

Medicinal and Computational Chemistry Dept., ChemDiv, Inc., 6605 Nancy Ridge Drive, San Diego, CA 92121 USA, Service: +1 877 ChemDiv, Tel: +1 858-794-4860, Fax: +1 858-794-4931, Email:

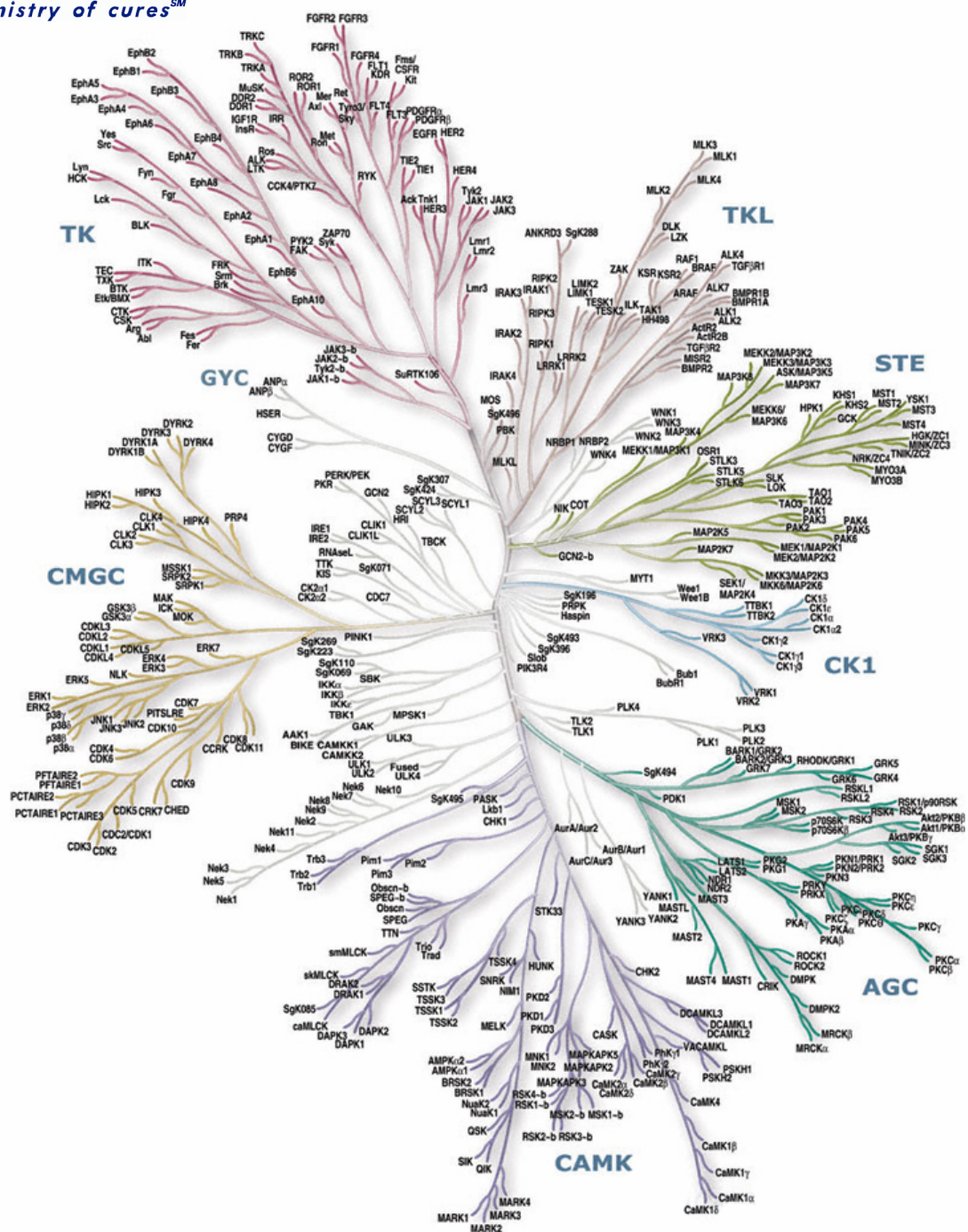
[ChemDiv@chemdiv.com](mailto:ChemDiv@chemdiv.com)

## Protein kinases

Protein kinases (PKs) are important mediators of normal cellular signal transduction. By adding phosphate groups to substrate proteins, they direct the activity, localization and overall function of many proteins, and serve to orchestrate the activity of almost all cellular processes. Protein kinases play a key role in virtually all physiological processes including proliferation, angiogenesis, migration, cell cycle, etc. The diversity of essential functions mediated by kinases is shown by the conservation of more than 50 distinct kinase families between yeast, invertebrate and mammalian kinomes. Of the 518 human protein kinases, 478 belong to a single superfamily whose catalytic domains are related in sequence. It is now recognized that abnormal phosphorylation of proteins mediated by kinases may result in diseases including cancer, diabetes, rheumatoid arthritis and hypertension, arteriosclerosis, psoriasis, and a large number of inflammatory responses [1]. The development of specific PK inhibitors as pharmacological tools and potential antiproliferative agents is an active and highly competitive area of research. The phylogenetic trees of the PK families, subfamilies and groups can be identified from the several databases [2]. Despite extensive efforts of pharmaceutical companies and academic groups, there are only a few small molecule inhibitors of protein kinases widely available as drugs. The reason for the scarcity of PK-targeted drugs is the stringent criteria required for a therapeutically useful small molecule inhibitor of these enzymes. Inhibitors need to be highly potent, selective among the closely related enzymes, and also possess adequate pharmacodynamic properties for the target of interest.

Protein kinases can be clustered into several distinct groups, families and sub-families, of increasing sequence similarity and biochemical function. The kinase dendrogram (Fig. 1) [3] shows the sequence similarity between these catalytic domains: the distance along the branches between two kinases is proportional to the divergence between their sequences. Seven major groups are labeled and colored distinctly. For instance, the tyrosine kinases (TKs) form a distinct group, whose members phosphorylate proteins on tyrosine residues, whereas enzymes in all other groups phosphorylate primarily serine and threonine residues.

Protein kinase inhibitors represent an important and still emerging class of targeted therapeutic agents. Drug discovery and development strategies have explored numerous approaches to target the inhibition of protein kinase signaling.



**Fig. 1.** The protein kinase dendrogram. Description: **TK** Tyrosine kinase; **AGC** Containing PKA, PKG, PKC families; **CAMK** Calcium/calmodulin-dependent protein kinase; **CK1** Casein kinase 1; **CMGC** Containing CDK, MAPK, GSK3, CLK families; **STE** Homologs of yeast Sterile 7, Sterile 11, Sterile 20 kinases; **TKL** Tyrosine kinase-like

