Published in final edited form as:

Curr Top Med Chem. 2016 ; 16(13): 1452–1462.

Cheminformatic-based Drug Discovery of Human Tyrosine Kinase Inhibitors

Terry-Elinor Reid^{1,2}, Joseph M. Fortunak², Anthony Wutoh³, and Xiang Simon Wang^{1,2,*}

¹Molecular Modeling and Drug Discovery Core Laboratory for District of Columbia Center for AIDS Research (DC CFAR), College of Pharmacy, Howard University, Washington, District of Columbia 20059, USA

²Department of Pharmaceutical Sciences, College of Pharmacy, Howard University, Washington, District of Columbia 20059, USA

³Center of Excellence; College of Pharmacy, Howard University, Washington, District of Columbia 20059, USA

Abstract

Receptor Tyrosine Kinases (RTKs) are essential components for regulating cell-cell signaling and communication events in cell growth, proliferation, differentiation, survival and metabolism. Deregulation of RTKs and their associated signaling pathways can lead to a wide variety of human diseases such as immunodeficiency, diabetes, arterosclerosis, psoriasis and cancer. Thus RTKs have become one of the most important drug targets families in recent decade. Pharmaceutical companies have dedicated their research efforts towards the discovery of small-molecule inhibitors of RTKs, many of which had been approved by the U.S. Food and Drug Administration (US FDA) or are currently in clinical trials. The great successes in the development of small-molecule inhibitors of RTKs are largely attributed to the use of modern cheminformatic approaches to identifying lead scaffolds. Those include the quantitative structure-activity relationship (QSAR) modeling, as well as the structure-, and ligand-based pharmacophore modeling techniques in this case. Herein we inspected the literature thoroughly in an effort to conduct a comparative analysis of major findings regarding the essential structure-activity relationships (SARs)/pharmacophore features of known active RTK inhibitors, most of which were collected from cheminformatic modeling approaches.

Keywords

Cheminformatics; ligand-based drug discovery; quantitative structure-activity relationship; pharmacophore modeling; tyrosine kinase inhibitors; vascular endothelial growth factor; epidermal growth factor receptor; platelet-derived growth factor receptor; hepatocyte growth factor receptor

*Address correspondence to this author at the Molecular Modeling and Drug Discovery Core Laboratory for District of Columbia Center for AIDS Research (DC CFAR), Department of Pharmaceutical Sciences, Howard University, Washington, District of Columbia 20059, USA; Tel: 1-202-806-6547; Fax: 1-202-806-7805; x.simon.wang@gmail.com.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

Among 518 protein kinases in the human genome, around 90 selectively catalyze the phosphorylation of tyrosine hydroxyl groups and are known as protein tyrosine kinases (PTKs) hence. Of these 90 PTKs, 32 are classified as non-receptor tyrosine kinase and the remaining 58 are classified as receptor tyrosine kinases (RTKs) due to the presence of a transmembrane domain. RTKs are grouped into 20 subfamilies and play key roles as regulators of cell-cell signaling and communication. They span the phospholipid membrane at the cell surface and can trigger a complex down-stream signaling cascade joined by non-receptor tyrosine kinases located in the cytoplasm of the cells. As a result RTKs are involved in moderating cell growth, proliferation, differentiation, survival and metabolism. [1,2] They had been found to be implicated in various human disorders such as immunological, neurological, metabolic and infectious diseases; many RTKs are being used for targeted therapies to treat various types of cancer.

The recurring theme regarding RTKs as drug targets is their implication in oncogenesis due to their role in promoting cell growth, survival, invasion and mortality of malignant tumor progression. In many cases it has been determined that the overexpression or genetic mutations of RTKs cause increased and uncontrolled cell proliferation and tumor formation. A study conducted by Paik *et al.* in 1990 confirmed that the HER2 (erbB-2, HER2/neu) gene was overexpressed in a subgroup of breast tumors.[3] After more than a decade researchers have gained a better understanding of HER2 signaling pathway and its role in oncogenesis. This signaling pathway is now considered to be one of the major pathways for cell proliferation and involved in around 85% of breast cancers. In addition, a study conducted by Salmon *et al.* revealed that the increased level of HER2 correlated directly with the aggressive progression of breast cancer.[4] Targeting HER2 signaling pathway provides therapeutic benefits as is seen with the approved drug trastuzumab.[5,6] Trastuzumab is a monoclonal antibody that binds to the juxtamembrane portion of HER2 receptor and inhibits HER2 signaling pathway thereby. There are multiple proposed mechanisms of action for Trastuzumab which include: (a) activation of antibody-dependent cellular cytotoxicity (ADCC)[7]; (b) inhibition of HER2 cleavage[8]; (c) induction of HER2 down-regulation; (d) inhibition of PI3K signaling pathway thus promoting apoptosis and arrest of proliferation; (e) inhibition of angiogenesis.[9]

In the 1990's it was observed that genetic mutations of RTKs resulted in deregulation of RTKs thereby inducing cell proliferation and eventually human cancers. This was particularly evident with epidermal growth factor receptor (EGFR) mutant which enhances the tumorigenic capacity in glioblastomas.[10] Mutations in EGFR are also common in lung carcinoma[11], colorectal cancer[12], squamous cell carcinoma of the head and neck (SCCHN)[13], gastrointestinal tumors[14], ovarian cancer [15] and colon cancer, just to name a few.[16] Other similar mutations on RTKs include that of fibroblast growth factor receptors (FGFR) which is associated with multiple carcinomas.[17-19] For example, human proto-oncogene RET mutation had been found to be implicated in multiple endocrine neoplasia type 2 cancer syndrome (MEN2).[20,21] More recently EPHA3 mutations have been correlated with human lung cancer because of the inactivation of this subtype of RTKs, which in turn suppresses tumor growth.[22]

It is not abnormal for multiple RTKs to be involved in the proliferation and progression of different carcinomas because tumor pathogenesis in human involves multiple genetic alterations of normal cells. There are six essential and consistent traits of cancer cells, suggested by Hanahan *et al.* and acquired via alterations in cell physiology include the following: (a) self-sufficiency in growth signals, (b) insensitivity to anti-growth signals, (c) tissue invasion and metastasis, (d) limitless replicative potential, (e) sustained angiogenesis and (f) evading apoptosis.[23] More recently two more traits have been added, i.e. (g) evading immune destruction and (h) reprogramming energy metabolism.[24] The alterations in cell physiology can result from an accumulation of genetic mutations of genes that encode oncoproteins, tumor suppressors as well as their regulators like RTKs. Hence RTKs become promising therapeutic targets for the development of tissue-selective anticancer drugs in comparison to conventional chemotherapies. The current review highlights modern cheminformatic tools, mainly the quantitative structure-activity relationship (QSAR) and pharmacophore modeling, towards the discovery of efficacious human RTKs inhibitors.

2. CHEMINFOMATIC APPROACHES TO DISCOVER RTK INHIBITORS

There are major routes the scientist can take in the query of RTKs inhibitors, part of which include the design of: (a) monoclonal antibodies; (b) small-molecule tyrosine kinase inhibitor; (c) immunotoxins; and (d) antisense inhibitors of growth factor receptors. Several small-molecule tyrosine kinase inhibitors have been approved by the U.S. Food and Drug Administration (US FDA) for treatment of multiple types of human carcinomas.

