang, J., P. L. Yang, and N. S. Gray, "Targeting Cancer with Small Molecule Kinase Inhibitors", *Nat Rev Cancer*, Vol. 9, No. 1, pp. 28-39, 2009.

56. Carpten, J. D., A. L. Faber, C. Horn, G. P. Donoho, S. L. Briggs, C. M. Robbins, G. Hostetter, S. Boguslawski, T. Y. Moses, S. Savage, M. Uhlik, A. Lin, J. Du, Y. W. Qian, D. J. Zeckner, G. Tucker-Kellogg, J. Touchman, K. Patel, S. Mousses, M. Bittner, R. Schevitz, M. H. Lai, K. L. Blanchard, and J. E. Thomas, "A Transforming Mutation in the Pleckstrin Homology Domain of Akt1 in Cancer", *Nature*, Vol. 448, No. 7152, pp. 439-44, 2007.

57. Kim, M. S., E. G. Jeong, N. J. Yoo, and S. H. Lee, "Mutational Analysis of Oncogenic Akt E17k Mutation in Common Solid Cancers and Acute Leukaemias", *Br J Cancer*, Vol. 98, No. 9, pp. 1533-5, 2008.

58. Malanga, D., M. Scrima, C. De Marco, F. Fabiani, N. De Rosa, S. De Gisi, N. Malara, R. Savino, G. Rocco, G. Chiappetta, R. Franco, V. Tirino, G. Pirozzi, and G. Viglietto, "Activating E17k Mutation in the Gene Encoding the Protein Kinase Akt1 in a Subset of Squamous Cell Carcinoma of the Lung", *Cell Cycle*, Vol. 7, No. 5, pp. 665-9, 2008.

59. Fresno Vara, J. A., E. Casado, J. de Castro, P. Cejas, C. Belda-Iniesta, and M. Gonzalez-Baron, "Pi3k/Akt Signalling Pathway and Cancer", *Cancer Treat Rev*, Vol. 30, No. 2, pp. 193-204, 2004.

60. Katso, R., K. Okkenhaug, K. Ahmadi, S. White, J. Timms, and M. D. Waterfield, "Cellular Function of Phosphoinositide 3-Kinases: Implications for Development, Homeostasis, and Cancer", *Annu Rev Cell Dev Biol*, Vol. 17, No. pp. 615-75, 2001.

61. Okkenhaug, K., "Signaling by the Phosphoinositide 3-Kinase Family in Immune Cells", *Annu Rev Immunol*, Vol. 31, No. pp. 675-704, 2013.

62. He, J., M. Vora, R. M. Haney, G. S. Filonov, C. A. Musselman, C. G. Burd, A. G. Kutateladze, V. V. Verkhusha, R. V. Stahelin, and T. G. Kutateladze, "Membrane Insertion of the Fyve Domain Is Modulated by Ph", *Proteins*, Vol. 76, No. 4, pp. 852-60, 2009.

63. Mahajan, K., D. Coppola, S. Challa, B. Fang, Y. A. Chen, W. Zhu, A. S. Lopez, J. Koomen, R. W. Engelman, C. Rivera, R. S. Muraoka-Cook, J. Q. Cheng, E. Schonbrunn, S. M. Sebti, H. S. Earp, and N. P. Mahajan, "Ack1 Mediated Akt/Pkb Tyrosine 176 Phosphorylation Regulates Its Activation", *PLoS One*, Vol. 5, No. 3, pp. e9646, 2010.

64. Thomas, C. C., M. Deak, D. R. Alessi, and D. M. van Aalten, "High-Resolution Structure of the Pleckstrin Homology Domain of Protein Kinase B/Akt Bound to Phosphatidylinositol (3,4,5)-Trisphosphate", *Curr Biol*, Vol. 12, No. 14, pp. 1256-62, 2002.

65. Bogoyevitch, M. A. and D. P. Fairlie, "A New Paradigm for Protein Kinase Inhibition: Blocking Phosphorylation without Directly Targeting Atp Binding", *Drug Discov Today*, Vol. 12, No. 15-16, pp. 622-33, 2007.

66. Yang, J., P. Cron, V. Thompson, V. M. Good, D. Hess, B. A. Hemmings, and D. Barford, "Molecular Mechanism for the Regulation of Protein Kinase B/Akt by Hydrophobic Motif Phosphorylation", *Mol Cell*, Vol. 9, No. 6, pp. 1227-40, 2002.

67. Du-Cuny, L., Z. Song, S. Moses, G. Powis, E. A. Mash, E. J. Meuillet, and S. Zhang, "Computational Modeling of Novel Inhibitors Targeting the Akt Pleckstrin Homology Domain", *Bioorg Med Chem*, Vol. 17, No. 19, pp. 6983-92, 2009.