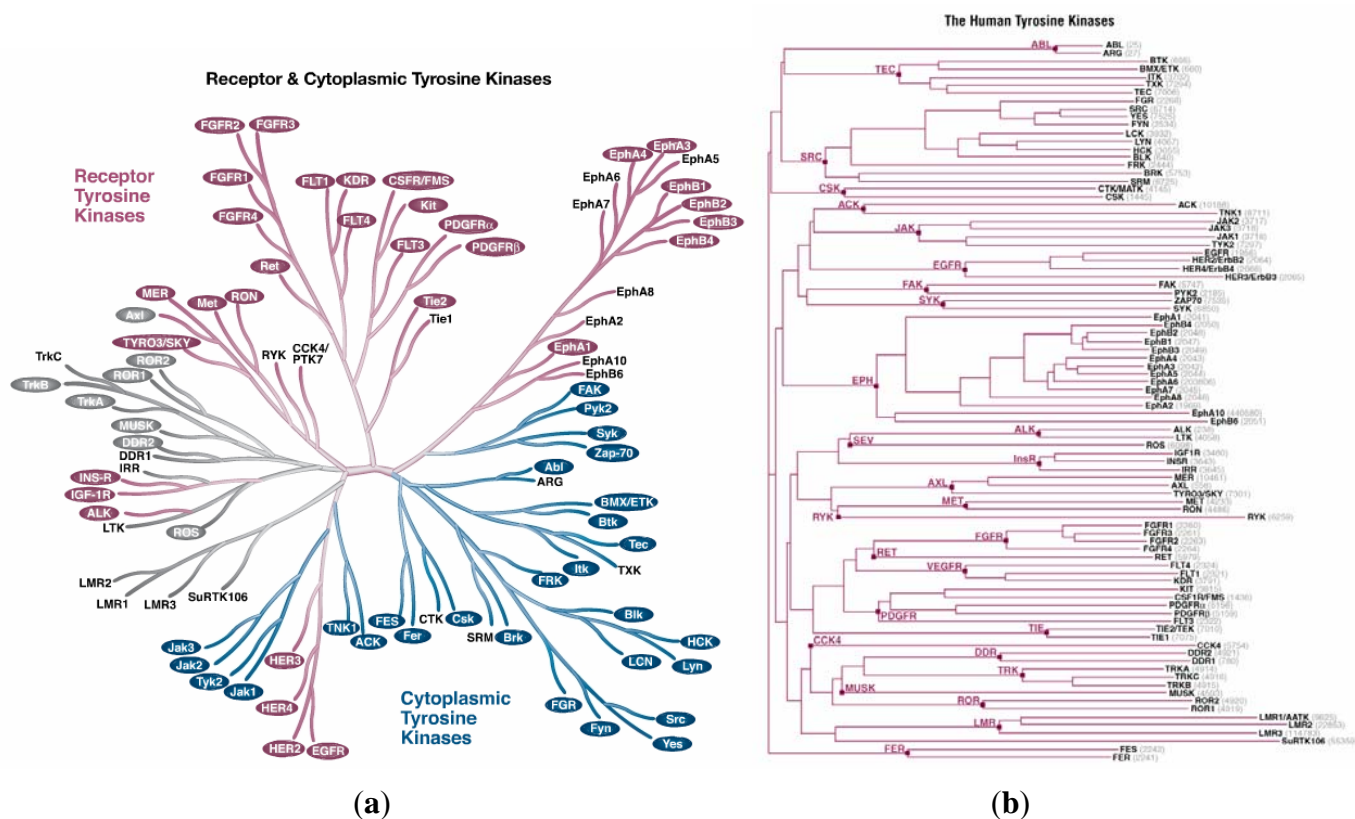
## Tyrosine kinase family

Among the PTs discovered to date tyrosine kinases seem to be the most attractive biological targets for cancer therapy, as quite often their abnormal signaling has been linked with tumor development and growth [4]. In addition, they play a critical role in other diseases, for example in inflammation [5] and rheumatoid arthritis [6].

Tyrosine kinases are known as key switches in many cellular signal transduction pathways and catalyze transfer of ATP  $\gamma$ -phosphate onto a protein substrate. Although tyrosine kinases vary in size, mechanism of activation, subunit composition, and subcellular localization, they all share a structurally

conserved A1P binding catalytic core [7], the main binding site for most of TK inhibitors. The conserved nature of this binding site represents a challenge for the selection of inhibitors. Most of TK ligands share specific steric, lipophilic, H-binding, and other parameters. The combination of these physico-chemical properties constitutes the basis for a statistical model discriminating between TK ligands and *non*-TK-active agents.

A large sub-family of TKs includes many groups which can be divided in two major classes in accordance to their localization and specificity: receptor tyrosine kinases and *non*-receptor (cytoplasmatic) tyrosine kinases (Fig. 2a,b).



**Fig. 2.** (a) tyrosine kinase dendrogram; (b) histogram which displays the genetic and morphological similarity within the TK-family enzymes

For example, three tyrosine kinase families, the Src, Tec and Syk kinase families are intimately involved in TLR signalling, the critical first step in cellular recognition of invading pathogens and tissue damage. Their activity results in changes in gene expression in affected cells. Key amongst these genes are the cytokines, which orchestrate both the duration and extent of inflammation. Tyrosine kinases also play important roles in cytokine function, and are implicated in signalling through both pro- and anti-inflammatory cytokines such as TNF, IL-6 and IL-10.

Among various groups of TKs, abl-kinase and grow-factor receptor tyrosine kinases, especially FGFR, EGFR, VGFR and IGF1R kinases, are the most promising targets particularly implicated in cancer grow and progression. Thus, constitutive activated TKs stimulate multiple signaling pathways responsible for DNA repair, apoptosis, and cell proliferation. During the last few years, thorough analysis of the mechanism underlying tyrosine kinase's activity led to novel cancer therapy using TKs blockers. These drugs are

remarkably effective in the treatment of various human tumors including head and neck, gastric, prostate and breast cancer and leukemias. The most successful example of kinase blockers is Imatinib (Imatinib mesylate, Gleevec, STI571), the inhibitor of bcr/abl oncoprotein, which has become a first-line therapy for chronic myelogenous leukemia. The introduction of STI571 for the treatment of leukemia in clinical oncology has had a dramatic impact on how this disease is currently managed. Others kinase inhibitors used recently in cancer therapy include Dasatinib (BMS-354825) specific for abl non-receptor cytoplasmic kinase, Gefitinib (Iressa), Erlotinib (OSI-774, Tarceva) and Sunitinib (SU 11248, Sutent) specific for VEGF receptor kinase, AMN107 (Nilotinib) and INNO-406 (NS-187) specific for c-KIT kinase. The following TK blockers for treatment of various human tumors are in clinical development: Lapatinib (Lapatinib ditosylate, Tykerb, GW-572016), Canertinib (CI-1033), Zactima (ZD6474), Vatalanib (PTK787/ZK 222584), Sorafenib (Bay 43-9006, Nexavar), and Leflunomide (SU101, Arava). In accordance with the examples just above, efficient tools are needed for the high-throughput search for novel candidates to be assayed as TK-targeted drugs. The key to harnessing the high therapeutic potential of TKs is in the design of high-quality small molecule libraries targeted against these proteins.

### Concept and Applications

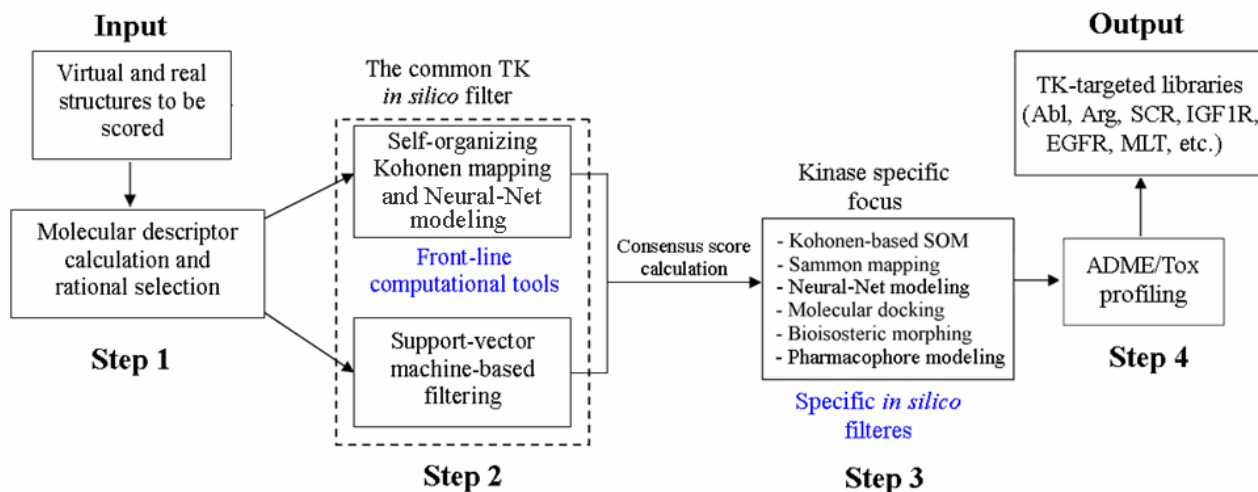
TK-targeted library design at CDL involves:

- *A combined profiling methodology that provides a consensus score and decision based on various advanced computational tools:*

1. Unique bioisosteric morphing and funneling procedures in designing novel potential TK ligands with high IP value. We apply CDL's proprietary Chemosoft<sup>TM</sup> software and commercially available solutions from Accelrys, MOE, Daylight and other platforms.
2. Neural Network tools for target-library profiling, in particular Self-organizing Kohonen maps, performed in SmartMining Software. We have also used the Sammon mapping and Support vector machine (SVM) methodology as more accurate computational tools to create our TK-focused library.
3. A molecular docking approach to focused library design.
4. Computational-based *in silico* ADME/Tox assessment for novel compounds includes prediction of human CYP P450-mediated metabolism and toxicity as well as many pharmacokinetic parameters, such as Brain-Blood Barrier (BBB) permeability, Human Intestinal Absorption (HIA), Plasma Protein binding (PPB), Plasma half-life time ( $T_{1/2}$ ), Volume of distribution in human plasma ( $V_d$ ), etc.