In fact, PTKs are now considered as the 2nd largest group of druggable targets after G protein-coupled receptors (GPCRs).[25] This has encouraged the search for novel small-molecule inhibitors that are target-specific and in some cases can bind a subset of targets simultaneously while exhibiting fewer adverse drug reactions (ADR). Efforts toward identifying new lead scaffolds for RTKs inhibitors have employed a rational drug discovery approach which includes the use of 2D- and 3D-QSAR, ligand- and structure-based pharmacophore modeling. These methods have proven to be effective, fast and cost-efficient. Both types of approaches can be taken, and have been taken in the absence of macromolecular target structures, to identify promising lead scaffolds for targeting RTKs. Case studies involving QSAR and pharmacophore modeling are to be discussed in the later section of this review. Sections 2.1 and 2.2 give brief synopses and the processes involved in these two cheminformatic approaches.

2.1. 2D and 3D-QSAR Modeling Methods

QSAR models are essentially the statistical models which try to correctly derive a relationship between the structural features (i.e. independent variables) of chemical compounds and their biological activities (i.e. dependent variables). The models derived after rigorous validation, can be used to accurately predict the activities of unknown compounds within a pre-determined applicability domain.[26] The essential and underlying principle behind QSAR modeling is that, compounds with similar structures exhibit similar activities.

There is a general consensus regarding the workflow for building robust and predictive QSAR models as outlined in (Fig. 1), which had been adopted by many researchers.[27-29] It should be noted that data curation is an important step in the QSAR workflow especially when the data is collected from public databases of chemicals.[30] An additional important step in the QSAR workflow is to validate the models using an external data set, which is the set that was not included in the model building process. One can move forward a step further to include a completely independent data set for the purpose of ultimate external validation.

There are many QSAR techniques now that can be applied to RTK inhibitors modeling. In general they can be classified into two main categories, i.e. continuous modeling and classification modeling, using basic regression methods such as multiple linear regression (MLR) and partial least square regression (PLS) to advanced machine learning methods like random forest (RF), support vector machine (SVM), *k*-Nearest Neighbor (*k*NN) and neural networks (NN), just to name a few.[31] The comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) are the classical 3D-QSAR techniques which seek to correlate steric, electrostatic interactions by Lennard-Jones and Coulomb potentials, to biological activities or binding affinities of ligand molecules. Moreover, the CoMSIA method tries to correlate not only steric and electrostatic interactions but also hydrophobic, hydrogen bond acceptor and hydrogen bond donor features.[32]

2.2. Ligand- and Structure-based Pharmacophore Modeling Methods

Ligand-based pharmacophore modeling is a powerful tool for which many successful stories in drug discovery have been reported.[33] IUPAC has described a pharmacophore as an ensemble of steric and electronic features which is necessary to ensure the optimal intermolecular interactions with a specific biological target and to trigger (or to block) its biological response.[34] The modeling involves creating a conformational space for the ligands and aligning them together to identify common pharmacophoric features to construct a model, or called the pharmacophore hypothesis. The general workflow for ligand-based pharmacophore modeling is illustrated in (Fig. 2) as shown below. The common algorithms for performing ligand- and structure-based pharmacophore modeling include HipHop/HypoGen from Accelrys Inc., PHASE from Schrodinger, LigandScout and MOE.[35]

3. CASE STUDIES

The cheminformatic approaches, comprised mainly of QSAR and pharmacophore modeling, has been applied extensively to assist in the discovery of lead small molecule inhibitors of RTKs such as vascular endothelial growth factor receptor (VEGFR), EGFR, FGFR, platelet-derived growth factor receptors (PDGFR) and hepatocyte growth factor receptor (HGFR). In addition to the discovery of lead compounds, the investigators were also able to explore the structure-activity relationships (SARs) of inhibitors which aid the lead optimization phase of the drug discovery pipeline.

3.1. VEGFR

There are three subtypes of the VEGFR family namely VEGFR-1, VEGFR-2 and VEGFR-3, which play crucial roles in haematopoietic cell development, endothelial cell development and lymphatic endothelial cell development, respectively. VEGFR-2 and its corresponding ligand, VEGF, are regarded to be one of the most important regulatory systems for tumor angiogenesis.[36,37] A consortium of small molecule inhibitors of VEGFR-2 have been developed and a significant number of them were approved with success for the treatment of human cancers. These include Semaxanib[38] followed by Sunitinib[39], the third drug candidates in the same chemical series of VEGFR inhibitors. Among the others there are Sorafenib[40] and Levatinib[41] approved by US FDA[42], plus the Vatalanib[43] which is still in clinical trials. So far there are many successful stories in the development/approval of novel inhibitors that target VEGFR, though a large number of candidates had been with-drawn from the clinical trials due to lack of potency and selectivity issue which causes side effects and toxicity. Thus, further research is actively being conducted toward the rational design of subtype-selective molecules with the aid of modern cheminformatics approaches.

Recently, a number of 3D pharmacophore modeling studies have been carried out in order to: (1) identify selective lead small molecule inhibitors of VEGFR; (2) identify a novel class of VEGFR inhibitors; and (3) gain insight into SARs. Two independent studies conducted by Yu *et al.*[44,45] identified similar pharmacophore features as the best hypothesis for VEGFR-2 inhibitors SARs. Those features include hydrogen bond acceptor (A), aromatic ring (R), hydrogen bond donor (D) and a hydrophobic group for binding as seen in Fig. (3).

The difference between the two studies was that a total of 10 biologically active molecules were used for the pharmacophore modeling in one study while a set of 28 more structurally diverse compounds were used in the second study. In addition, the modeling work was conducted using HipHop algorithm in one study while Poling algorithm in another, both of which are from CATALYST modeling platform. Virtual screening was carried out using this hypothesis which resulted in the identification of an experimentally determined active lead compound that gave 100% inhibition at the IC₅₀ of 50 μ M. From the docking results of the active the consensus of both studies was found: (1) the hydrogen bond donor moiety is necessary for activity because of the H-bond interaction between NH of the ligand and the carboxylate oxygen of the Glu883; (2) the hydrogen bond acceptor moiety is also important due to its interaction with NH at backbone of Cys917 and, (3) the hydrophobic group enhances the activity by residing in the hydrophobic pocket comprising of residues Ile886, Leu887, Ile890, Val896 and Leu1017 that all contain hydrophobic side chains. These results are also consistent with experimentally determined structures as well as other studies that employ structure-based approaches.[46,47] Pharmacophore models can also be constructed using structure-based technique as outlined in a study conducted by Lee *et al.*[48] The models were built from two known VEGFR-2 inhibitors on the market, i.e. Sunitinib and Sorafenib that were complexed with VEGFR-2 crystal structure. With the aid of the pharmacophore modeling as well as molecular docking, Lee selected 100 hits among which 10 compounds inhibited VEGFR-2 activities with IC₅₀ ranging from 1 μ M to 10 μ M. These studies also identified important pharmacophoric features such as the presence of

hydrophobic groups, a hydrogen bond acceptor and two hydrogen bond donors, which are consistent with the findings of Yu *et al.*'s study. Most importantly, these features are part of the interactions with the side chains of those residues mentioned above.

