2D, 3D-QSAR and Docking studies on inhibitors of IR and IGF-1R: Exploration of binding mode at Tyrosine Kinase Domain

Mustafa Kamal Pasha

NUST201463251MRCMS64014F

RESEARCH CENTRE FOR MODELING & SIMULATION

2016

2D, 3D-QSAR and Docking studies on inhibitors of IR and IGF-1R: Exploration of binding mode at Tyrosine Kinase Domain

Mustafa Kamal Pasha

NUST201463251MRCMS64014F

Research Centre for Modeling & Simulation

A thesis submitted to the

National University of Sciences & Technology

In partial fulfillment of the requirement for the degree of Masters of Science

2016

STATEMENT OF ORIGINALITY

I hereby declare that work embodied in this thesis is the result of original research. I have conducted this work and have accomplished this thesis entirely on the basis of my personal efforts and under guidance of my supervisor, **Dr. Ishrat Jabeen.** If any part of this thesis is proved to be copied out from any source or found to be reproduction of the same scholastic work, I shall stand by the consequences. No portion of the work presented in my dissertation has been submitted in support of any other degree or qualification of this or any other University or Institute of learning.

_____   _____

Date   Mustafa Kamal Pasha

Acknowledgement

Primarily, I would like to enunciate my sincere gratitude and indebtedness to my Supervisor **Dr. Ishrat Jabeen** for the constant support of my M.S studies and related research, for her patience, motivation, and extensive knowledge. Her guidance supported me throughout the research phase and writing of the thesis. I could not thank enough to have an advisor and mentor on such terms.

Besides my advisor, the rest of the committee: **Dr. Farooq Amed Kiani** and **Dr. Shah Rukh abbas's** efforts are also noteworthy to have imparted their insightful comments and encouragement during the journey of thesis. It not only provoked me to widen my research but also their support in learning was admirable.

Also, my sincere Thank goes to **Pharmacoinformatics** group which provided me with the opportunity to join their team as researcher. They always stimulated discussions and ideas to mull over. Without their inestimable support, the research would not be possible to be conducted.

Apart from the research team, I would like thank my close friends exclusively **Hafiz Suliman Munawar** whose kind support has always ensured my true potential, **Junaid Nasir, Mahnoor Nadeem, Maqsood Ahmed** and **Captain Minhas Alam**, without their support it would have been a difficult task to complete.

Abstract

Insulin like growth factor-1 receptor (IGF-1R) along with highly homologous insulin receptor IR is integral component of cell proliferation, growth and survival and apoptosis protection. However, overexpression of IGF-1R and IR-A, and increased bioavailability of insulin-like growth factors (IGF-1 and IGF-2) in various cancers is a well-known phenomenon of tumorogenesis. Various strategies have been proposed in the past to evade cell proliferation during cancer. It includes targeting IGF-IR with selective antibodies and inhibition of its tyrosine kinase domain with small molecules. However, the results of these investigations remained gloomy and only few compounds reached in the clinical trials that ultimately failed to inhibit growth of human cancers. It is might be due to the fact that compensatory cross talk between IGF-1R and IR compromise drug efficacy during selective inhibition of IGF-1R. Recent molecular dynamics simulations of various co-crystalized ligand revealed that majority of the inhibitors bind to the basal state of the kinase hinge region thus, preventing the kinase activation. Therefore, development of small molecular antagonists impeding the tyrosine kinase domains of IGF-IR and IR-A has been advocated as a promising concept to evade cell proliferation and drug resistance in cancer chemotherapy. Therefore, present thesis, various *in silico* tools has been utilized to explore the binding hypothesis and molecular basis of interaction of tyrosine kinase domain of IR and IGF-1R with small molecular inhibitors. Briefly, combined ligand (3D QSAR, GRIND) as well as structure based techniques (Molecular Docking) have been utilized to identify dual 3D structural features of tyrosine kinase domain inhibitors.

Our results demonstrated the importance of two hydrogen bond acceptors, two hydrogen bond donors at a certain distance from each other as well as from molecular boundaries plays

significant role in dual inhibition of IR and IGF-IR. Our Pharmacophore models were used for the virtual screening of ChemBridge and WDI databases that resulted in eight potential hits as dual inhibitors of IR and IGF-1R. The current study will pave the way toward designing potential drug candidates with better Administration, Distribution, Metabolism and Elimination (ADME) properties and reduced toxicity against cancer.

Table of Contents

List of Figures

List of Tables

Introduction

The insulin like growth factor system comprises the IGF-I, IGF-II, their respective receptors and six IGF binding proteins (IGFBP) that regulate the availability of IGF-I and IGF-II. The IGFBP play important role in the regulation of IGFs once released in the bloodstream, this in turn is controlled by the concentration of insulin [22]. The IGFBP binds to the IGFs in order to increase their availability, by means of elevating the half-life of these molecules so that there is a regulated amount of IGFs available for binding to their respective receptors. Studies have shown that the administration of insulin results in a decrease in the serum levels of IGFBPs.

Insulin-like Growth Factors (IGF) belong to the family of mitogens that have a crucial role in terms of various cell signaling pathway including the regulation of cellular growth, proliferation, differentiation, transformation and apoptosis [23]. The IGF is secreted by the liver and is triggered by the action of growth hormone. As the name implies, the Insulin-like growth factor is roughly 48% similar to pro-insulin in terms of sequence. In the early developmental stages, both insulin and insulin like growth factors act as mitogens and promote cellular growth and proliferation. However, in the postnatal stages of life, insulin acts as a hormone involved in glucose metabolism while the IGFs retain their function as mitogens [24].

The IGF-1R is a transmembrane receptor comprising of two α subunits that are associated with ligand binding and two β subunits that mediate the intracellular signaling pathway. The IGF-1R is similar to insulin receptor in terms of structure and holds a significant sequence homology. The extent of similarity is quite high and has been reported to be around 45-65% in the ligand binding domains and approximately 60-85% homology exists in the tyrosine kinase domains [25, 26]. Studies have shown that IGF-1R and the IR have evolved from common

ancestral genes, although their functions vary but through the timeline of evolution, the structure has remained highly conserved in both receptors [27].The α subunit of the tetrameric receptors is extracellular and is rich in cysteine residues and has potential N-glycosylation sites (Asn-X-Ser). The cysteine residues of the IGF-IR are specific for the binding of IGF, whereas the N and C terminals of the IR are more dedicated towards the binding of insulin [28]. The IGF-1R as well as IR comprise of several structural domains. Towards the N terminus lies the L1 domain, which is the leucine rich repeat domain comprising the residue numbers 1-157, the second domain is the cysteine rich domain comprising the region of residues 158-310 and having seven disulphide molecules. The third domain is the L2 region, made up of residues 311-470. These domains constitute the ectodomain regions of the receptors and are followed by juxtamembrane domains and the intracellular C terminal ends that are associated with the internal signaling cascade [29]. Similar to the IGF-1R, the IR has a tyrosine kinase domain towards the intracellular side which is flanked by two regulatory regions. The kinases have a binding cleft where ATP can bind. This cleft comprises the activation loop (A loop) of the receptor. The said loop has crucial phosphorylation sites; Tyr 1158, 1162 and 1163. The phosphorylation of these sites is central in activation as it gives rise to the subsequent cascade and signaling pathways [30]. One of the flanking regions plays a role in binding the IR substrates (IRS), this region is also known to be involved in the process of receptor internalization [31]. The binding of respective ligands to the extracellular domain of IGF-1R and the IR results in the autophosphorylation of the tyrosine residue receptors, particularly, Tyr 1158, 1162,1163 are important in this regard [32]. The phosphorylation of these tyrosine residues results in stabilizing the activation loop [33]. When activation has not occurred, the Asparagine of the conserved Asp-Phe-Gly (DFG) motif points away from the ATP binding loop; this is termed as the DFG-out conformation. When activated,

the Asp of DFG motif points towards the ATP binding cleft and the conformation is termed as DFG-in[18]. The phosphotyrosines are competitive binding site analogues for the insulin receptor (IR), insulin receptor substrate (IRS) and in some cases IGF1R. Upon binding, these substrates are phosphorylated and lead to the activation of the PI3K and Ras-MAPK pathways. A significant similarity exists in the pathways activated by IGF-IR and the IR, since they share a number of components pertaining to the said pathways [34]. Binding of insulin to the IR receptor mediates a complex signaling pathway that in turn is responsible for controlling various cellular processes. Two main signaling pathways induced by the action of insulin are the phosphatidylinositol 3-kinase (PIK3) / protein kinase B (PKB) pathway responsible for the metabolic actions of insulin and the second pathway is the Ras- activated protein kinase (MAPK) pathway which is responsible for control of cell growth and differentiation [35].
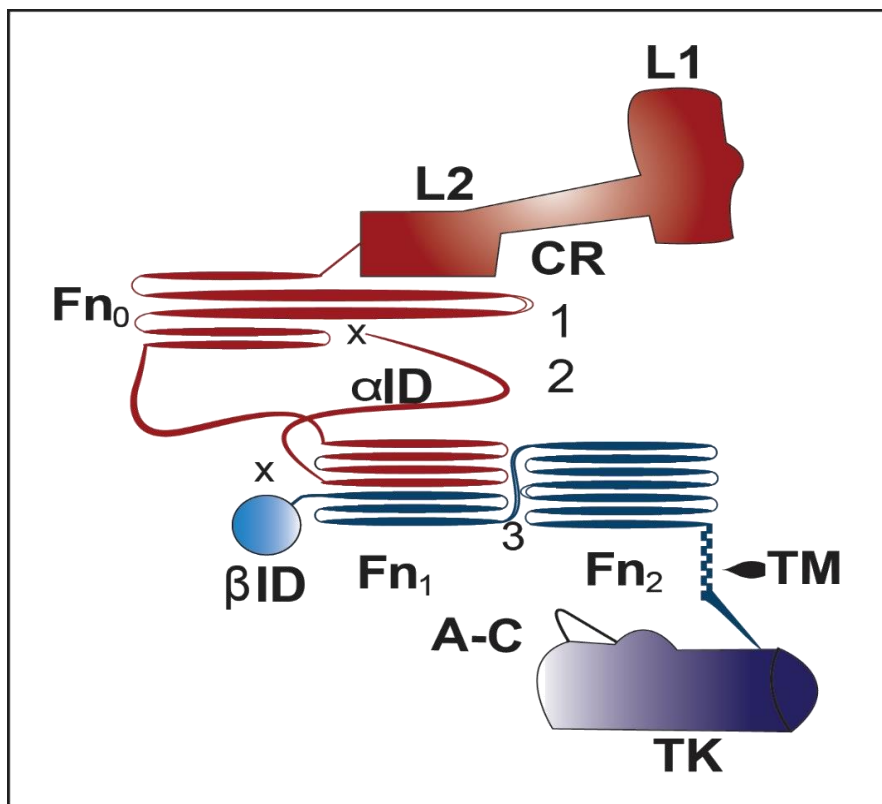
Figure 1: *Structural organization of IGF1R and IR structures adopted from [36-38]*
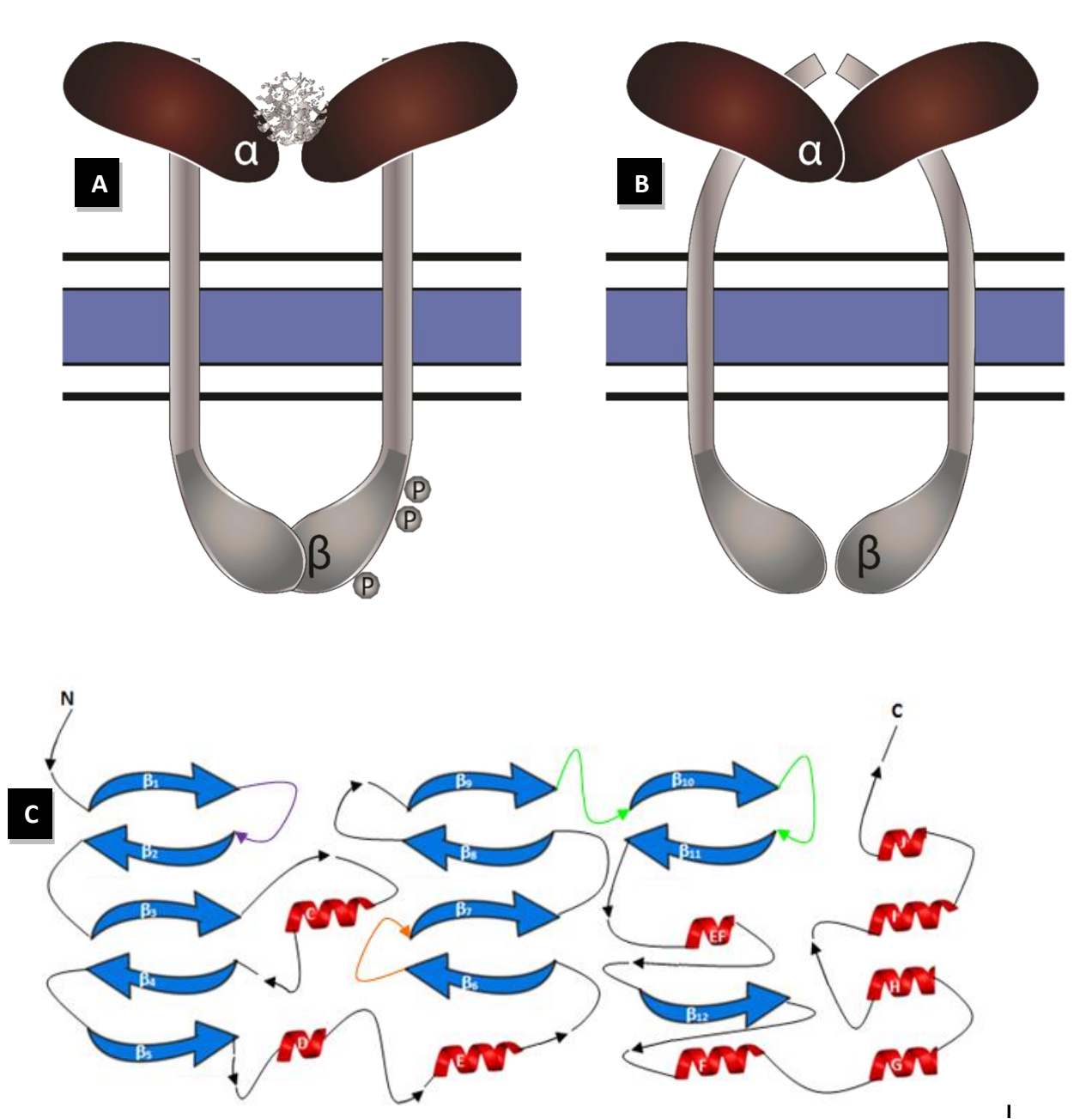
Figure 2: *A: Attachment of insulin results in structural change which brings β domain closer and Phosphorylation takes place, B: The Apo state (or basal) where the is no phosphorylation is done prior to attachment of insulin[36], C: A zoomed in Tyrosine kinase domain illustration, Purple: Nucleotide binding loop, Orange: catalytic loop, Green: Activation loop [1, 39, 40]*

The IR has two isoforms, IR-A and IR-B. The existence of two isoforms is due to the alternate splicing of exon 11 in the IR, which results in a shorter form, the IR-A form, which lacks a stretch of 12 amino acids and the IR-B form. The reason behind the presence of two isoforms of the insulin receptor is speculated to be the different affinities of insulin binding as per requirement at the organ/system level [41]. The two isoforms are differentially expressed on plasma membrane and regulate the insulin signaling pathway by activating varying classes of phosphatidylinositol 3-kinase (PIK3). The abundance of these two receptor variants is controlled in a tissue specific manner. The IR-A is associated with fetal tissues and has also been observed in cancer cells. Whereas the IR-B is predominantly associated with the tissues specific for insulin metabolism [26].

Since the IGF-1R does not possess the exon 11, its structure is similar to the IR-A. The similarity between IR-A and IGF-1R has been studied in detail and a number of facts have been elucidated [42]. The IGF-1R is activated after binding to IGF-I whereas the IR is specific for binding of insulin. Although the ligands for each receptor are different and are involved in different signaling pathways however, there is well established data about binding of IGF-II to the IR-A leading to its activation. IGF-II, which is normally required for the proper development of embryo has an unclear role in the post-natal period [43], it has been positively associated with the hyper proliferation of cells and establishment of malignancies. Furthermore, there is evidence of a high number of IGF-IR present in transformed cells [44]. Physiologically, the IR is responsible for binding of insulin and should ideally be present in the hepatocytes and the skeletal muscles but it has been observed that IR is present in other tissues including the brain, heart, monocytes, granulocytes, pancreatic acini, vascular endothelium, kidney and fibroblasts.

This observation suggests that IR plays a role not only in the insulin related metabolism but also has functional roles in other systems[45].

High levels of insulin induce a higher level of serum IGF-I which in turn have properties of mitogen and promote anti-apoptotic behavior in cells. Hence chronic hyperinsulinemia has been associated with carcinogenesis. The IGF-IR is not only involved in the process of transformation but has also been found to play an important role in the maintenance of the transformed state [46]. Therefore these receptors are associated with aggressive tumors. Over the last two decades, studies have provided in depth evidence that the IGF-1R is present in various types of cancers and that the blockage of this receptor provides a means of blocking the signaling pathways that follow, particularly the Ras-Raf pathway which is important in providing resistance to the tumors [47]. An increase in the components of the Ras-Raf pathway leads to phosphorylation of the components involved in the Programmed cell death (PCD). Moreover, the p53 cascade is affected negatively. The p53 down regulation is linked directly to the prevention of PCD [48].

IGF-1R is involved in the proliferation of transformed cells, it is also involved in transformation by oncogenic viruses [46]. As of recent studies, IR-A has also been seen in the course of tumor genesis [42]. In cells where the IGF-1R and the IR are co-expressed, they form a hybrid (IR-A/IGF-1R). The function of these hybrids is not exactly clear but their expression has been observed in cancer cells. Studies on different myosarcoma cell lines revealed that IR-A was predominantly expressed in sarcoma cells [49]. Most of the cancers express genes for the insulin receptor (IR) as well as the genes encoding Insulin like growth factor 1 receptor (IGF-1R). The IGF-1R has been indicated as important in cancers and its expression is positively associated with the progression of tumor [50]. One of the key features noted in the expression of IR-A in

cancers is associated with its high affinity for IGF-II, which leads to signaling pathways that are involved in transformation and in evading apoptosis [51]. Targeting of IR-A alone would result in hyperglycemia, leading to metabolic complications [52].

Since both, the IR-A and IGF-IR have been observed in various cancers, there is a need to develop potential inhibitors that could block these receptors and in turn control the transformation of cells. Previously, monoclonal antibodies have been targeted to block the IGR-IR receptor. Additionally, small molecule inhibitors are being explored for their ability to bind to the receptor and in turn block the binding of the original substrates in order to stop the rigorous signaling followed by the receptor binding [3]. Small molecule inhibitors that can bind to the tyrosine kinase domain have been reported to bind to the catalytic domain responsible for kinase activity. One key point is the targeting of these receptors in their inactive state because the inactivated ATP binding pocket is less conserved and hence a more reliable target. The activated state is conserved among the kinase family and hence proves a promiscuous target. Therefore, a better approach is to target these receptors by small molecule inhibitors when they are in the inactivated or the DGF-out conformation [53].

The IGF-IR and IR both belong to receptor tyrosine kinases (RTKs) and targeting any one of them has shown to result in tumor resistance. The protein kinase domain of these receptors provides a docking site for the protein kinase C (PKC) and protein kinase C related proteins (PRP). The catalytic domain of kinases has two lobed; the N and C lobes. The catalytic cleft is present in between these two lobes. Glycine is predominantly present in the N lobe. This provides a potential for exploiting the binding pocket and directing small molecule inhibitors against the RTKs [54].