The fundamentals for these applications are described in a series of our recent articles on the design of exploratory small molecule chemistry for bioscreening [for related data visit ChemDiv, Inc. online source: [www.chemdiv.com](http://www.chemdiv.com)]. Our multi-step *in silico* approach to TK-focused library design is schematically illustrated in Fig. 3.





**Fig. 3.** Multi-step computational approach to TK-targeted libraries design

This common approach was effectively applied for the developing of our TK-focused, in particular for abl, VGFRs, Src, YES, ErbB, Met and IGF1R kinases.

• *Synthesis, biological evaluation and SAR study for the selected structures:*

1. High-throughput synthesis with multiple parallel library validation. Synthetic protocols, building blocks and chemical strategies are available.
2. Library activity validation via bioscreening; SAR is implemented in the next library generation.

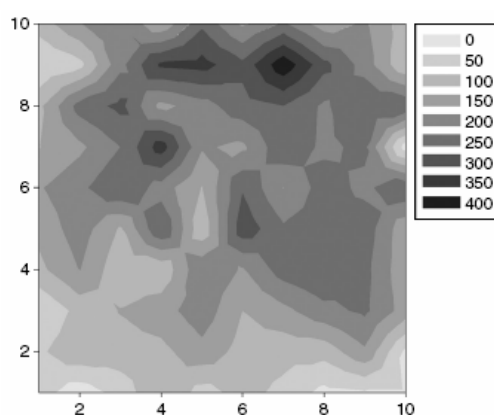
### Virtual screening on TK-specific activity

#### *The common TK-filter*

At the initial stage of our TK-targeted library design, we have collected a 22,110-compound database of known drugs and compounds entered into preclinical or clinical trials; their structures and assignments were obtained from Prous Science Integrity [8]. Each compound in this database is characterized by a defined profile of target-specific activity, focused against 1 of more than 100 different protein targets. The database was filtered based on MW (not more than 800). Molecular features encoding the relevant physicochemical and topological properties of compounds were calculated from 2D molecular representations and selected by PCA (Step 1, Fig. 3). These molecular descriptors encode the most significant molecular features, such as molecular size, lipophilicity, H-binding capacity, flexibility, and molecular topology. Taken in combination, they define both pharmacokinetic and pharmacodynamic behavior of compounds and are effective for property-based classification of target-specific groups of active agents. However, it should be noted that for each particular target-specific activity group, another, more optimal set of descriptors can be found, which provides better classification ability. As shown in Fig. 3, ‘front-line’ computational tools include Kohonen-based SOM generation as well as Neural-Net- and SVM-based modeling; these algorithms have been effectively used across the Step 2, decoded in Fig. 3.

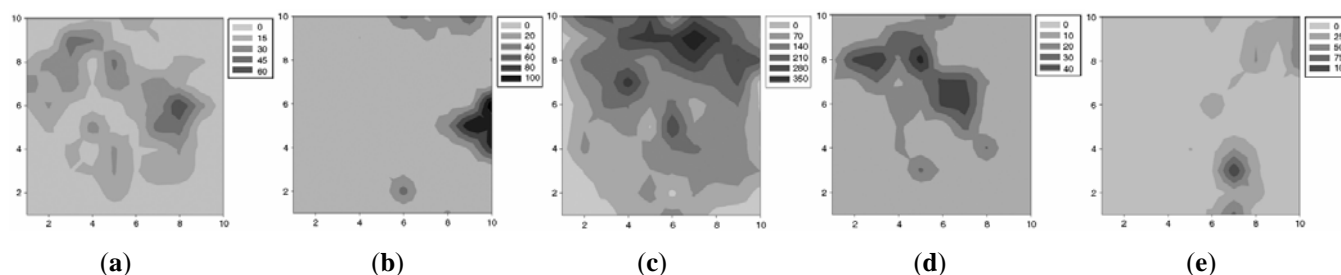
#### *Self-organizing Kohonen mapping*

A Kohonen SOM of 22,110 pharmaceutical leads and drugs generated as a result of the unsupervised learning procedure is depicted in Fig. 4. It shows that the studied compounds occupy a wide area on the map, which can be characterized as the area of druglikeness. Distribution of various target-specific groups of ligands in the Kohonen map demonstrates that most of these groups have distinct locations in specific regions of the map (Fig. 5a-e). A possible explanation of these differences is in that, as a rule, receptors of one type share a structurally conserved ligand-binding site. The structure of this site determines molecular properties that a receptor-selective ligand should possess to properly bind the site. These properties include specific spatial, lipophilic, and H-bonding parameters, as well as other features influencing the pharmacodynamic characteristics. Therefore, every group of active ligand molecules can be characterized by a unique combination of physicochemical parameters differentiating it from other target-specific groups of ligands. Another explanation of the observed phenomenon can be related to different pharmacokinetic requirements to drugs acting on different biotargets.



**Fig. 4.** Property space of 22,110 pharmaceutical leads and drugs visualized using the Kohonen map (the data have been smoothed)

The described algorithm represents an effective procedure for selection of target-focused compound subsets compatible with high throughput *in silico* evaluation of large virtual chemical space. Whenever a large enough set of active ligands is available for a particular receptor, the quantitative discrimination function can be generated allowing selection of a series of compounds to be assayed against the target. It is important to note that focusing on physicochemical rather than structural features makes this approach complementary to any available ligand structure similarity technique.

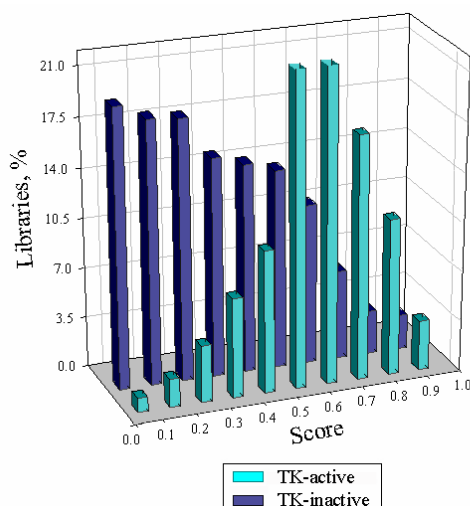


**Fig. 5.** Distribution of 5 large target-specific groups of pharmaceutical agents on the Kohonen map: (a) tyrosine kinase inhibitors (1423 compounds); (b) nuclear receptor agonists/antagonists (1122 compounds); (c) GPCR agonists/antagonists (12,711 compounds); (d) potassium channel activators (1143 compounds); (e) calcium channel antagonists (1321 compounds)

The predictive ability of the model constructed towards TK-active agents was approx. 77%. Therefore, this model can be satisfactorily used for targeted-library design and rational compound selection.

### Neural-Net modeling

Using the same knowledgebase we have further developed a property-based neural network (NN) algorithm for effective discrimination between TK inhibitors and compounds belonging to *non*-kinase activity classes. Following our strategy, 1423 known TK ligands belonging to different TK classes were used as a positive training set, TK(+). A subset of over 8592 compounds, representing over 200 various *non*-kinase based active ligands was used as a negative training set, TK(-). Using a special feature selection procedure, a 19-descriptor set was chosen for NN experiments. These descriptors encode significant molecular properties, such as lipophilicity, charge distribution, topological features, steric and surface parameters. The back-propagated feed-forward nets were constructed and trained with the molecular descriptors as input values and activity scores as output values. To assess the predictive ability of the NN models generated, we used three independent randomizations within the reference dataset which included tree groups of compounds (training, cross-validation and test group). The resulting histogram is shown in Fig. 6.



