We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

5,400
Open access books available

133,000
International authors and editors

165M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Computational Strategies in Cancer Drug Discovery

Gabriela Mustata Wilson[1] and Yagmur Muftuoglu[2]
[1]Health Services and Health Administration,
College of Nursing and Health Professions,
University of Southern Indiana, Evansville
[2]Department of Pharmacology, Yale University, New Haven
USA

1. Introduction

Over the last 40 years since "the war on cancer" began, cancer death rates have been significantly declining. Major advances in molecular and cellular biology have led to several breakthroughs in the field of cancer research. One of the most important advances in this area was probably the identification of genes that are closely involved in cancer initiation, progression, invasion, and angiogenesis, particularly those that cause cancer, those that suppress it, and those that promote or inhibit programmed cell death (apoptosis). As a result, the rates of new diagnoses and the rates of death from all cancers combined continue to decline. The National Cancer Institute's Cancer Trends Report for 2009/2010 highlights the fact that the four most common cancers – of the prostate, breast, lung, and colorectal, specifically – have dropped considerably in the past few years (National Cancer Institute - Cancer Trends Progress Report - 2009/2010).

Toward the study and treatment of such cancers, there are 680 genes, 545 proteins and 3 RNAs associated with 102 different types of cancer that have been identified to date (National Cancer Institute - Cancer Trends Progress Report - 2009/2010). Targeting these, a total of 1370 drugs, out of which 1056 are small molecules and 314 are biologics, are either in preclinical or clinical trials or are already FDA approved (National Cancer Institute - Cancer
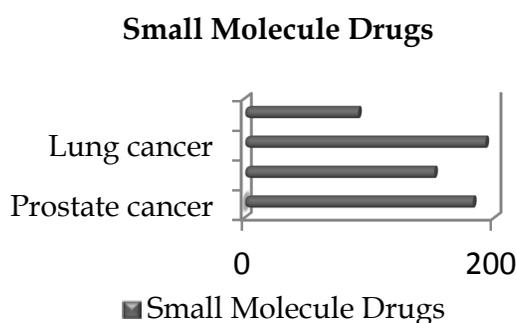

Fig. 1. Small molecule drugs developed for the four most common cancers: prostate, breast, lung, and colorectal.

Trends Progress Report - 2009/2010). As shown in Fig.1, the number of small molecule drugs developed for the four most common cancers mentioned earlier is rather impressive. This dramatic resurgence reflects not only our increasing understanding of the genes and pathways that are responsible for the initiation and progress of cancer, but also the fact that we now have access to a much more powerful range of drug discovery technologies.

The development of new anticancer drugs proves to be a very intricate, costly, and time-consuming process. The advancement and broad application of experimental methods hold enormous promise as foundations for the rapid discovery of new anticancer therapeutics. Towards this goal, computer-aided drug discovery is becoming increasingly important, given the advantage that much less investment in technology, resources, and time is required. Due to the dramatic increase of information available on genomics, small-molecules, and protein structures, computational tools are now being integrated at almost every stage in the discovery and development pipeline for drug discovery (see Fig.2).
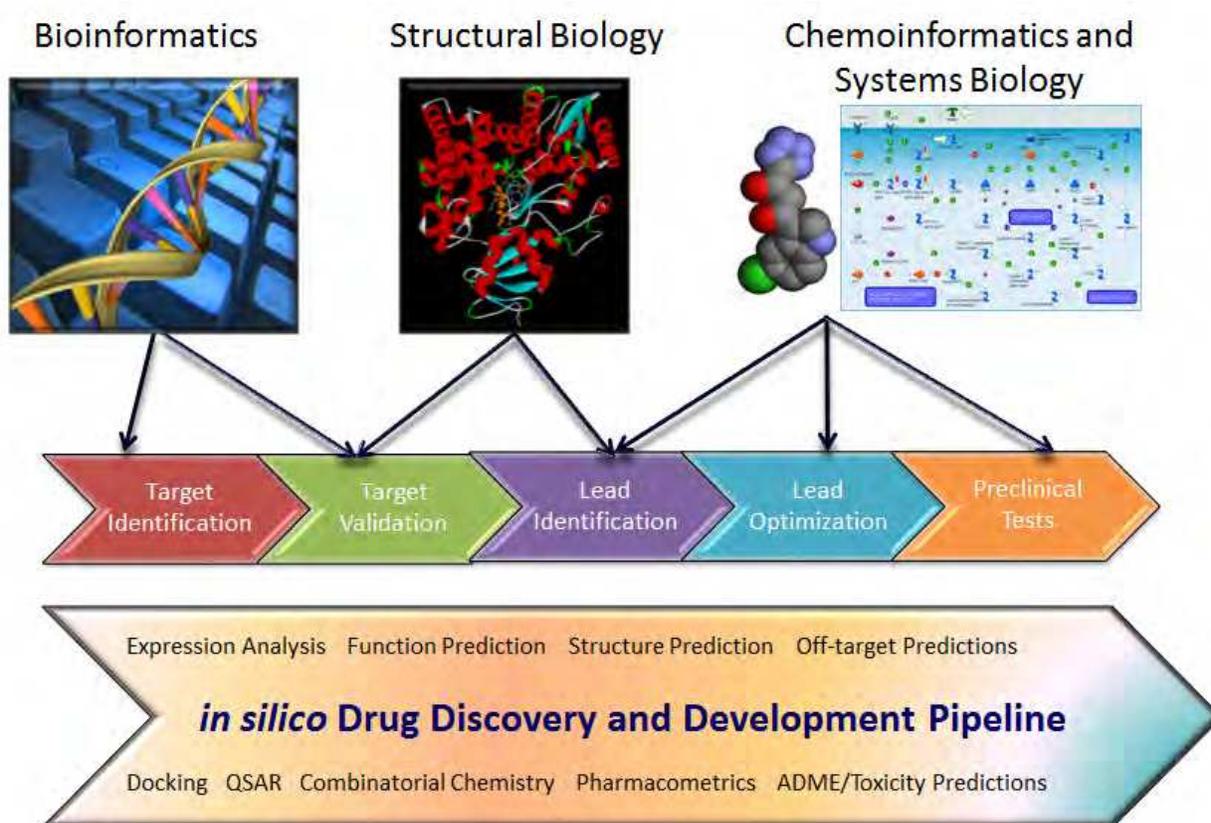

Fig. 2. Schematic diagram showing how *in silico* tools are being integrated at almost every stage in the discovery and development pipeline for drug discovery.

Acknowledging that the area of cancer therapeutics is a complex and time-consuming process, this chapter gives an overview of the computational methodologies used for rational drug design, such as ligand-based (LB) and structure-based (SB) approaches, as well as systems biology modeling. Key principles will be illustrated through case studies that explore the field of anticancer drug design to demonstrate that research advances, with the aid of *in silico* drug design and computational systems biology, have the potential to create novel anticancer drugs that will give hope to millions of cancer patients.

2. Computational strategies

The application of computational tools to drug discovery, including cancer research, has grown steadily for the past couple of years. *In silico* drug design consists of a collection of tools that help to make rational decisions at the different steps of the drug discovery process, such as the identification of a biomolecular target of therapeutical interest, the selection or the design of new lead compounds, and their modification to obtain better affinities, as well as pharmacokinetic and pharmacodynamic properties.

2.1 Structure-based approaches

The first step in developing a drug is to find a small molecule that will bind to a target protein and alter its biological functions. Target structure availability provides a good starting point for modeling target-ligand interactions using structure-based (SB) approaches (Anderson 2003; Gane and Dean 2000; Klebe 2000). Efficacious computational approaches may release the heavy burdens traditionally placed on experimental work. The goals of such methods include identifying effective modifications of existing lead compounds, as well as discovering novel lead compounds (*de novo* design). In recent years, several cases of successful applications of structure-based drug design have been reported (Combs 2007; Coumar et al. 2009; Khan et al. 2010; van Montfort and Workman 2009). Given the three-dimensional structure of a target molecule, chemical compounds having potentially high affinity for this target can be designed rationally with the aid of computational methods. Based on a binding site-derived pharmacophore model, a pattern of putative interaction sites, the results consist of a collection of virtual ligands complementary to a three-dimensional structure of the binding pocket.