Initial studies revolved around screening and development of small molecule inhibitors against the IGF-IR only but there exist evidences for the role of IR-A in transformation and maintenance of tumors. Targeting IGF-1R alone can have undesirable consequences and may lead to implications that may strengthen malignancies [55]. There is increasing evidence that repressing any one of the RTKs is not enough, the other receptor, still functional, will do enough to compensate and would lead to maintenance of the tumor. Elevated levels of insulin result in higher signaling by IR-A. Also, there is evidence of IGF-II binding to IR-A resulting in an increased in the mitogenic signaling in the transformed cells [56]. Therefore, a better approach would be co-targeting both the receptors in order to make sure that the mitogenic signaling cascade activated by both these receptors is down-regulated so that proliferative pathways that lead to tumor formation can be stopped [57].

The prime reason for targeting the mentioned receptors is their established role in cancers. Since cancer is among the leading cause of death globally, finding possible therapeutic agents is of great importance. Studies have shown that targeting either IGF-1R or IR alone is not enough, as there exists the liability of the counter-part being activated [55].

- Targeting the tyrosine kinase domain of IR-A and IGF-1R with small molecules.

- Screen dual/selective inhibitor of tyrosine kinase domain of IR-A and IGF-1R.

- Identify novel, potent and more promising inhibitors of IR-A and IGF-1R by the development of *in silico* approaches.

- Determine the 3D structural features of inhibitors of IR-A and IGF-1R.

- To understand molecular basis of interaction of small molecules with IR-A and IGF-1R

Finding novel small molecule inhibitors that can bind to the tyrosine kinase domain common of both the receptors is a direction that can help bring down the ill effects that are associated with the mitogenic pathways activated as a consequence of ligand binding to the respective receptor. Therefore our approach is to search for the dual inhibitors that can simultaneously target both, IGF-1R and IR in order to increase the efficacy of chemotherapeutic agents being designed to curb the casualties of a fatal disease.

Literature Review

Both, the IGF-1R and the IR have an extracellular domain having an α subunit and a β subunit that is the transmembrane domain and has a tyrosine kinase activity. The insulin-like growth factor receptor (IGF-1R) and the Insulin Receptor (IR) have a phylogenetic relation that traces back to a common ancestral receptor. Although the structural similarities are striking but the ligands for both the receptors are different. Insulin is the ligand specific for binding and activation of IR while the IGF-1R is activated by the binding of IGF-I and IGF-II [58].

Extensive research has been done on searching probable inhibitors of IGF-1R due to its role in various cancers. But more recently, the role of IR has also been observed in certain cancers. In fact, many tumors show an elevation of both these receptors [42]. Since the establishment of the fact that both receptors are responsible for cancer progression and can be targeted simultaneously [59, 60], a deluged amount of studies have been performed as SAR, X-ray crystallographic as well as docking and pharmacophoric studies.

Drug binding site:

It has been assented that the affinity of binding pocket of IGF1R and IR vary [61].Pocket has been identified by Ward *et al*. They declared that CT peptide residues are critical in ligand binding, particularly F714 in case of IR and F701 in IGF-1R. It has also been established that any change in this particular peptide region is known to nullify the ligand binding property of the respective receptor [62]. Another key point has been noted that the tyrosine kinase residues 1158, 1162 and 1163 are important in the activation loop and they serve as potential docking sites. Later, The catalytic loop region from 1103 – 1112, activation loop region ranging from 1122 –

1144and also the phosphate binding loop between β1 and β2 (Fig: 2 (C)) have been proposed by Munshi *et al.* to play an important role in the activation of the domain and ATP binding. Furthermore, it has also been elucidated that this pocket arranges itself into two forms as "DFG out" in the in-activated and "DFG in" in the activated states because of the motif's direction towards or away from the ATP pocket [40]. Interestingly, it has been reported that the ATP binding site of both IR and IGF-1R have a high degree of sequence similarity and this pocket has been targeted in a number of studies[59].

Structure Activity Relation (SAR) studies:

Structure Activity Relation (SAR) studies are based on the principle that similarity of compounds will lead to similar binding properties to the respective receptor. Keeping this in view, compounds against IGF-1R have been identified by SAR studies. Taha *et al.* have considered QSAR modeling for Human protein tyrosine phosphatase inhibitors from a diverse subset of 154 training compounds and concluded that pharmacophore modeling is a better strategy to mine against virtual libraries and have identified few inhibitors which match the pharmacophore query. On the other hand, the QSAR modeling multilinear regression equation demonstrated the positive contribution of logp suggesting that hydrophobic ligand have superior inhibitory properties. Similarly, overall topology, sum of all energy states and inhibitor's length plays a vital role in its inhibitory property [63].The Receptor guided 3D QSAR studies have been delineated to have potential of obtaining more potent ligands against the IGF-1R receptor. For this, Muddassar *et al* concluded by comparing receptor guided and structure guided CoMFA and CoMSIA studies on a series of 54 1, 3-disubstituted imidazole [1,5-α] pyrazine derivatives that steric and electrostatic field play important role in receptor ligand affinity. They further

concluded that receptor guided QSAR is more accurate in predicting the binding residues interactions[64] while IR has not yet been modeled as 2D/3D.

As many as 50 or more compounds from different classes of inhibitors have been reported to bind to the said receptors [3]. All these compounds have a common core structure but their pattern of binding is different. In such cases, 2D and 3D SAR studies are required in order to determine the different peripheral groups that are involved in protein binding [65, 66]. Keeping this in view, Sperandio *et al.* performed SAR studies on a class of 4,6-bis-anilino-1H-pyrrolo[2,3-d]pyrimidine for IGF1R. This was executed on a query compound without the protein structure to see which peripheral groups are rather available for protein binding to show steric hindrance of the compounds [65]. The principle of aligning common parts has been exercised for QSAR studies which revealed the positive charge and electron density played important role in depicting activity of the compounds. Likewise, comparative QSAR studies for different RTKs on experimental data have also been done to elucidate the important factors for the binding pocket and the conclusion revealed diverse results for hydrophobicity, electronic and steric effects [67, 68].

A study was performed by Liu X. *et al.* for SAR studies on thiazolidinediones (TZD) as inhibitors of IGF1R. They showed that their optimization of the series led to discovery of two potent analogues from Novel 5-benzylidenethiazolidine-2,4-dione (compound 5) and 5- (furan-2-ylmethylene) thiazolidine-2,4-dione (compound 6) with IC50 of 57 and 61 nM. The purpose of their study was to exemplify the improved efficacy of novel compounds on IGF1R than IR, which was believed that this potency was due to the H-bonds present on TZD moiety as it played a pivotal role in anchoring the pose into the receptor. They further validated their results via hierarchical virtual screening and pharmacophore modeling [2].

In silico drug design and SAR studies for the binding of small molecule inhibitors related to bis-azaindoles were performed and a 3D homology model was constructed using the 3-phophorylated IGF-1R structures that were ATP-bound. A number of modifications and ligand screenings led to a series of potent azaindolesanalogues that showed binding modes for IGF-1R. The overall tenfold selectivity was observed by N-4 at indole moiety [52]. CoMFA and CoMSIA require structural 3D alignment according to a template to work.

In addition to the bis-azaindoles, a class of compoundsbelonging to the group of 4-amino-1H-pyrazolo[3,4-d] pyrimidineinhibitors were studied and they showed that steric bulk is unfavorable for its low activity which was in agreement with their docking studies stated elsewhere. The cross validation of the 3D- QSAR model gave a high cross correlation of $q^2_{pred.}=$ 0.590 [69].

A 3D-QSAR study on traditional Chinese medicines (TCM) has been performed by Chang *et al.* which revealed 3- (2-carboxyphenyl) -4 (3H) -quinazolinone from *Isatisindigotica*, (+) -1 (R) -Coclaurine from *NelumbonuciferaGaertn and* (+) -N-methyllaurotetanine from *Lindera aggregate* potent inhibitors of IGF1R. The purpose of their study was to develop memory enhancing candidates but because of CoMFA and CoMSIA, Multiple linear regressions, Bayesian networks, support vector machine and MD results of these candidates gave evidence for stable complexes with IGF1R for its inhibition. The predicted Bioactivity also suggested that these candidates had drug like properties [70].

In line with different classes of compounds being explored as potential inhibitors, a novel class of azapeptide's sub monomers was explored by Kurian *et al.* The preliminary results of the SAR from the series of azapeptides adopted folded structures in solution; the Ac-DIazaYET-NH2 displayed a stable β-turn geometry that was found to translate to 5-fold increase in IRTK

inhibitory activity relative to the parent pentapeptide at 400 µM. It was confirmed via assays that it was able to inhibit ~70% of the activity of RTK [71].

Molecular Modeling:

The molecular modeling started when 21 homology models of IGF1R were built in order to assess the ability to differentiate between active and inactive via high throughput docking. All the models built yielded different enrichment factor from few folds to 12 folds. The performance could also be observed from good to worse, compared with crystal structures [72].

Later, IGF-1R along with Focal Adhesion Kinase (FAK) was modeled, which showed an important role of IGF-1R in the progression of cancer. The kinase domain of both these receptors were shown to interact and hold a potential site for the binding of small molecule inhibitors [73].

It was also reported that studies on minima-mining based on simulated annealing approach for 3 unbound Insulin receptor structures wereprepared by applying weak RMSD constraints. In spite of conventional protocol, this study ran many successive short cycles lasting certain picoseconds. After having many minima, clustering and connecting centroid structures based on similarity produced structures [74].

An attempt was made by Hung *et al.* from TCM for inhibition of RTK of IGF1R. For this, they executed molecular dynamics (MD) for stable bound complex and found that Leu975 and Gly1055 or Asp1056 were key ligand binding residues[70]. In addition to MD Group based quantitative structure activity relationship (GQSAR) was performed and the models generated showed site-specific clues for the design of potential IGF-1R inhibitors [75].

The literature is silent about the modeling and SAR studies of Insulin receptor for the purpose of inhibiting the tyrosine kinase domain. Fewer studies have been found regarding the IR *in silico* approaches.

Molecular Docking:

Taha*et al.*demonstrated docking upon the models they build via QSAR modeling. Their docking results suggested that pyrimidinedione head of 157 was hydrogen bonded to the Arginine221 which truly maps with the selected pharmacophore model's HBA feature. Besides HBA the pentylene chain was fitted against the hydrophobic feature and acid moiety was consistent with negative ionizeable feature of the pharmacophore[63]. Furthermore, a high throughput docking (HTS) on 21; template built homology models of IGF1R was carried with GLIDE of GOLD by Ferrara *et al.* They deduced that size of the pocket matters and few regions in the pocket are more conserved than any other. This study was performed on three crystal structures; two having ATP (1JQH) and NVP-AEW541, while the third structure was in inactive state (IP4O). Moreover, the rigidity of the receptor was found to be the key element in HTS. Further, Docking performed on smaller size of the pocket gave worse results than bigger pocket [72].

Another study took IGF-1R coordinates from protein databank (1JQH) for molecular docking. The ATP binding site/active site was used for docking with a radius of 6.5$^{°}$. With the generation of 100 poses for ligand (structure A) they understood the receptor-ligand interactions and also showed key interactions where the ligand could bind for its inhibition which revealed several key residues i.e.M-1082, K-1033, D-1086, G-1006, and L-1005. It was evident from their

study that hydrophobic region of ligand binds with the hydrophobic pocket of receptor. Moreover, two HBA with K1033 and M1082 and one HBD with E1080 were proposed [64].

The first known inhibitor of IGF1R by NCI was developed by docking 1-aza-9-oxafluorene ring derivative with 4-benzylamino substituent via GOLD. The docking model showed H-bonding with Met1082 (IGF-1R) found in the hinge region of the ATP pocket [76].This also correlated with the residues found by Muddassar *et al.*

In wake of the above study, aryl-heteroaryl urea (AHU) scaffold was used by Engen *et al.* for identification of small molecular inhibitors of IGF-1R. They chose the hetero-aryl ring system which contained the 4-aminoquinoline framework that holds the potential of enhanced binding with the ATP binding cleft of the receptor kinase domain. They practiced modifications of the basic ring system framework and docked ligands into the active site of IGF-1R. Their results include a series of ligands that consistently bind to the active site of IGF-1R [77].

Li *et al.* have performed docking using genetic algorithm of GOLD V3.0.1of 25 compounds of class 4-amino-1H-pyrazolo[3,4-d] pyrimidines inhibitors to generate binding affinities and positions of the said class in the active site of IGF1R (3LWO). It has been concluded that, although the class has larger molecules than the already present inhibitor found in the structure they have used. Still the conformations, binding free energies and predictive activity had a good correlation with already docked inhibitor CCX [69]. Subsequently, DiallyDisulphide (DADS), a compound of garlic was docked against the IGF-1R and showed 7 possible interactions [78]. Meanwhile, Huang *et al*. executed docking protocol on insulin receptor with hexapeptideGDYMNM and deduced that there are two docking pathways for this receptor's tyrosine kinase domain. From these two pathways one was able to bring hexapeptide closer to ATP and produced a conformation able to phosphorylate, while, the other couldn't.

16

Their structural analysis showed that activation loop having hydrophilic residues along with few others might play a crucial role in peptide binding kinetics [74].

Mahajanakatti *et al.* performed multiple receptor docking to elucidate the binding ability and inhibitory potential of curcumin against the kinase receptors known to be up regulated in cancers [79]. This study showed binding potential of selected natural compounds to a number of kinase receptors associated with cancers, including the IGF-1R receptor.

Other than that, high throughput virtual screening and extra precision docking were performed by Abdullahi *et al.* on 164 inhibitors extracted from Binding-DB of BMS-754607, 2, 44-disubstituted pyrrolo [1, 2-f][1, 2, 3] triazine dual inhibitor and a non ATP-competitive inhibitor with bioactive values in order to identify ligands that had the best binding ability towards the IGF-1R [75].

Supplementary to above, docking of a dataset of 68 compounds from IBS natural library (In-terBio Screen Ltd) and anticancer nature compound on IGF1R and IR was done with the help of GLIDE from Schrodinger. From this study a compound named EGCG (epigallocatechingallate) was found to be active and reliable against IGF1R but not with IR. Further investigation revealed that it reduced the cell proliferation, and mRNA expression, oxidative stress of IGF1R. It also showed interaction pattern with Val983, Leu975, Ala1001, V1033, M1049, L1051, M1052, M1112 and F1124. Moreover, interactions like lipophilicity, hydrogen bonding and π-π stacking played a pivotal role in receptor- ligand active site[80].

Furthermore, recent studies demonstrated that V1010,A1028, V1060, M1076, L1078, M1079, A1080 and F1151 are involved in the hydrophobic interactions of IGF1R and IR with Benzo α pyroneflavonoid [81].

Pharmacophore Modeling:

Taha *et al.* elucidated that two HBAs and hydrophobic areas with one negative ionize able feature are important for inhibitors to occupy the selective features using HipHop-REFINE software [63].

Liu *et al*. developed a pharmacophore model of IGF-1R that characterized three hydrogen bonds of the hinge region backbone with the inhibitors. Among these, two are hydrogen bond donors target the backbone E1050 and M1052 along with one hydrogen bond acceptor that target M1052 amide group. They used this model to find pharmacophore contrived by TZD, based potential hits in order to minimize the chances of any errors or mishit being included in targeting ligand against the IGF-1R [2]. Ramdhave *et al.* built 7 different pharmacophore models against 7 different inactive co-crystal structures of IGF1R (PDB IDs; 2ZM3, 3F5P, 2OJ9, 3D94, 3I81, 3O23 and 3LVP) and did virtual screening against SWEETLEAD database. All the pharmacophores built had positive ionizable group on the hinge Nitrogen atom except 3D94. They identified 9 hits after visualizing S-score and docking into inactive IGF1R co-crystal structure which also satisfied key residue interaction patterns. Their work needed experimental validations based on their study [82].

Dual inhibitors of IGF-1R and IR:

Benzimidazole based dual inhibitors of IGF-1R and IR have been studied and have shown favorable results in inhibition of the mitogenic cascade. Pyrazolopyramindines have also been tested and have shown good inhibition of both receptors, their binding mode involves a hydrogen bond interaction with the residues Glu1050 and Met1052 of the hinge loop [59]. This interaction has been pointed as a key feature in disrupting the P loop and the Active loop of the

receptors which results in an overall inhibition of the target receptor. The pyrrolopyrimidine ligands interact with the hydrophobic back pocket and their binding induces significant reduction in the level of phosphorylation and hence the subsequent signaling pathways.

The dual inhibitor strategy is relatively new and has paved a wide area of studies as there exist a number of binding conformations for different compounds. Understanding the binding modes of different compounds shall require a context dependent approach in order to gain an insight into the varying drug selectivity profiles.

Co-targeting of IGF-1R and IR:

Although, the insulin receptor (IR) is involved in the metabolic pathways in routine but there exist evidences that point towards its high expression in cancer. Particularly, when small molecule inhibitors are targeted against IGF-1R, it has been observed that the tumors switch towards IR as an alternative to keep up with the proliferative pathways and the mitogenic cascades. Co-targeting was started as early as in 2006, when Haluska *et al.* gave in vivo and in vitro proves of dual inhibition of both receptors with BMS-554417 [83]. Recently, it has been suggested that dual targeting of receptors could be helpful to evade from compensatory crosstalk as a survival mechanism generated in the presence of solo target i.e. IGF1R, thence, giving bigger therapeutic window and beneficial response. This being blocked, the unwanted proliferation and subsequent tumor survival will subside [59].