More recently, Rajagopalan *et al.* exploited pharmacophore modeling in the analysis of VEGFR-2 inhibitors.[49] This study generated pharmacophore hypothesis based on 81 N-Phenyl-N'-{4-(4-quinolyloxy) phenyl} urea analogues and 4-aminopyrimidine-5-carbaldehyde oxime analogues, using the module of PHASE in Schrodinger package. The algorithm of PHASE is unique in that it can identify six features, i.e. hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively charged group (N), positively charged group (P) and aromatic ring (A). In the end, a five-feature hypothesis of AADRR was identified as the best hypothesis which is consistent with the models built by Yu *et al.* and Lee *et al.* The authors also mentioned the significance of distance between individual pharmacophoric features, which is crucial for favorable binding of VEGFR-2 inhibitors.

This is evident when comparing the features of CMP1 to CMP28 of an IC₅₀ of 31 μ M in (Fig. 4), the least active compound in the data set. CMPD28 does satisfy the hydrophobic feature as well as the hydrogen bond donor and aromatic ring, though there is no hydrogen bond acceptor plus the molecule is more rigid. In addition, the distance between the aromatic ring and the hydrophobic group is much smaller in comparison to CMP1, which means that this interaction in the active site was lost and its binding affinity was affected.

The 3D-QSAR modeling studies have also been conducted on VEGFR inhibitors using CoMFA and/or CoMSIA methods with the aid of molecular docking in some cases. The CoMFA and CoMSIA methods are powerful tools for providing insights into not only the various hydrophobic and hydrogen bond interactions but also how steric and electrostatic properties of molecules play a role in rendering the inhibitory activities.[50,51] The results of CoMFA/CoMSIA modeling from Du *et al.*[52], Zeng *et al.*[53] and Wu *et al.*[54] all report the similar contributions of electrostatic forces, hydrophobic interactions as well as hydrogen bond donor interactions to the binding of inhibitors to VEGFR (Table 1).

As outlined before from the results of a number of pharmacophore modeling studies, it is understood now about the significance of hydrophobic, hydrogen bond donor/acceptor interactions. Furthermore, it is evident from the results of multiple CoMFA and CoMSIA studies that steric and electrostatic interactions also contribute to the inhibitory activities, in which case electrostatic does more than steric factors (Table 2). The general consensus from the studies was that sterically-bulky, hydrophobic groups are located in proximity to the hydrophobic pocket of VEGFR-2; and the electrostatic fields or regions are located on the opposite end of the molecules.

3.2. EGFR

One of the most extensively studied RTKs is the EGFR, including subtypes ErbB-1, ErbB-2, ErbB-3, and ErbB-4. The overexpression of EGFR had been found to cause the brain, lung, bladder, colon, breast, head and neck tumors.[55] Erlotinib, Gefitinib, and Lapatinib are EGFR inhibitors which have been approved by US FDA for the treatment of metastatic non-

small cell lung cancer (NSCLC) and breast cancers.[39,56,57],[42,58] Although there are a large number of known EGFR inhibitors, their clinical efficacy and expectance for approval are limited due to mutations conferred to EGFR that result in drug resistance.[59] Hence the search for novel EGFR inhibitors that are less susceptible to mutation is still ongoing. In this effort, researchers have been employing rational drug design approaches like pharmacophore and QSAR modeling to guide the discovery process. A plethora of QSAR and pharmacophore modeling studies have been published in recent years to the query of novel EGFR inhibitors. Herein we attempt to compile as much information as possible to provide a complete picture of the SARs for the diverse sets of EGFR inhibitors.

Anilinoquinazolines (cf. Fig. 5) have been widely recognized for the inhibition of EGFR. [60-62] Due to the limited success of reversible inhibitors of EGFR, irreversible inhibitors are favorably being considered since they can form covalent bonds with a cysteine residue (Cys797) within the active site of EGFR. This will render a much stronger binding to the target thus make an efficacious drug.[63] Recently Xiang *et al.* successfully applied pharmacophore modeling to 138 2nd-generation EGFR inhibitors. The common features of these inhibitors included a hydrogen bond donor, an aromatic hydrophobic ring and two hydrogen bond acceptors (DRAA). In addition there has to be a function group in the molecule for a covalent bond to occur, in which case the β -carbon of acrylamide group or propargyl amide group can undergo a Michael addition reaction with the receptor residues. [64]

The following studies to be discussed utilized quinazoline analogues for modeling. A comparative analysis of the methodology and outcomes of four recent studies by Du *et al.* [65], Nandi *et al.*[66], Noolvi *et al.*[67] and Assefa *et al.*[68] provided conclusive and insightful SARs information. All studies employed regression models; one included CoMFA and CoMSIA techniques from which contour maps aided significantly in interpreting the QSAR results. The importance of the choice of descriptors in model generation was evident from study to study. The top model built by Du *et al.* with $R^2 = 0.7835$ used nine descriptors to generate a hypothesis, four of which anti-correlates the IC₅₀ values. These include: (1) *MEERCH*, which is the Min electron-electron repulsion for C-H bond; (2) *PMIC*, which provides 3D coordinate information regarding the molecule; (4) *R4U*, which is the autocorrelation coefficient lag 4; (5&6) *HATS3u* and *HATS5u*, which takes into account molecular 3D geometry; (6) *SIC1*, which reflect molecular polarity and polarizability; (7) *ANRIC*, a quantum chemical descriptor of the C atom nucleophile; (8) *LPI*, an index of molecular branching and (9) *EEig09d*, a topological molecular descriptor. From those nine descriptors selected, the authors concluded that the chemical features of quinazoline derivatives essential for favorable activities comprise of polarizability, activation energy, mass distribution, connectivity and branching. The polarizability of quinazolines is indeed important for stabilization of the hydrogen bond formed between amide nitrogen donor of the ligand and the carboxyl group of Met769 at the hinge region of EGFR. It is also a factor that influences the interactions with both hydrophobic and hydrophilic regions of the active site of the protein. In addition, it was determined that the connectivity and branching information, accounted by *EEig09d* and *LPI*, affected the ability of the ligands to interact favorably in the active site.