68. Wu, Z., J. C. Hartnett, L. A. Neilson, R. G. Robinson, S. Fu, S. F. Barnett, D. Defeo-Jones, R. E. Jones, A. M. Kral, H. E. Huber, G. D. Hartman, and M. T. Bilodeau, "Development of Pyridopyrimidines as Potent Akt1/2 Inhibitors", *Bioorg Med Chem Lett*, Vol. 18, No. 4, pp. 1274-9, 2008.

69. Ashwell, M. A., J. M. Lapierre, C. Brassard, K. Bresciano, C. Bull, S. Cornell-Kennon, S. Eathiraj, D. S. France, T. Hall, J. Hill, E. Kelleher, S. Khanapurkar, D. Kizer, S. Koerner, J. Link, Y. Liu, S. Makhija, M. Moussa, N. Namdev, K. Nguyen, R. Nicewonger, R. Palma, J. Szwaya, M. Tandon, U. Uppalapati, D. Vensel, L. P. Volak, E. Volckova, N. Westlund, H. Wu, R. Y. Yang, and T. C. Chan, "Discovery and Optimization of a Series of 3-(3-Phenyl-3h-Imidazo[4,5-B]Pyridin-2-Yl)Pyridin-2-Amines: Orally Bioavailable, Selective, and Potent Atp-Independent Akt Inhibitors", *J Med Chem*, Vol. 55, No. 11, pp. 5291-310, 2012.

70. Rollinger, J. M., T. M. Steindl, D. Schuster, J. Kirchmair, K. Anrain, E. P. Ellmerer, T. Langer, H. Stuppner, P. Wutzler, and M. Schmidtke, "Structure-Based Virtual Screening for the Discovery of Natural Inhibitors for Human Rhinovirus Coat Protein", *J Med Chem*, Vol. 51, No. 4, pp. 842-51, 2008.

71. Rollinger, J. M., H. Stuppner, and T. Langer, "Virtual Screening for the Discovery of Bioactive Natural Products", *Prog Drug Res*, Vol. 65, No. pp. 211, 213-49, 2008.

72. Ajay, A., W. P. Walters, and M. A. Murcko, "Can We Learn to Distinguish between "Drug-Like" and "Nondrug-Like" Molecules?", *J Med Chem*, Vol. 41, No. 18, pp. 3314-24, 1998.

73. Ballester, P. J. and J. B. Mitchell, "A Machine Learning Approach to Predicting Protein-Ligand Binding Affinity with Applications to Molecular Docking", *Bioinformatics*, Vol. 26, No. 9, pp. 1169-75, 2010.

74. Ballester, P. J., I. Westwood, N. Laurieri, E. Sim, and W. G. Richards, "Prospective Virtual Screening with Ultrafast Shape Recognition: The Identification of Novel Inhibitors of Arylamine N-Acetyltransferases", *J R Soc Interface*, Vol. 7, No. 43, pp. 335-42, 2010.

75. Rester, U., "From Virtuality to Reality - Virtual Screening in Lead Discovery and Lead Optimization: A Medicinal Chemistry Perspective", *Curr Opin Drug Discov Devel*, Vol. 11, No. 4, pp. 559-68, 2008.

76. Sun, H., "Pharmacophore-Based Virtual Screening", *Curr Med Chem*, Vol. 15, No. 10, pp. 1018-24, 2008.

77. Rastelli, G., "Emerging Topics in Structure-Based Virtual Screening", *Pharm Res*, Vol. 30, No. 5, pp. 1458-63, 2013.

78. Chen, H., P. D. Lyne, F. Giordanetto, T. Lovell, and J. Li, "On Evaluating Molecular-Docking Methods for Pose Prediction and Enrichment Factors", *J Chem Inf Model*, Vol. 46, No. 1, pp. 401-15, 2006.

79. Cavasotto, C. N. and A. J. Orry, "Ligand Docking and Structure-Based Virtual Screening in Drug Discovery", *Curr Top Med Chem*, Vol. 7, No. 10, pp. 1006-14, 2007.

80. Jones, G., P. Willett, and R. C. Glen, "A Genetic Algorithm for Flexible Molecular Overlay and Pharmacophore Elucidation", *J Comput Aided Mol Des*, Vol. 9, No. 6, pp. 532-49, 1995.

81. Jones, G., P. Willett, R. C. Glen, A. R. Leach, and R. Taylor, "Development and Validation of a Genetic Algorithm for Flexible Docking", *J Mol Biol*, Vol. 267, No. 3, pp. 727-48, 1997.