*Table 1 : Insulin Receptor (IR) Tyrosine Kinase Domain; Data for Structural Comparison of the Binding Pocket*

S.no	PDB ID	Year	Origin	Residual Count	Resolution	Reference
1	1GAG	1/17/2001	Homo sapiens	319	2.70 Å	[84]
2	5E1S	10/14/2015	Homo sapiens	308	2.26 Å	[85]
3	1I44	3/7/2001	Homo sapiens	306	2.40 Å	[86]
4	4XLV	3/25/2015	Homo sapiens	328	2.30 Å	[87]
5	4OGA	8/27/2014	Homo sapiens	616	3.50 Å	[88]
6	4IBM	8/21/2013	Homo sapiens	612	1.80 Å	[18]
7	3LOH	4/28/2010	Homo sapiens	1790	3.80 Å	(Obsolete)
8	3ETA	5/26/2009	Homo sapiens	634	2.60 Å	[89]
9	3EKN	12/30/2008	Homo sapiens	307	2.20 Å	[16]
10	3EKK	12/23/2008	Homo sapiens	307	2.10 Å	[14]
11	1P14	7/22/2003	Homo sapiens	306	1.90 Å	[90]
12	1IRK	1/7/1998	Homo sapiens	324	1.90 Å	[39]
13	1RQQ	12/30/2003	Homo sapiens	876	2.60 Å	[91]
14	2AUH	11/1/2005	Homo sapiens	365	3.65 Å	[92]
15	2B4S	11/15/2005	Homo sapiens	1208	2.30 Å	[93]
16	2Z8C	8/12/2008	Homo sapiens	309	3.25 Å	[94]
17	3BU5	2/19/2008	Homo sapiens	321	2.10 Å	[95]
18	5HHW	4/13/2016	Homo sapiens	307	1.79 Å	[96]

Table 2: IGF-1R Tyrosine Kinase Domain; Data for Structural Comparison of the Binding Pocket

| S.no	PDB ID	Year	Origin	Residual Count	Resolution	Reference
1	5HZN	4/6/2016	Homo Sapiens	304	2.2 Å	[96]
2	4D2R	24/22/2015	Homo Sapiens	302	2.1 Å	[97]
3	3O23	5/4/2011	Homo Sapiens	305	2.1 Å	[98]
4	3QQU	4/20/2011	Homo Sapiens	1204	2.9 Å	[12]
5	3LW0	9/29/2010	Homo Sapiens	1216	1.79 Å	[99]
6	3NW5	7/28/2010	Homo Sapiens	307	2.14 Å	[100]
7	3LVP	7/21/2010	Homo Sapiens	1344	3.0 Å	[52]
8	3I81	12/22/2009	Homo Sapiens	315	2.08 Å	[101]
9.	3F5P	12/30/2008	Homo Sapiens	4928	2.9 Å	[15]
10.	3D94	7/29/2008	Homo Sapiens	301	2.3 Å	[102]
11.	2ZM3	6/10/2008	Homo Sapiens	1232	2.5 Å	[19]
12.	2OJ9	5/1/2007	Homo Sapiens	307	2.0 Å	[103]
13.	1P4O	4/292003	Homo Sapiens	644	1.5 Å	[40]
14.	1JQH	4/19/2002	Homo Sapiens	924	2.1 Å	[104]
15.	1K3A	11/28/2001	Homo Sapiens	313	2.1 Å	[105]

Methodology

Since it has been established that co-targeting both receptors would be the best policy, the dataset of total 33 dual inhibitors of IGF1R and IR was extracted from the literature. This dataset was diverse and belonged to different classes of compounds i.e. imidazopyrazine, imidazopyridine, pyrrolopyramidine, quinolines, cyanoquinolines, isoquinolinedione series and several others[1, 3, 6, 9, 11-15, 17-19, 21, 106-109]. This record proved to be insufficient for building any reliable hypothesis because of its diversity. A total of 14 more modulators were extracted from literature [2, 4, 5, 10] which were more selective towards IGF1R than IR. This divided the data into two halves; 'selective' and the 'dual'. The Combined sum of 47 modulators was then used in different methodologies in search of results on said hypothesis. All the IC_{50} values from the database were converted into pIC_{50} for data simplification by taking the negative natural log of molar units of the potency. The following structures show the comprehensive database made for evaluation.
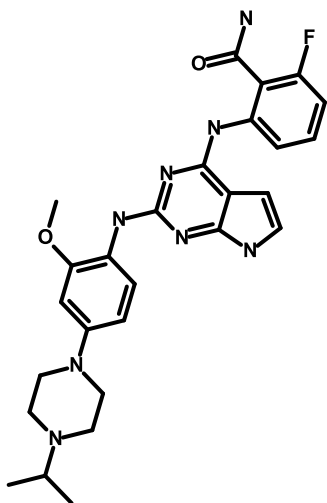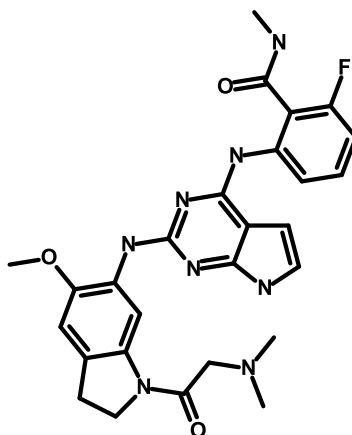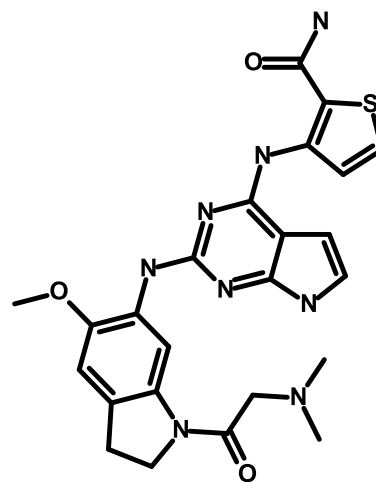
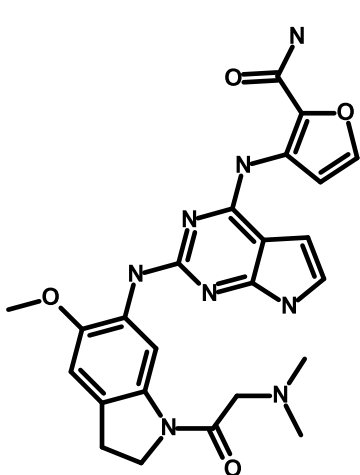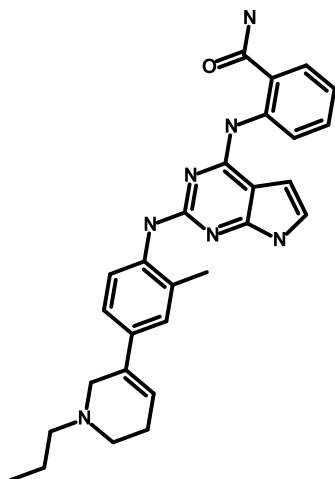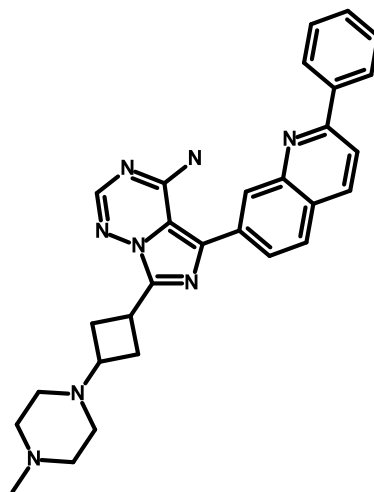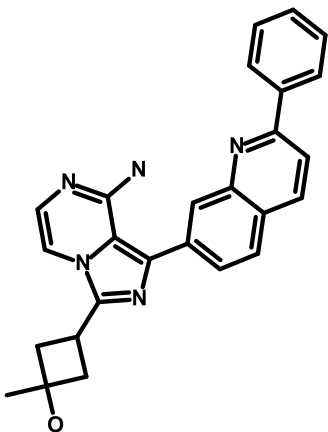|		
PIC_{50} (nM) = IGF1R/ IR= 8.6/ 8.7 [16]	PIC_{50} (nM) = IGF1R/ IR= 8.6/ 8.7 [16]	PIC_{50} (nM) = IGF1R/ IR= 9.5/ 9.3 [20]

PIC$_{50}$ (nM) = IGF1R/ IR= 9.0/ 9.3 8.7 [20]
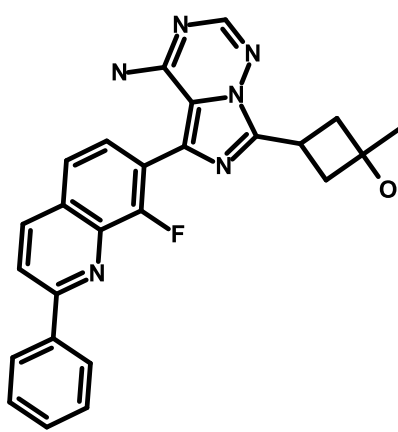
PIC$_{50}$ (nM) = IGF1R/ IR= 8.6/ 8.7 [14]

PIC$_{50}$ (nM) = IGF1R/ IR= 8.3/ 8.2 [1, 2]

PIC$_{50}$ (nM) = IGF1R/ IR= 6.9/ 6.5 [6]

PIC$_{50}$ (nM) = IGF1R/ IR= 7.4/ 7.1 [6]
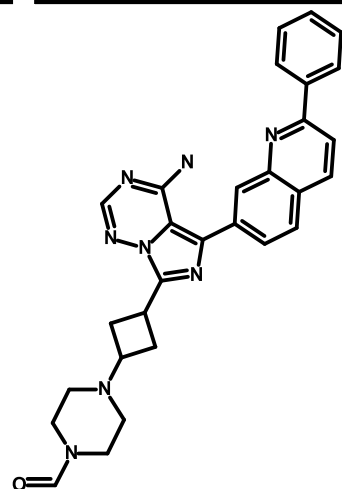
PIC$_{50}$ (nM) = IGF1R/ IR= 7.7/ 7.1 [6]
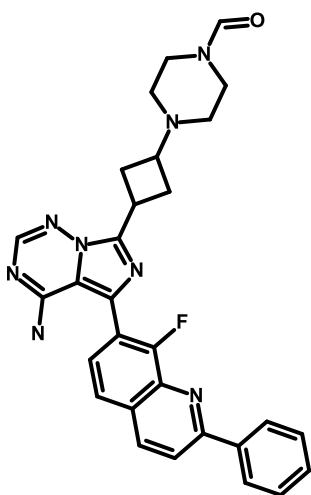
PIC$_{50}$ (nM) = IGF1R/ IR= 8.0 / 7.3 [6]
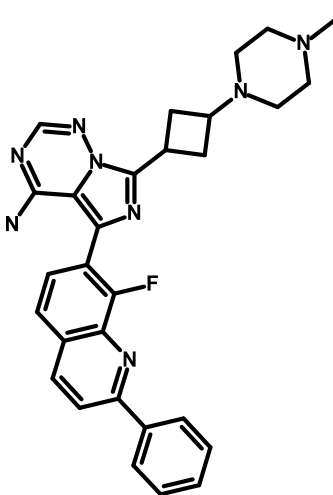
PIC$_{50}$ (nM) = IGF1R/ IR= 7.8/ 7.1 [6]
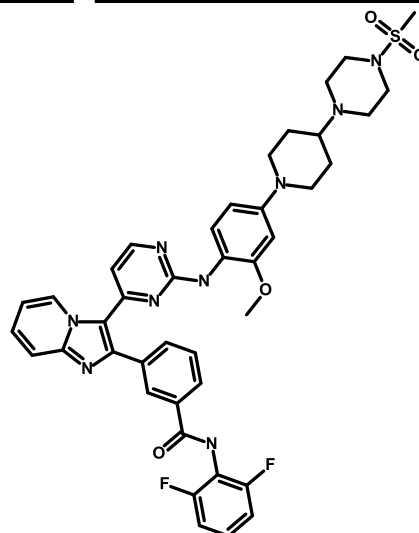
PIC$_{50}$ (nM) = IGF1R/ IR= 8.3/ 8.3 [13]
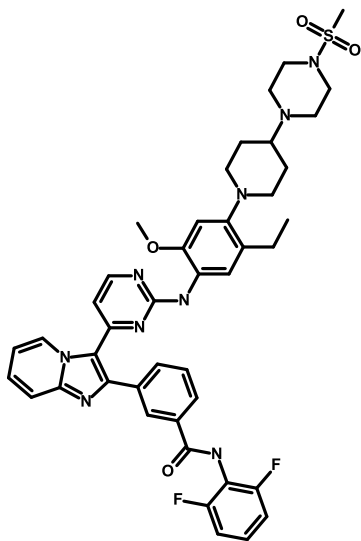
PIC$_{50}$ (nM) = IGF1R/ IR= 8.1/ 8.1 [13]
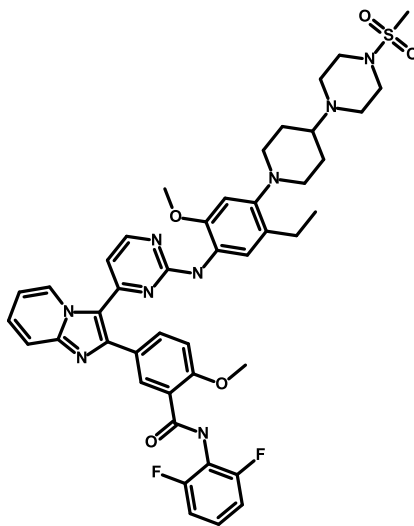
PIC$_{50}$ (nM) = IGF1R/ IR= 7.5/ 7.6[13]
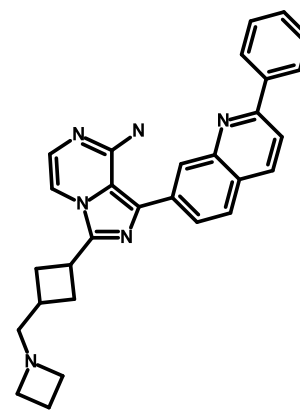
PIC$_{50}$ (nM) = IGF1R/ IR= 7.4/ 7.1 [11]

PIC$_{50}$ (nM) = IGF1R/ IR= 7.9/ 7.8 [8]

PIC$_{50}$ (nM) = IGF1R/ IR= 7.6/ 7.4 [12]

PIC$_{50}$ (nM) = IGF1R/ IR= 7.6/ 7.6 [9]

PIC$_{50}$ (nM) = IGF1R/ IR= 7.1/ 7.2 [3]