A successful example of structure-based pharmacophore modeling can be found in the identification of PUMA inhibitors (Mustata et al. 2011). PUMA, the p53 upregulated modulator of apoptosis, is induced by a wide range of apoptotic stimuli through both p53-dependent and -independent mechanisms. This cancer treatment target is central in mitochondria-mediated cell death by interacting with all known antiapoptotic Bcl-2 family members (Yu and Zhang 2009). Over the years, it has become increasingly apparent that apoptosis acts as a barrier against oncogenesis. Deregulated apoptosis contributes to tumor formation, tumor progression, and impaired responsiveness to anticancer therapies, and recent studies suggest that the function of PUMA is compromised in cancer cells (Yu and Zhang 2009).

Based on the binding of BH3-only proteins with Bcl-2-like proteins, a number of approaches have been used to identify small molecules that can modulate these interactions and, therefore, inhibit apoptosis. Most of the efforts have focused on the development of Bcl-2 family inhibitors that mimic the actions of the proapoptotic BH3 domains (Cory and Adams 2002; Fesik 2005). A number of such compounds have been identified through a variety of methods, including computational modeling, structure-based design, and high-throughput screening of natural product and synthetic libraries (Zhang et al. 2007). Most notably, this approach led to the development of a potent and specific Bcl-2/Bcl-XL small-molecule inhibitor called ABT-737. Derivatives of ABT-737 are being tested in clinical trials, and several have already demonstrated effective antitumor effects in preclinical models (Bruncko et al. 2007). ABT-737 has extremely high affinity for Bcl-XL, Bcl-2 and Bcl-w, with a dissociation constant (Ki) below 1nM for each of them, but it binds poorly to Mcl-1 and A1 (Oltersdorf et al. 2005). Consequently, ABT and its analogs are found to induce apoptosis in a variety of cancer cells with overexpression of Bcl-XL and Bcl-2 in cell culture and in mice.

Motivated by the success of ABT-737 and derivatives, Mustata et al. (Mustata et al. 2011) used a structure-based approach to identify small molecular BH3 decoys or inhibitors that mimic the conserved bind surface (interactions) provided by the Bcl-2 like proteins to sequester PUMA, and therefore prevent its binding to Bcl-2-like proteins, thereby preventing apoptosis. The authors used the 3D structure of PUMA BH3 domain in complex with Mcl-1, a member of the pro-survival Bcl2-family (Day et al. 2008), to visualize and derive the most relevant protein-protein interactions and also to determine if these interactions are conserved among other Bcl-2 family members. They identified two conserved salt bridges (between Arg142 (PUMA BH3) - Asp237 (human Mcl-1) and Asp146 (PUMA BH3) – Arg244 (human Mcl-1)) and one conserved hydrophobic interaction (Leu141 (PUMA BH3) – Phe251 (human Mcl-1)).

These interactions were 'translated' into pharmacophoric features, resulting in a structure-based pharmacophore model. The model was used to screen ZINC8.0 database, which resulted in 48 hits from which they selected the 13 most promising based on *in silico* ADME/Toxicity profiling and favorable binding energies. *In vitro* and *in vivo* biological analyses concluded that ten of these inhibited PUMA-induced apoptosis at 25 µM, and eight of these inhibited PUMA-induced growth suppression in DLD1 cells using an adenovirus expressing PUMA. It was also found that, in HCT116 cells deficient in cyclin-dependent kinase (CDK) inhibitor p21 (p21-KO cells), three unreported compounds outside of the 13 originally purchased reduced PUMA-induced apoptosis and growth suppression in a significant manner when all three were added at 25 µM 15 minutes following irradiation. In this way, the authors of this study were able to identify a handful of PUMA inhibitors that displayed inhibitory activity *in vitro* and antiapoptotic activity in several relevant cell lines through the generation of an SB pharmacophore model.
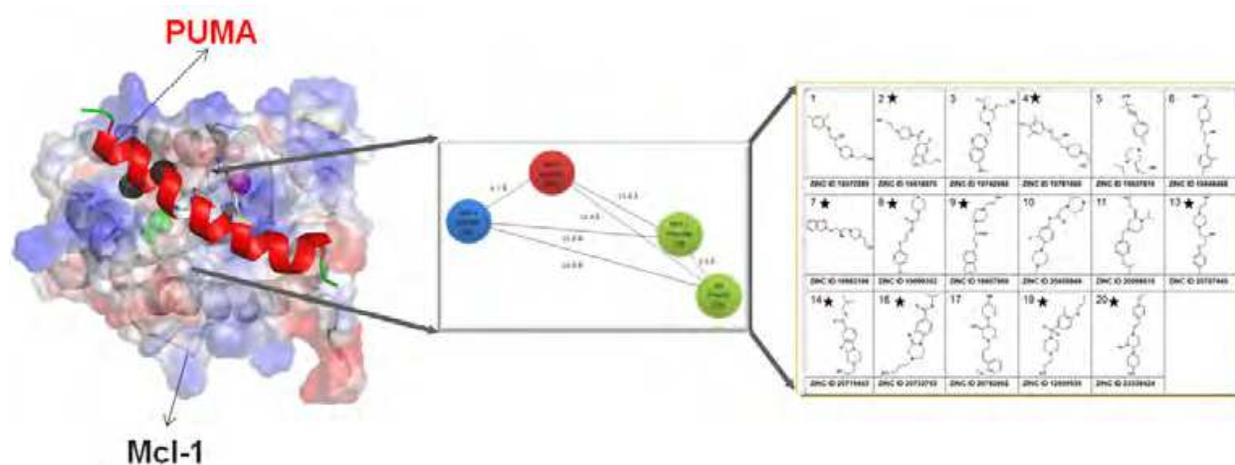

Fig. 3. Computational strategy employed towards the identification of PUMA-Bcl-2 disruptors. The complex between MPUMA and one of the Bcl2 proteins, Mcl1, is represented on the left side. The two-dimensional pharmacophore is represented in the middle, together with the inter-residue distances. Key conserved interactions derived from sequence and structural data include an Asp-Arg salt-bridge interaction (PUMA Asp146.$O_{\delta1}$ with MCl-1 Arg244.N_ε and A-1 Arg88.N_ε, blue feature), an Arg-Asp salt-bridge interaction (PUMA Arg142.NH_1 with MCl-1 Asp237.$O_{\delta2}$, red feature), and a Leu-Phe hydrophobic interaction (PUMA Leu141.$C_{\delta1}$ with MCl-1 Phe251.C_β and A1 Phe95.$C_{\delta1}$, green features). The structures of the 20 small-molecule candidates are shown on the right.

2.2 Ligand-based approaches

The information gained from structure-based studies provides a good starting point for optimizing select ligands using ligand-based approaches. Ligand-based drug design exploits information about known active compounds (and possibly about inactive compounds, as well) to discover new actives (Stahura and Bajorath 2005). Ligand-based approaches rely on the central similarity-property principal, which states that similar molecules should exhibit similar properties (Johnson and Maggiora 1990). Therefore, the activity prediction of a compound or a set of compounds will be done based on the similarity or distance to a set of reference ligands with known bioactivity to a protein target (Rognan 2007). Different types of two- and three-dimensional molecular descriptors, features, and substructures in combination with a variety of classification schemes – such as recursive partitioning, Bayesian statistics, neural networks, or other machine-learning methods – have been used for this purpose.

Pharmacophore-based virtual screening can be viewed as the intersection between structure-based and ligand-based approaches, as either the protein structure or known ligands can be used as references to build the models (i.e. receptor-based pharmacophore and ligand-based pharmacophore). One advantage of similarity searching over a pharmacophore-based search is that it does not require a set of structurally unrelated compounds of similar biological activity to derive a model. Thus, similarity-based virtual screening has proven very convenient, as it is computationally inexpensive and requires relatively less information (Sheridan and Kearsley 2002).

For similarity searching, even one active molecule can be used to search a database for related compounds. It mostly uses 2D descriptors, also called topological descriptors, which are derived from the connectivity table of the molecule and take into account distances among atoms in terms of number of bonds in the shortest path between them. The most commonly used descriptors are topological fingerprints (Hert et al. 2004), which encode the presence or absence of substructural fragments in molecules in a binary fingerprint without taking into account the number of occurrences of the feature.