Nandi *et al.* used 62 anilinoquinazoline derivatives to build partial least square QSAR regression models based on topological and physiochemical descriptors. Of the models generated the top-ranked one showed an external predictability of $R^2 = 0.678$ with five types of descriptors. Of those five descriptors, the *chiV3Cluster* and *XAMostHydrophobicHydrophilicDistance* showed a positive correlation with biological activities (IC₅₀). *XAMostHydrophobicHydrophilicDistance* describes the distance between the most hydrophobic and the most hydrophilic point on the vdW surface of the molecule; while the *chiV3Cluster* describes the valence of molecular connectivity index. The increase in the value of *XAMostHydrophobicHydrophilicDistance* favors the EGFR inhibition. The general pharmacophore features for quinazoline derivatives are outlined in (Fig. 6), which includes the quinazoline ring that resides in the hydrophobic pocket of EGFR and the hydrogen bond acceptor N that forms a hydrogen bond with Met769. Those results of the QSAR modeling in the above studies are consistent in that, though the descriptors chosen in Du's and Nandi's studies, i.e. *SIC1* vs. *XAMostHydrophobicHydrophilicDistance*, were different, they both represent the significance of hydrophobicity and hydrophilicity of the ligands for their biological activities. The same is also true in Xiang's study since it also identified the features such as aromatic ring and hydrogen bond acceptor within their pharmacophore hypothesis. Consistently, Assefa *et al.*'s CoMSIA model showed that the electrostatic feature ranks high in significance to activities of quinazolines, second to HBD and followed by hydrophobic feature contributing. It should be pointed out that the difference between Assefa's models and the previously mentioned studies lie in that the former utilized both quinazoline derivatives and quinoline derivatives. The feature that contributes most significantly to the CoMSIA models for quinoline derivatives is HBA followed by hydrophobic features. The CoMFA contour maps for quinazoline analogs revealed that bulky substituents are favored at the meta (R10) and R1 positions of the aniline ring (cf. Fig. 6), though there exists a limit. On the other hand, tertiary amines and bulky groups are not favored at the N of the aniline due to its potential to form hydrogen bonding interactions. The contour maps also show a region around the R1 and R2 position that favors negative electrostatic potential. Similar contour maps were generated for quinoline derivatives which reveal hydrophobic substituents are favored on the aniline ring and HBA groups are favored at the para position of the quinolone ring thereby facilitating hydrogen bonding interactions.

A previous 2D-QSAR study performed by Nandi *et al.*[69] also identified electrostatic features as major contributing factor to biological activities. More descriptive information can be drawn from Noolvi *et al.* study which utilizes MLR, Principal Component Regression (PCR), and PLS techniques. The R^2 for all the models were ranged from 0.799 to 0.912. Some of the important descriptors present within multiple models included the numbers of bromine and chlorine counts, which contribute positively to activities of the 4-anilino quinazoline in the presence of electron donating group at the C-6 position of quinazoline ring (cf. Fig. 7). This was consistent with the results from Assefa *et al.* which indicates an increase in activity while the substituent at the C-3 position of the anilino ring becomes larger, e.g., the activities increased with substituent sequence of H>F>I>Cl>Br. Also, the presence of a negative electrostatic potential substituent at the C-6 and C-7 positions of the quinazoline ring is favored. Although not present in all three models, the descriptor of

T_2_N_7 was present in both PLS and PCR models. It describes the number of double-bond atoms separating any pair of nitrogen atoms by seven bonds' distance. This particular descriptor showed a positive contribution to biological activities in both types of models. Because of the above, Noolvi *et al.* postulated that the presence of an amino group within the aniline moiety is essential for interaction with EGFR.

The general consensus deduced from the above-mentioned QSAR and pharmacophore studies is that a bulky substitution group at C-3 of the anilino ring in the presence of an electron donating group on C-6, C-7 or both, contributes positively towards the activities of EGFR inhibitors. In addition, the nitrogen atoms on the quinoline ring and the amine group of the aniline play significant roles by forming hydrogen bond interactions that enhance the inhibitory activities. The hydrophobicity of ligands is also important pharmacophore feature as can be seen from the ring system of both quinoline and aniline moieties. They render essential hydrophobic interactions in the lipophilic sub-pockets of the EGFR active site.

3.3. PDGFR

The PDGFR has been divided into two subfamilies, i.e. PDGFR α and PDGFR β , each of which plays different role in signal transduction. The PDGFR α -associated signaling controls gastrulation and the development of many organs and neuroprotective tissues, while the PDGFR β -associated signaling is vital to early hematopoiesis and blood vessel formation. [70] Deregulation of PDGFR-associated signaling causes fibrotic diseases, various vasculo-proliferative diseases, and cancer.[71] There are multiple ligand scaffolds that had been investigated for their inhibitory effects to PDGFR, some of which include 1-phenylbenzimidazoles [72], piperazinylquinazolines[73], and pyrimidine[74] analogues.

During the past decade QSAR approaches have been employed in aiding the discovery of PDGFR inhibitors, mainly via the deciphering of their SARs. Katritzky *et al.* and Shen *et al.* utilized MLR as well as artificial neural network (ANN) to model the 1-phenylbenzimidazole analogs (Fig. 8). Though the descriptor types included in both studies are different, they all contained quantum chemical descriptors, charge related descriptors and thermodynamic descriptors.

The best multilinear QSAR model for phenylbenzimidazole derivatives obtained by Katritzky *et al.* ($R^2=0.71$) contained seven descriptors, while their classification models contained five descriptors. The common descriptors among those two models were *min e-e repulsion* for C-C bond, which was also the most statistically significant descriptor in MLR models. Together with the *relative negative charge* descriptor, both are involved in electrostatic interaction of the ligand to receptor, while the former governs the energy of the electrostatic interaction that is important in determining conformational changes of the ligand. The relative negative charge descriptor refers to the charge distribution of the ligand and is directly related to electrostatic interactions between ligands and PDGF receptor. From this study one can conclude about the importance of electrostatic features for improved inhibitory activities of phenylbenzimidazole derivatives against PDGF receptor.

Shen *et al.* on the other hand used evolution algorithm to generate MLR QSAR models for phenylbenzimidazole derivatives, and their top-ranked model contains four descriptors.

They include the spatial descriptor *B54,7*, a quantum mechanical descriptor *Qm*, an indicator variable descriptor *I5* and an E-State index descriptor *S-ssO*, all of which are frequently chosen features. Two spatial descriptors *B54,7* and *I5* signify the importance of steric effects to activities, and the negative correlation of *B54,7* with activity indicates that bulkier groups at C-4 and C-5 of 1-phenyl-benzimidazole in (Fig. 8) reduce its biological potency. In the meantime, the positive correlation of *I5* indicates that substitutions at C-5, which includes -OCH- and -SH groups to name a few, increase the biological activities. Similarly, the positive correlation of *Qm* (electronic effects of substituent on the meta position), shows that electron donating groups at C-5 would enhance the inhibitory activities. A frequently appearing feature not included in the top models is the number of rotatable bonds, i.e. *rotbonds*. It does show up in the 2nd top-ranked MLR model and show the positive correlation, an indicative of a positive effect on activity with the increasing number of rotatable bonds.

QSAR modeling studies were also carried out on piperazinylquinazolines (cf. Fig. 9) by Guha *et al.*[75] and Khadikar *et al.*[76], both of whom utilized the same data set for modeling. The difference between those two studies lie in the number and type of descriptors used as well as the modeling technique employed. Khadikar *et al.* employed linear regression with topological descriptors only for QSAR modeling while Guha *et al.* included linear regression as well as advanced methods like random forest and neural network, together with a wide variety of geometric, topological and electronic descriptors. The use of multiple QSAR techniques and different families of descriptors allowed a more detailed analysis of the SARs, which is the case by Guha *et al.* When comparing the common descriptors selected during model building for each method, SURR-5 (describes the hydrophobicity of a surface) appears to be the most common by all selected models. Furthermore, they created a plot of variable importance which indicated SURR-5 as the most important descriptor for this particular data set.