PIC$_{50}$ (nM) = IGF1R/ IR= 7.0/ 7.1 [17]
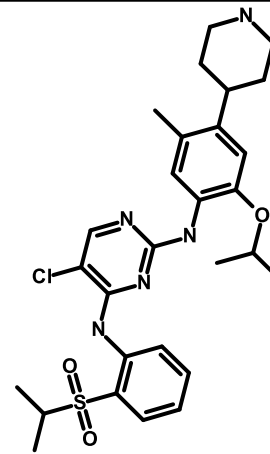
PIC$_{50}$ (nM) = IGF1R/ IR= 8.0/ 8.1 [21]
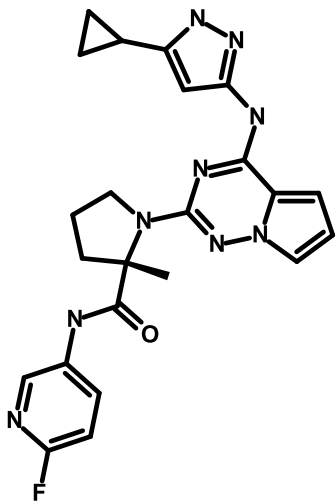
PIC$_{50}$ (nM) = IGF1R/ IR= 8.7/ 8.7 [21]
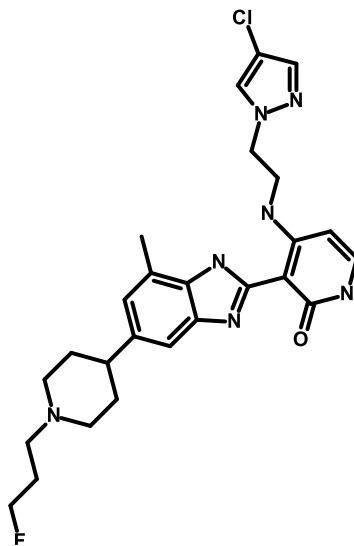
PIC$_{50}$ (nM) = IGF1R/ IR= 7.4/ 7.1 [3]
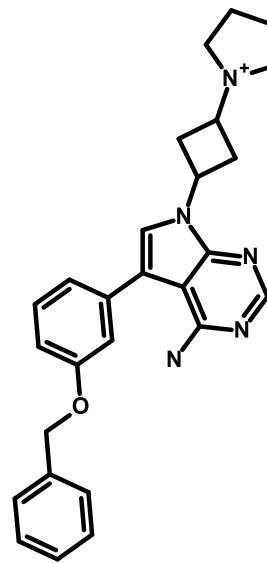
PIC$_{50}$ (nM) = IGF1R/ IR= 7.7/ 5.5 [3]
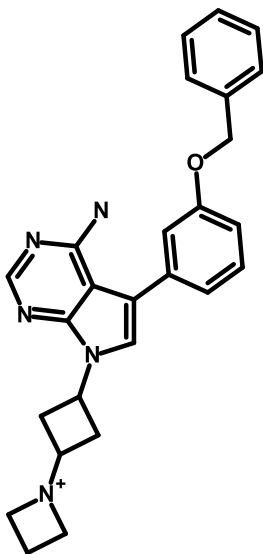
PIC$_{50}$ (nM) = IGF1R/ IR= 7.0/ 5.6 [7]
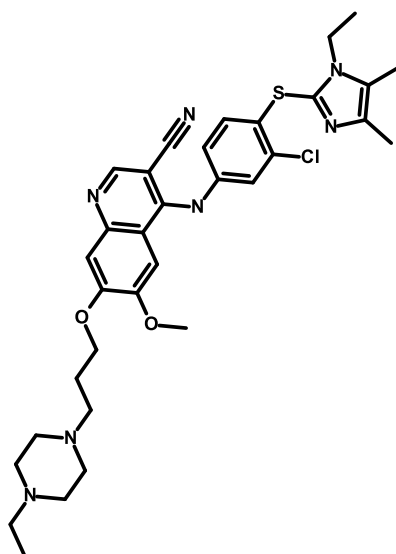
PIC$_{50}$ (nM) = IGF1R/ IR= 8.0/ 8.6 [15]
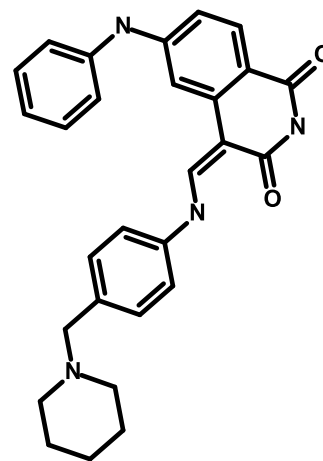
PIC$_{50}$ (nM) = IGF1R/ IR= 6.3/ 6.4 [19]
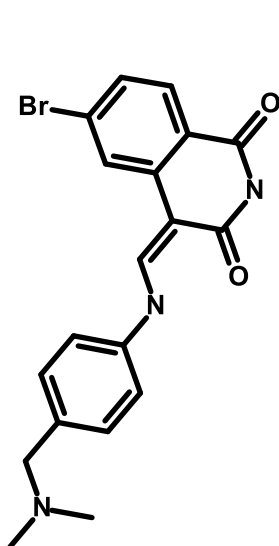
PIC$_{50}$ (nM) = IGF1R/ IR= 5.1/ 5.3 [19]
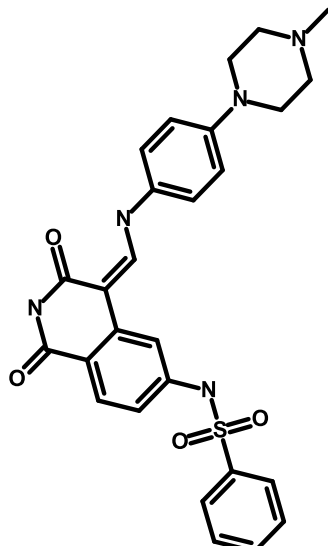
PIC$_{50}$ (nM) = IGF1R/ IR= 6.0/ 6.1[19]
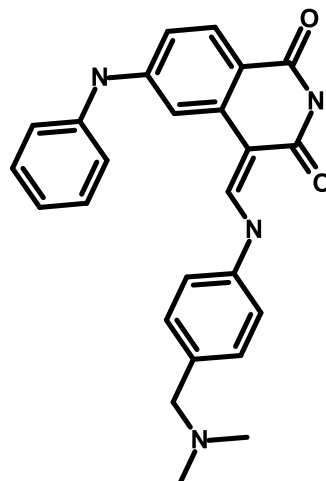
PIC$_{50}$ (nM) = IGF1R/ IR= 6.0/ 6.2 [19]
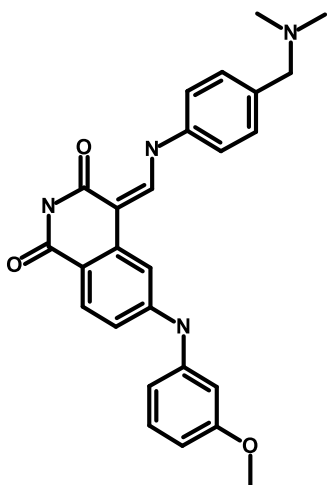
PIC$_{50}$ (nM) = IGF1R/ IR= 5.8/ 6.1 [19]
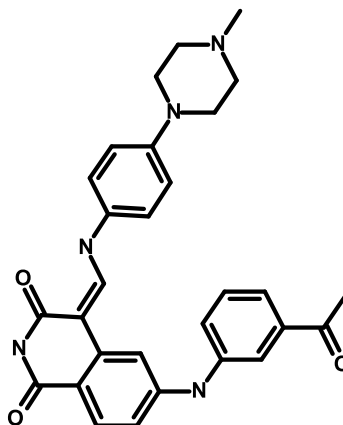
PIC$_{50}$ (nM) = IGF1R/ IR= 6.4/ 6.2 [19]
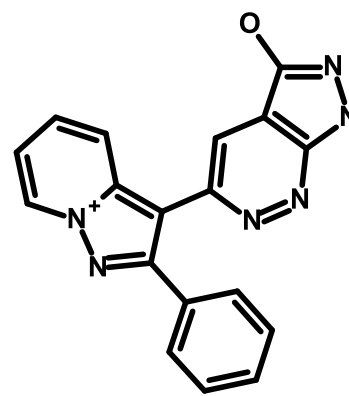
PIC$_{50}$ (nM) = IGF1R/ IR= 4.9/ 5.7 [18]
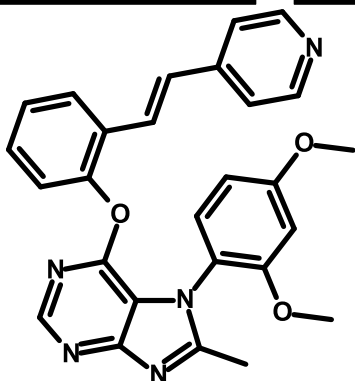
PIC$_{50}$ (nM) = IGF1R/ IR= 6.9/ 5.5 [10]
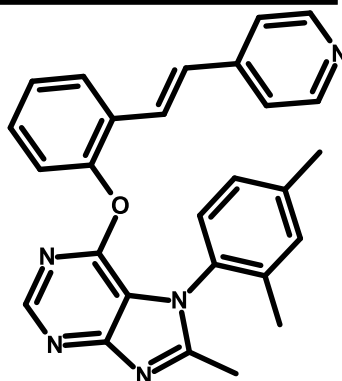
PIC$_{50}$ (nM) = IGF1R/ IR= 6.6/ 5.4 [10]
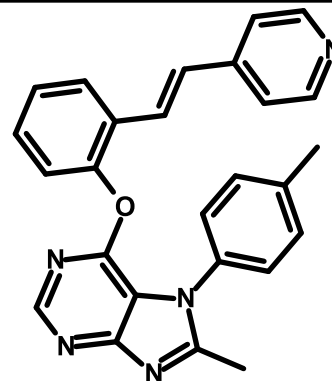
PIC$_{50}$ (nM) = IGF1R/ IR= 7.4/ 5.5 [10]
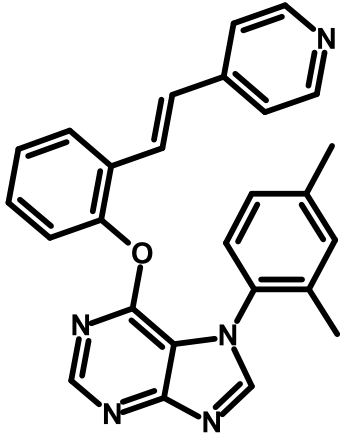
PIC$_{50}$ (nM) = IGF1R/ IR= 7.3/ 5.5 [10]
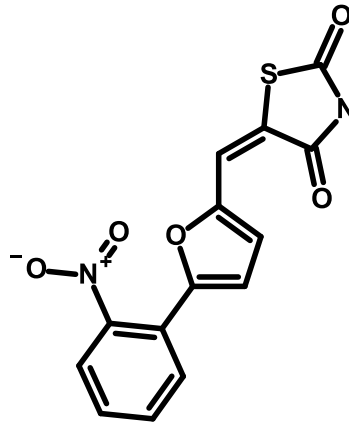
PIC$_{50}$ (nM) = IGF1R/ IR= 5.7/ 4.4 [2]
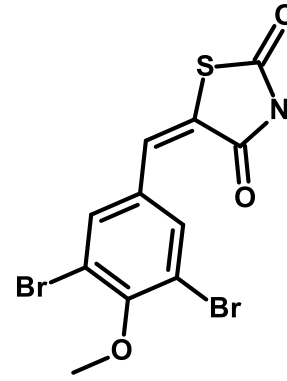
PIC$_{50}$ (nM) = IGF1R/ IR= 5.1/ 4.1 [2]
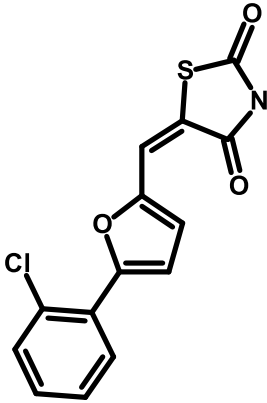
PIC$_{50}$ (nM) = IGF1R/ IR= 5.0/ 4.0 [2]
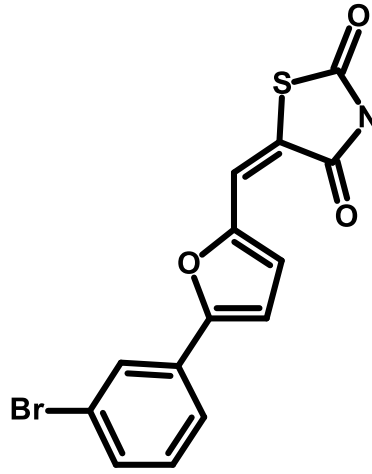
PIC$_{50}$ (nM) = IGF1R/ IR= 7.2/ 5.5 [2]
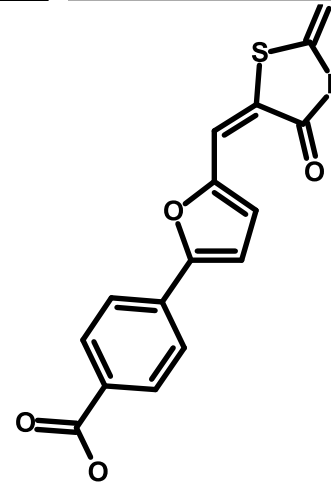
PIC$_{50}$ (nM) = IGF1R/ IR= 7.2/ 4.7 [2]
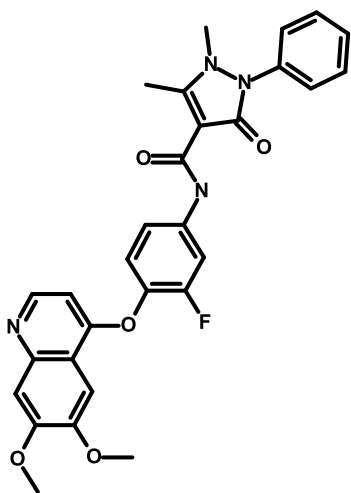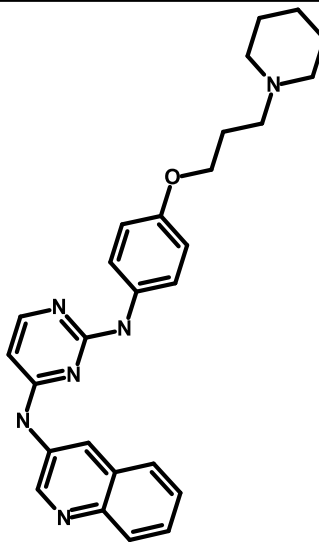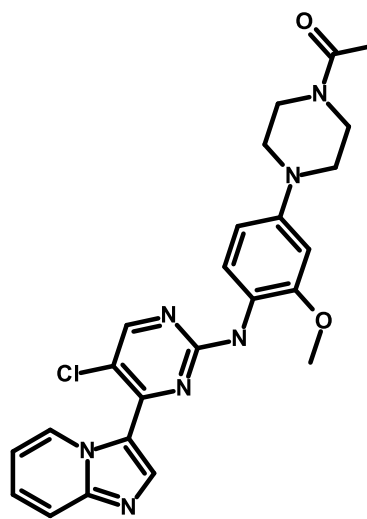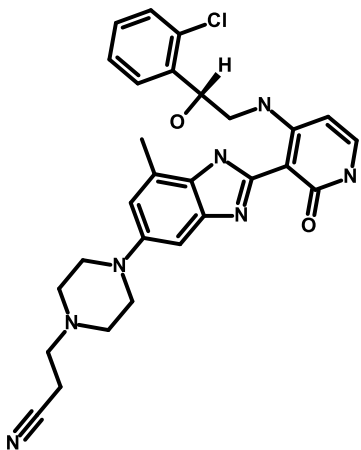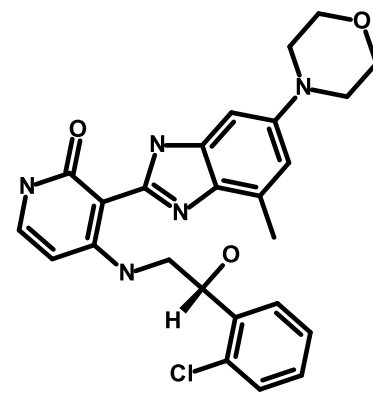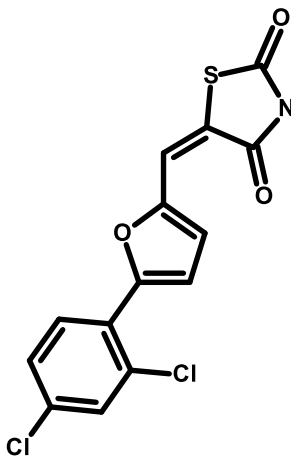
PIC$_{50}$ (nM) = IGF1R/ IR= 5.7/ 4.5 [2]
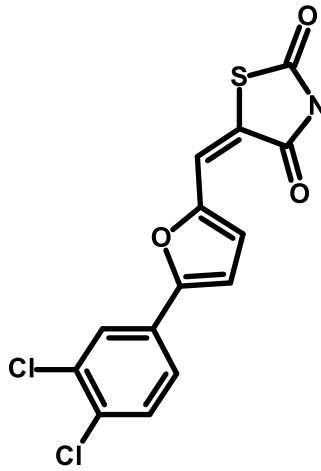
PIC$_{50}$ (nM) = IGF1R/ IR= 5.5/ 4.2 [2]
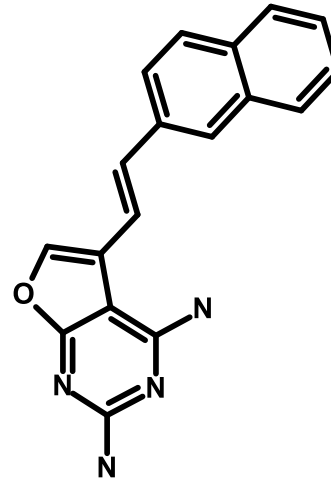
PIC$_{50}$ (nM) = IGF1R/ IR= 5.1/ 4.6 [5]

27
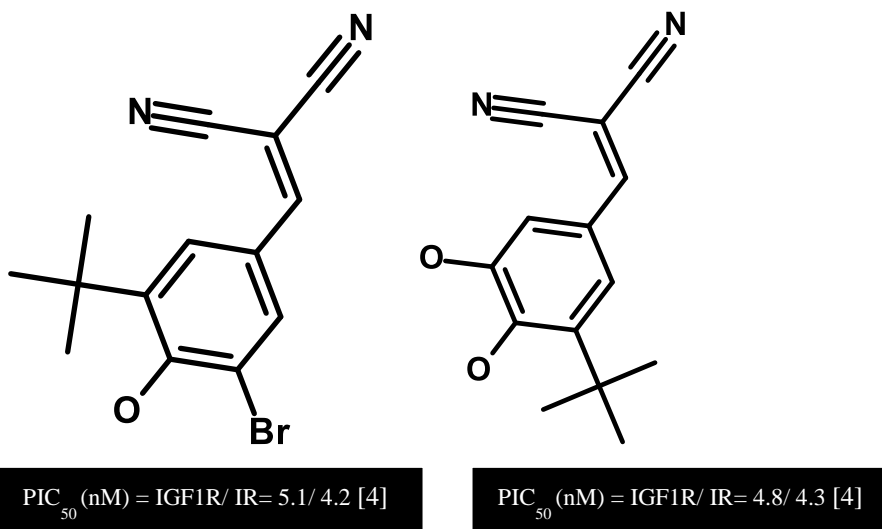
Figure 3: Complete 2D structural Database of dual and selective inhibitors with their –ve natural log of Molar concentrations extracted from literature.

PIC$_{50}$ (nM) = IGF1R/ IR= 5.1/ 4.2 [4]

PIC$_{50}$ (nM) = IGF1R/ IR= 4.8/ 4.3 [4]

2D QSAR

A traditional multivariate analysis known as Hansch analysis [110] was performed using Molecular Operating Environment (MOE 2007 v10). Physiochemical molecular descriptors were computed after building 3D structures and energy minimization of the inhibitors of interest in MOE and the descriptors were short listed based on QSAR-contingency table analysis [111]. The QSAR equation was derived from Multi linear regression (MLR) or Partial Least Square analysis (PLS) to establish the relationship between the descriptors and the biological activities. Biological activities of the test set that was excluded from the model development can be predicted by cross validation of model. This was done by taking q^2 by classical "leave one out" (LOO) method based on. The q^2 derived was built with n-1 compounds and the nth compound was predicted. The q^2 must be greater than 0.5 to be considered as significant.

Another application of the 2D-QSAR was the prediction of important descriptors that contributed in the biological activity of the compounds. "The rule of thumb" was used to

calculate or discretize the descriptors up to 5 per molecule. This was done by removing each parameter according to their relative importance and by visualizing autocorrelation matrix i.e. high interrelationship between parameters produced by MOE as its built-in function.

GRIND Models (3D-QSAR)

The sketched inhibitors in MOE had 3D conformations. From these conformations certain types of 3 dimensional conformations were generated and along with their biological activities they were imported into Pentacle v 1.06 [112]. Then, Molecular Interaction Fields (MIFs) were computed using probes i.e. hydrophobic interactions (DRY), Hydrogen bond donor (O), Hydrogen bond acceptor (N1) as H-bond acceptor and Topology representing the molecular edges (TIP). The maps of 3D interaction energies were made automatically via software between molecule and chemical probes. The MIFs were calculated by replacing each probe iteratively through the GRID and total energy at each node was computed. Total energy at each node is the sum of Lennard-Jones energy (E_{lj}), electrostatic energy (E_{el}), and hydrogen bond energies (E_{hb}) at that point.

$$\mathbf{E_{xyz} = \Sigma E_{lj} + \Sigma E_{el} + \Sigma E_{hb}}$$

Most relevant regions from the calculations of energy nodes were extracted by a built-in algorithm called AMANDA [113]. Default cutoff values for probes were used to discretize the MIFs. Nodes with energy below the cutoff values were discarded. Consistently large auto and cross correlation (CLACC) [113] algorithm was used for encoding the pre-filtered nodes into GRIND thus producing consistent sets of variables. The calculated values were represented via correlograms plots. This correlogram was made by plotting the products of node-node energies

vs. the distances separating the nodes. The auto and cross correlogram had following pairs of nodes: DRY–DRY, O–O, N1–N1 and TIP–TIP as auto-correlograms and DRY–O, DRY–N1, DRY–TIP, O–N1, O–TIP, and N1–TIP as cross-correlograms. A 3D-QSAR model is built against Partial Least Square (PLS) and the assessment of the quality is cross validated by q^2 and Standard error of prediction (SDEP).

Structural comparison

Besides searching for small molecular inhibitors, a 3D receptor structural search was also performed which gave different structures that could be used in the analysis. Specifically human origin structures were chosen in activated (2P, 3P) and inactivated (0P) forms. Following two tables (Table 5 & 6) are the complete list of IGF1R and IR co-crystal structures found. From this crystallographic database of receptors, the structural comparison of the tyrosine kinase domain of IR (Table 1) and IGF1R (Table 2) was done. After matching the pocket residues and loop comparison, it was critical to choose IGF1R and IR structures for analysis based on their geometry and 3D coordinates. For this, these structures were aligned and on the basis of RMSD, that used blossom62 [114] matrix calculation methodology. Moreover, it was also seen whether the Co-crystal structures are recently published and they have higher resolution. On account of above comparison, PDB ID 5HZN of IGF1R [96] and PDB ID 5ES1 of IR [115] were selected.

3D conformational analysis

All the molecules from the database had no energy minimized conformations for the reason that they were either downloaded from the web portals or built using MOE software. Number of techniques was used to produce native conformations against database entries. The intent was to find out the best fit, optimal, stable and reliable conformation for hypothesis. It was understood

that these conformations could interact with macromolecules under observation with optimal binding affinity. Different algorithms for the production of native poses were used which could be found as following:

- **Stochastic search -** The conformational import was done via MOE's inherent application. Molecular force field function deployed in MOE as MMFF94 was used for energy minimization. Fragment based methodology was also used where molecules were divided into overlapping fragments whose conformations were considered independent before the complete assembly of the molecule. This process included the initialization of bond rotations and random inversion of all the chiral centers following bond rotations to randomized dihedral angles (possibly multiples of 30° or 60°). Later, perturbation of all the atomic positions was done. The Cartesian coordinates were energy minimized at the end [116].

- **Extended conformations** -Concerning the degree of molecular freedom or extension was optional to achieve geometrically or extended conformation which had been observed to impact on biological activity[117]. An online web based portal "CORINA" [118] was used for the generation of extended conformations. The database was processed to produce such poses rather than compact frames. Input was given in the form of smiles. By default it was robust and handled the stereo chemical information of the molecules to generate low energy conformations via neutralization of formal charges, orientation of the 3D structures with reference to their moment of inertia and removal of counter ions in salts if any.

- **Docked poses** - Docking runs were performed against IGF1R protein for the generation of poses that could be considered as bioactive conformations. First, the co-crystal ligand

was re-docked into the receptor with different combinations of placement and scoring functions. Later, finalized pair of scoring function and placement method was used to dock the complete database. This resulted in formation of large databases of poses for each ligand.

- **Energy Minimization**- An MOE based application for energy minimization of database was performed which used molecular mechanics and force field based algorithms. This application calculated atomic coordinates that are local minima of the energy function. These dynamic simulations and vibrational analysis were done using the supported energy function in the package i.e. AMBER, CHARMM, MMFF94, etc. [119].

- **Complex minimized poses**- Each finalized conformation from the docked poses was then energy minimized taking receptor surroundings into account using the same force field function as it was used in "Energy Minimization" methodology.

Pharmacophore Modeling

The Database of dual as well as selective inhibitors was imported into MOE and a stochastic conformational search was executed to obtain energy minimized conformations which could be said to be stable for molecular interactions. It was considered that these conformations were bioactive conformations. These conformations were packed and merged with the original database. Finally, the comprehensive database was used to build Pharmacophore [120, 121]. The features were built against abstract conformations. Typical pharmacophoric features that were used to build the model were:

- Hydrophobic centroids
- Aromatic rings

- H-bond acceptor

- H-bond Donor

- Cations &

- Anions

Two types of pharmacophores were built; structure based and ligand based. For ligand based pharmacophore, a ligand with highest potency was chosen. Later, pharmacophore mapping was done by identifying and placing the common binding features deemed to be important for the biological activity. After the identification it was then screened against the packed abstract representation of the conformations to generate "hits" which could be considered as active entities from the distinct bioactive conformations. It was assumed that the pharmacophore holding the most potent ligand's features was able to identify novel compounds that bind to the same site in a similar fashion as the known compounds do. For the sake of model hypothesis validation, the Matthew's correlation coefficient was derived from the True positives (TP), True negative (TN), False positives (FP) and False negatives (FN) hit rates. The threshold set for the quality assessment was 0.5 or greater.

For a 3D structure based, the strategy was to obtain a pharmacophore based on the binding site of the receptor [122]. It was hypothesized that the inhibitors from the database binds in a similar fashion with receptor, then common group (features) could interact with the same protein residue. This information was extracted from the region defined by the automated program. A similar fashion of search for hits and model assessment was performed as it was done on ligand based pharmacophore.

Docking and Pose analysis

Dock application from MOE was used to produce optimal fit configurations of the inhibitors and macromolecular target (IGF-1R & IR). For each inhibitor from the database, a set of conformations were generated and scored in an effort to determine favorable binding mode. Prior to docking the database, a site was required to be found for ligands to bind. For this, A grid representation of the molecular volume by Hendlich *et al.* [123] to locate the binding site for ligand was performed. A built-in application of the MOE known as "Site Finder" based on Alpha Shapes was used to determine the binding pocket automatically[124].