These fingerprints can be pre-calculated and compared, usually by means of their Tanimoto distance, in a very fast and efficient manner to any reference set. The encoded substructures can either be a predefined list common to all sets of molecules analyzed or a list that depends on the analyzed set, in which all the encountered substructures up to a certain path length are considered. Contrary to similarity searching, pharmacophore searching has perhaps proven the most widely applied virtual screening method in terms of novel lead discovery, with hit rates for selected data sets of 1 to 20% (Good et al. 2000). Conversely, substructure-based fingerprints methods provide poorer scaffold hopping, as they are based in common substructure searching.

An interesting example of small-molecules designed using a ligand-based approach is the case of tubulin inhibitors (Chiang et al. 2009). Tubulin polymerization, an essential component of cell cycle progression and cell division, represents an important target in anticancer therapy. Several antimitotic agents – like vinblastine and colchicine – have already been discovered and are clinically used, despite often low bioavailability, significant toxicity, rapid acquired resistance, and the resulting overexpression of drug-resistant pumps that eject these antimitotic inhibitors from the cell. However, due to these unfavorable properties, researchers have devoted substantial effort to discover new agents with more tolerable and effective properties, especially since it is believed that antimitotic agents could work to diminish blood supply to cancerous tumors. The authors of this study based their

model generation on a set of 21 indole-derivatives synthesized originally (Liou et al. 2006) for potential tubulin inhibition by this research group and used structure-activity relationship (SAR) analysis to drive it. These compounds were chosen such that their inhibitory half-maximal concentration (IC_{50}) values spanned over three orders of magnitude, from 1.2 nM to 6 µM (Liou et al. 2006).

Based on the chemical similarities of these compounds, the authors selected four common pharmacophoric features, including a hydrogen bond donor (HBD), a hydrogen bond acceptor (HBA), hydrophobic group (HY), and a hydrophobic aromatic group (HYA). They also used the HypoRefine feature to generate excluded volumes (EX) in an attempt to separate inactive compounds from active ones (Fig. 4). Following validation of their most significant pharmacophore hypothesis, the authors then used the hypothesis to screen both the ChemDiv database and an in-house database of approximately 130,000 compounds. Although the authors do not report the resulting number of hits, they note that the top 1000 hits were then made subject to visual inspection, ruling out all but 142 of them.

These compounds were then biologically tested using the human oral squamous carcinoma KB cell line. From among these 142 biologically tested compounds, four, shown in Fig. 4, were found to inhibit the KB cell line with IC_{50} values of 187 nM, 2.0 µM, 3.0 µM, and 5.7 µM, respectively. The most potent compound in this set of four active molecules was also found to inhibit the proliferation of other cancer cell lines like MCF-7, NCI-H460, and SF-268, giving IC_{50} values of 236 nM, 285 nM, and 319 nM.
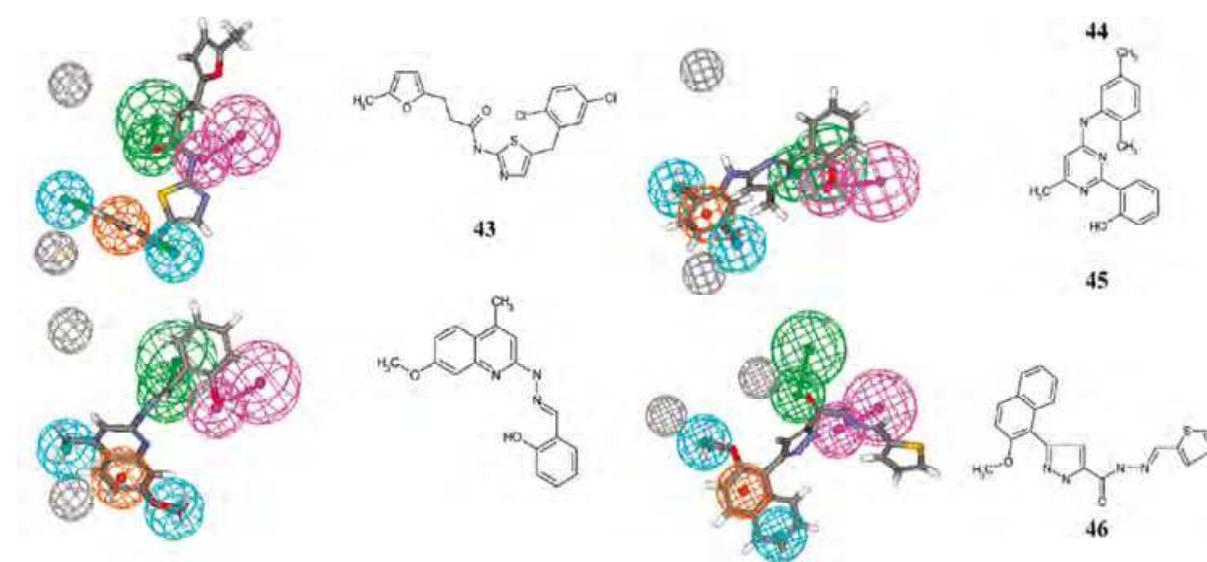

Fig. 4. The mapping configuration of four hit compounds found to inhibit the KB cell line onto the pharmacophore model along with their chemical structures. The pharmacophore model is represented by the colored spheres: HBD (magenta), HBA (green), HY (cyan), HYA (dark orange), and EX (black). Reprinted from the Journal of Medicinal Chemistry, 2009, 52, 4221-4233, Copyright 2009 American Chemical Society.

Another example of an LB approach is the development of a 3-dimensional pharmacophore model generated utilizing a set of known inhibitors of c-Myc-Max heterodimer formation (Mustata et al. 2009). c-Myc is a member of the basic helix-loop-helix leucine zipper protein

family (bHLH-ZIP). Dimerization with another bHLH-ZIP protein, Max, controls biological functions like apoptosis, transcriptional activation, and cellular transformation. Inhibition of the interactions between these two important proteins therefore represents yet another means for potential cancer treatment. This is especially true since deregulation of the c-Myc oncogene is a hallmark of cancer-related abnormalities leading to especially aggressive tumors in the cervix, colon, lung, breast, and hematopoietic organs (Nesbit et al. 1998).

Previously, the authors of this study showed through nuclear magnetic resonance (NMR) studies that inhibitors of the c-Myc-Max dimerization function by binding to the inherently disordered c-Myc protein, altering its structure and making it incapable of dimerization with the Max protein (Hammoudeh et al. 2009).