**Fig. 6.** Distribution of TK-active and TK-inactive compounds from the test set. An average predictive accuracy was 76%

The classification quality was approximately the same in each of these three independent cycles: up to 82% of TK ligands and 70% of *non*-TK ligands were correctly classified in the corresponding test sets. We carried out a wet lab experimental validation of the developed model via highthroughput screening of 5,000 compounds from the CDL corporate compound database against *abl*-kinase (*see below*). The experimental activity data (hit rate) was consistent with the expected from NN calculations, which demonstrates a high utility of NNs in designing TK-specific combinatorial libraries. The model demonstrated an enhanced level of discrimination between “active” and “inactive” libraries.

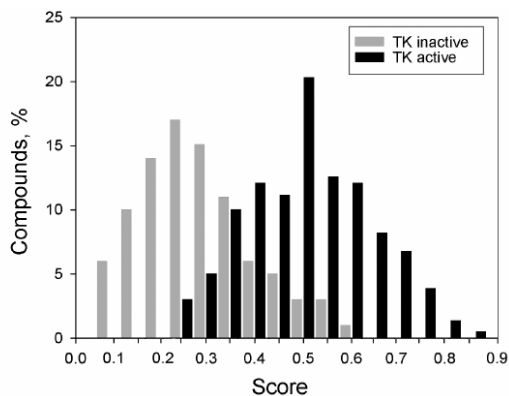
### SVM-based modeling

Recently, a so-called Support Vector Machines (SVM) [<sup>9</sup>] method has become popular as an alternative method. At least as powerful and versatile as ANNs, SVM approach is being adjusted for various application, from genomics to face recognition, including drug design [<sup>10</sup>]. Recently, we tested SVM as a classification tool in several drug-discovery programs and found it typically outperforming other approaches, in particular, ANNs [<sup>11</sup>]. Here, we used SVM algorithm for selection of compounds for primary and secondary screening against TKs.

The main parameters of the SVM-based classification model are similar to that used in NN-modeling (*see above*). Thus, as a training set, we used 1423 known TK ligands from different classes (positive training set, TK(+)), and a set of over 8592 compounds, representing over 200 various *non*-kinase active ligands (negative training set, TK(-)). All molecules were additionally filtered for molecular weight range (200–600) and atom type content (only C, N, O, H, S, P, F, Cl, Br, and I were permitted).

For the entire database of TK-active and TK-inactive structures, we have calculated sixty five molecular descriptors encoding such molecular properties as lipophilicity, charge distribution, topological features, steric and surface parameters, using ChemoSoft<sup>TM</sup>. Low-variability and highly correlated ( $R > 0.9$ ) descriptors were removed reducing the set to 39. A sensitivity analysis [<sup>12</sup>] was applied to further reduce the number of the redundant descriptors. The resulted 8 molecular descriptors (logP, no. of H-bond acceptors, no. of H-bond donors, no. of rotatable bonds, molecular refractivity, density, Zagreb index, relative positive surface area), were used for generation of the SVM classification model [<sup>13</sup>]. Before modeling each descriptor was scaled to [0;1] range (by training set; scaled values for other subsets were derived using train set scaling factors). SVM classifiers were based on linear or nonlinear (Radial Basis Functions, RBF) kernel. In our experiments, the nonlinear RBF kernel provided the best classification ability. The goodness of the model has been evaluated using an internal validation procedure. The whole set of all 10015 compounds was divided into three parts: training set (for building the SVM model), validation set (for checking model quality while generating SVM models; this set was used to check SVM models instead of leave-one-out crossvalidation, as the latter is too slow for large data sets), and the test set (for checking prediction quality of the best models). The resulting figure 7 illustrates the distributions of calculated SVM scores for compounds in TK(+) and TK(-) test sets, correspondingly. In order to assess the classification quality of the trained SVM model, we calculated percent of correctly classified compounds in each set at different threshold scores. With the threshold score 0.4, the model correctly classified up to 70% of TK(+) and 80% of TK(-) compounds). Further, we have also validated our model by calculating the SVM scores for the set of several known abl-kinase inhibitors, present neither in the training nor in the cross-validation set. It is seen, that our model correctly assigned all these active agents to the category of potential tyrosine kinase actives, as the predicted scores were in the range of 0.4-0.9.





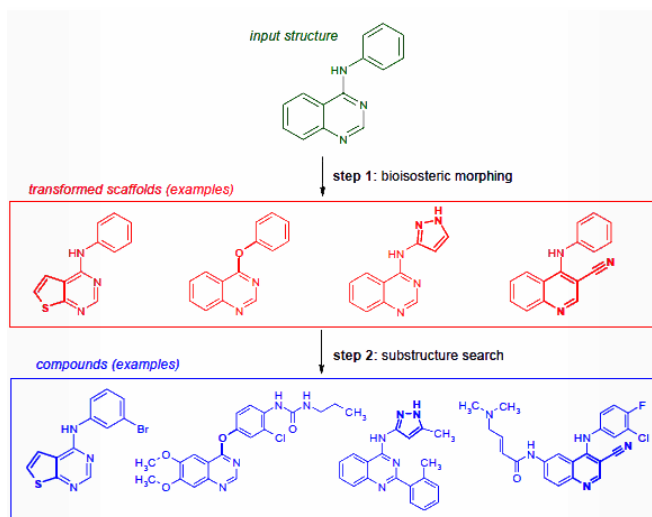
**Fig. 7.** SVM score distribution of the test set compounds. An average predictive accuracy was 75%

After the models were developed and successfully validated we have further classified the structures from our virtual library through this common *in silico* filter. Thus, based on the outputs outputted from these models we have calculated a consensus score for each compound tested. As a result, a large set of high-score structures (5,000 compounds) was collected and further evaluated using specific computational models (Step 3, see below).

#### *Specific in silico filters*

The set of the compounds selected were further expanded and tested using specific computational approaches including bioisosteric morphing, molecular docking, Sammon mapping, etc. The concept of bioisosterism is central in drug design and development [14]. The term refers to the compounds or substructures that share similar shapes, volumes, electronic distributions and physicochemical properties and have similar biological activity [15]. Bioisosteric approach is useful for morphing the marginal chemotypes. Thus, bioisosteric transformations within TK-group are illustrated in Fig. 8, in which a 4-anilinoquinazoline scaffold representing a core fragment of many potent inhibitors of the receptor tyrosine kinases, is used as an input structure.

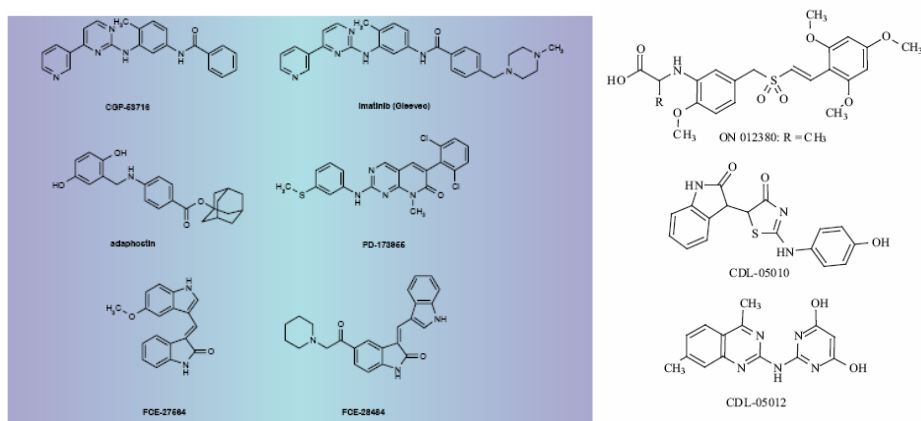
It should be particularly noted that following the original concept of diversity-oriented compound library design we have effectively applied three computational methods (*see above*) which were based solely on physicochemical descriptors, and so they provide various structures of high diversity. In turn, bioisosteric morphing generally operates within the defined and relatively narrow scope of the core/template structure of active compound. Thus, the final set included two main groups: structures which were obtained at the output of front-line filters (5,000 compounds) as well as structures generated by bioisosteric transformations within TK-active compounds. These groups were combined and gave 10,000 unique structures which were further evaluated.



**Fig. 8.** Two-step procedure for generation and search of bioisosterically transformed analogs of a 4-anilinoquinazoline scaffold (the transformed fragments are highlighted in bold). The bioisosteric morphing of active compounds was carried out with the use of the Bioisoster module of the ChemoSoft<sup>TM</sup> software.