The first choice for docking was to perform runs consecutively with each member of ensemble by using rigid-receptor docking [125-128]. The docking protocol was optimized in such a way that it could be used for different stages which could be integrated into the framework. The dock algorithm [129]automatically generated 3D conformations which seemed to be optimal fit into the pocket with the help of 'Placement method' for ligand placement. Following placement methods were used in the optimization of protocol

- Alpha Triangle

- Triangle Matcher

- Alpha PMI

After the placement of the ligand into the pocket it was then scored having stress on favorable ionic, hydrophobic and H-bond contacts. The variability of the scoring functions as determined by Corbeil *et al.*[130] was considered. The lower the score, the better the pose was considered. Following scoring functions were used

- London dG Scoring

- Affinity dG Scoring

- Alpha HB Scoring

Upon several combinations of placement methods and scoring functions for docking; Alpha PMI and Alpha HB was considered final for further dockings and analyzation. Once Docking runs were performed the generated databases of optimal conformations for bioactive interactions between molecules and receptors were further analyzed via consensus scoring. It was then used for GRIND, QSAR, Pharmacophore and virtual Screening.

Consensus scoring

After the dock application run for database (Table 1) a large conformational space was produced via particular scoring function and placement method. The whole of the information was transferred into excel spread for analysis. It has been reported by Charifson *et al.*[131] that 3-4 scoring functions are sufficient for deducing good results. The poses generated were sorted according to their relative scores (rank-by-number). Different scoring functions were used again for ranking the same conformational space. After all the scoring functions were utilized for the same space they were ranked according to their scores assigned later (rank-by-rank). Finally, the average was taken for each entry and then the ranking was performed again on the basis of average voting for top 10% to produce final rank upon which decision could be made (rank-by-vote).
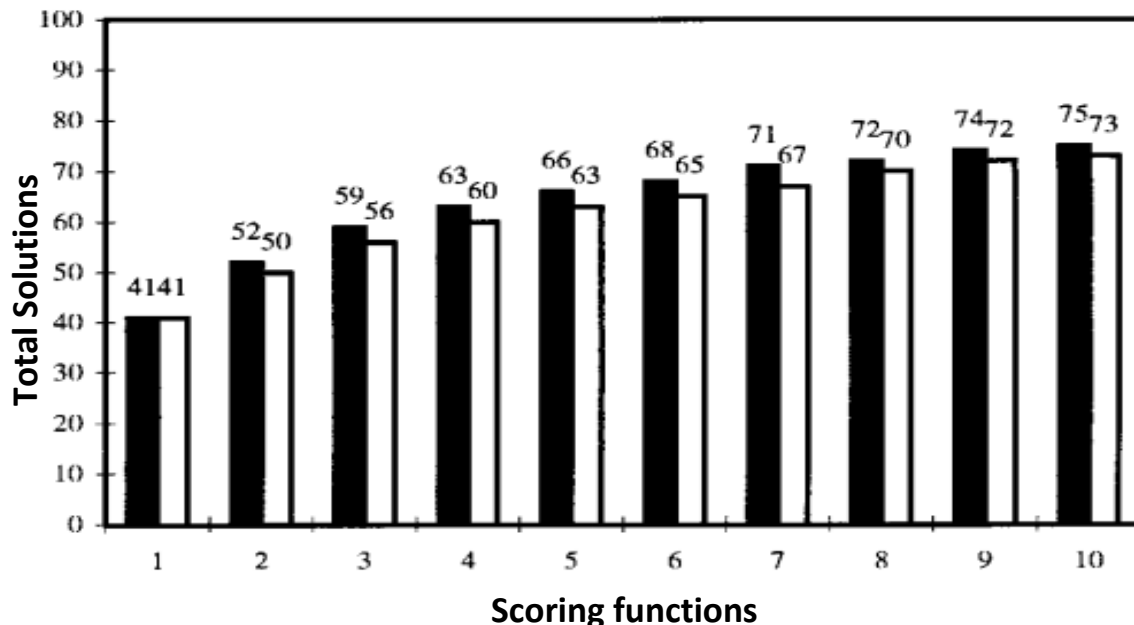
Figure 4: Relationship between scoring functions used in consensus scoring (X-axis) and the total solution (Y-axis) [132]. Solid bars were for rank-by-number and hollow bars were for rank-by-rank strategy.

Following assumptions were taken into account before the rank analysis;

- All the compounds were docked perfectly
- For convenience, it was ruled that the each scoring function had accuracy at the same level for the database.
- Each scoring function was independent to each other.

Virtual screening

A reliable, cost-effective and time saving technique for the discovery of leads and hits generation virtual screening technique was used. It was used more likely in a sense to produce successful clinical candidates.

The Pharmacophore generated was screened against "World Drug Index" [133] and "ChemBridge" [134] databases. This performance was done on the basis of structural based virtual screening. The GUI used was of MOE where software first checked whether the compounds from the databases to be screened had functional groups required by the pharmacophore or not. Following this, the algorithm checked for the 3D spatial arrangements of these compound that matches the query [124].

Results and Discussions

2D-QSAR

In line with previous QSAR studies performed our 2D-QSAR models on smaller data of dual inhibitors showed topological surface area (TPSA) as important descriptor. Nonethless, the statistics was not promising upon which any conclusions could be made. Further analysis of the subject series of dual inhibitors revealed three outliers for each receptor (IGF1R and IR). The Isoquinilinedione compound 33, BMS-754807 and Cyanoquinoline compound 29having tertiary groups, 6-floro-3-pyridilsubstituent and piprazine moiety were responsible for their behaviour as outlier.

A totall of 47 inhibitors were incorporated into MOE v2007.09 as binary input. The dual inhibitors for IGF1R and IR were considered "1" and selective inhibitors for IGF1R were considered "0". Physico-chemical descriptors were calculated against the database. After removing the unneccessary descriptors the overall statistics delineated a_don and vdw_vol asimportant descriptors. The statisctics testified a promising trend towards the acceptance of the model.While the decision based on p- value was also acceptable because with various other descriptors i.e. TPSA, logp, mr, SMR, Energy, etc. the p-value was higher.To further verify the model it was tested via classical Leave one out (LOO) method, which demonstrated a drastic increase in the p-value. This study was in line with pharmacophoric model [2] which stated two H-bond donor play vital role in the interactions.

Table 3: Binary model of IGF1R and IR obtained from Hansch analysis

Descriptors	Total Accuracy	Validated Accuracy	P-value	Validated P-value
• a_don • vdw_vol	0.808	0.765	0.00002	0.0003

Structural comparison

The co-crystal structure of human origin Insulin like growth factor receptor-1 (IGF1R) and Insulin receptors (IR) were compared for structural comparison. Several structures (table 1 & 2) were aligned togeather to gain insight about the sequence similarities and structural deviations.

Table 4: Comparison of co-crystal structures of IGF1R and IR

IR	IGF1R	RMSD	Chain similarity [%]
4IBM [18]	3D94 [102]	3.535	77.9/76.5
4IBM [18]	3I81 [101]	2.575	77.4/ 77.9
4IBM [18]	3WL0 [99]	1.824	77.4/ 76.9
4IBM [18]	4D2R [97]	2.539	78.3/79.9
5ES1 [85]	3I81 [101]	2.183	75.7/78.9
5ES1 [85]	3WL0 [99]	2.036	75.7/77.8
5HHW [96]	3I81 [101]	2.388	79.4/76.8
5HHW [96]	3WL0 [99]	2.065	80.5/76.8

On the basis of above mentioned comparison of IGF1R and IR chains, 4IBM (IR) and 3WL0 (IGF1R) were chosen finally. This selection was made to opt the docking protocol that best suits the co-crystal structures of both receptors. Further, refinement was obtained via

comparison of similar co-crystal structures of IGF1R. This comparison prioritized the utilization of 5HZN for the docking protocol when compared with 3WL0 which was previously selected.

Table 5: Comparison of co-crystal structures of IGF1R

IGF1R	IGF1R	RMSD	Chain similarity [%]
4D2R [97]	**3WL0 [99]**	2.045	98.3/95.7
4D2R [97]	**5HZN [96]**	4.075	95.9/94.7
3WL0 [99]	**5HZN [96]**	3.618	95.9/97.3

Docking/ pose selection

Initially docking was performed on a smaller dataset of 33 dual inhibitors. The 3D conformations of inhibitors were intially minimized by calculating atomic coordinates that are local minima of the energy function using molecular mechnics and forcefield based algorithms [119].The pdb structure 5HZN (IGF1R) used as atarget was also minimized in a similar fashion.

Table 6: Placement methods and Scoring functions combinations used in docking

S.No	Placement Method	Scoring Fn.	E_Score	RMSD (Å)	Rank	Residues involved
1.	α PMI	Affinity dG	-5.0194	2.0704	1	**Lys 1030**
2.	α PMI	α HB (Chosen)	-116.528	2.3095	6	**Lys 1030, Phe 1151, Glu 1077**
3.	α PMI	London dG	-15.0063	2.3095	1	**Lys 1030, Phe 1151, Glu 1077**
4.	α Triangle	Affinity dG	-3.7146	2.4203	1	**Lys 1030**
5.	α Triangle	α HB	-133.5868	1.4658	3	**Lys 1030**
6.	α Triangle	London dG	-16.1993	2.0046	56	**Lys 1030, Phe 1151, Gly 1152**

7.	Triangular matcher	Affinity dG	-3.1628	2.1076	7	**Lys 1030**
8.	Triangular matcher	α HB	-113.7775	2.6610	5	**Lys 1030, Glu 1077**
9.	Triangular matcher	London dG	-112.6802, -67.2396	2.2002, 2.4583	1, 41	**Lys 1030, Arg 1136---Lys 1030**

Dock application[129] from MOE was used to produce optimal fit configurations of co-crystallized inhibitor (NVP-AEW541) and macromolecular target (5HZN). A set of conformations were generated and scored in an effort to determine favorable binding mode. Three different placement methods and scoring functions were utilized in an effort to produce optimal fit that best matches with the co-crystallized conformation and also the

interactions. The co-crystal information (Fig 4) was matched against the final pose obtained via docking (fig 5) of the similar compound. In addition to Lys 1030, new interactions i.e. Phe 1151, Glu 1077forming hydrophobic interactions were found. Almost 900 conformations with different placement methods i.e. Alpha Triangle, Triangle Matcher, Alpha PMI, were generated which were found to be optimally fit into the pocket of the receptor. Later, these conformations were scored via scoring function algorithm based on favorable ionic, hydrophobic and H-bond contacts and the variability was considered [130]. Poses having lower scores were considered for further analysis.
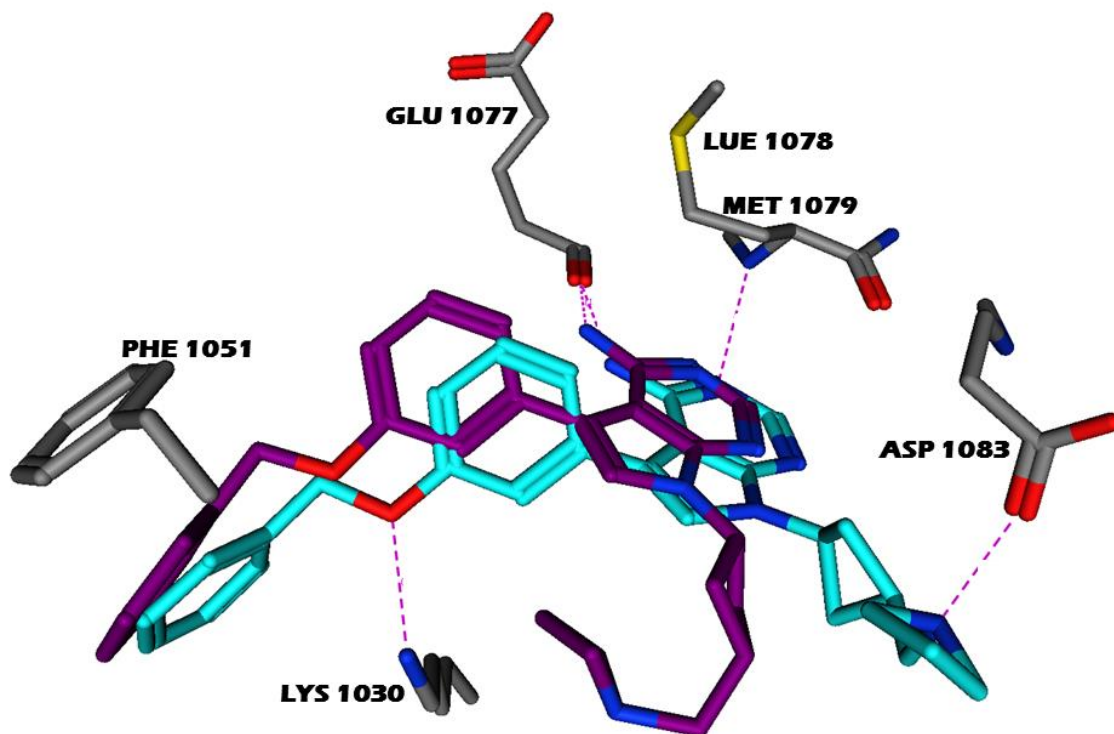
Figure 5: *Co-crystal (Blue) and docked (Purple) conformation interaction pattern*

Upon several combinations of placement methods and scoring functions for docking; Alpha PMI and Alpha HB was considered final for further dockings and analyzation.

This combination was used to perform docking runs for the database of dual inhibitors. The pose generation threshold for each inhibitor was set to be 100. The generated databases of optimal conformations for bioactive interactions between molecules and receptors were further analyzed via consensus scoring. The product of docking runs was rescored in order to extract the information of the best bioactive conformation. It was then used for GRIND, QSAR, Pharmacophore and virtual Screening.

Consensus Scoring

An excel spread analysis was performed on the solutions produced by the docking optimization. It was already published approach to use 3-4 scoring functions. Ranking was

performed according to the relative scores via rank-by- number approach. Later, rank-by-rank approach is used for the second rescoring function [131]. Finally, the ranking on top 10% was performed via rank-by-vote for each entry based on the consensus produced by the average taken against each entry.
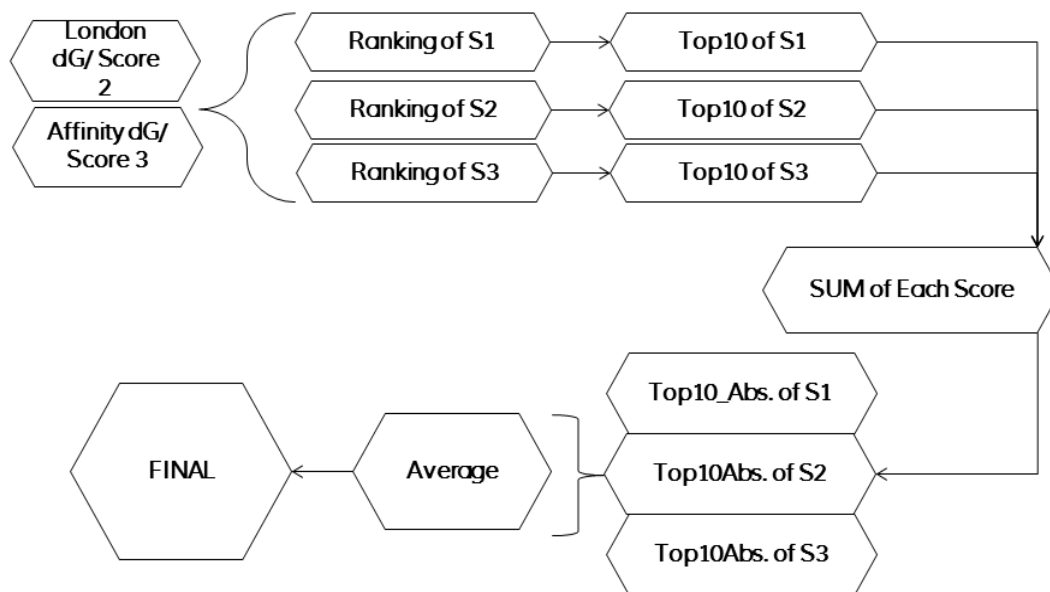
Figure 6: Consensus scheme used to produce best solutions.

S1*= scoring scheme named α-HB used previously while redocking the co-crystal ligand.

S2*= London dG.

S3*= Affinity dG.

A total of 1187 poses were generated for the curated database by assigning threshold of 100 poses per ligand. They were then ranked according to their least energy. Top 10 poses were given "1" and "0" was assigned to rest of solutions. A total sum of top 10 solutions was taken and poses having rank by vote as 3/3 were considered most suitable. A consensus was built upon average of the poses having maximum votes and least rank. These solutions obtained from the consensus scoring were then used for further analysis i.e. Pharmacophore and GRIND.

3D-QSAR

The predictive discernment of 3D-QSAR was further utilized to gain a better understanding of the database behaviour. Classical GRIND models were build for this purpose which initially used extended 3D conformations generated by CORINA CLASSIC [118]. The conformations were exported in package Pentacle V1.05 along with their biological activities against IGF1R and IR. Two separate models were executed after the calcualtion of descriptors, descretization and encoding for analysis. The statistical model built against IGF1R delineated non-consistency in correlogram peaks. However, after running one cycle of fractional factorial design (FFD), the model statistics improved. The correllogram encoded demonstrated the possitive contribution of two variables of N1-N1 (H-bond acceptor) at a distance of 13.6 – 14.0 Å and another set of the same variable with relatively lower importance showed the distance of the same variables at 10.8 – 11.2 Å which is a closer distance as compared to the first peak. This shows that the two H-bond acceptor should be present at the said distance for the inhbitory effects of IGF1R inhibitors. On the other hand, H-bond acceptor descriptor as reported eearlier have shown biased involvement in the activity. It was acceptable to state that these variables present at smaller distance i.e 6.00- 6.40 have the strongest peak and demonstrated its negative effect towards activity. Contrary to it, if similar variables found at larger distance i.e 13.60- 14.00, their role could be taken as postive towards their inhibitory effect towards IGF1R.While, hydrophobic and H-bond acceptor variables such as; DRY- N1 at a distance of 2.00– 2.40 Å have also indisputeably downright contribution towards the inhibitory effects of the database used. Other variables i.e. TIP- TIP, DRY- O, O- N1, O- TIP, N1- TIP, the correlogram plots have shown the inconsistency of the models built against these variables. For this reason these variables are not used in building any definite conclusions. In conjunction with IR model build

with similar descriptors has shown improved statistics without FFD. The draw back of the model was its inconsistency of correlograms.

Energy minimized conformations produced via MOE V2007.09 were used. Two separate models were executed in the similar fashion. Initital statistics for IGF1R were bad enough to decide upon any descriptor. For this reason two cycles of FFD were ran and correlogram correlation was observed. N1- N1 and DRY- N1 variables have shown positive contributions towards the inhibitory effects at the distances of 11.20- 11.60 Å and 2.00- 2.40 Å respectively. On the other hand, IR model has demonstrated promising statistics. The correlogram peaks only made an impression of N1- N1 variablesthat were able enough to differentiate the data of actives and inactives. While rest of the variables have shown inconsistant contribution in the results.

While, the model executed on docking solutions of 33 inihibitors showed the worst correllograms and statistcis. No two variables at a certain distance could be found to be important for the inhibitory effect or contributed negavtively towards their activity. Although, after one FFD, the statistics improved but the overall fashion of the correlogram conistency remained worse. Except the O-O peak which was able to differentiate the data at a distance of 13.20- 13.60 Å. In conjunction with IGF1R, the IR model showed similar statistics and correlogram peaks with no impact on the activity. When complex minimized inhibitors were placed in the model, no significant results could be demonstrated for both receptors.