For this study, they aimed to develop an LB model based on previous work in an effort to discover new classes of potential c-Myc-Max inhibitors. The authors made use of the Genetic Algorithm with Linear Assignment for Hypermolecular Alignment of Datasets (GALAHAD) in SYBYL 8.0 software (Shepphird and Clark 2006) to develop a ligand-based pharmacophore model that takes into account ligand flexibility, steric overlaps, and strain energies. Model generation was based on a set of six c-Myc-Max inhibitors that these authors had previously reported – composed of the original compound, which they called 10058-F4, and five compounds derived from it – and gave 20 pharmacophore model hypotheses. The model that received the best overall score was found to contain two hydrophobic features, two acceptor atoms, and one donor atom; this is shown in Fig.5.
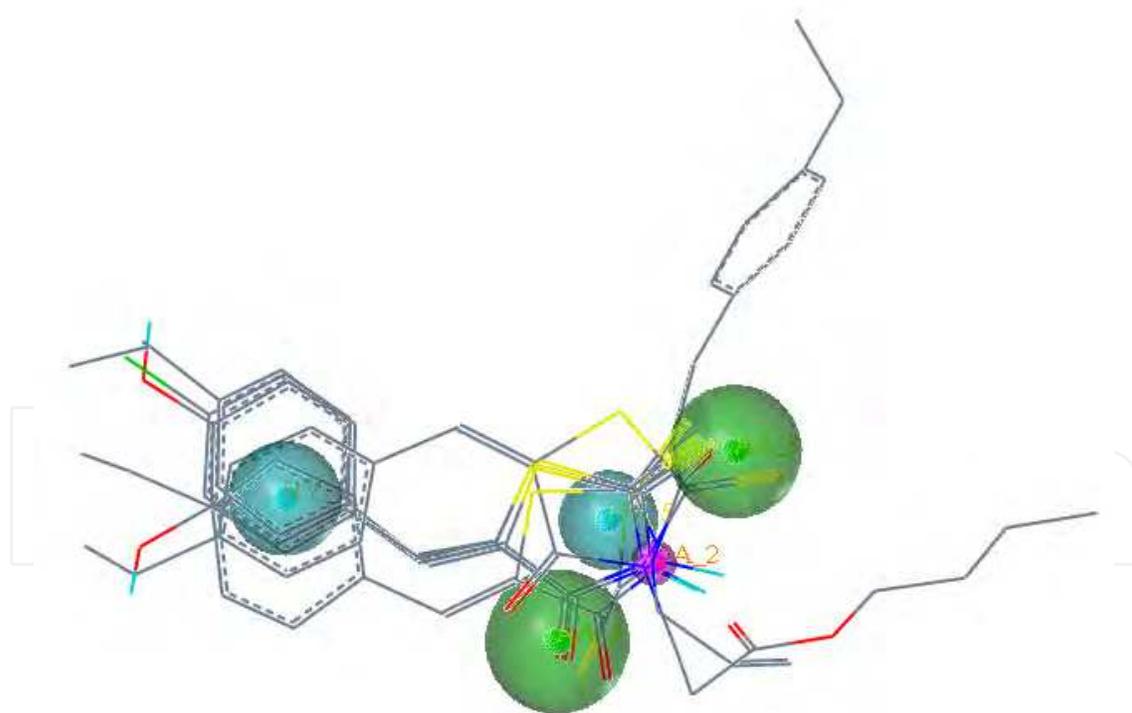

Fig. 5. GALAHAD model obtained from six compounds in the biological data. It includes two hydrophobes (light blue), one donor atom (purple), and two acceptor atoms (green)., and the sphere sizes indicate query tolerances.

By making use of two inactive analogues of the original compound, 10058-F4, the authors of this study were also able to refine the highest-scoring pharmacophore model generated

through their work. This was accomplished using the Tuplets function in SYBYL 8.0 (Shepphird and Clark 2006) software. The refined LB model was then tested using a set of ten compounds composed of four inactive analogues of the original compound 10058-F4 and the six compounds described above. This validation step resulted in one false negative, one false positive, and eight correct identifications.

The refined and validated LB model was then used to screen the set of approximately five million drug-like compounds of the ZINC 7.0 database, resulting in 15,822 hits, which amounts to about 0.31% of the ZINC 7.0 drug-like molecule database. Since the Tanimoto score of this set was found to be 0.50, it was determined that it was, indeed, composed of structurally-diverse molecules based on similar biological activity data referenced by this study.

The 100 top-ranking compounds among the 15,822 hits that resulted from the database screening were then tested by adsorption, distribution, metabolism, excretion, and toxicity (ADME/Tox) analyses using ADME Boxes version 4.0 software (Mustata et al. 2009) in an attempt to rationally devonvolute the extensive hit-list. The highest-ranking 30 compounds that passed the ADME/Tox filtering steps were then analyzed for their predicted probability of being an inhibitor of or metabolite for the cytochrome P450 isoform CYP3A4. The authors rationalized that this was necessary since the original compound 10058-F4 was previously found to be extensively metabolized, resulting in rapid clearance, and that the CYP3A4 is believed to metabolize more than 50% of drugs in the human body.

From among the compounds that passed the ADME/Tox filters, nine were purchased from ChemBridge and tested *in vitro*. At a concentration of 200 µM, four of these compounds were found to completely inhibit c-Myc-Max, three were found to partially inhibit it, and the remaining two were found to be inactive. The four highly effective compounds identified through this initial test were then determined to have IC_{50} values about two- to ten-fold lower than that of the original parent compound 10058-F4. These compounds were also tested using fluorescence polarization, and this indicated that they were able to bind to the c-Myc protein at the same location as 10058-F4. The authors also tested these four compounds in HL60 cells in a manner described in their previous work, finding compounds 5360134 and 6370870 to be more active than 10058-F4, with IC_{50} values of 23, 16.7, and 35 µM, respectively.

This study respresents the first report of a pharmacophore model that provides a hypothetical picture of the main chemical features responsible for the activity of c-Myc-Max heterodimer disruptors. The authors successfully identified a set of structurally diverse compounds that showed affinities in the micromolar range and inhibitory activity against the growth of c-Myc-overexpressing cells, therefore demonstrating the applicability of ligand-based pharmacophore modeling to the identification of novel and potentially more puissant inhibitors of the c-Myc oncoprotein (Mustata et al. 2009).

The same group also recently identified the binding site and determined the conformation by which the parental compound, 10058-F4, binds c-Myc and stabilizes the intrinsically disordered monomer over the highly ordered c-Myc-Max heterodimer (Follis et al. 2008; Follis et al. 2009). Docking of the top two newly identified Myc-Max heterodimer disruptors revealed similar binding mode to the 10058-F4 parental compound (Fig.5), centered around residues 402-412 (Follis et al. 2008; Follis et al. 2009). This information should allow the authors to define the major structural determinants of affinity/specificity.
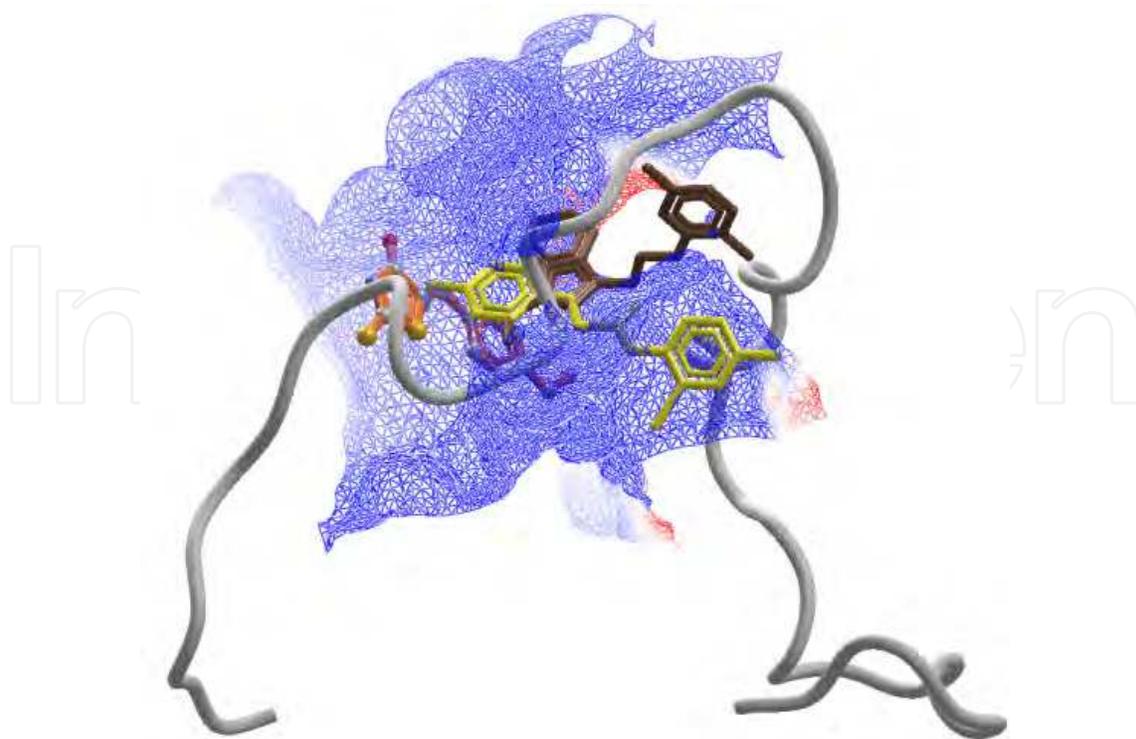
Fig. 6. Docking of the two most active compounds, 5360134 and 6370870, to the c-Myc fragment that binds the parenteral compound 10058-F4. The electrostatic interaction surface at the binding site region is displayed and colored red for negative charge and blue for positive charge. Docking simulations were performed using Molegro Virtual Docker, taking into account side chain flexibility for all residues (Thomsen and Christensen 2006).