### Abl kinase-targeted library

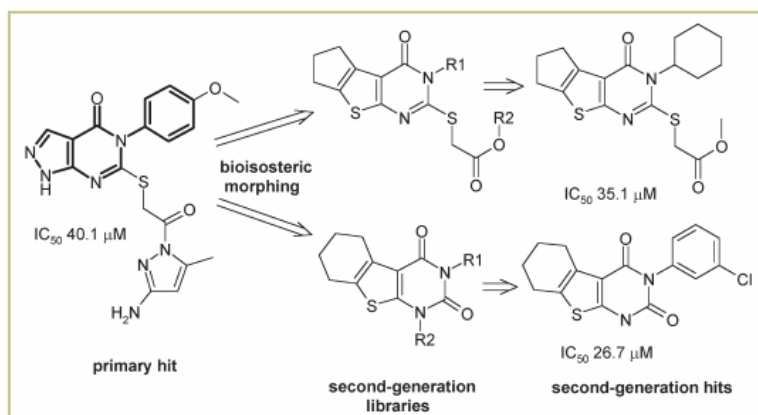
In recent work, we have developed a hands-on methodology for selection of initial library for screening against abl kinase, a therapeutically significant enzyme from TK family, and optimization of active compounds [16]. For example, Abl kinase was found to be the key regulator of chronic myeloid leukemia (CML) grow and progression. Gleevec, highly potent bcr-abl TK inhibitor, seems to be the most successful effort to date for treatment of CML. It is believed that binding of Gleevec prevents phosphorylation of Tyr393 residue in the active site, thus inactivating the enzyme with IC<sub>50</sub> of 38 nM. It is further noted that Gleevec is a potent inhibitor of two additional kinases, namely c-kit and PDGFR (both in  $\alpha$  and  $\beta$  isoforms). The structure of Gleevec and several reported bcr/abl kinase inhibitors are shown in Figure 9.



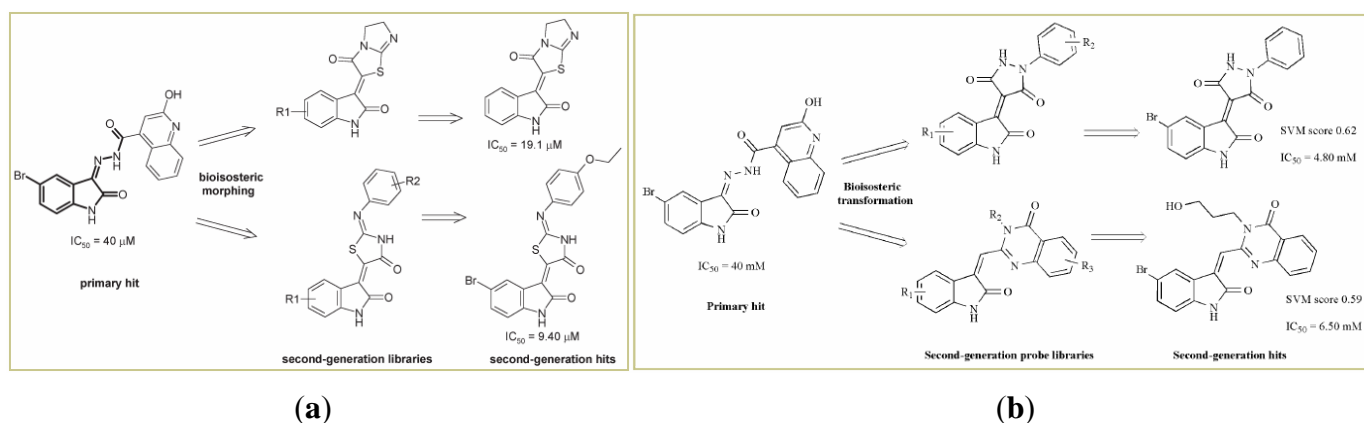
**Fig. 9.** Structures of representative abl kinase inhibitors

As we communicated in [17], a unique *in silico* classification model which was based on the differential properties of known TK inhibitors and therapeutic molecules active against other targets was developed and successfully applied for abl kinase-targeted library design. This model was primarily based on SVM classifier described above and on a computational strategy for bioisosteric morphing. We also used the later approach to design a second-generation series against the Abl tyrosine kinase. This series featured lower

overall toxicity and improved IP potential compared to the initial hits outputted from our biological screening. Some of the primary hits with a low IP potential (Fig. 10) or that contain an undesired *N*-aroylhydrazone group (Fig. 11a,b) are shown. In all cases, novel bioisosteric analogs were generated around the initial scaffolds. As a result, the overall confirmed hit rate of the secondary focused libraries was 5-10%, substantially higher than the primary hit rate (0.5%).



**Fig. 10.** Bioisosteric transformations of the primary pyrazolopyrimidine scaffold and novel active chemotypes. The queried substructure within the initial hit is shown in bold

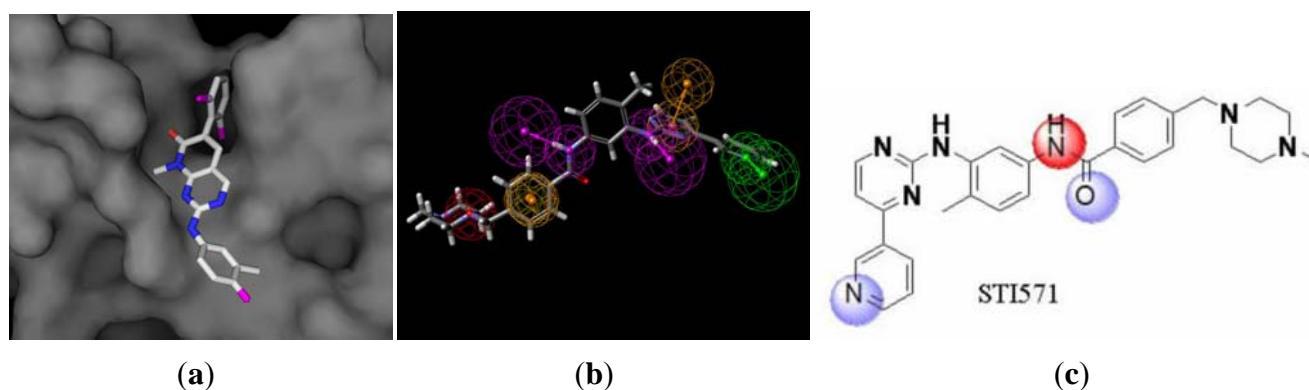


**Fig. 11.** Bioisosteric transformations of the primary 3-(*N*-aroylhydrazone)indol-2-one scaffold and novel active chemotypes. The queried substructure within the initial hit is shown in bold

Following the methodology applied, an initial round of our virtual screening against *abl* kinase was accomplished for a set of 100,000 compounds selected from our collection at Chemical Diversity Labs. Based on the structure of known *abl* inhibitors (see Fig. 9) and using a data mining algorithm which discriminates compounds according to their kinase inhibitory potential (SVM-classifier), we yielded 12,000 high-score compounds. They were further scored against the kinase using target-based approaches (*see below*) giving us a set of 550 ‘hit’ compounds for biological testing. In addition, a subset of 10,000 compounds was randomly selected from the same initial database of 100,000. As shown in figures 9 and 10, biological trials have revealed several highly potent hits which were reasonably regarded as potential lead-compounds.