From the results of all regression models generated, the higher absolute value of SURR-5 correlates with increased hydrophobicity, and in turn higher activities. Another descriptor common to both linear regression and advanced models, RNHS-3, is a measure of the relative hydrophilic surface area of a molecule. Though less significant than SURR-5, it emphasizes the importance of the hydrophobic nature of the molecule to inhibitory activities. The last descriptor to discuss is MDEN-23, i.e. molecular distance edge vector between secondary and tertiary nitrogen, which is the highest-weighted descriptor to contribute to activities positively in regression model. In particular, the increased number of secondary nitrogen and large nitrogen-containing rings seem to boost the increase activities of PDGFR inhibitors.

3.4. HGFR

HGFR, also known as c-MET, is one of two members in the MET subfamily of RTKs. It is an oncoprotein that is implicated in a variety of malignancies including colon, bladder, kidney, lung, prostate, gastric and breast cancers.[77] c-MET is not only a viable drug target but also a biologic marker for malignancies. The drug discovery for c-MET kinase inhibitors is an attractive area for major pharmaceutical companies such as GlaxoSmithKline, Merck,

Bristol-Myers Squibb and Pfizer who are currently conducting clinical trials on promising drug candidates.[78]

Two 3D pharmacophore/QSAR studies on c-MET inhibitors conducted by Huang *et al.*[79] and Tai *et al.*[80] have been reported. Albeit the fact that the algorithms/software used by both groups were different and their data sets were structurally diverse and independent of each other, the final conclusions of both appear to be similar. Tai *et al.*'s study generated a model with correlation coefficient of 0.983, including pharmacophoric features of two hydrogen bond acceptors, one hydrophobic group and one aromatic ring. In comparison, Huang *et al.* built a model with correlation coefficient of 0.962, whose pharmacophoric features include two hydrogen bond acceptors, one hydrophobic group and two excluded volumes. One hydrogen bond acceptor is found to be essential for hydrogen bonding interaction with NH at the backbone of Met1160 and another forms a hydrogen bond with NH of backbone of Asp1222, which can be validated by the crystal structure of c-MET (PDB code: 3CE3). The hydrophobic feature is essential for hydrophobic interactions to the hydrophobic core comprised of active site residues Ile1084, Gly1163 and Met1211. Excluded volume is the feature that is inaccessible by other molecules but only the original one. The aromatic ring in Tai *et al.*'s model can be considered as being in the excluded volume region that can form π - π stacking interaction with Phe1223. Hence both studies seem to be consistent to each other. Once again the distances between each pharmacophoric feature are important as shown from both studies.

A more recent study by Yuan *et al.*[81] also employed CoMFA and CoMSIA 3D-QSAR modeling for HGFR inhibitors. Unlike the previous studies of Huang and Tai *et al.* which utilized a diverse data set of multiple molecular scaffolds, this study incorporated only the triazolopyrazine derivatives. Their results indicate that hydrogen bond donor (27.8%) and hydrophobic (24.4%) features exert a large influence on bindings followed by electrostatic (21.9%), hydrogen bond acceptor (17%) and lastly steric effect (6.8%). All of above chemical features identified from CoMFA and CoMSIA are consistent with the findings of Huang and Tai. Based on those studies we can hypothesize that pharmacophore features such as hydrogen bond acceptors and donors, hydrophobic groups and aromatic ring in a unique set of spatial arrangements are essential for HGFR inhibitors' activities.

CONCLUSIONS

This review summarizes recent cheminformatic studies, mainly comprising of QSAR and pharmacophore modeling in this case, to generating well-defined SARs for RTK inhibitors. It would lend insightful structural information for the rational design of novel and selective inhibitors against these important receptor subfamilies. It should be noted that many studies only include analogues of a single molecular scaffold, thus the SARs derived may not be generalizable for all RTK inhibitors. This highlights the importance of conducting a larger scale cheminformatic modeling studies. Also as can be seen from this review that most of the computational work is being done on only four RTKs so far thus much remains to be done on the remaining subfamilies which may be associated with other disease types as well as disease states. Overall, the RTKs are becoming as important as the GPCRs in terms of

druggable protein targets while modern cheminformatic tools would assist substantially in the discovery of more RTK inhibitors in this new field.

ACKNOWLEDGEMENTS

This work was supported in part by District of Columbia Developmental Center for AIDS Research (P30AI087714), National Institutes of Health Administrative Supplements for U.S.-China Biomedical Collaborative Research (5P30A1087714-02), the National Institute on Minority Health and Health Disparities of the National Institutes of Health under Award Number G12MD007597. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We also thank the Alzheimer's Drug Discovery Foundation (ADDF) for the grant support (20130503).

Biography

X.S. Wang

REFERENCES

1. Robinson DR, Wu YM, Lin SF. The protein tyrosine kinase family of the human genome. *Oncogene*. 2000; 19(49):5548–5557. [PubMed: 11114734]
2. Zwick E, Bange J, Ullrich A. Receptor tyrosine kinases as targets for anticancer drugs. *Trends Mol. Med*. 2002; 8(1):17–23. [PubMed: 11796262]
3. Paik S, Hazan R, Fisher ER, Sass RE, Fisher B, Redmond C, Schlessinger J, Lippman ME, King CR. Pathologic findings from the National Surgical Adjuvant Breast and Bowel Project: prognostic significance of erbB-2 protein overexpression in primary breast cancer. *J. Clin. Oncol*. 1990; 8(1): 103–112. [PubMed: 1967301]
4. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*. 1987; 235(4785):177–182. [PubMed: 3798106]
5. Gutierrez C, Schiff R. HER2: biology, detection, and clinical implications. *Arch. Pathol. Lab Med*. 2011; 135(1):55–62. [PubMed: 21204711]
6. Hudis CA. Trastuzumab--mechanism of action and use in clinical practice. *N. Engl. J. Med*. 2007; 357(1):39–51. [PubMed: 17611206]
7. Cooley S, Burns LJ, Repka T, Miller JS. Natural killer cell cytotoxicity of breast cancer targets is enhanced by two distinct mechanisms of antibody-dependent cellular cytotoxicity against LFA-3 and HER2/neu. *Exp. Hematol*. 1999; 27(10):1533–1541. [PubMed: 10517495]
8. Molina MA, Codony-Servat J, Albanell J, Rojo F, Arribas J, Baselga J. Trastuzumab (herceptin), a humanized anti-Her2 receptor monoclonal antibody, inhibits basal and activated Her2 ectodomain cleavage in breast cancer cells. *Cancer Res*. 2001; 61(12):4744–4749. [PubMed: 11406546]
9. Valabrega G, Montemurro F, Aglietta M. Trastuzumab: mechanism of action, resistance and future perspectives in HER2-overexpressing breast cancer. *Ann. Oncol*. 2007; 18(6):977–984. [PubMed: 17229773]
10. Nishikawa R, Ji XD, Harmon RC, Lazar CS, Gill GN, Cavenue WK, Huang HJ. A mutant epidermal growth factor receptor common in human glioma confers enhanced tumorigenicity. *Proc. Natl. Acad. Sci. U. S. A*. 1994; 91(16):7727–7731. [PubMed: 8052651]
11. Shigematsu H, Gazdar AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. *Int. J. Cancer*. 2006; 118(2):257–262. [PubMed: 16231326]