Another successful example of LBDD is the targeting of I Kappa B Kinase β (Noha et al. 2011). The NF-κB signaling pathway, which is activated by tumor necrosis factor-α (TNF-α), stimulates through a complex signaling cascade that leads to the transcription of proinflammatory target genes, including I kappa B kinase β (IKK-β), the expression of which may promote tumor growth in the human body. As a result, IKK-β, a key player in this pathway, represents yet another potential target for the treatment of cancer, in addition to inflammation . Since previous attempts at molecular docking that have used the active site of homology models for IKK-β were riddled with uncertainty due to the lack of an accurate three-dimensional X-ray crystal structure (Noha et al. 2011) the authors of this study decided to use LB pharmacophore modeling to identify new compounds with affinity to IKK-β.

The LB pharmacophore model for this study was based on a set of five compounds with high activity (i.e. IC_{50} values of 100 nM or less) and at least a several-fold difference in selectivity for IKK-β over IKK-α in an attempt to develop an IKK-β inhibitor-specific pharmacophore model. Pharmacophore model hypotheses were built using the "HipHop algorithm" of the Catalyst 4.11 (Patel et al. 2002) software built into the DiscoveryStudio 2.1 package. The highest-ranking LB model hypothesis was iteratively refined with shape constraints and exclusion volume spheres (XVOLs) using the "refinement algorithm" available in DiscoveryStudio as "steric refinement with excluded volumes" based on two biologically inactive compounds from the literature. The final LB model contained one hydrogen-bond acceptor (HBA), one hydrophobic (H), one aromatic ring (RA), and two

hydrogen-bond donor (HBD) features. The model was further refined using a dataset extracted from the literature of 44 biologically inactive compounds, 128 active compounds, and 12,775 diverse random decoy compounds.

The refined model was then used to screen the National Cancer Institute (NCI) compound database using the "FAST" algorithm on DiscoveryStudio with the "fast flexible" search function to generate a maximum of 100 conformations per molecule in the database. Out of 247041 compounds in the database, 1860, comprising 0.8% of the entire set, were identified as hits. In an attempt to select for the most relevant compounds from among these extensive hits, the Rapid Overlay of Chemical Structures (ROCs) algorithm (Moffat et al. 2008) was used to analyze the compounds extracted from the NCI database. A combined scoring function, composed of the ROCS "color force field" score and the shape Tanimoto coefficient, was used in this form of three-dimensional similarity-based hit ranking. Two very active and structurally-diverse compounds from the literature training dataset were chosen for the shape-based screening and were passed through the three-dimensional geometry generation capabilities of CORINA from the Molecular Networks program (Sadowski et al. 2003). The top ten high-scoring compounds were tested *in vitro*, and it was found that the most potent inhibitor from among these ten, compound NSC 719177, could inhibit IKK-β with an IC_{50} value of approximately 6.95 μM (Noha et al. 2011). Cell-based analyses were also conducted to test the ability of compound NSC 719177 to inhibit NF-κB activation in HEK293 cells stably transfected and carrying a luciferase reporter gene activated by a promoter composed of multiple copies of the NF-κB response element. Compound NSC 719177 was found to have a cell-based assay IC_{50} value of approximately 5.85 μM and exhibited dose-dependent activity in inhibiting TNF-α-induced luciferase activity. Therefore, Noha and colleagues (Noha et al. 2011) were able to demonstrate the successful application of ligand-based approaches to the identification of low micromolar inhibitors against IKK-β.

2.3 Combined structure-based and ligand-based approaches

The studies discussed earlier are only a few, and by no mean the most, representative examples for cancer research. Nevertheless, it is clear that structure-based and ligand-based drug design approaches have a great impact on the discovery of anticancer drugs, and the combination provided by these complementary computational methods are even more valuable. One recent example of how both methods could be integrated in cancer research is the development of a "merged" pharmacophore model for aromatase reported by Muftuoglu and Mustata (Muftuoglu and Mustata 2010) toward the identification of better breast cancer drugs.

Excluding cancers of the skin, breast cancer is the most frequently diagnosed cancer in women (2011) and ranks second as a cause of cancer death (after lung cancer). Currently, one of eight American women has the chance of having invasive breast cancer some time during her life. Approximately two-thirds of breast cancer tumors are hormone dependent, requiring estrogens to grow (Brueggemeier et al. 2005). One approach in treating hormone-dependent cancer is to interfere with endogenous hormone production. Aromatase has always been considered the most promising target for the endocrine treatment of breast cancer (Meunier et al. 2004; Meunier et al. 2004) because, by inhibiting the aromatase enzyme, the estrogen production is decreased and the tumor growth stopped or reduced.

The recent study reported by Muftuoglu and Mustata (Muftuoglu and Mustata 2010) is the first study aimed to develop a pharmacophore model, based on both ligand and structural

information, to screen for new classes of effective AIs. The model developed was generated through a merging of a ligand-based model and a structure-based model.

The structure-based pharmacophore model was generated using the recent X-ray structure deposited in the Protein Data Bank (PDB) [pdb code: 3EQM] (Ghosh et al. 2009), which contains the chemical features important for androstenedione–enzyme interaction, while the LB model was developed from the most comprehensive list of non-steroidal aromatase inhibitors. Development of the final model, named the Merged Model (see Fig.7), required a database of both active AIs and inactive compounds for use in the training and testing sets, in addition to structural information based on the X-ray crystal structure of aromatase. As such, two test sets were created: a training set, to develop the LB model, and a test set, to validate the LB, SB, and Merged Models. The training set was composed of 20 of the most active AIs found in the literature.

The test set, on the other hand was composed of 36 slightly less active AIs and nine inactive compounds, also found in the literature. Both sets, however, represented a collection of structurally very diverse compounds. The LB model, as shown in Fig.7, was generated using the Conformation Import function and the PCHD scheme from the Pharmacophore Elucidation function in MOE (Chemical Computing Group 2008), and resulted in four pharmacophoric features: two hydrophobic/aromatic groups, one hydrogen bond acceptor, and one hydrogen bond acceptor projection. In addition, an SB model was also generated based on the x-ray crystal structure of aromatase using LigandScout software (Wolber and Langer 2005), which included excluded volume areas to reflect possible steric hindrances. The resulting SB model, shown in Fig.7, was composed of three chemical features – one hydrophobic/aromatic group and one hydrogen bond acceptor – in addition to 11 excluded volume spheres. The two models, LB and SB, were then merged into one, named the "Merged Model" (see Fig.7) to capture both types of information (i.e., known active and inactive inhibitors as well as the structure of the enzyme).
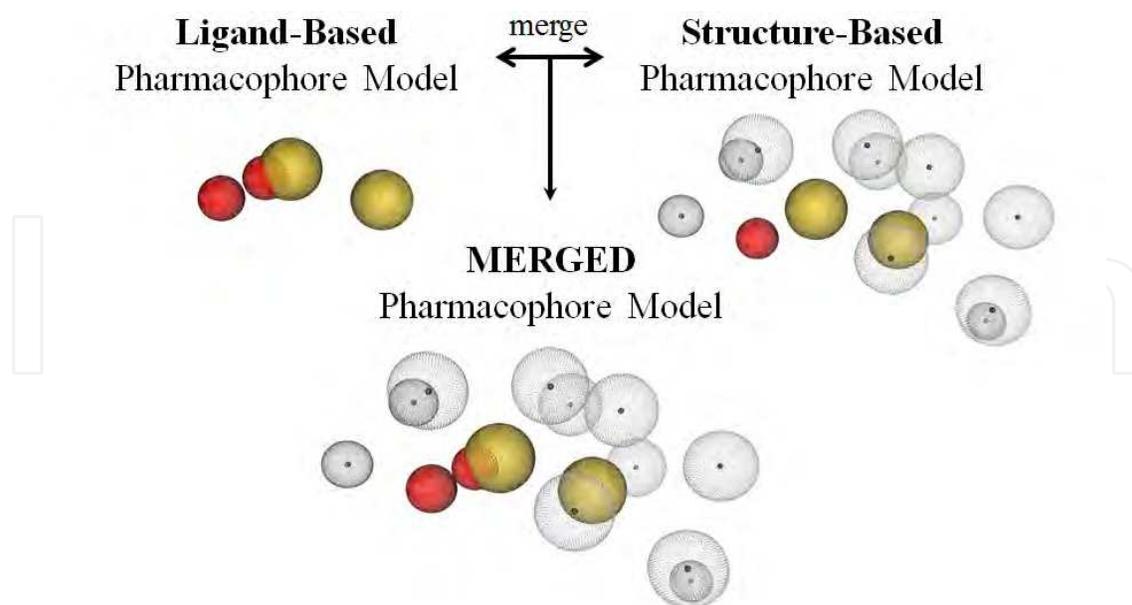

Fig. 7. Schematic diagram depicting the general methodology for the development of the Merged Model based off of the original SB and LB models. Hydrophobic groups are shown as light yellow spheres, hydrogen-bond acceptors are shown as red spheres, and excluded volumes are shown as gray spheres.