Lastly, the inhibitors having binary assignation were imported into Pentacle V1.05. The proportioned correlograms built against the data were true enough to not only differentiate between actives and inactive but also between the negative and positive contribution in the data by different variables. After running two cycles of FFD the statisctics were improved and the correolgram having inconsistant variables were removed. The results produced were

interestingly consistant as well as inconsistant with previous models executed along with biological IC_{50} values. A set of four new variables H-bond donors O- O (Fig 8 (1)) and N1-N1 (Fig 8 (2)), at a distance of 16.80- 17.20 Å and 12.8- 13.20 Å respectively, have completely differentiated the inhibitors as selective (0) and dual (1). The consistency with previous models was taken into account by variable sets of DRY- N1 which happened to be present at the distance of 18.00- 19.20 Å. While O-N1 (Fig 8 (3)) and O- Tip (Fig 8 (4)) at distacne of 18.4- 18.8 were found concurrently, have also contributed positively towards the inhibitory trends.
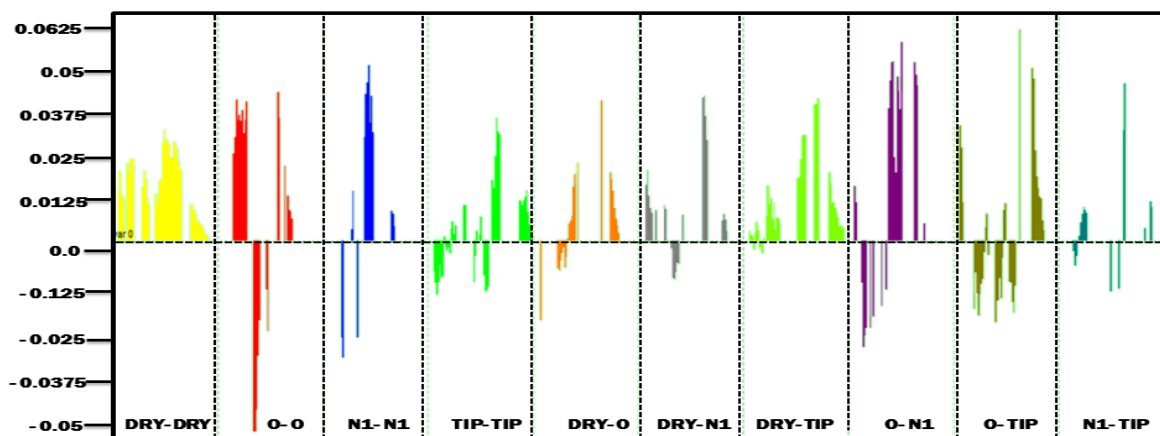
Figure 7: Binary_extended Correlogram delineating important descriptors

Anyhow,The overall trend seen for H-bonding relied on the electronegativoity of the atoms. Atoms i.e. Sulpher, Chlorine, Boron, Flourine have shown least inclination towards H- bonding except the fact that flourine is most electronegative. In copmarative analysis with previouse published reports [64][65] [67, 68][69] our preliminary models remained consistent with electrostatic potential and hydrophobicity except steric bulk. More over the final model has revealed steric bulk along with H-bond acceptor; N1- Tip to play a pivotal role (distance). Along with steric and eletrostatatic descriptor, two sets of O-N1 and N1- N1 variables imparted their presence as obligatory descriptors at the distance of 14.80- 15.20 Å and 12.40- 14.00 Å.
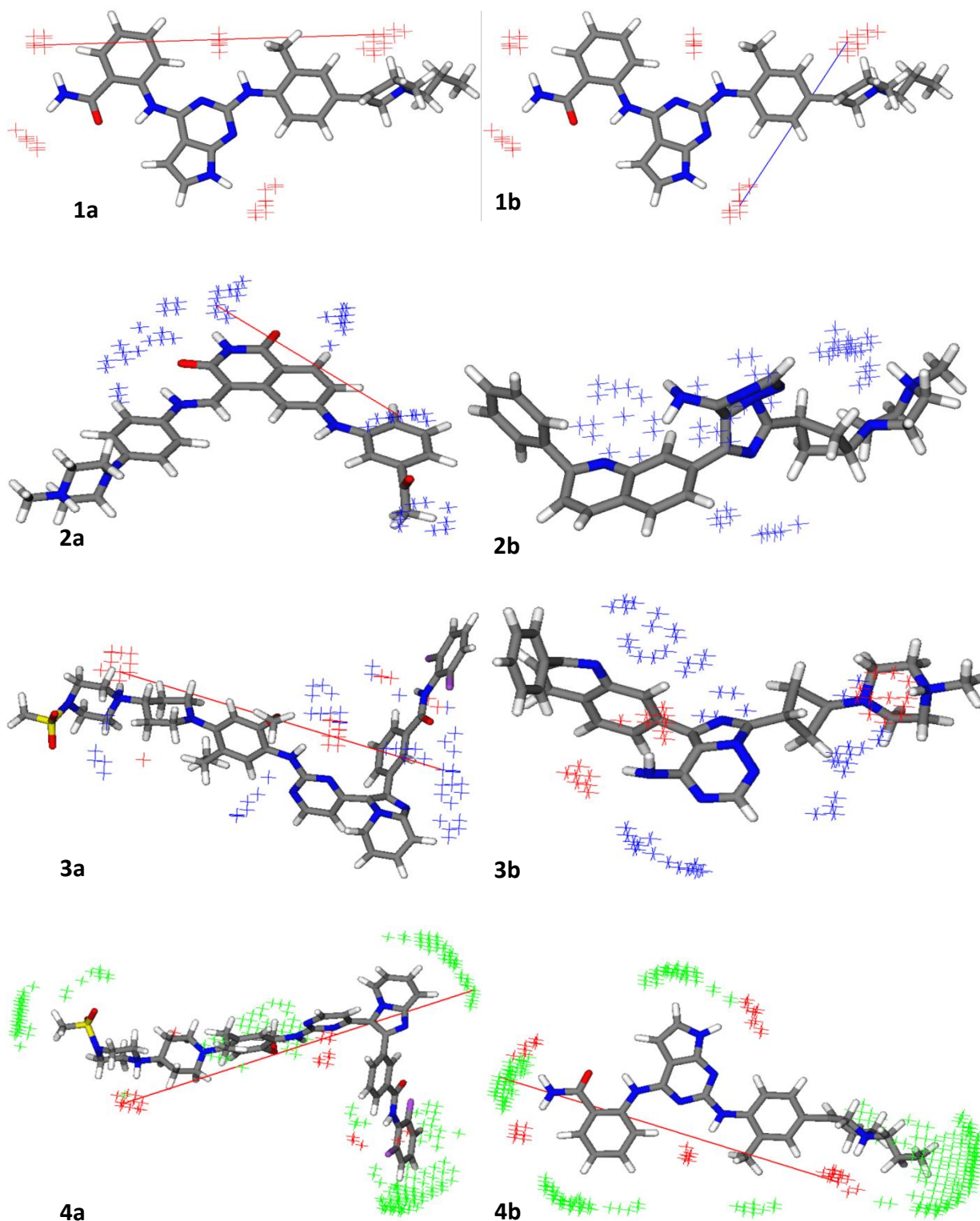
Figure 8: 1a) Distance between H-bond donors (red contours) 1b) lesser distance of 10 found between inactive compounds 2a) Distance between two H-bond acceptors (blue contours) 2b) no distance found in inactive compounds 3) Distance between H- bond donor (red) and acceptor (blue) 4) Optimal distance between H-bond donor (red) and steric bulk (green).

Table 7: GRIND Models against IGF1R and IR

Conformations	IGF1R			IR		
	R2	Q2	FFD/ LV	R2	Q2	FFD/ LV
Extended	0.74	0.60	1/ 2	0.98	0.73	0/ 5
Energy minimized	0.73	0.58	2/ 2	0.84	0.61	0/3
Docked Solutions (33)	0.86	0.62	1/3	0.93	0.62	1/3
Complex Minimized (33)	0.92	0.58	2/3	0.93	0.60	1/3
Extended-Binary	**R2= 0.760, Q2= 0.60, FFD/LV= 2/ 2**					

Pharmacophore Modeling

The database devised of initial 33 dual inhibitors of IGF1R and IR were exhaustively searched for their native conformation via FF function of MOE V2007.09. These conformations were packed and merged with their respective original database. A small molecular inhibitor (NVP-AEW541) was used as a reference inhibitor to build a pharmacophore. The reason for chosing NVP-AEW541 was its co-crystal structural importance that demonstrated several key interactions.Typical pharmacophoric features were built against abstract conformations i.e. hydrophobic centroids, Aromatic rings, H-bond acceptor, H-bond Donor, Cations & Anions.
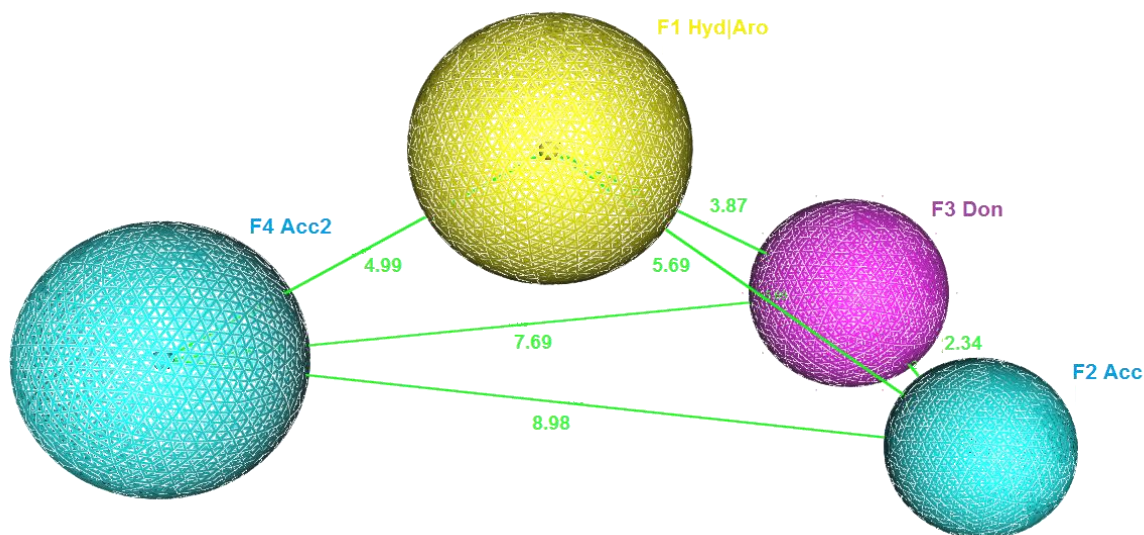
Figure 9: Pharmacophore model built against 33 dual inhibitors

Below are the mean distances of each descriptor from its center to other descriptors

Table 8: Pharmacophore built against stochastic searched conformations

Descriptors	Distances
Hyd|Aro- Don	3.87 Å
Hyd|Aro- Acc	2.34 Å
Hyd|Aro- Acc2	5.69 Å
Acc- Acc2	8.98 Å
Acc- Don	2.34 Å
Don- Acc2	5.69 Å

The model contained two H-bond acceptors with the radius of 0.9 Å and 1.6 Å and two HYD|ARO and H-bond donor with a radius of 1.4 Å and 1.3 Å respectively. Whenthis pharmacophore model was screened against the abstract information obtained from the stochastic

search with the threshold of 150 nM inhibitory concenteration, it efficiently marked the most potent inhibitors. The hits were then used to build a GRIND model which was explained in the GRIND section (table 5). Despite of having a good Mathew's correlation factor of 0.68, it was clear from the analysis that this model was failed to produce any reliable denouement. Once it was confirmed that the pharmcophore model is reliably efficient to differentiate between actives and inactives, the extended database (fig. 3), defined else where in the text was then treated to build pharmacophore. Alike prameters and protocol was exercised for the extended database to obtain the abstract information. Several models were built against the treated information. The best representative among several was chosen on the basis of Mathew's correlation factor 0.872, procuring 21 true postives (TP) and only 2 false positives (FP). When the data was assigned binary numbers: "1" for dual and "0" for selective, the denouement obtained was in similar fashion as it was with data having bioactivities against each entry. The abstract conformation was unpacked later to procure energy minimized conformation to utilize them for GRIND models (table 7). Previous published reports [2, 63, 82] have completely matched with the present pharmacophore model irrespective of the numbers of the pharmacophoric features. Unlike published reports our model failed to correspond to negative ionizeable feature.
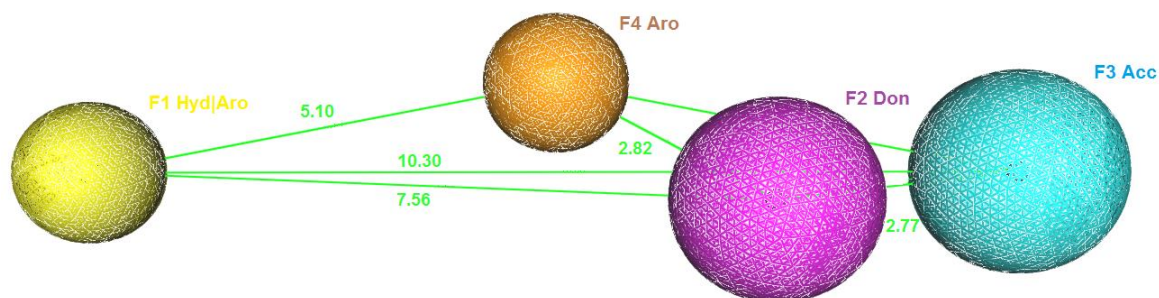
Figure 10 :*Pharmacophore built against extended database*

Below are the mean distances of each descriptor from its center to other descriptors

Table 9: Distances of Descriptors (Binary_extended)

Descriptors	Distances
Hyd|Aro- Don	5.10 Å
Hyd|Aro- Acc	10.30 Å
Hyd|Aro- Aro	7.56 Å
Aro- Acc	5.31 Å
Aro- Don	2.82 Å
Don- Acc	2.77 Å

Virtual screening

A reliable, cost-effective and time saving technique for the discovery of leads and hits, virtual screening technique was used. It was used more likely in a sense to produce successful clinical candidates.

The final Pharmacophore generated was screened against World Drug Index (WDI) and ChemBridge databases [133, 134] having several hundreds of thousands entries each. This performance was done on the basis of structural based virtual screening. MOE pharmacophoric search query was devised in such a way that it screened against both libraries. Initially the both libraries were exhaustively searched for the native conformations of each entry via conformational import algorithm of the MOE package. Millions of conformations generated were packed against both libraries. Similar to already explained pharmacophore query search, an exhaustive search to map functional features and 3D spatial arrangements of the pharmacophore against the libraries was performed. Interestingly, both libraries gave different number of hits i.e. 725 hits were extracted from World Drug Index and 19773 from ChemBridge data base.

The filtered hits were then subjected to online database of CYP filters which further reduced the number of hits i.e. 143 for WDI and 1540 for ChemBridge. To further reduce the number of hits, drug like descriptor comparison was executed on the filtered results. This further narrowed down the list of hits. A consensus was built on excel spread sheet for final entities which finally, revealed a handful of compounds. Eventually, these compounds were taken as external set for GRIND model and validated across internal set upon which the first reliable pharmacophore and GRIND models were built. This was the last nail in the coffin, producing the final settlement of the compounds i.e. 18 for WDI (table 10) and 37 (table 11) for Chembridge. The compounds predicted near "0" were taken as selective inhibitors for IGF1R while those which were near "1" were taken as dual inhibitors (table 8) which could be further validated via MD simulations and experimental validations. The column of LV 2 shows the predicted values of the WDI drugs against their inhibitors. All the entities in this table are above 0.5 which are considered to be active against dual receptors as they are predicted near one. For our comparison, to tighten the criteria, we have increased the threshold to 0.60 and scrutinized the results which revealed only 4 FDA approved drugs. Among them DB00775, DB01297 and DB09075 were under predicted with 0.62300003, 0.63910002 and 0.76560003 respectively, while DB01051 was slightly over-predicted with 1.2187999 (Table 12) at latent variable two upon which binary_extended GRIND model relied.

Table 10: Predicted biological activity values of World Drug Index using final GRIND model

WDI_ID	Smiles	Predicted pIC$_{50}$	FDA Comment	Generic_Name
DB01061	S1[C@H]2N([C@@H](C(O)=O)C1(C)C)C(=O)[C@H]2NC(=O)[C@H](NC(=O)N1CCNC1=O)c1ccccc1	0.50	approved	Azlocillin
DB08108	S(=O)(=O)(N[C@@H](CCC(O)=O)C(O)=O)c1cc2c(cc(OCc3ccc(cc3)C#N)cc2)cc1	0.52	experimental	N-({6-(4-Cyanobenzyl)Oxy]Naphthalen-2-Yl}Sulfonyl)-D-Glutamic Acid
DB00948	S1[C@H]2N([C@@H](C(O)=O)C1(C)C)C(=O)[C@H]2NC(=O)[C@H](NC(=O)N1CCN(S(=O)(=O)C)C1=O)c1ccccc1	0.53	approved	Mezlocillin
DB09042	P(OC[C@@H]1OC(=O)N(C1)c1cc(F)c(cc1)-c1ccc(nc1)-c1nn(nn1)C)(O)(O)=O	0.54	approved	Tedizolid Phosphate
DB04590	Fc1c(cc(OCC)cc1O[C@@H]1CCOC1)[C@@H](Nc1ccc(cc1)C(N)=N)C(O)=O	0.54	experimental	(2r)-({4-[Amino(Imino)Methyl]Phenyl}Amino){5-Ethoxy-2-Fluoro-3-[(3r)-Tetrahydrofuran-3-Yloxy]Phenyl}Aceticacid
DB09073	O=C1N(c2nc(ncc2C(C)=C1C(=O)C)Nc1ncc(N2CCN	0.55	approved	Palbociclib