### Molecular docking and pharmacophore-constrained screening

High-resolution structural information is available for numerous kinase targets. These data are invaluable for discovery of ligands with both diverse chemotypes and binding modes. Protein kinase inhibitors typically bind at the highly conserved nucleotide-binding pocket of the catalytic domain. Specific protein kinase inhibitors take advantage of limited sequence variation surrounding the ATP-binding site, as well as conformational differences between inactive and active forms of kinases. We have used a guided pharmacophore-constrained structure-based screening strategy for our focused-library design. Thus, the crystal structure of the abl kinase/PD-173955 complex was used for docking study and pharmacophore modeling (Fig. 12a). Crystallographic waters were removed and bound ligand was used to define the active site. We also assumed that no significant induced fit effects occur upon the binding and the receptor is rigid to a good approximation. Initially, we generated an active site map for abl tyrosine kinase. We subsequently produced the respective 3D pharmacophore space available to conduct virtual screening and to prioritize compounds (Fig. 12b,c). Using sets of overlapping spheres derived from the protein-ligand complex crystallographic data, the active site of a receptor can be modeled. Sphere centers were used to define atom positions of a potential ligand as well as excluded volumes. Our XCGen program generated 3D molecular conformations using standard stereochemical rules and molecular mechanics refinements. Generated conformers were used as starting points for iterative modification of molecular geometry to obtain better fit for a previously generated 3D pharmacophore. Results of this analysis were prioritized and 550 compounds with the best fit were selected for further biological screening. Notably, the program did not find a pharmacophore fit solution for more than 10,000 compounds. Very likely, these compounds are unable to bind to the active site of abl enzyme.

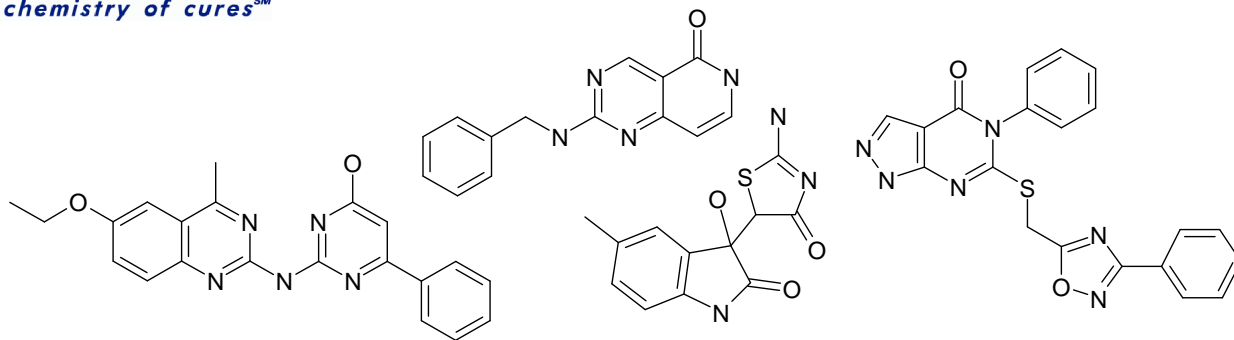


**Fig. 12.** (a) PD-173955 in the active site of abl kinase (this data was used for docking evaluation); pharmacophore models constructed based on two known abl kinase inhibitors: (b) Glivec and (c) ST1571

In addition, we have also used other computational techniques for abl kinase-focused library design. These include Sammon mapping and NN modeling. These methods are very effective for the more detailed analysis of the initially reduced set of compounds obtained after the second step of our virtual screening.

Representative compounds from our abl kinase-targeted sublibrary are shown in Fig. 13.





**Fig. 13.** Representative structures from abl kinase-targeted sublibrary

### VEGF-targeted library

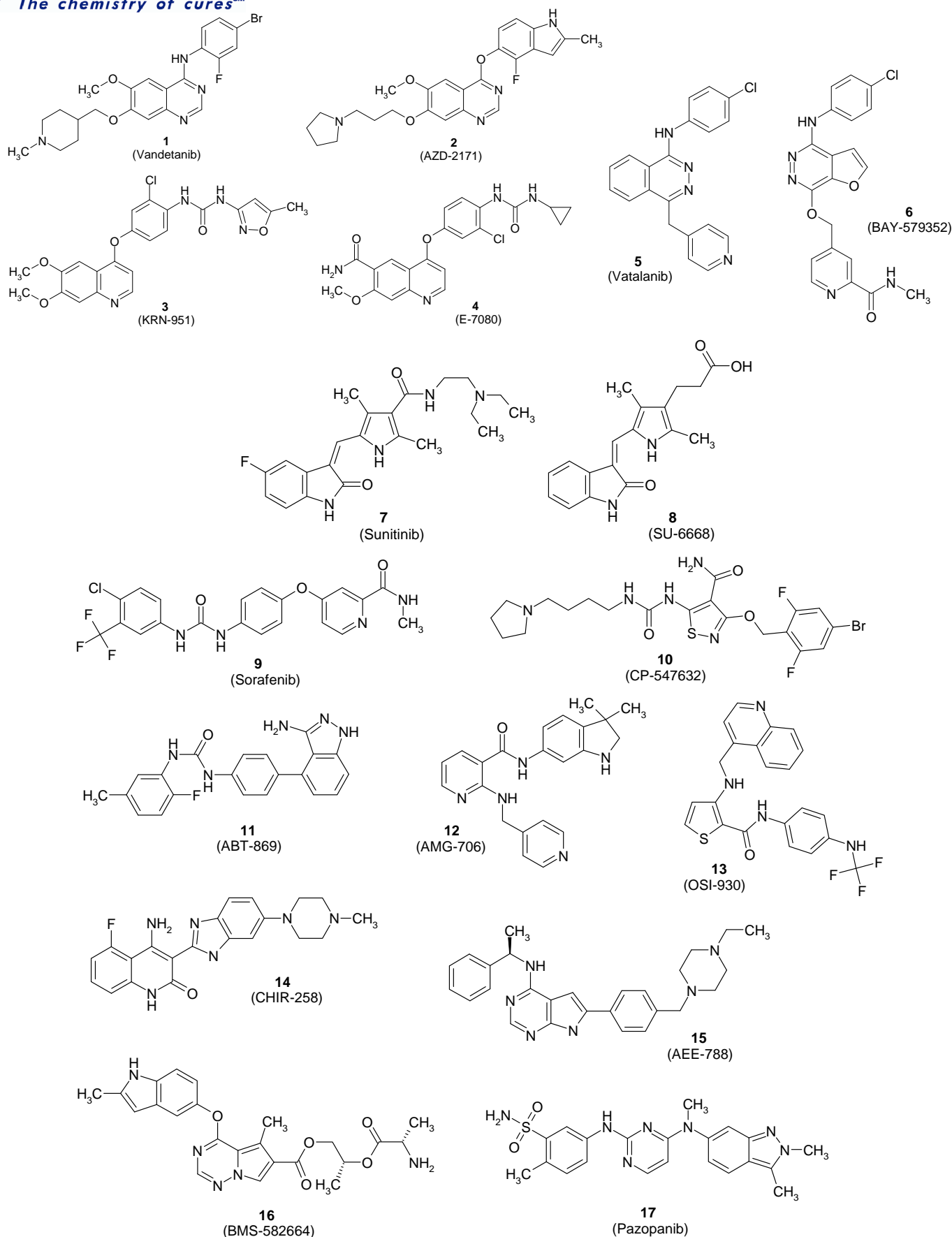
Vascular endothelial growth factors (VEGFs) and a respective family of tyrosine kinases receptors (VEGFRs) are key proteins modulating angiogenesis, the formation of new vasculature from an existing vascular network. There has been a considerable *in vivo* evidence, including clinical observations, that abnormal angiogenesis is implicated in a number of malignancies, which include rheumatoid arthritis, inflammation, cancer, psoriasis, degenerative eye conditions and others. Anti-angiogenic therapies based on inhibition of VEGF/VEGFR signaling were reported to be powerful clinical strategies in oncology and ophthalmology [18]. Current efforts have yielded promising clinical data for several anti-angiogenic therapeutics. In total, there are 4 launched drugs and more than 30 agents in development that were reported to antagonize this pathway. Fierce competition further highlights the level of interest in pharmaceutical industry to development of VEGF/VEGFR-targeting drugs. There are two main groups of such drugs, namely drugs based on biological macromolecules and small-molecule inhibitors [19]. Despite the promising clinical data obtained for therapies based on biologics, high manufacturing cost and in some instances, rapid metabolic degradation limit their clinical potential. In addition, complex structure of biological macromolecules poses challenges to their optimization. To overcome these difficulties, extensive studies of small-molecule VEGFR inhibitors have been performed in the past decade; their structures are disclosed within figure 14. As a result, two drugs were launched and a series of advanced clinical candidates are under development (Table 1). Binding of VEGFs to their receptors followed by formation of VEGFR homo- and heterodimers induces tyrosine kinase activity of VEGFR. This leads to autophosphorylation of an intracellular tyrosine residues and initiates signaling. All small-molecule agents reported to-date target VEGF signaling by inhibiting VEGFR receptor tyrosine kinase (RTKs) activity. opposed to biological therapies homed at the extracellular region of the receptors, all advanced synthetic molecules target *intracellular* ATP-binding pocket of the VEGFRs [20]. Due to the structural conservation of the ATP-binding pockets in protein kinases, these agents display high affinity for the additional members of kinome including PDGFR, Raf-kinase, ErbB family of receptors and other targets [21]. This “dual” inhibitor profile offers an intriguing possibility for disruption of several independent biological pathways vital for tumor proliferation and metastasis in the clinical setting [22]. Although in general, small molecules lack potency and specificity associated with biologics, ultimately, they may prove to be the modality of choice in achieving good balance between therapeutic window, tumor resistance, PK profile and manufacturing costs. To further illustrate the

point, several reports commented on insufficient efficacy displayed by a mono-therapies that block single angiogenic pathway [23].