All three models were also computationally validated using the active and inactive compounds of the test set. In the end, it was found that the LB model is more discriminating than the SB Model, while the SB Model performs better in identifying active inhibitors. The Merged Model, however, was found to combine the strengths of each, proving superior to both original models. The authors of this study, therefore, would choose the Merged Model for screening of virtual libraries, although this piece of their work is not yet published.

The authors do, however, discuss another area of analysis they pursued – virtual docking – by explaining the theory behind a three-stage process that focuses in on the most stable binding conformation for a ligand based on optimization and individualized pose selection. Stage One governs the entry of the ligand into the binding pocket, where all AIs are believed to bind and some are experimentally proven to be binding. Stage Two allows for the prediction of each ligand's coordinating heteroatom, since it is known that non-steroidal AIs bind in the active site of aromatase by heteroatom coordination as the sixth ligand to the iron atom of the aromatase heme moiety. And lastly, Stage Three refines the binding conformation predictions made in Stage Two. The authors of this study refer to this new method as Refined Virtual Docking and implement the protocol for two aromatase inhibitors (Fig.8).
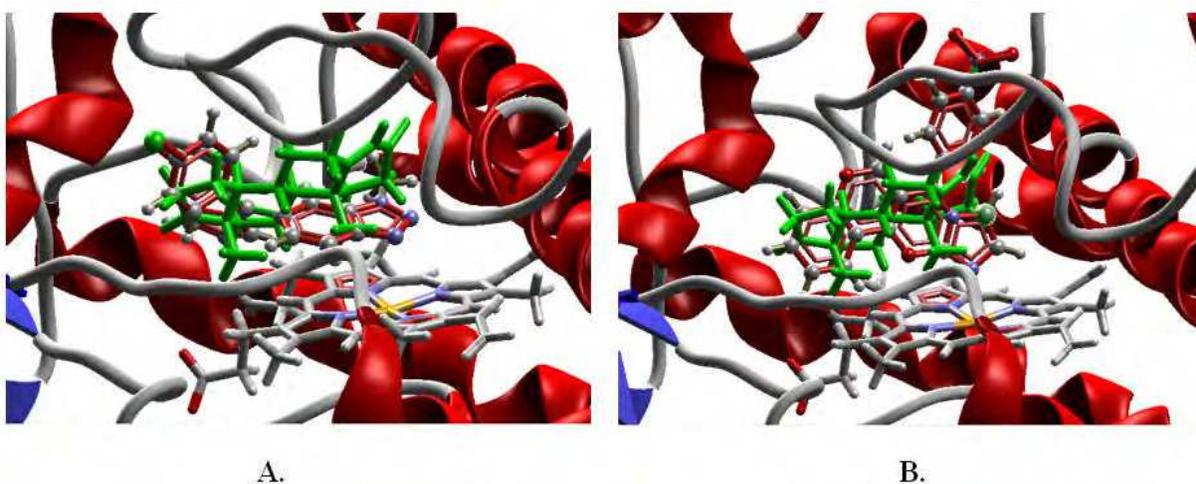

Fig. 8. Predicted binding conformations of (A) vorozole (B) 3-imidazolyl based on the Refined Docking Protocol, showing the heme moiety (stick, colored by atom), the preferred adrostenedione substrate (stick, green), and the aromatase inhibitors (ball-and-stick). Reprinted from Bioorganic Medicinal Chemistry Letters, 2010, 20(10), 3050-3064 , Copyright (2010), with permission from Elsevier.

In summary, the authors of this study successfully developed a powerful pharmacophore model for the identification of new classes of AI based on both LB and SB pharmacophore modeling. Additionally, they validated the models, proving the Merged Model to be more powerful and specific than the original models, and also developed a new docking protocol specific to the structural characteristics of their target enzyme, aromatase. Through this study, the authors demonstrated that both models and both types of information (i.e., known active and inactive inhibitors as well as the structure of the target enzyme) are essential to the development of a successful pharmacophore model, and subsequently, to the identification of novel, potent, highly specific, and potentially less toxic aromatase inhibitors.

2.4 Computational systems biology

Systems Biology is directly applicable in the medical sciences, as the ability to therapeutically target complex diseases such as cancer will be greatly improved by a global understanding of the unified signaling and regulatory network that integrates all environmental signals into a net outcome or phenotype. One example of such application is the identification of new uses for existing molecular targets through a systems biology analysis of the differences between normal and diseased samples. A second example can be found in the elucidation of complex signaling relationships or network analysis, which would allow for targeting of the most appropriate region of a signaling cascade for the development of more efficient and safe therapeutic agents. This could be achieved through studies that characterize perturbations of the biological system caused by small molecules, including existing therapeutic agents. The information gained through such studies could be extremely valuable since it would enable us to discriminate between cellular changes associated with therapeutic benefits and cellular changes associated with side-effects.

A dynamic model of such a network could be adapted to describe its changes in the context of a disease, for example by incorporating genetic mutations as changes in the node states or in interactions. The model would be able to predict the phenotypes associated with mutations and their combination with diverse environmental signals, and it would provide strategies for reversing a diseased phenotype into a healthy one. This emerging understanding would enable fundamental advances in cancer treatment by allowing clinicians to prescribe treatments specifically targeted to individuals and their present conditions. Treatments then may be tailored so that the least invasive intervention yields the greatest system-wide benefit, maximizes the body's self-healing abilities, and minimizes side effects. In this way, linking molecular characterizations to clinical phenotypes in a causal manner will be a key challenge of systems medicine, and several promising steps have already been made in this direction.

For example, such a methodology can be applied to prostate cancer, one of the most frequent malignancies and the second-leading cause of cancer mortality in North American males (Jemal et al. 2009). Although prostate cancer rates are progressing slowly, there are cases in which tumors behave aggressively, resulting in poor prognosis and eventually in the death of the patient. Disruptions in the balance of the insulin-like growth factor (IGF) axis and downstream signaling proteins have been attributed a critical role in the establishment and maintenance of the transformed phenotype in prostate cancer.

A recent study performed by Vellaichamy and collaborators from University of Michigan together with GeneGo, Inc. (Vellaichamy et al. 2010) demonstrates in a very elegant manner how computational systems biology approaches can be applied to better understand the biology and biochemistry of prostate cancer in order to establish new prognostic markers and to detect cellular functions suitable for therapeutic interference. These authors applied a topological scoring approach to investigate the response of LNCap prostate cancer cells, a well-studied model system for prostate cancer progression, to treatment with synthetic androgen (R1881). Their computational method combines disease- or condition-specific, high-throughput molecular data with the global network of protein interactions to identify nodes which occupy significant network positions with respect to differentially expressed genes or proteins in the molecular dataset.