	CC2)cc1)C1CCCC1			
DB07132	O=C1N=C2C(C=C(NC(=O)N)C=C2)=C1[C@@H](C)c1[nH]ccc1	0.56	experimental	1-{2-Oxo-3-[(1r) -1- (1h-Pyrrol-2-Yl) Ethyl]-2h-Indol-5-Yl}Urea
DB03782	O=C1N=C2C(C=C(NC(=O)N)C=C2)=C1[C@@H](C)c1[nH]ccc1	0.57	experimental	N-(1-Adamantyl)-N'-(4-Guanidinobenzyl) Urea
DB00775	S(=O)(=O)(N[C@@H](Cc1ccc(OCCCCC2CCNCC2)cc1)C(O)=O)CCCC	0.62	**approved**	**Tirofiban**
DB04644	Clc1cc(Cl)ccc1C(=O)NC(=O)Nc1ccc(OCCCC(O)=O)c(C)c1C	0.63	experimental	4-{4-[3-(2,4-Dichloro-Benzoyl)-Ureido]-2,3-Dimethyl-Phenoxy}-Butyric Acid
DB03702	FC(F)(F)C(=O)[C@H](NC(=O)[C@@H]1N(CCC1)C(=O)[C@H](NC(=O)c1ccc(cc1)C(=O)NCC(O)=O)C(C)C)C(C)C	0.63	experimental	2-[4-[[(S)-1-[[(S-2-[[(Rs)-3,3,3-Trifluoro-1-Isopropyl-2-Oxopropyl]Aminocarbonyl]Pyrrolidin-1-Yl-]Carbonyl]-2-Methylpropyl]Aminocarbonyl]Benzoylamino]Acetic Acid
DB01297	O(CC(O)CNC(C)C)c1ccc(NC(=O)C)cc1	0.63	**approved**	**Practolol**
DB05009	O1[C@H](C(=O)NCC)C(O)C(O)[C@@H]1n1c2nc(nc(N)c2nc1)C#CCC1CCC(CC1)C(OC)=O	0.69	investigational	Bms068645
DB09075	Clc1ccc(nc1)NC(=O)C(=O)N[C@H]1CC[C@@H](C[C@H]1NC(=O)c1sc2c(	0.76	**approved**	**Edoxaban**
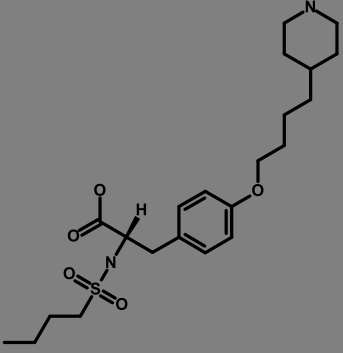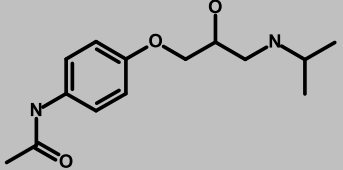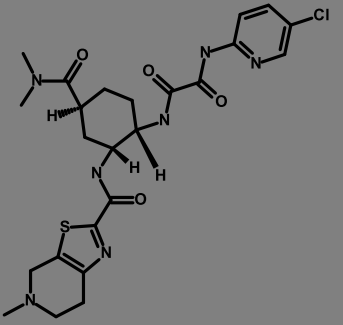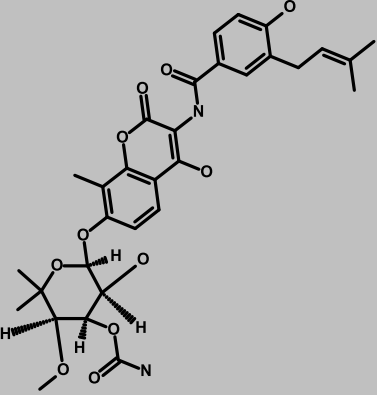
	n1)CCN(C2)C)C(=O)N(C)C			
DB07413	S(C[C@H]1O[C@@H](n2c3ncnc(N)c3nc2)[C@H](O)[C@@H]1O)CCNCCCCCCCCC	0.77	experimental	5'-S-[2-(Decylamino)Ethyl]-5'-Thioadenosine
DB06866	O(C)c1cc(ccc1)-c1cc2C[C@H](O)[C@@H](c2cc1)[C@@H](CCCCC(N)=N)C(O)=O	0.85	experimental	6-Carbamimidoyl-2-[2-Hydroxy-5-(3-Methoxy-Phenyl)-Indan-1-Yl]-Hexanoic Acid
DB07102	O(C)c1cc(ccc1)-c1cc2C[C@H](O)[C@@H](c2cc1)[C@@H](CCCCC(N)=N)C(O)=O	1.01	experimental	(2S)-2-Amino-5-Oxo-5-[(4-Phenylmethoxyphenyl) Amino]Pentanoic Acid
DB01051	O1[C@@H](Oc2ccc3c(OC(=O)C(NC(=O)c4cc(CC=C(C)C)c(O)cc4)=C3O)c2C)[C@@H](O)[C@@H](OC(=O)N)[C@@H](OC)C1(C)C	1.21	**approved**	**Novobiocin**

The compounds having FDA approval were Chosen for further analysis and are considered final hits. They satisfied atomic count, Molecular weight, and other properties for a drug like entity.

Table 11: Predicted biological activity values of ChemBridge database using final GRIND model

S. No.	Mol ID	Smiles	LV2/ Predicted
1.	M1256521	s1c(SCC(=O)NCC(OCC)=O)nnc1SCC(=O)NCC(OCC)=O	0.550
2.	M10240677	S=C1NC(=O)C(=CNc2ccc(cc2)C(OC)=O)C(=O)N1	0.556
3.	M11953929	S(=O)(=O)(Nc1nc(ccn1)C)c1ccc(NC=C2C(=O)NC(=O)NC2=O)cc1	0.803
4.	M3518425	O(C)c1cc(ccc1OC)CCNC(=O)CC(=O)N\N=C\c1ccc(cc1)C(O)=O	0.727
5.	M1069594	O(CC(=O)NNC(=O)c1cc(O)ccc1)c1ccc(OC)cc1	0.647
6.	M17122063	O(C(=O)c1cc(cc([N+](=O)[O-])c1)C(=O)NCCCN(C)C)C	0.503
7.	M14449951	O(C)c1cccc(OC)c1C(=O)Nc1cc(ccc1)C(=O)Nc1ccccc1C(O)=O	0.569
8.	M49568874	O1CCN(CC1)c1cc(NCCN)ccc1[N+](=O)[O-]	0.685
9.	M3600099	O=C1N(C)C(=O)N(c2nc([nH]c12)N\N=C(/C(=O)N)\C#N)C	0.505
10.	M11672011	O(C(=O)c1ccc(NC(=O)C(=O)NCCCN(CC)CC)cc1)C	0.607
11.	M2621576	O(C(=O)c1ccc(NC(=O)C(=O)NCCN(CC)CC)cc1)C	0.887
12.	M8998670	s1c(nnc1SCC(=O)Nc1ccc(cc1)C(OC)=O)N	0.650
13.	M1071130	O(C(=O)c1ccc(NC(=O)C(=O)NCc2ncccc2)cc1)CC	0.565
14.	M2739781	s1c2CCCCc2c2c1N=C(SCCCC(=O)NCCO)NC2=O	0.546
15.	M3498828	O=C1NC(=O)NC=C1NC(=O)c1cc2c(cc1)C(=O)N(CC(C)C)C2=O	0.689
16.	M1968087	s1c2CCCCc2c2c1N=C(SCC(=O)NCCCO)NC2=O	0.514
17.	M14563607	O=C(Nc1cc(NC(=O)C)ccc1)C(=O)Nc1cc(NC(=O)C)ccc1	0.624
18.	M3629374	O(CC(=O)Nc1ccc(cc1)C(=O)Nc1ccccc1C(O)=O)c1cc(C)c(cc1)C	0.655
19.	M3462950	O=C1CC(CC(=O)C1=CNCCN1CCN(CC1)CC(=O)Nc1cc(cc(c1)C)C)(C)C	0.799
20.	M2676403	s1c2CCCCc2c2c1N=C(SCC(=O)NCCCN1CCOCC1)NC2=O	0.789
21.	M84344774	O(CC(O)CN1C(=O)C(NC1=O)(C)C)c1cc(OCC(O)CN2C(=O)C(NC2=O)(C)C)ccc1	0.894
22.	M3518032	Brc1ccccc1C(=O)Nc1cc(ccc1)C(=O)Nc1ccccc1C(O)=O	0.517
23.	M17235919	O(c1cc(C(O)=O)c(cc1)C(O)=O)c1cc(NC(=O)COc2cc([N+](=O)[O-])ccc2)ccc1	0.515
24.	M15408305	O(C)c1ccc(NC(=O)CN2CCN(CC2)CCNC=C2C(=O)CC(CC2=O)(C)C)cc1	0.868
25.	M8537675	s1c2N=C(SCC(=O)NCCCN3CCOCC3)NC(=O)c2c(C)c1C	0.725
26.	M1599654	S(CC(=O)Nc1ccc(cc1)C(OC)=O)c1[nH]ncn1	0.542
27.	M85056799	S1C(CC(=O)Nc2ccc(NC(=O)CC3SC(=O)NC3=O)cc2)C(=O)NC1=O	0.509
28.	M12020562	S(=O)(=O)(NNC(=O)CCC(=O)Nc1ccc(cc1C)C)c1ccc(NC(=O)C)cc1	0.737
29.	M85056885	O(CC(O)CN1C(=O)CCC1=O)c1ccc(OCC(O)CN2C(=O)CCC2=O)cc1	0.511
30.	M1979596	s1c2CCCCc2c2c1N=C(SCC(=O)NCCN1CCOCC1)NC2=O	0.738
31.	M2628799	Clc1cc(ccc1)C(=O)Nc1cc(ccc1)C(=O)Nc1ccccc1C(O)=O	0.621
32.	M85054213	S1\C(\NC(=O)C1CC(O)=O)=N\N=C\c1ccc(OS(=O)(=O)c2ccc(NC(=O)C)cc2)cc1	0.540
33.	M15074693	O1CCN(CC1)CCCNC(=O)C(=O)Nc1ccc(OCC)cc1	0.502
34.	M2857162	S(=O)(=O)(N1CCOCC1)c1ccc(cc1)C(=O)NCCN(C)C	0.538
35.	M11737390	O(CC)c1ccc(NC(=O)C(=O)NCCN(C)C)cc1	0.591
36.	M8212322	S1\C(=C\c2cc(OCC(=O)N)ccc2)\C(=O)N=C1Nc1cc(ccc1)C(O)=O	0.668
37.	M3594593	S(CC(=O)NC=1C(=O)N(N(C)C=1C)c1ccccc1)c1nc2N(C)C(=O)N(C)C(=O)c2n1CC	0.567

Table 12: Finalized Compounds (Hits)for future investigations

Generic Names	2D- Structure	Class	Target	Indication
Tirofiban		phenylpropanoic acids	Integrin β-3 & Integrin α-IIb	Acute Coronary syndrome
Practolol		Acetanilides, Benzenoids	Beta-1 adrenergic receptor	Cardiac Arrhythmia Emergency treatment
Edoxaban		Novel Oral Anti-Coagulants (NOACs) class of drugs	Coagulation Factor X	Stroke risk and systemic embolis in patients with non-valvular atrial fibrillation (NVAF)
Novobiocin		Coumarin glycosides	DNA Gyrase sub Unit B & DNA topoisomerase 1	Infections due to Staphylococci and other susceptible organisms

Conclusions

The Pharmacophore and GRIND models concurrently delineated two H-bond donors, two H-bond acceptor, overall Topology and vdw_vol as important descriptors for dual inhibition of IGF-IR and IR.

Our GRIND model showed two H-bond donor at the distance of 168-17.2 Å and two H-bond acceptors at the distance of 12.8- 13.2 Å respectively. It also predicted steric bulk along with H- bond donor at 18-22 Å. Furthermore, our pharmacophore model delineated the distances between the descriptors like hydrophobic region, H- bond acceptors and donor region to be important. Mathew's correlation factor of our pharmacophore model was 0.87 based on TP hit rate. Since the Mathew's correlation factor is quite acceptable for the pharmacophore, its query against WDI and ChemBridge resulted in potential hits against IGF-IR and IR which belonged to diverse classes. Further experimental validation of identified hits could pave the way towards successful development of chemotherapeutic agents against various cancers.

References:

1.   Ji, Q.-s., et al., *A novel, potent, and selective insulin-like growth factor-I receptor kinase inhibitor blocks insulin-like growth factor-I receptor signaling in vitro and inhibits insulin-like growth factor-I receptor–dependent tumor growth in vivo.* Molecular cancer therapeutics, 2007. **6**(8): p. 2158-2167.

2.   Liu, X., et al., *Discovery and SAR of thiazolidine-2, 4-dione analogues as insulin-like growth factor-1 receptor (IGF-1R) inhibitors via hierarchical virtual screening.* Journal of medicinal chemistry, 2010. **53**(6): p. 2661-2665.

3.   Li, R., A. Pourpak, and S.W. Morris, *Inhibition of the insulin-like growth factor-1 receptor (IGF1R) tyrosine kinase as a novel cancer therapy approach.* Journal of medicinal chemistry, 2009. **52**(16): p. 4981-5004.

4.   Párrizas, M., et al., *Specific inhibition of insulin-like growth factor-1 and insulin receptor tyrosine kinase activity and biological function by tyrphostins.* Endocrinology, 1997. **138**(4): p. 1427-1433.

5.   Li, W., et al., *Inhibition of insulin-like growth factor I receptor autophosphorylation by novel 6-5 ring-fused compounds.* Biochemical pharmacology, 2004. **68**(1): p. 145-154.

6.   Jin, M., et al., *Discovery of an orally efficacious imidazo [5, 1-f][1, 2, 4] triazine dual inhibitor of IGF-1R and IR.* ACS medicinal chemistry letters, 2010. **1**(9): p. 510-515.

7.   García-Echeverría, C., et al., *In vivo antitumor activity of NVP-AEW541—a novel, potent, and selective inhibitor of the IGF-IR kinase.* Cancer cell, 2004. **5**(3): p. 231-239.

8.   Liu, L., et al., *Structure-based design of novel class II c-Met inhibitors: 2. SAR and kinase selectivity profiles of the pyrazolone series.* Journal of medicinal chemistry, 2012. **55**(5): p. 1868-1897.

9.   Ducray, R., et al., *Discovery of novel imidazo [1, 2-a] pyridines as inhibitors of the insulin-like growth factor-1 receptor tyrosine kinase.* Bioorganic & medicinal chemistry letters, 2011. **21**(16): p. 4698-4701.

10. Wood, E.R., et al., *Discovery of an inhibitor of insulin-like growth factor 1 receptor activation: implications for cellular potency and selectivity over insulin receptor.* Biochemical pharmacology, 2009. **78**(12): p. 1438-1447.

11. Mulvihill, M.J., et al., *Novel 2-phenylquinolin-7-yl-derived imidazo [1, 5-a] pyrazines as potent insulin-like growth factor-I receptor (IGF-IR) inhibitors.* Bioorganic & medicinal chemistry, 2008. **16**(3): p. 1359-1375.

12. Buchanan, J.L., et al., *Discovery of 2, 4-bis-arylamino-1, 3-pyrimidines as insulin-like growth factor-1 receptor (IGF-1R) inhibitors.* Bioorganic & medicinal chemistry letters, 2011. **21**(8): p. 2394-2399.

13. Emmitte, K.A., et al., *Discovery and optimization of imidazo [1, 2-a] pyridine inhibitors of insulin-like growth factor-1 receptor (IGF-1R).* Bioorganic & medicinal chemistry letters, 2009. **19**(3): p. 1004-1008.

14. Chamberlain, S.D., et al., *Discovery of 4, 6-bis-anilino-1H-pyrrolo [2, 3-d] pyrimidines: Potent inhibitors of the IGF-1R receptor tyrosine kinase.* Bioorganic & medicinal chemistry letters, 2009. **19**(2): p. 469-473.

15. Miller, L.M., et al., *Lead identification to generate 3-cyanoquinoline inhibitors of insulin-like growth factor receptor (IGF-1R) for potential use in cancer treatment.* Bioorganic & medicinal chemistry letters, 2009. **19**(1): p. 62-66.

16.   Chamberlain, S.D., et al., *Optimization of 4, 6-bis-anilino-1H-pyrrolo [2, 3-d] pyrimidine IGF-1R tyrosine kinase inhibitors towards JNK selectivity.* Bioorganic & medicinal chemistry letters, 2009. **19**(2): p. 360-364.

17.   Wittman, M., et al., *Discovery of a 1 H-Benzoimidazol-2-yl)-1 H-pyridin-2-one (BMS-536924) inhibitor of insulin-like growth factor I receptor kinase with in vivo antitumor activity.* Journal of medicinal chemistry, 2005. **48**(18): p. 5639-5643.

18.   Anastassiadis, T., et al., *A highly selective dual insulin receptor (IR)/insulin-like growth factor 1 receptor (IGF-1R) inhibitor derived from an extracellular signal-regulated kinase (ERK) inhibitor.* Journal of Biological Chemistry, 2013. **288**(39): p. 28068-28077.

19.   Mayer, S.C., et al., *Lead identification to generate isoquinolinedione inhibitors of insulin-like growth factor receptor (IGF-1R) for potential use in cancer treatment.* Bioorganic & medicinal chemistry letters, 2008. **18**(12): p. 3641-3645.

20.   Chamberlain, S.D., et al., *Optimization of a series of 4, 6-bis-anilino-1h-pyrrolo [2, 3-d] pyrimidine inhibitors of igf-1r: elimination of an acid-mediated decomposition pathway.* Bioorganic & medicinal chemistry letters, 2009. **19**(2): p. 373-377.

21.   Marsilje, T.H., et al., *Synthesis, structure–activity relationships, and in vivo efficacy of the novel potent and selective anaplastic lymphoma kinase (ALK) inhibitor 5-Chloro-N 2-(2-isopropoxy-5-methyl-4-(piperidin-4-yl) phenyl)-N 4-(2-(isopropylsulfonyl) phenyl) pyrimidine-2, 4-diamine (LDK378) currently in phase 1 and phase 2 clinical trials.* Journal of medicinal chemistry, 2013. **56**(14): p. 5675-5690.

22.   Clemmons, D.R., *Insulin-like growth factor binding proteins and their role in controlling IGF actions.* Cytokine & growth factor reviews, 1997. **8**(1): p. 45-62.

23.   Jones, J.I. and D.R. Clemmons, *Insulin-Like Growth Factors and Their Binding Proteins: Biological Actions*.* Endocrine reviews, 1995. **16**(1): p. 3-34.

24.   Fernandez, R., et al., *The Drosophila insulin receptor homolog: a gene essential for embryonic development encodes two receptor isoforms with different signaling potential.* The EMBO Journal, 1995. **14**(14): p. 3373.

25.   Torres, A.M., et al., *Solution structure of human insulin-like growthfactor II. Relationship to receptor and binding protein interactions.* Journal of molecular biology, 1995. **248**(2): p. 385-401.

26.   Belfiore, A., et al., *Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease.* Endocrine reviews, 2009. **30**(6): p. 586-623.

27.   Chitnis, M.M., et al., *The type 1 insulin-like growth factor receptor pathway.* Clinical Cancer Research, 2008. **14**(20): p. 6364-6370.

28.   Kjeldsen, T., et al., *The ligand specificities of the insulin receptor and the insulin-like growth factor I receptor reside in different regions of a common binding site.* Proceedings of the National Academy of Sciences, 1991. **88**(10): p. 4404-4408.

29.   McKern, N.M., et al., *Structure of the insulin receptor ectodomain reveals a folded-over conformation.* Nature, 2006. **443**(7108): p. 218-221.

30.   Hubbard, S.R., *Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog.* The EMBO journal, 1997. **16**(18): p. 5572-5581.

31.   Taniguchi, C.M., B. Emanuelli, and C.R. Kahn, *Critical nodes in signalling pathways: insights into insulin action.* Nature reviews Molecular cell biology, 2006. **7**(2): p. 85-96.

32.   Rosen, O.M., et al., *Phosphorylation activates the insulin receptor tyrosine protein kinase.* Proceedings of the National Academy of Sciences, 1983. **80**(11): p. 3237-3240.

33.   Russo, A.A., P.D. Jeffrey, and N.P. Pavletich, *Structural basis of cyclin-dependent kinase activation by phosphorylation.* Nature Structural & Molecular Biology, 1996. **3**(8): p. 696-700.

34.   Werner, H., D. Weinstein, and I. Bentov, *Similarities and differences between insulin and IGF-I: structures, receptors, and signalling pathways.* Archives of physiology and biochemistry, 2008. **114**(1): p. 17-22.

35.   Avruch, J., *Insulin signal transduction through protein kinase cascades*, in *Insulin Action*. 1998, Springer. p. 31-48.

36.   De Meyts, P. and J. Whittaker, *Structural biology of insulin and IGF1 receptors: implications for drug design.* Nature Reviews Drug Discovery, 2002. **1**(10): p. 769-783.

37.   Luo, R.Z.-T., et al., *Quaternary structure of the insulin-insulin receptor complex.* Science, 1999. **285**(5430): p. 1077-1080.

38.   Ottensmeyer, F., et al., *Mechanism of transmembrane signaling: insulin binding and the insulin receptor.* Biochemistry, 2000. **39**(40): p. 12103-12112.

39.   Hubbard, S.R., et al., *Crystal structure of the tyrosine kinase domain of the human insulin receptor.* Nature, 1994. **372**(6508): p. 746-754.

40.   Munshi, S., et al., *Structure of apo, unactivated insulin-like growth factor-1 receptor kinase at 1.5 Å resolution.* Acta Crystallographica Section D: Biological Crystallography, 2003. **59**(10): p. 1725-1730.