**Table 1.** Marketed and late clinical development small-molecule inhibitors of VEGF signaling.\*

Drug name	Originator	Highest development phase	Diseases	Cellular target
Sorafenib	Bayer	Launched-2005	Brain, breast, colorectal, lung, endocrine, ovarian, liver, female reproductive system cancers	Flt3, C-KIT, PDGFR, Raf kinase, VEGFR-2/3
Sunitinib	Pfizer, Sugan	Launched - 2006	Breast, colorectal, endocrine, gastric, liver, non-small cell lung, prostate and renal cancers, myeloid leukemia	c-FMS, Flt3, C-KIT, PDGF, VEGFR-1/2/3
Vatalanib	Novartis	Phase III	Breast, colorectal, pancreatic, lung, prostate cancers, glioblastoma, Kaposi's sarcoma, multiple myeloma, myeloid leukemia, solid tumors	C-KIT, PDGF, VEGFR-1/2/3
Vandetanib	AstraZeneca	Phase III	Brain, breast, endocrine cancers, NSCLC, solid tumors	EGFR, FGFR, RET, VEGFR-1/2/3
AZD-2171	AstraZeneca	Phase II/III	Colorectal cancer and NSCLC	VEGFR-1/2/3
SU-6668	Sugan	Phase II	Breast and liver cancers, solid tumors	FGFR, PDGFR, VEGFR-2
CP-547632	OSI, Pfizer	Phase II	NSCLC and ovarian cancer	EGFR, PDGFR, VEGFR-2
Pazopanib	GlaxoSmithKline	Phase II	Psoriasis, multiple myeloma, ovarian and renal cancers, sarcoma	VEGFR-2
AMG-706	Amgen	Phase II	Gastrointestinal cancer and NSCLC	PDGFR, VEGFR-1/2/3
AEE-788	Novartis	Phase I/II	Glioblastoma	EGFR, HER2, VEGFR-2
E-7080	Eisai	Phase I	Solid tumors	VEGFR-2
CHIR-258	Chiron	Phase I	Multiple myeloma, myeloid leukemia and solid tumors	FGFR3, PDGFR, VEGFR-1/2
OSI-930	OSI	Phase I	Tumors	C-KIT, PDGFR, VEGFR-2
BAY-579352	Bayer	Phase I	Tumors	PDGFR, VEGFR-2
ABT-869	Abbott	Phase I	Tumors	CSF1R, ERK, Flt3, PDGFR, VEGFR-2
BMS-582664	Bristol-Myers Squibb	Phase I	Digestive/gastrointestinal cancer	FGFR1, VEGFR-2
KRN-951	Kirin Brewery	Phase I	Tumors and AMD	VEGFR-2

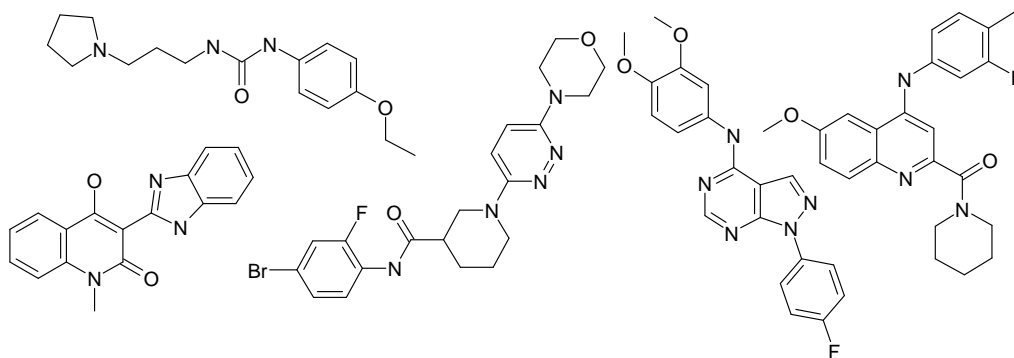
\*data as of June 2006.



**Fig. 14.** Potent small-molecule inhibitors of VEGFRs currently evaluated in advanced biological trials

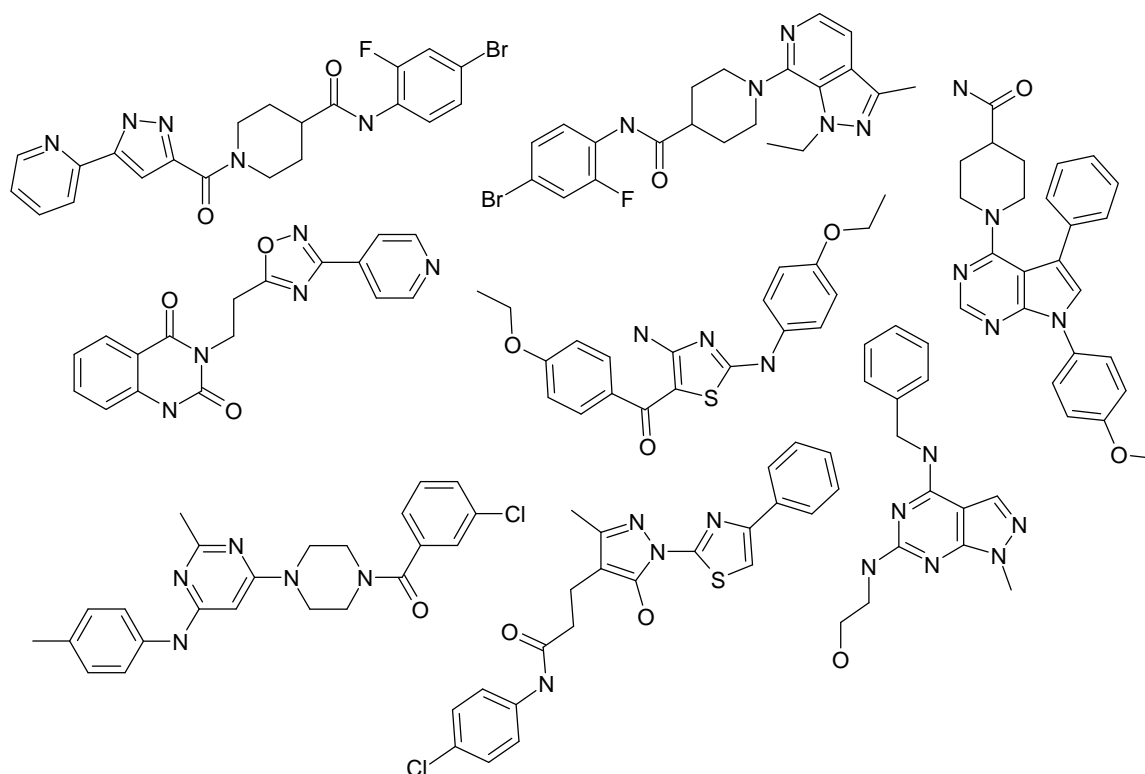
Using the methodology described above for abl kinase we have designed a unique VEGFR-targeted sublibrary based on the data collected. This data includes the structures of the whole class of TK inhibitors and especially VEGFR-active agents obtained from Prous Science Integrity database. Thus, we have

successfully applied *in silico* approaches assigned to Step 2 and Step 3 (see Fig. 3) to produce this focused sublibrary (see Fig. 15 for representative examples).