Using such analysis, they were able to identify individual signaling cascades leading to the top transcriptional regulators revealing that PI3K signaling is supported by consistently high topological scores derived from both proteomics and microarray datasets. Fig.9 shows this cascade in the context of IGF signaling, with the PI3K cascade highlighted by the red line, where all of the elements that achieve high topological scores with respect to both sets are marked by red boxes. Through this study, the authors determined the central role of this pathway in regulating events that follow androgen treatment: through inhibition of GSK3 kinase and its ability to phosphorylate c-Myc and cyclin D (Fig.9).
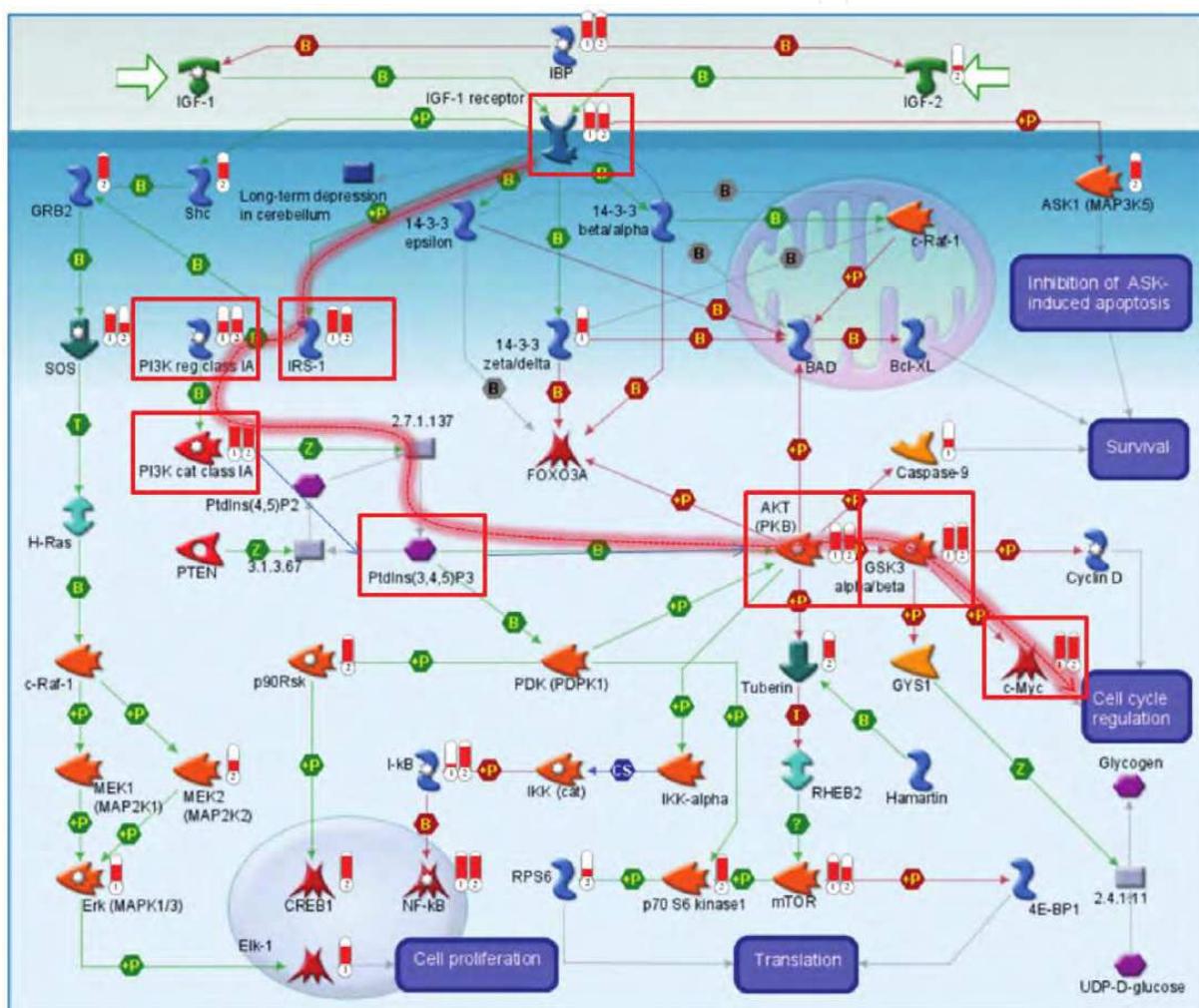

Fig. 9. Map for IGF signaling showing topologically significant genes identified from microarrays and iTRAQ proteomics. Red level in the ''thermometers'' represents relative rank (percentile) of a protein in the corresponding list of topologically significant proteins. Red boxes and highlighted path illustrate signaling cascade with strongest support from both sets. Adapted from Vellaichamy et al., PLoS One (2010), vol 5 (6), 1-10 with permission fromPloS ONE.

3. Conclusions

Undoubtedly, computational approaches have had a major impact on the design of anticancer drugs and drug candidates over the years and have provided fruitful insights into cancer in general. Nevertheless, to be useful to a biologist or a physician, computational models generated using these approaches should produce useful predictions that match experimental results, allow experiments to be performed *in silico* to save time and cost, and facilitate the understanding of how a system or process works. In that sense, computational approaches have room for improvement in all of the major areas: virtual screening techniques, ADME/Toxicity predictions, and ligand docking protocols.

The optimization of these techniques can be represented by its own theoretical feedback signaling system: the biology informs the computational approaches, which, in turn, inform the biology, and so on, creating immense room for both fundamental breakthroughs and incredible application-based advancements. The field is continuously evolving and challenges still remain, but we expect to see accelerated activity in this area as compounds continue to move through clinical trials and as the science and technology continue to develop. We hope this chapter will stimulate researchers to adopt and apply computational tools to the discovery of future cancer drugs.

4. References

National Cancer Institute - Cancer Trends Progress Report - 2009/2010.

MetaDrug, GeneGo Inc., St. Joseph, MI, http://www.genego.com

Anderson AC (2003) The process of structure-based drug design. Chem Biol 10: 787-797

Gane PJ and Dean PM (2000) Recent advances in structure-based rational drug design. Curr Opin Struct Biol 10: 401-404

Klebe G (2000) Recent developments in structure-based drug design. J Mol Med 78: 269-281

Combs AP (2007) Structure-based drug design of new leads for phosphatase research. IDrugs 10: 112-115

Coumar MS, Leou JS, Shukla P, Wu JS, Dixit AK, Lin WH, Chang CY, Lien TW, Tan UK, Chen CH, Hsu JT, Chao YS, Wu SY, and Hsieh HP (2009) Structure-based drug design of novel Aurora kinase A inhibitors: structural basis for potency and specificity. J Med Chem 52: 1050-1062

Khan A, Prakash A, Kumar D, Rawat A, Srivastava R, and Srivastava R (2010) Virtual screening and pharmacophore studies for ftase inhibitors using Indian plant anticancer compounds database. Bioinformation 5: 62-66

van Montfort RL and Workman P (2009) Structure-based design of molecular cancer therapeutics. Trends Biotechnol 27: 315-328

Mustata G, Li M, Zevola N, Bakan A, Zhang L, Epperly M, Greenberger JS, Yu J, and Bahar I (2011) Development of small-molecule PUMA inhibitors for mitigating radiation-induced cell death. Curr Top Med Chem 11: 281-290

Yu J and Zhang L (2009) PUMA, a potent killer with or without p53. Oncogene 27: S71-S83

Cory S and Adams JM (2002) The Bcl2 family: regulators of the cellular life-or-death switch. Nat Rev Cancer 2: 647-656

Fesik SW (2005) Promoting apoptosis as a strategy for cancer drug discovery. Nat Rev Cancer 5: 876-885

Zhang L, Ming L, and Yu J (2007) BH3 mimetics to improve cancer therapy; mechanisms and examples. Drug Resist Updat 10: 207-217

Bruncko M, Oost TK, Belli BA, Ding H, Joseph MK, Kunzer A, Martineau D, McClellan WJ, Mitten M, Ng SC, Nimmer PM, Oltersdorf T, Park CM, Petros AM, Shoemaker AR, Song X, Wang X, Wendt MD, Zhang H, Fesik SW, Rosenberg SH, and Elmore SW (2007) Studies leading to potent, dual inhibitors of Bcl-2 and Bcl-xL. J Med Chem 50: 641-662