41.   Mosthaf, L., et al., *Functionally distinct insulin receptors generated by tissue-specific alternative splicing.* The EMBO journal, 1990. **9**(8): p. 2409.

42.   Buck, E., et al., *Compensatory insulin receptor (IR) activation on inhibition of insulin-like growth factor-1 receptor (IGF-1R): rationale for cotargeting IGF-1R and IR in cancer.* Molecular cancer therapeutics, 2010. **9**(10): p. 2652-2664.

43.   DeChiara, T.M., A. Efstratiadis, and E.J. Robertsen, *A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting.* 1990.

44.   Christofori, G., P. Naik, and D. Hanahan, *A second signal supplied by insulin-like growth factor II in oncogene-induced tumorigenesis.* Nature, 1994. **369**(6479): p. 414-418.

45.   Kaplan, S.A., *Cell receptors.* American Journal of Diseases of Children, 1984. **138**(12): p. 1140-1146.

46.   Yu, H. and T. Rohan, *Role of the insulin-like growth factor family in cancer development and progression.* Journal of the National Cancer Institute, 2000. **92**(18): p. 1472-1489.

47.   Benvenuti, S., et al., *Oncogenic activation of the RAS/RAF signaling pathway impairs the response of metastatic colorectal cancers to anti–epidermal growth factor receptor antibody therapies.* Cancer research, 2007. **67**(6): p. 2643-2648.

48.   Vincent, A.M. and E.L. Feldman, *Control of cell survival by IGF signaling pathways.* Growth hormone & IGF research, 2002. **12**(4): p. 193-197.

49.   Sciacca, L., et al., *In IGF-I receptor-deficient leiomyosarcoma cells autocrine IGF-II induces cell invasion and protection from apoptosis via the insulin receptor isoform A.* Oncogene, 2002. **21**(54): p. 8240-8250.

50.   Pollak, M., *The insulin and insulin-like growth factor receptor family in neoplasia: an update.* Nature Reviews Cancer, 2012. **12**(3): p. 159-169.

51.   Denley, A., et al., *Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR.* Molecular Endocrinology, 2004. **18**(10): p. 2502-2512.

52.   Nemecek, C., et al., *Design of Potent IGF1-R Inhibitors Related to Bis-azaindoles.* Chemical biology & drug design, 2010. **76**(2): p. 100-106.

53.   Liu, Y. and N.S. Gray, *Rational design of inhibitors that bind to inactive kinase conformations.* Nature chemical biology, 2006. **2**(7): p. 358-364.

54.   Frödin, M., et al., *A phosphoserine/threonine-binding pocket in AGC kinases and PDK1 mediates activation by hydrophobic motif phosphorylation.* The EMBO journal, 2002. **21**(20): p. 5396-5407.

55.   Christopoulos, P.F., P. Msaouel, and M. Koutsilieris, *The role of the insulin-like growth factor-1 system in breast cancer.* Molecular cancer, 2015. **14**(1): p. 1.

56.   Denley, A., et al., *The insulin receptor isoform exon 11-(IR-A) in cancer and other diseases: a review.* Hormone and Metabolic Research, 2003. **35**(11/12): p. 778-785.

57.   Huang, F., et al., *IRS2 copy number gain, KRAS and BRAF mutation status as predictive biomarkers for response to the IGF-1R/IR inhibitor BMS-754807 in colorectal cancer cell lines.* Molecular cancer therapeutics, 2015. **14**(2): p. 620-630.

58.   Ullrich, A., et al., *Insulin-like growth factor I receptor primary structure: comparison with insulin receptor suggests structural determinants that define functional specificity.* The EMBO journal, 1986. **5**(10): p. 2503.

59.   Jin, M., et al., *Small-molecule ATP-competitive dual IGF-1R and insulin receptor inhibitors: structural insights, chemical diversity and molecular evolution.* Future medicinal chemistry, 2012. **4**(3): p. 315-328.

60.   Vincent, E.E., et al., *Targeting non-small cell lung cancer cells by dual inhibition of the insulin receptor and the insulin-like growth factor-1 receptor.* PLoS One, 2013. **8**(6): p. e66963.

61.   Yamaguchi, Y., et al., *Ligand-binding properties of the two isoforms of the human insulin receptor.* Endocrinology, 1993. **132**(3): p. 1132-1138.

62.   Ward, C.W. and M.C. Lawrence, *Ligand-induced activation of the insulin receptor: a multi-step process involving structural changes in both the ligand and the receptor.* Bioessays, 2009. **31**(4): p. 422-434.

63.   Taha, M.O., et al., *Discovery of new potent human protein tyrosine phosphatase inhibitors via pharmacophore and QSAR analysis followed by in silico screening.* Journal of Molecular Graphics and Modelling, 2007. **25**(6): p. 870-884.

64.   Muddassar, M., et al., *Receptor guided 3D-QSAR: a useful approach for designing of IGF-1R inhibitors.* BioMed Research International, 2008. **2008**.

65.   Sperandio, O., M. Petitjean, and P. Tufféry, *wwLigCSRre: a 3D ligand-based server for hit identification and optimization.* Nucleic acids research, 2009. **37**(suppl 2): p. W504-W509.

66.   Kim, K.H., G. Greco, and E. Novellino, *A critical review of recent CoMFA applications*, in *3D QSAR in drug design*. 1998, Springer. p. 257-315.

67.   Kurup, A., R. Garg, and C. Hansch, *Comparative QSAR study of tyrosine kinase inhibitors.* Chemical reviews, 2001. **101**(8): p. 2573-2600.

68.   Sheridan, R.P., et al., *QSAR models for predicting the similarity in binding profiles for pairs of protein kinases and the variation of models between experimental data sets.* Journal of chemical information and modeling, 2009. **49**(8): p. 1974-1985.

69.   Li, Y.-S., L. Zhou, and X. Ma, *Molecular docking and 3D QSAR studies of substituted 4-amino-1H-pyrazolo [3, 4-d] pyrimidines as insulin-like growth factor-1 receptor (IGF1R) inhibitors.* Medicinal Chemistry Research, 2012. **21**(10): p. 3301-3311.

70.   Hung, I.-C., et al., *Memory enhancement by traditional Chinese medicine?* Journal of Biomolecular Structure and Dynamics, 2013. **31**(12): p. 1411-1439.

71.   Kurian, L.A., T.A. Silva, and D. Sabatino, *Submonomer synthesis of azapeptide ligands of the Insulin Receptor Tyrosine Kinase domain.* Bioorganic & medicinal chemistry letters, 2014. **24**(17): p. 4176-4180.

72.   Ferrara, P. and E. Jacoby, *Evaluation of the utility of homology models in high throughput docking.* Journal of Molecular Modeling, 2007. **13**(8): p. 897-905.

73.   Zheng, D., et al., *Targeting of the protein interaction site between FAK and IGF-1R.* Biochemical and biophysical research communications, 2009. **388**(2): p. 301-305.

74.   Huang, Z. and C.F. Wong, *Simulation reveals two major docking pathways between the hexapeptide GDYMNM and the catalytic domain of the insulin receptor protein kinase.* Proteins: Structure, Function, and Bioinformatics, 2012. **80**(9): p. 2275-2286.

75.   Abdullahi, A.D., et al., *Application of Group-Based QSAR and Molecular Docking in the Design of Insulin-Like Growth Factor Antagonists.* Tropical Journal of Pharmaceutical Research, 2015. **14**(6): p. 941-951.

76.   Krug, M., et al., *Discovery and selectivity-profiling of 4-benzylamino 1-aza-9-oxafluorene derivatives as lead structures for IGF-1R inhibitors.* Bioorganic & medicinal chemistry letters, 2010. **20**(23): p. 6915-6919.

77.   Engen, W., et al., *Synthesis of aryl-heteroaryl ureas (AHUs) based on 4-aminoquinoline and their evaluation against the insulin-like growth factor receptor (IGF-1R).* Bioorganic & medicinal chemistry, 2010. **18**(16): p. 5995-6005.

78.   Arunkumar, R., et al., *Effect of diallyl disulfide on insulin-like growth factor signaling molecules involved in cell survival and proliferation of human prostate cancer cells in vitro and in silico approach through docking analysis.* Phytomedicine, 2012. **19**(10): p. 912-923.

79.   Mahajanakatti, A.B., et al., *Exploring inhibitory potential of Curcumin against various cancer targets by in silico virtual screening.* Interdisciplinary Sciences: Computational Life Sciences, 2014. **6**(1): p. 13-24.

80.   Singh, P. and F. Bast, *Screening of multi-targeted natural compounds for receptor tyrosine kinases inhibitors and biological evaluation on cancer cell lines, in silico and in vitro.* Medical Oncology, 2015. **32**(9): p. 1-18.

81.   Singh, P. and F. Bast, *Screening and biological evaluation of myricetin as a multiple target inhibitor insulin, epidermal growth factor, and androgen receptor; in silico and in vitro.* Investigational new drugs, 2015. **33**(3): p. 575-593.

82.   Ramdhave, A.S. and M. Nandave, *REDISCOVERY OF IGF-1R TARGETED INHIBITORS FROM AN APPROVED DRUGS DATABASE THROUGH RECEPTOR-BASED PHARMACOPHORE MODELLING AND DOCKING.* Indo American Journal of Pharmaceutical Research, 2016. **6**(1): p. 4295-4301.

83.   Haluska, P., et al., *In vitro and in vivo antitumor effects of the dual insulin-like growth factor-I/insulin receptor inhibitor, BMS-554417.* Cancer Research, 2006. **66**(1): p. 362-371.

84.   Parang, K., et al., *Mechanism-based design of a protein kinase inhibitor.* Nature Structural & Molecular Biology, 2001. **8**(1): p. 37-41.

85.   Sanderson, M.P., et al., *BI 885578, a Novel IGF1R/INSR Tyrosine Kinase Inhibitor with Pharmacokinetic Properties That Dissociate Antitumor Efficacy and Perturbation of Glucose Homeostasis.* Molecular cancer therapeutics, 2015. **14**(12): p. 2762-2772.

86.   Till, J.H., et al., *Crystallographic and Solution Studies of an Activation Loop Mutant of the Insulin Receptor Tyrosine Kinase INSIGHTS INTO KINASE MECHANISM.* Journal of Biological Chemistry, 2001. **276**(13): p. 10049-10055.

87.   Cabail, M.Z., et al., *The insulin and IGF1 receptor kinase domains are functional dimers in the activated state.* Nature communications, 2015. **6**.

88.   Menting, J.G., et al., *Protective hinge in insulin opens to enable its receptor engagement.* Proceedings of the National Academy of Sciences, 2014. **111**(33): p. E3395-E3404.

89.   Patnaik, S., et al., *Discovery of 3, 5-disubstituted-1H-pyrrolo [2, 3-b] pyridines as potent inhibitors of the insulin-like growth factor-1 receptor (IGF-1R) tyrosine kinase.* Bioorganic & medicinal chemistry letters, 2009. **19**(11): p. 3136-3140.

90.   Li, S., et al., *Structural and biochemical evidence for an autoinhibitory role for tyrosine 984 in the juxtamembrane region of the insulin receptor.* Journal of Biological Chemistry, 2003. **278**(28): p. 26007-26014.

91.   Hu, J., et al., *Structural basis for recruitment of the adaptor protein APS to the activated insulin receptor.* Molecular cell, 2003. **12**(6): p. 1379-1389.

92.   Depetris, R.S., et al., *Structural basis for inhibition of the insulin receptor by the adaptor protein Grb14.* Molecular cell, 2005. **20**(2): p. 325-333.

93.   Li, S., et al., *Crystal structure of a complex between protein tyrosine phosphatase 1B and the insulin receptor tyrosine kinase.* Structure, 2005. **13**(11): p. 1643-1651.

94.   Katayama, N., et al., *Identification of a key element for hydrogen-bonding patterns between protein kinases and their inhibitors.* Proteins: Structure, Function, and Bioinformatics, 2008. **73**(4): p. 795-801.

95.   Wu, J., et al., *Structural and biochemical characterization of the KRLB region in insulin receptor substrate-2.* Nature Structural & Molecular Biology, 2008. **15**(3): p. 251-258.

96.   Stauffer, F., et al., *Identification of a 5-[3-phenyl-(2-cyclic-ether)-methylether]-4-aminopyrrolo [2, 3-d] pyrimidine series of IGF-1R inhibitors.* Bioorganic & medicinal chemistry letters, 2016. **26**(8): p. 2065-2067.

97.   Kettle, J.G., et al., *Discovery and Optimization of a Novel Series of Dyrk1B Kinase Inhibitors To Explore a MEK Resistance Hypothesis.* Journal of medicinal chemistry, 2015. **58**(6): p. 2834-2844.

98.   Lesuisse, D., et al., *Discovery of the first non-ATP competitive IGF-1R kinase inhibitors: Advantages in comparison with competitive inhibitors.* Bioorganic & medicinal chemistry letters, 2011. **21**(8): p. 2224-2228.

99.   Heinrich, T., et al., *Allosteric IGF-1R inhibitors.* ACS medicinal chemistry letters, 2010. **1**(5): p. 199-203.

100.   Sampognaro, A.J., et al., *Proline isosteres in a series of 2, 4-disubstituted pyrrolo [1, 2-f][1, 2, 4] triazine inhibitors of IGF-1R kinase and IR kinase.* Bioorganic & medicinal chemistry letters, 2010. **20**(17): p. 5027-5030.

101.   Wittman, M.D., et al., *Discovery of a 2, 4-disubstituted pyrrolo [1, 2-f][1, 2, 4] triazine inhibitor (BMS-754807) of insulin-like growth factor receptor (IGF-1R) kinase in clinical development.* Journal of medicinal chemistry, 2009. **52**(23): p. 7360-7363.

102.   Wu, J., et al., *Small-molecule inhibition and activation-loop trans-phosphorylation of the IGF1 receptor.* The EMBO journal, 2008. **27**(14): p. 1985-1994.

103.   Velaparthi, U., et al., *Discovery and initial SAR of 3-(1H-benzo [d] imidazol-2-yl) pyridin-2 (1H)-ones as inhibitors of insulin-like growth factor 1-receptor (IGF-1R).* Bioorganic & medicinal chemistry letters, 2007. **17**(8): p. 2317-2321.

104.   Pautsch, A., et al., *Crystal structure of bisphosphorylated IGF-1 receptor kinase: insight into domain movements upon kinase activation.* Structure, 2001. **9**(10): p. 955-965.

105.   Favelyukis, S., et al., *Structure and autoregulation of the insulin-like growth factor 1 receptor kinase.* Nature Structural & Molecular Biology, 2001. **8**(12): p. 1058-1063.

106.   Stanley, D., et al., *O ptimization of a series of 4, 6-bis-anilino-1 H-pyrrolo [2, 3- d] pyrimidine inhibitors of IGF-1R: Elimination of an acid-mediated decomposition pathway.* Bioorg. Med. Chem. Lett, 2009. **19**: p. 373-377.

107.   Finlay, M.R.V., et al., *Discovery of a potent and selective EGFR inhibitor (AZD9291) of both sensitizing and T790M resistance mutations that spares the wild type form of the receptor.* J. Med. Chem, 2014. **57**(20): p. 8249-8267.

108.   Mulvihill, M.J., et al., *Discovery of OSI-906: a selective and orally efficacious dual inhibitor of the IGF-1 receptor and insulin receptor.* Future medicinal chemistry, 2009. **1**(6): p. 1153-1171.

109. Carboni, J.M., et al., *BMS-754807, a small molecule inhibitor of insulin-like growth factor-1R/IR.* Molecular cancer therapeutics, 2009. **8**(12): p. 3341-3349.

110. Nogrady, T. and D.F. Weaver, *Medicinal chemistry: a molecular and biochemical approach*. 2005: Oxford University Press.

111. Hogg, R. and E. Tanis, *Nonparametric methods.* Probability and Statistical Inference, 4th edn. MacMillan Publishing Co., New York, USA, 1993: p. 589-646.

112. Pastor, M., et al., *GRid-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors.* Journal of medicinal chemistry, 2000. **43**(17): p. 3233-3243.

113. Durán, A.n., G.C. Martínez, and M. Pastor, *Development and validation of AMANDA, a new algorithm for selecting highly relevant regions in molecular interaction fields.* Journal of chemical information and modeling, 2008. **48**(9): p. 1813-1823.

114. Henikoff, S. and J.G. Henikoff, *Amino acid substitution matrices from protein blocks.* Proceedings of the National Academy of Sciences, 1992. **89**(22): p. 10915-10919.

115. Sack, J.S., et al., *Crystal structure of microtubule affinity-regulating kinase 4 catalytic domain in complex with a pyrazolopyrimidine inhibitor.* Acta Crystallographica Section F: Structural Biology Communications, 2016. **72**(2): p. 129-134.

116. Ferguson, D.M. and D.J. Raber, *A new approach to probing conformational space with molecular mechanics: random incremental pulse search.* Journal of the American Chemical Society, 1989. **111**(12): p. 4371-4378.

117. Diller, D.J. and K.M. Merz Jr, *Can we separate active from inactive conformations?* Journal of computer-aided molecular design, 2002. **16**(2): p. 105-112.

118. {Sadowski, S., Jens, J. Gasteiger, and G. Klebe, *Comparison of automatic three-dimensional model builders using 639 X-ray structures.* Journal of chemical information and computer sciences, 1994. **34**(4): p. 1000-1008.

119. Gill, P.E., W. Murray, and M.H. Wright, *Practical optimization.* 1981.

120. Deschênes, A. and E. Sourial, *Ligand Scaffold Replacement using MOE Pharmacophore Tools.* J. Chem. Comp. Group, 2007.

121. Grimshaw, S., *Scaffold Replacement in MOE.*

122. Wang, R., Y. Gao, and L. Lai, *LigBuilder: a multi-purpose program for structure-based drug design.* Molecular modeling annual, 2000. **6**(7-8): p. 498-516.

123. Hendlich, M., F. Rippmann, and G. Barnickel, *LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins.* Journal of Molecular Graphics and Modelling, 1997. **15**(6): p. 359-363.

124. Labute, P. and M. Santavy, *Locating binding sites in protein structures.* Journal of Chemical Computing Group, 2007.

125. Knegtel, R.M., I.D. Kuntz, and C. Oshiro, *Molecular docking to ensembles of protein structures.* Journal of molecular biology, 1997. **266**(2): p. 424-440.

126. Österberg, F., et al., *Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock.* Proteins: Structure, Function, and Bioinformatics, 2002. **46**(1): p. 34-40.

127. Pang, Y.-P. and A.P. Kozikowski, *Prediction of the binding sites of huperzine A in acetylcholinesterase by docking studies.* Journal of computer-aided molecular design, 1994. **8**(6): p. 669-681.

128. Bold, G., et al., *New anilinophthalazines as potent and orally well absorbed inhibitors of the VEGF receptor tyrosine kinases useful as antagonists of tumor-driven angiogenesis.* Journal of medicinal chemistry, 2000. **43**(12): p. 2310-2323.

129. Edelsbrunner, H., *Weighted Alpha Shapes, 1992.* Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, Illinois, 1992. **61810**.

130. Corbeil, C.R., C.I. Williams, and P. Labute, *Variability in docking success rates due to dataset preparation.* Journal of computer-aided molecular design, 2012. **26**(6): p. 775-786.

131. Charifson, P.S., et al., *Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins.* Journal of medicinal chemistry, 1999. **42**(25): p. 5100-5109.

132. Wang, R. and S. Wang, *How does consensus scoring work for virtual library screening? An idealized computer experiment.* Journal of chemical information and computer sciences, 2001. **41**(5): p. 1422-1426.

133. Wishart, D., et al., *DrugBank: A comprehensive resource for in silico drug discovery and explorat.*

134. Groom, C.R., et al., *The Cambridge structural database.* Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials, 2016. **72**(2): p. 171-179.