12. Nagahara H, Mimori K, Ohta M, Utsunomiya T, Inoue H, Barnard GF, Ohira M, Hirakawa K, Mori M. Somatic mutations of epidermal growth factor receptor in colorectal carcinoma. *Clin. Cancer Res.* 2005; 11(4):1368–1371. [PubMed: 15746034]
13. Lee JW, Soung YH, Kim SY, Nam HK, Park WS, Nam SW, Kim MS, Sun DI, Lee YS, Jang JJ, Lee JY, Yoo NJ, Lee SH. Somatic mutations of EGFR gene in squamous cell carcinoma of the head and neck. *Clin. Cancer Res.* 2005; 11(8):2879–2882. [PubMed: 15837736]
14. Kwak EL, Jankowski J, Thayer SP, Lauwers GY, Brannigan BW, Harris PL, Okimoto RA, Haserlat SM, Driscoll DR, Ferry D, Muir B, Settleman J, Fuchs CS, Kulke MH, Ryan DP, Clark JW, Sgroi DC, Haber DA, Bell DW. Epidermal growth factor receptor kinase domain mutations in esophageal and pancreatic adenocarcinomas. *Clin. Cancer Res.* 2006; 12(14 Pt 1):4283–4287. [PubMed: 16857803]
15. Schilder RJ, Sill MW, Chen X, Darcy KM, Decesare SL, Lewandowski G, Lee RB, Arciero CA, Wu H, Godwin AK. Phase II study of gefitinib in patients with relapsed or persistent ovarian or primary peritoneal carcinoma and evaluation of epidermal growth factor receptor mutations and immunohistochemical expression: a Gynecologic Oncology Group Study. *Clin. Cancer Res.* 2005; 11(15):5539–5548. [PubMed: 16061871]
16. Riely GJ, Politi KA, Miller VA, Pao W. Update on epidermal growth factor receptor mutations in non-small cell lung cancer. *Clin. Cancer Res.* 2006; 12(24):7232–7241. [PubMed: 17189394]
17. Kannan K, Givol D. FGF receptor mutations: dimerization syndromes, cell growth suppression, and animal models. *IUBMB. Life.* 2000; 49(3):197–205. [PubMed: 10868910]
18. Cappellen D, De OC, Ricol D, de MS, Bourdin J, Sastre-Garau X, Chopin D, Thierry JP, Radvanyi F. Frequent activating mutations of FGFR3 in human bladder and cervix carcinomas. *Nat. Genet.* 1999; 23(1):18–20. [PubMed: 10471491]
19. Jang JH, Shin KH, Park JG. Mutations in fibroblast growth factor receptor 2 and fibroblast growth factor receptor 3 genes associated with human gastric and colorectal cancers. *Cancer Res.* 2001; 61(9):3541–3543. [PubMed: 11325814]
20. Edery P, Eng C, Munnich A, Lyonnet S. RET in human development and oncogenesis. *Bioessays.* 1997; 19(5):389–395. [PubMed: 9174404]
21. Chappuis-Flament S, Pasini A, De VG, Segouffin-Cariou C, Fusco A, Attie T, Lenoir GM, Santoro M, Billaud M. Dual effect on the RET receptor of MEN 2 mutations affecting specific extracytoplasmic cysteines. *Oncogene.* 1998; 17(22):2851–2861. [PubMed: 9879991]
22. Zhuang G, Song W, Amato K, Hwang Y, Lee K, Boothby M, Ye F, Guo Y, Shyr Y, Lin L, Carbone DP, Brantley-Sieders DM, Chen J. Effects of cancer-associated EPHA3 mutations on lung cancer. *J. Natl. Cancer Inst.* 2012; 104(15):1182–1197. [PubMed: 22829656]
23. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000; 100(1):57–70. [PubMed: 10647931]
24. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011; 144(5):646–674. [PubMed: 21376230]
25. Cohen P. Protein kinases--the major drug targets of the twentyfirst century? *Nat. Rev. Drug Discov.* 2002; 1(4):309–315. [PubMed: 12120282]
26. Yee, LC.; Wei, YC. Current Modeling Methods Used in QSAR/QSPR.. In: Dehmer, M.; Varmuza, K.; Bonchev, D., editors. *Statistical Modelling of Molecular Descriptors in QSAR/QSPR.* first ed.. Wiley-VCH; 2012. p. 1-31.
27. Peterson YK, Wang XS, Casey PJ, Tropsha A. Discovery of geranylgeranyltransferase-I inhibitors with novel scaffolds by the means of quantitative structure-activity relationship modeling, virtual screening, and experimental validation. *J. Med. Chem.* 2009; 52(14):4210–4220. [PubMed: 19537691]
28. Tang H, Wang XS, Huang XP, Roth BL, Butler KV, Kozikowski AP, Jung M, Tropsha A. Novel inhibitors of human histone deacetylase (HDAC) identified by QSAR modeling of known inhibitors, virtual screening, and experimental validation. *J. Chem. Inf. Model.* 2009; 49(2):461–476. [PubMed: 19182860]
29. Tropsha A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol. Inf.* 2010; 29:476–488.