Oltersdorf T, Elmore SW, Shoemaker AR, Armstrong RC, Augeri DJ, Belli BA, Bruncko M, Deckwerth TL, Dinges J, Hajduk PJ, Joseph MK, Kitada S, Korsmeyer SJ, Kunzer AR, Letai A, Li C, Mitten MJ, Nettesheim DG, Ng S, Nimmer PM, O'Connor JM, Oleksijew A, Petros AM, Reed JC, Shen W, Tahir SK, Thompson CB, Tomaselli KJ, Wang B, Wendt MD, Zhang H, Fesik SW, and Rosenberg SH (2005) An inhibitor of Bcl-2 family proteins induces regression of solid tumours. Nature 435: 677-681

Day CL, Smits C, Fan FC, Lee EF, Fairlie WD, and Hinds MG (2008) Structure of the BH3 domains from the p53-inducible BH3-only proteins Noxa and Puma in complex with Mcl-1. J Mol Biol 380: 958-971

Stahura FL and Bajorath J (2005) New methodologies for ligand-based virtual screening. Curr Pharm Des 11: 1189-1202

Johnson M and Maggiora G (1990) Concepts and applications of molecular similarity. John Wiley & Sons, Inc., New York,

Rognan D (2007) Chemogenomic approaches to rational drug design. Br J Pharmacol 152: 38-52

Sheridan RP and Kearsley SK (2002) Why do we need so many chemical similarity search methods? Drug Discov Today 7: 903-911

Hert J, Willett P, Wilton DJ, Acklin P, Azzaoui K, Jacoby E, and Schuffenhauer A (2004) Comparison of fingerprint-based methods for virtual screening using multiple bioactive reference structures. J Chem Inf Comput Sci 44: 1177-1185

Good AC, Krystek SR, and Mason JS (2000) High-throughput and virtual screening: core lead discovery technologies move towards integration. Drug Discov Today 5: 61-69

Chiang YK, Kuo CC, Wu YS, Chen CT, Coumar MS, Wu JS, Hsieh HP, Chang CY, Jseng HY, Wu MH, Leou JS, Song JS, Chang JY, Lyu PC, Chao YS, and Wu SY (2009) Generation of ligand-based pharmacophore model and virtual screening for identification of novel tubulin inhibitors with potent anticancer activity. J Med Chem 52: 4221-4233

Liou JP, Mahindroo N, Chang CW, Guo FM, Lee SW, Tan UK, Yeh TK, Kuo CC, Chang YW, Lu PH, Tung YS, Lin KT, Chang JY, and Hsieh HP (2006) Structure-activity relationship studies of 3-aroylindoles as potent antimitotic agents. ChemMedChem 1: 1106-1118

Mustata G, Follis AV, Hammoudeh DI, Metallo SJ, Wang H, Prochownik EV, Lazo JS, and Bahar I (2009) Discovery of novel Myc-Max heterodimer disruptors with a three-dimensional pharmacophore model. J Med Chem 52: 1247-1250

Nesbit CE, Grove LE, Yin X, and Prochownik EV (1998) Differential apoptotic behaviors of c-myc, N-myc, and L-myc oncoproteins. Cell Growth Differ 9: 731-741

Hammoudeh DI, Follis AV, Prochownik EV, and Metallo SJ (2009) Multiple independent binding sites for small-molecule inhibitors on the oncoprotein c-Myc. J Am Chem Soc 131: 7390-7401

Shepphird JK and Clark RD (2006) A marriage made in torsional space: using GALAHAD models to drive pharmacophore multiplet searches. J Comput Aided Mol Des 20: 763-771

Follis AV, Hammoudeh DI, Wang H, Prochownik EV, and Metallo SJ (2008) Structural rationale for the coupled binding and unfolding of the c-Myc oncoprotein by small molecules. Chem Biol 15: 1149-1155

Follis AV, Hammoudeh DI, Daab AT, and Metallo SJ (2009) Small-molecule perturbation of competing interactions between c-Myc and Max. Bioorg Med Chem Lett 19: 807-810

Thomsen R and Christensen MH (2006) MolDock: a new technique for high-accuracy molecular docking. J Med Chem 49: 3315-3321

Noha SM, Atanasov AG, Schuster D, Markt P, Fakhrudin N, Heiss EH, Schrammel O, Rollinger JM, Stuppner H, Dirsch VM, and Wolber G (2011) Discovery of a novel IKK-beta inhibitor by ligand-based virtual screening techniques. Bioorg Med Chem Lett 21: 577-583

Patel Y, Gillet VJ, Bravi G, and Leach AR (2002) A comparison of the pharmacophore identification programs: Catalyst, DISCO and GASP. J Comput Aided Mol Des 16: 653-681

Moffat K, Gillet VJ, Whittle M, Bravi G, and Leach AR (2008) A comparison of field-based similarity searching methods: CatShape, FBSS, and ROCS. J Chem Inf Model 48: 719-729

Sadowski J, Schwab C, and Gasteiger J (2003) CORINA: 3D Structure Generator.

Muftuoglu Y and Mustata G (2010) Pharmacophore modeling strategies for the development of novel nonsteroidal inhibitors of human aromatase (CYP19). Bioorg Med Chem Lett 20: 3050-3064

Brueggemeier RW, Hackett JC, and az-Cruz ES (2005) Aromatase inhibitors in the treatment of breast cancer. Endocr Rev 26: 331-345

Meunier B, de Visser SP, and Shaik S (2004) Mechanism of oxidation reactions catalyzed by cytochrome p450 enzymes. Chem Rev 104: 3947-3980

Meunier B, de Visser SP, and Shaik S (2004) Mechanism of oxidation reactions catalyzed by cytochrome p450 enzymes. Chem Rev 104: 3947-3980

Ghosh D, Griswold J, Erman M, and Pangborn W (2009) Structural basis for androgen specificity and oestrogen synthesis in human aromatase. Nature 457: 219-223

Chemical Computing Group (2008) MOE.

Wolber G and Langer T (2005) LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J Chem Inf Model 45: 160-169

Jemal A, Siegel R, Ward E, Hao Y, Xu J, and Thun MJ (2009) Cancer statistics, 2009. CA Cancer J Clin 59: 225-249

Vellaichamy A, Dezso Z, JeBailey L, Chinnaiyan AM, Sreekumar A, Nesvizhskii AI, Omenn GS, and Bugrim A (2010) "Topological significance" analysis of gene expression and proteomic profiles from prostate cancer cells reveals key mechanisms of androgen response. PLoS One 5: e10936
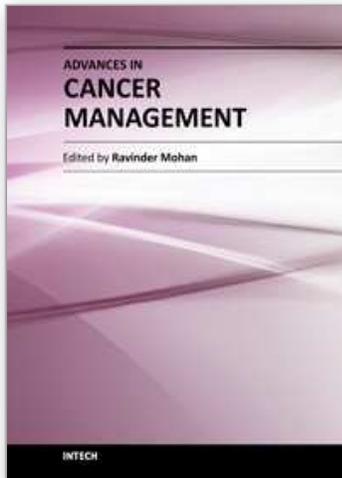
Advances in Cancer Management

Edited by Prof. Ravinder Mohan

ISBN 978-953-307-870-0

Hard cover, 278 pages

Publisher InTech

Published online 27, January, 2012

Published in print edition January, 2012

Cancer is now the most common cause of death in the world. However, because of early diagnosis, better treatment, and advanced life expectancy, many cancer patients frequently live a long, happy, and healthy life after the diagnosis- and often live as long as patients who eventually do not die because of cancer. This book presents newer advances in diagnosis and treatment of specific cancers, an evidence-based and realistic approach to the selection of cancer treatment, and cutting-edge laboratory developments such as the use of the MALDI technique and computational methods that can be used to detect newer protein biomarkers of cancers in diagnosis and to evaluate the success of treatment.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Gabriela Mustata Wilson and Yagmur Muftuoglu (2012). Computational Strategies in Cancer Drug Discovery, Advances in Cancer Management, Prof. Ravinder Mohan (Ed.), ISBN: 978-953-307-870-0, InTech, Available from: http://www.intechopen.com/books/advances-in-cancer-management/computational-strategies-in-cancer-drug-discovery

INTECH
open science | open minds