**Fig. 15.** Representative structures from VEGFR-targeted sublibrary

Finally, we have used the same *in silico* strategy to develop additional TK-focused libraries including Src, YES, ErbB, Met and IGF1R kinase-targeted sets. The representative examples of high-score structures entered in these libraries are shown within the figure below. As a result, we have selected a set of 15,000 structures which can be reasonably regarded as potential Tyrosine Kinase inhibitors (see Fig. 16).



**Fig. 16.** Representative structures from the common TK-targeted library

In summary, we have developed and effectively applied a multi-step computational approach to design of our TK-targeted library. In particular, we have successfully validated this strategy towards abl kinase and VEGFRs. The related biological trials have revealed several highly potent inhibitors, and we can confidently conclude that described *in silico* pathway represents an effective method for TK-targeted libraries design. Moreover, we provide rapid and efficient tools for follow-up chemistry on discovered hits, including single isomer chemistry, stereoselective synthesis and racemic mixture separation. The developed



libraries are updated quarterly based on a "cache" principle. Older scaffolds/compounds are replaced by templates resulting from our in-house development (unique chemistry, literature data, computational approaches) while the overall size of the library remains the same (ca. 15-26K compounds). As a result, the libraries are renewed each year, proprietary compounds comprising 50-75% of the entire set. Clients are invited to participate in the template selection process prior to launch of our synthetic effort.

## References

- <sup>1</sup> (a) Cohen P. *Curr. Opin. Chem. Biol.* 1999, 3, 459-465; (b) Shawver L.K. et al. *Cancer Cell.* 2002, 1, 117-123; (c) Hao D. et al. *Cancer Invest.* 2002, 20, 387-404; (d) Haluska P., Adjei A.A. *Curr. Opin. Investig. Drugs.* 2001, 2, 280-286; (e) Cohen, P. *Nat. Rev. Drug. Disc.*, 2002, 1, 309.
- <sup>2</sup> For example: (a) <http://www.istech.info/IDPKP/>, (b) Krupa A. et al. *Nucl. Acids Res.* 2004, 32, D153-155
- <sup>3</sup> Manning et al. (*Science*, 298, 1912-1934).
- <sup>4</sup> Pytel D, Sliwinski T, Poplawski T, Ferriola D, Majsterek I. Tyrosine kinase blockers: new hope for successful cancer therapy. *Anticancer Agents Med Chem.* 2009 Jan;9(1):66-76.
- <sup>5</sup> Page TH, Smolinska M, Gillespie J, Urbaniak AM, Foxwell BM. Tyrosine kinases and inflammatory signalling. *Curr Mol Med.* 2009 Feb;9(1):69-85.
- <sup>6</sup> Tristano AG. Tyrosine kinases as targets in rheumatoid arthritis. *Int Immunopharmacol.* 2009 Jan;9(1):1-9
- <sup>7</sup> Cohen, P. *Curr. Opin. Chem. Biol.*, 1999, 3, 459.
- <sup>8</sup> Prous Science, URL: <http://www.prous.com>.
- <sup>9</sup> Vapnik, V. *Statistical Learning Theory*, Wiley: New York, 1998.
- <sup>10</sup> (a) Burbidge, R.; Trotter, M.; Buxton, B.; Holden, S. *Comput. Chem.*, 2001, 26, 5; (b) Trotter, M.; Buxton, B.; Holden, S.B. *Measurement and Control*, 2001, 34, 235; (c) Warmuth, M.K.; Liao, J.; Ratsch, G.; Mathieson, M.; Putta, S.; Lemmen, C. *J. Chem. Inf. Comp. Sci.*, 2003, 43, 667.
- <sup>11</sup> Zernov, V.V.; Balakin, K.V.; Ivaschenko, A.A.; Savchuk, N.P.; Pletnev, I.V. *J. Chem. Inf. Comput. Sci.*, 2003, published on web, October 2003.
- <sup>12</sup> Bigus, J.P. *Data Mining with Neural Networks*, McGraw-Hill, 1996.
- <sup>13</sup> Burbidge, R.; Trotter, M.; Buxton, B.; Holden, S. *Comput. Chem.*, 2001, 26, 5.
- <sup>14</sup> (a) Burger, A. *Prog. Drug Res.*, 1991, 37, 287; (b) Patani, G.A.; LaVoie, E.J. *Chem. Rev.*, 1996, 96, 3147; (c) Olesen, P.H. *Curr. Opin. Drug Discov. Devel.*, 2001, 4, 471; (d) Chen, X.; Wang, W. *Ann. Reports Med. Chem.*, 2003, 38, 333.
- <sup>15</sup> (a) Patani, G.A.; LaVoie, E.J. *Chem. Rev.*, 1996, 96, 3147; (b) Chen, X.; Wang, W. *Ann. Reports Med. Chem.*, 2003, 38, 333.
- <sup>16</sup> (a) Sergey E. Tkachenko, Ilya Okun, Konstantin V. Balakin, Charles E. Petersen, Yan A. Ivanenkov, Nikolay P. Savchuk, Andrey A. Ivashchenko. Efficient Optimization Strategy for Marginal Hits Active Against abl Tyrosine Kinases. *Current Drug Discovery Technologies*, 2004, 1, 201-10; (b) Balakin K.V. , Tkachenko S.E. , Okun I. , Skorenko A.V. , Ivanenkov Y.A. , Savchuk N.P. , Ivashchenko A.A. , Nikolsky Y. Bioisosteric morphing in primary hit optimization *Chemistry Today* 22. 1. 15-18.

<sup>17</sup> Sergey E. Tkachenko, Ilya Okun, Konstantin V. Balakin, Charles E. Petersen, Yan A. Ivanenkov, Nikolay P. Savchuk, Andrey A. Ivashchenko. Efficient Optimization Strategy for Marginal Hits Active Against *abl* Tyrosine Kinases. *Current Drug Discovery Technologies*, 2004, 1, 201-10.

<sup>18</sup> A. Kiselyov, K.V. Balakin and S.E. Tkachenko, VEGF/VEGFR signaling as a target for inhibiting angiogenesis, *Expert Opin Investig Drugs* 16 (2007), 83–107.

<sup>19</sup> (a) KERBEL RS: A cancer therapy resistant to resistance. *Nature* (1997) 390:335-336; (b) GASTL G, HERMANN T, STEURER M et al.: *Oncology* (1997) 54:177-184; (c) SAARISTO A, KARPANEN T, ALITALO K: Mechanisms of angiogenesis and their use in the inhibition of tumor growth and metastasis. *Oncogene* (2000) 19:6122-9; (d) CARDONES AR, BANEZ LL: VEGF inhibitors in cancer therapy. *Curr. Pharm. Des.* (2006) 12:387-94.

<sup>20</sup> MANNING G, WHYTE DB, MARTINEZ R et al.: The protein kinase complement of the human genome. *Science* (2002) 298:1912-1934.

<sup>21</sup> COHEN P: The development and therapeutic potential of protein kinase inhibitors. *Curr. Opin. Chem. Biol.* (1999) 3:459-465.

<sup>22</sup> (a) RELF M, LEJEUNE S, SCOTT PA et al.: Expression of the angiogenic factors vascular endothelial cell growth factor, acidic and basic fibroblast growth factor, tumor growth factor beta-1, platelet derived endothelial cell growth factor, placenta growth factor, and pleiotrophin in human primary breast cancer and its relation to angiogenesis. *Cancer Res.* (1997) 57:963-969; (b) TAYLOR AP, OSORIO L, CRAIG R et al.: Tumor-specific regulation of angiogenic growth factors and their receptors during recovery from cytotoxic therapy. *Clin. Cancer Res.* (2002) 8:1213-1222.

<sup>23</sup> (a) ADAMS J, HUANG P, PATRICK D: A strategy for the design of multiplex inhibitors for kinase-mediated signalling in angiogenesis. *Current Opinion in Chemical Biology* (2002) 6:486-492 (author comment: a useful review describing compounds which simultaneously inhibit multiple kinase pathways); (b) CHERRINGTON JM, STRAWN LM, SHAWVER LK: New paradigms for the treatment of cancer: the role of anti-angiogenesis agents. *Adv. Cancer Res.* (2000) 79:1-38.