30. Fourches D, Muratov E, Tropsha A. Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research. *J. Chem. Inf. Model.* 2010; 50(7): 1189–1204. [PubMed: 20572635]
31. Varnek A, Baskin I. Machine learning methods for property prediction in chemoinformatics: Quo Vadis? *J. Chem. Inf. Model.* 2012; 52(6):1413–1437. [PubMed: 22582859]
32. Klebe G, Abraham U, Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* 1994; 37(24):4130–4146. [PubMed: 7990113]
33. Kubinyi, H. Success Stories of Computer-aided Design.. In: Ekins, S., editor. *Computer Applications in Pharmaceutical Research and Development*. John Wiley & Sons, Inc.; Hoboken: NJ, USA: 2006. p. 377-424.
34. Wermuth CG, Ganellin CR, Lindberg P, Mitscher LA. Glossary of terms used in medicinal chemistry. IUPAC Recommendations. 1998
35. Yang SY. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discov. Today.* 2010; 15(11-12):444–450. [PubMed: 20362693]
36. Hicklin DJ, Ellis LM. Role of the vascular endothelial growth factor pathway in tumor growth and angiogenesis. *J. Clin. Oncol.* 2005; 23(5):1011–1027. [PubMed: 15585754]
37. Holmes K, Roberts OL, Thomas AM, Cross MJ. Vascular endothelial growth factor receptor-2: structure, function, intracellular signalling and therapeutic inhibition. *Cell Signal.* 2007; 19(10): 2003–2012. [PubMed: 17658244]
38. Mendel DB, Schreck RE, West DC, Li G, Strawn LM, Tanciongco SS, Vasile S, Shawver LK, Cherrington JM. The angiogenesis inhibitor SU5416 has long-lasting effects on vascular endothelial growth factor receptor phosphorylation and function. *Clin. Cancer Res.* 2000; 6(12): 4848–4858. [PubMed: 11156244]
39. FDA Approves New Treatment for Gastrointestinal and Kidney Cancer . U.S.Food and Drug administration; 6-18-2009
40. Wilhelm SM, Adnane L, Newell P, Villanueva A, Llovet JM, Lynch M. Preclinical overview of sorafenib, a multikinase inhibitor that targets both Raf and VEGF and PDGF receptor tyrosine kinase signaling. *Mol. Cancer Ther.* 2008; 7(10):3129–3140. [PubMed: 18852116]
41. Matsui J, Funahashi Y, Uenaka T, Watanabe T, Tsuruoka A, Asada M. Multi-kinase inhibitor E7080 suppresses lymph node and lung metastases of human mammary breast tumor MDA-MB-231 via inhibition of vascular endothelial growth factor-receptor (VEGF-R) 2 and VEGF-R3 kinase. *Clin. Cancer Res.* 2008; 14(17):5459–5465. [PubMed: 18765537]
42. Arora A, Scholar EM. Role of tyrosine kinase inhibitors in cancer therapy. *J. Pharmacol. Exp. Ther.* 2005; 315(3):971–979. [PubMed: 16002463]
43. Wood JM, Bold G, Buchdunger E, Cozens R, Ferrari S, Frei J, Hofmann F, Mestan J, Mett H, O'Reilly T, Persohn E, Rosel J, Schnell C, Stover D, Theuer A, Towbin H, Wenger F, Woods-Cook K, Menrad A, Siemeister G, Schirner M, Thierauch KH, Schneider MR, Dreys J, Martiny-Baron G, Totzke F. PTK787/ZK 222584, a novel and potent inhibitor of vascular endothelial growth factor receptor tyrosine kinases, impairs vascular endothelial growth factor-induced responses and tumor growth after oral administration. *Cancer Res.* 2000; 60(8):2178–2189. [PubMed: 10786682]
44. Yu H, Wang Z, Zhang L, Zhang J, Huang Q. The discovery of novel vascular endothelial growth factor receptor tyrosine kinases inhibitors: pharmacophore modeling, virtual screening and docking studies. *Chem. Biol. Drug Des.* 2007; 69(3):204–211. [PubMed: 17441906]
45. Yu H, Wang Z, Zhang L, Zhang J, Huang Q. Pharmacophore modeling and in silico screening for new KDR kinase inhibitors. *Bioorg. Med. Chem. Lett.* 2007; 17(8):2126–2133. [PubMed: 17306530]
46. Sun J, Li DD, Li JR, Fang F, Du QR, Qian Y, Zhu HL. Design, synthesis, biological evaluation, and molecular modeling study of 4-alkoxyquinazoline derivatives as potential VEGFR2 kinase inhibitors. *Org. Biomol. Chem.* 2013; 11(44):7676–7686. [PubMed: 24108070]
47. Miyazaki Y, Matsunaga S, Tang J, Maeda Y, Nakano M, Philippe RJ, Shibahara M, Liu W, Sato H, Wang L, Nolte RT. Novel 4-amino-furo[2,3-d]pyrimidines as Tie-2 and VEGFR2 dual inhibitors. *Bioorg. Med. Chem. Lett.* 2005; 15(9):2203–2207. [PubMed: 15837294]

48. Lee K, Jeong KW, Lee Y, Song JY, Kim MS, Lee GS, Kim Y. Pharmacophore modeling and virtual screening studies for new VEGFR-2 kinase inhibitors. *Eur. J. Med. Chem.* 2010; 45(11): 5420–5427. [PubMed: 20869793]
49. Rajagopalan M, Balasubramanian S, Ramaswamy A, Prakash MP. Pharmacophore based 3D-QSAR modeling and free energy analysis of VEGFR-2 inhibitors. *J. Enzyme Inhib. Med. Chem.* 2012
50. Cramer RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* 1988; 110(18):5959–5967. [PubMed: 22148765]
51. Bohm M, St r. J. Klebe G. Three-dimensional quantitative structure-activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. *J. Med. Chem.* 1999; 42(3):458–477. [PubMed: 9986717]
52. Du J, Lei B, Qin J, Liu H, Yao X. Molecular modeling studies of vascular endothelial growth factor receptor tyrosine kinase inhibitors using QSAR and docking. *J. Mol. Graph. Model.* 2009; 27(5):642–654. [PubMed: 19081278]
53. Zeng H, Zhang H. Combined 3D-QSAR modeling and molecular docking study on 1,4-dihydroindeno[1,2-c]pyrazoles as VEGFR-2 kinase inhibitors. *J. Mol. Graph. Model.* 2010; 29(1): 54–71. [PubMed: 20471293]
54. Wu X, Wu S, Chen WH. Molecular docking and 3D-QSAR study on 4-(1H-indazol-4-yl) phenylamino and aminopyrazolopyridine urea derivatives as kinase insert domain receptor (KDR) inhibitors. *J. Mol. Model.* 2012; 18(3):1207–1218. [PubMed: 21695506]
55. Kurup A, Garg R, Hansch C. Comparative QSAR study of tyrosine kinase inhibitors. *Chem. Rev.* 2001; 101(8):2573–2600. [PubMed: 11749389]
56. Cohen MH, Johnson JR, Chen YF, Sridhara R, Pazdur R. FDA drug approval summary: erlotinib (Tarceva) tablets. *Oncologist.* 2005; 10(7):461–466. [PubMed: 16079312]
57. Gefitinib (marketed as Iressa) Information. Food and Drug Administration; 2009. <http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm110473.htm>
58. Burris HA III. Dual kinase inhibition in the treatment of breast cancer: initial experience with the EGFR/ErbB-2 inhibitor lapatinib. *Oncologist.* 2004; 9(Suppl 3):10–15. [PubMed: 15163842]
59. Spaans JN, Goss GD. Trials to Overcome Drug Resistance to EGFR and ALK Targeted Therapies - Past, Present, and Future. *Front Oncol.* 2014; 4:233. [PubMed: 25221748]
60. Wissner A, Mansour TS. The development of HKI-272 and related compounds for the treatment of cancer. *Arch. Pharm. (Weinheim).* 2008; 341(8):465–477. [PubMed: 18493974]
61. Blair JA, Rauh D, Kung C, Yun CH, Fan QW, Rode H, Zhang C, Eck MJ, Weiss WA, Shokat KM. Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. *Nat. Chem. Biol.* 2007; 3(4):229–238. [PubMed: 17334377]
62. Tsou HR, Overbeek-Klumpers EG, Hallett WA, Reich MF, Floyd MB, Johnson BD, Michalak RS, Nilakantan R, Discafani C, Golas J, Rabindran SK, Shen R, Shi X, Wang YF, Upeslaciis J, Wissner A. Optimization of 6,7-disubstituted-4-(arylamino)quinoline-3-carbonitriles as orally active, irreversible inhibitors of human epidermal growth factor receptor-2 kinase activity. *J. Med. Chem.* 2005; 48(4):1107–1131. [PubMed: 15715478]
63. Zhou W, Ercan D, Chen L, Yun CH, Li D, Capelletti M, Cortot AB, Chirieac L, Iacob RE, Padera R, Engen JR, Wong KK, Eck MJ, Gray NS, Janne PA. Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature.* 2009; 462(7276):1070–1074. [PubMed: 20033049]
64. Xiang M, Lei K, Fan W, Lin Y, He G, Yang M, Chen L, Mo Y. In silico identification of EGFR-T790M inhibitors with novel scaffolds: start with extraction of common features. *Drug Des. Devel. Ther.* 2013; 7:789–839.
65. Du H, Hu Z, Bazzoli A, Zhang Y. Prediction of inhibitory activity of epidermal growth factor receptor inhibitors using grid search-projection pursuit regression method. *PLoS. One.* 2011; 6(7):e22367. [PubMed: 21811593]
66. Nandi S, Bagchi MC. In silico design of potent EGFR kinase inhibitors using combinatorial libraries. *Mol. Simulation.* 2011; 37:196–209.