82. Horvath, D., "Pharmacophore-Based Virtual Screening", *Methods Mol Biol*, Vol. 672, No. pp. 261-98, 2011.

83. Loving, K., N. K. Salam, and W. Sherman, "Energetic Analysis of Fragment Docking and Application to Structure-Based Pharmacophore Hypothesis Generation", *J Comput Aided Mol Des*, Vol. 23, No. 8, pp. 541-54, 2009.

84. Salam, N. K., R. Nuti, and W. Sherman, "Novel Method for Generating Structure-Based Pharmacophores Using Energetic Analysis", *J Chem Inf Model*, Vol. 49, No. 10, pp. 2356-68, 2009.

85. *Schrödinger Suite Protein Preparation Wizard*, version 2011, Schrödinger, LLC, New York, NY, 2012

86. *Ligprep*, version 2.5, Schrödinger, LLC, New York, NY, 2012

87. *Confgen*, version 2.3, Schrödinger, LLC, New York, NY, 2012

88. Watts, K. S., 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", *J Chem Inf Model*, Vol. 50, No. 4, pp. 534-46, 2010.

89. *Qikprop*, version 3.5, Schrödinger, LLC, New York, NY, 2012

90. *Phase*, version 3.4, Schrödinger, LLC, New York, NY, 2012

91. Dixon, S. L., A. M. Smondyrev, E. H. Knoll, S. N. Rao, D. E. Shaw, and R. A. Friesner, "Phase: A New Engine for Pharmacophore Perception, 3d Qsar Model Development, and 3d Database Screening: 1. Methodology and Preliminary Results", *J Comput Aided Mol Des*, Vol. 20, No. 10-11, pp. 647-71, 2006.

92. Dixon, S. L., A. M. Smondyrev, and S. N. Rao, "Phase: A Novel Approach to Pharmacophore Modeling and 3d Database Searching", *Chem Biol Drug Des*, Vol. 67, No. 5, pp. 370-2, 2006.

93. *Glide*, version 5.8, Schrödinger, LLC, New York, NY, 2012

94. Friesner, R. A., 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", *J Med Chem*, Vol. 47, No. 7, pp. 1739-49, 2004.

95. *Epik*, version 2.3, Schrödinger, LLC, New York, NY, 2012

96. Irwin, J. J., T. Sterling, M. M. Mysinger, E. S. Bolstad, and R. G. Coleman, "Zinc: A Free Tool to Discover Chemistry for Biology", *J Chem Inf Model*, Vol. pp. 2012.

97. Lipinski, C. A., F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings", *Adv Drug Deliv Rev*, Vol. 46, No. 1-3, pp. 3-26, 2001.

98. Wu, Z., R. G. Robinson, S. Fu, S. F. Barnett, D. Defeo-Jones, R. E. Jones, A. M. Kral, H. E. Huber, N. E. Kohl, G. D. Hartman, and M. T. Bilodeau, "Rapid Assembly of Diverse and Potent Allosteric Akt Inhibitors", *Bioorg Med Chem Lett*, Vol. 18, No. 6, pp. 2211-4, 2008.

99. Parker, K. C., "Scoring Methods in Maldi Peptide Mass Fingerprinting: Chemscore, and the Chemapplex Program", *J Am Soc Mass Spectrom*, Vol. 13, No. 1, pp. 22-39, 2002.

100. *Macromodel,* version 9.9, Schrödinger, LLC, New York, NY, 2012

101. Backman, T. W., Y. Cao, and T. Girke, "Chemmine Tools: An Online Service for Analyzing and Clustering Small Molecules", *Nucleic Acids Res*, Vol. 39, No. Web Server issue, pp. W486-91, 2011.

102. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic Local Alignment Search Tool", *J Mol Biol*, Vol. 215, No. 3, pp. 403-10, 1990.

103. Henderson-Smart, D. J. and P. Steer, "Methylxanthine Treatment for Apnea in Preterm Infants", *Cochrane Database Syst Rev*, Vol. 3, pp. CD000140, 2001.

104. Barr, R. G., B. H. Rowe, and C. A. Camargo, "Methylxanthines for Exacerbations of Chronic Obstructive Pulmonary Disease", *Cochrane Database Syst Rev*, Vol. 2, pp. CD002168, 2003.

105. Tilley, S., "Methylxanthines in Asthma", *Methylxanthines*, Vol. 200, No.1, pp. 439-456, 2011.

106. Cappelli, A., G. Giuliani, M. Anzini, D. Riitano, G. Giorgi, and S. Vomero, "Design, Synthesis, and Structure-Affinity Relationship Studies in Nk1 Receptor Ligands Based on Azole-Fused Quinolinecarboxamide Moieties", *Bioorg Med Chem*, Vol. 16, No. 14, pp. 6850-9, 2008.