67. Noolvi MN, Patel HM, Bhardwaj V. 2D QSAR STUDIES ON A SERIES OF 4-ANILINO QUINAZOLINE DERIVATIVES AS TYROSINE KINASE (EGFR) INHIBITOR: AN APPROACH TO DESIGN ANTICANCER AGENTS. *Dig. J. Nanomater. Biostruct.* 2010; 5(2): 387–401.
68. Assefa H, Kamath S, Buolamwini JK. 3D-QSAR and docking studies on 4-anilinoquinazoline and 4-anilinoquinoline epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors. *J. Comput. Aided Mol. Des.* 2003; 17(8):475–493. [PubMed: 14703120]
69. Nandi S, Bagchi MC. 3D-QSAR and molecular docking studies of 4-anilinoquinazoline derivatives: a rational approach to anticancer drug design. *Mol. Divers.* 2010; 14(1):27–38. [PubMed: 19330460]
70. Shim AH, Liu H, Focia PJ, Chen X, Lin PC, He X. Structures of a platelet-derived growth factor/propeptide complex and a platelet-derived growth factor/receptor complex. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107(25):11307–11312. [PubMed: 20534510]
71. Ostman A. PDGF receptors-mediators of autocrine tumor growth and regulators of tumor vasculature and stroma. *Cytokine Growth Factor Rev.* 2004; 15(4):275–286. [PubMed: 15207817]
72. Palmer BD, Kraker AJ, Hartl BG, Panopoulos AD, Panek RL, Batley BL, Lu GH, Trump-Kallmeyer S, Hollis Showalter HD, Denny WA. Structure-activity relationships for 5-substituted 1-phenylbenzimidazoles as selective inhibitors of the platelet-derived growth factor receptor. *J. Med. Chem.* 1999; 42(13):2373–2382. [PubMed: 10395478]
73. Pandey A, Volkots DL, Seroogy JM, Rose JW, Yu JC, Lambing JL, Hutchaleelaha A, Hollenbach SJ, Abe K, Giese NA, Scarborough RM. Identification of orally active, potent, and selective 4-piperazinylquinazolines as antagonists of the platelet-derived growth factor receptor tyrosine kinase family. *J. Med. Chem.* 2002; 45(17):3772–3793. [PubMed: 12166950]
74. Boschelli DH, Wu Z, Klutchko SR, Showalter HD, Hamby JM, Lu GH, Major TC, Dahring TK, Batley B, Panek RL, Keiser J, Hartl BG, Kraker AJ, Klohs WD, Roberts BJ, Patmore S, Elliott WL, Steinkampf R, Bradford LA, Hallak H, Doherty AM. Synthesis and tyrosine kinase inhibitory activity of a series of 2-amino-8H-pyrido[2,3-d]pyrimidines: identification of potent, selective platelet-derived growth factor receptor tyrosine kinase inhibitors. *J. Med. Chem.* 1998; 41(22): 4365–4377. [PubMed: 9784112]
75. Guha R, Jurs PC. Development of linear, ensemble, and nonlinear models for the prediction and interpretation of the biological activity of a set of PDGFR inhibitors. *J. Chem. Inf. Comput. Sci.* 2004; 44(6):2179–2189. [PubMed: 15554688]
76. Khadikar PV, Shrivastava A, Agrawal VK, Srivastava S. Topological designing of 4-piperazinylquinazolines as antagonists of PDGFR tyrosine kinase family. *Bioorg. Med. Chem. Lett.* 2003; 13(18):3009–3014. [PubMed: 12941323]
77. Birchmeier C, Birchmeier W, Gherardi E, Vande Woude GF. Met, metastasis, motility and more. *Nat. Rev. Mol. Cell Biol.* 2003; 4(12):915–925. [PubMed: 14685170]
78. Abidoye O, Murukurthy N, Salgia R. Review of clinic trials: agents targeting c-Met. *Rev. Recent Clin. Trials.* 2007; 2(2):143–147. [PubMed: 18473999]
79. Huang D, Zhu X, Tang C, Mei Y, Chen W, Yang B, Han J, Qian H, Huang W. 3D QSAR Pharmacophore Modeling for c-Met Kinase Inhibitors. *Med. Chem.* 2012; 8(6):1117–1125. [PubMed: 22779800]
80. Tai W, Lu T, Yuan H, Wang F, Liu H, Lu S, Leng Y, Zhang W, Jiang Y, Chen Y. Pharmacophore modeling and virtual screening studies to identify new c-Met inhibitors. *J. Mol. Model.* 2012; 18(7):3087–3100. [PubMed: 22203475]
81. Yuan H, Zhuang J, Hu S, Li H, Xu J, Hu Y, Xiong X, Chen Y, Lu T. Molecular Modeling of Exquisitely Selective c-Met Inhibitors through 3D-QSAR and Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 2014; 54(9):2544–2554. [PubMed: 25181449]

Fig. (1).
The general workflow for QSAR model building, internal and external validation.

Fig. (2).
The general workflow for ligand-based pharmacophore modeling.

Fig. (3).
The pharmacophore model generated by CATALYST HypoGen with hydrophobic (blue); ring aromatic (orange), hydrogen bond donor (purple), and hydrogen bond acceptor (green). [45]

Fig. (4). (Left) The illustration of pharmacophoric features identified from the top-ranked HypoGen model for CMPD1 (IC₅₀=0.003 μ M), the most active in the data set. (Right) The pharmacophoric features aligned to the CMPD28 (IC₅₀=31 μ M), the least active compound in the data set.[45]

Fig. (5).
The illustration of pharmacophoric features identified as essential for irreversible EGFR inhibition.

Fig. (6).
The common pharmacophore elements of quinazoline derivatives found for enhanced inhibitory activity.

Fig. (7).
The chemical scaffold of 4-anilino quinazoline analogs.

Fig. (8).
The chemical scaffold of 1-phenyl-benzimidazole analogs.

Fig. (9).
The chemical scaffold of piperazinyquinazoline analogs.

Table 1

The results of CoMFA/CoMSIA studies from three different research groups.

Research Group	CoMFA %contrib.		CoMSIA %contrib.				CoMSIA		CoMFA	
	<i>S</i>	<i>E</i>	<i>S</i>	<i>E</i>	<i>H</i>	<i>D</i>	<i>A</i>	<i>q</i> ²	<i>r</i> ²	<i>r</i> ²
Du <i>et al.</i> (2009)	46.5	53.5	7.4	25	28.4	20.4	18.8	0.715	0.961	0.546
Zeng <i>et al.</i> (2010)	52.8	47.2	15.2	21.6	22.4	20.5	20.3	0.569	0.942	0.600
Wu <i>et al.</i> (2012)	57.4	42.6	25.1	42.2	32.7			0.595	0.947	

*r*²: correlation coefficient; *q*²: cross-validated correlation coefficient; S, E, H, A, D: steric, electrostatic, hydrophobic as well as hydrogen bond acceptor and donor fields, respectively.