Rational drug design of antineoplastic agents using 3D-QSAR, cheminformatic, and virtual screening approaches

Jelica Vucicevic[1], Katarina Nikolic*[1], and John B.O. Mitchell*[2]

[1]Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Belgrade, Vojvode Stepe 450, 11000 Belgrade, Serbia

[2]EaStCHEM School of Chemistry and Biomedical Sciences Research Complex, University of St Andrews, St Andrews KY16 9ST, UK

Corresponding Authors:

Katarina Nikolic, PhD

Assistant Professor

Department of Pharmaceutical Chemistry

Faculty of Pharmacy, University of Belgrade

Vojvode Stepe 450, 11000 Belgrade, Serbia

e-mail: knikolic@pharmacy.bg.ac.rs

tel: +381-11-3951-259

cell: +381-63-84-30-677

Dr John B. O. Mitchell,

EaStCHEM School of Chemistry and Biomedical Sciences Research Complex,

University of St Andrews,

North Haugh,

St Andrews,

Scotland,

KY16 9ST.

United Kingdom.

e-mail: jbom@st-andrews.ac.uk

+44-(0)1334-467259

Abstract

Background: Computer-Aided Drug Design has strongly accelerated the development of novel antineoplastic agents by helping in the hit identification, optimization, and evaluation. **Results:** Computational approaches such as cheminformatic search, virtual screening, pharmacophore modeling, molecular docking and dynamics have been developed and applied to explain the activity of bioactive molecules, design novel agents, increase the success rate of drug research, and decrease the total costs of drug discovery. Similarity searches and virtual screening are used to identify molecules with an increased probability to interact with drug targets of interest, while the other computational approaches are applied for the design and evaluation of molecules with enhanced activity and improved safety profile. **Conclusion:** In this review are described the main *in silico* techniques used in rational drug design of antineoplastic agents and presented optimal combinations of computational methods for design of more efficient antineoplastic drugs.

Keywords: antineoplastic agents; pharmacophore; QSAR; rational drug design; cheminformatics; virtual screening; virtual docking

Introduction

Since cancer is a leading cause of death worldwide, discovery of novel, more potent antineoplastic agents is one of the most important and active drug discovery fields [1,2]. Even though drug research is a challenging, time consuming and expensive process, the number of compounds available to consider in the lead discovery stages of the drug discovery pipeline has significantly increased due to cheminformatics and Computer-Aided Drug Design (CADD) methodologies. *In silico* drug design mainly involves ligand-based methods, such as Quantitative Structure Activity Relationship (QSAR), Ligand-Based Pharmacophore Modeling, and structure-based methods, such as Virtual Docking (VD), Structure-Based

Pharmacophore Modeling, and Molecular Dynamics (MD) [3]. It is now widely accepted that rational drug design provides essential molecular understanding of drug-target interactions that is not easily and completely accessible from experimental techniques. Application of the *in silico* techniques has had a great impact on the efficient discovery of novel drug candidates [4-7]. Structure-based drug design requires the three-dimensional-structure of target or ligand-target complex to have been previously obtained by crystallography, Nuclear Magnetic Resonance (NMR) spectroscopy, or homology modeling. These 3D-structures are used as templates for performing virtual screening and virtual docking. Such structure-based studies have been employed for selection of molecular determinants for ligands binding and rational drug design. Finally, MD simulations of ligand-target complexes have been used to analyze the target structure conformations and to evaluate the affinity and stability of ligand-target complexes.

The 3D-structures of ligand-target complexes have been employed for identifying structural origins of selectivity and chemical scaffolds of novel ligands as drug candidates with improved efficacy and safety profiles. In the rational drug design of antineoplastic agents, these structure-based methodologies are very useful and powerful tools in the early phases of the drug discovery process [8]. However, ligand-based drug design methods, which provide information about essential structural characteristics for biological activity based on the ligands alone, have also shown themselves to be efficient techniques for the identification of potential lead structures, and some have argued that they are more efficient than methods based on target protein structures [9]. A recently developed method that combines structure and pharmacophore activity relationship studies into a single drug-target interaction evaluation represents an exciting novel approach in drug discovery [10]. Finally, combination of results obtained from ligand-based and structure-based virtual screening and integrated

with virtual docking and QSAR data has been claimed to be the most comprehensive rational drug discovery procedure with the highest success rate [10-13].

This review will describe and discuss some of the most recent investigations in which these *in silico* tools were used for the optimization and rational design of novel antineoplastic drugs.

QSAR studies in the rational design of antineoplastic drugs

QSAR methods have been widely applied in medicinal chemistry in order to gain insight into the critical structural requirements for compounds to interact with the biomolecules of interest. This computational tool uses information from molecular structures of ligands, correlates this with corresponding activities by employing different mathematical algorithms, determines the most important structural features controlling the bioactivity, and facilitates their application for design of novel drug candidates with enhanced activity on the drug target [11-13]. Furthermore, pharmacophores selected in QSAR studies can be used as search queries to screen the databases, leading to the identification of novel potential antineoplastic drugs [14-16]. Such QSAR models can predict the bioactivities of novel designed compounds and prioritize them for *in vitro* evaluation.

In order to increase the interpretability and reduce the risk of overfitting, 3D-QSAR models are usually built by using only a few descriptors [17]. Those commonly used include topological indices (TIs), geometrical, constitutional and physicochemical descriptors. Constitutional descriptors including number of atoms, bond count, atom type counts, ring count, and molecular weight (MW) are simple features that reflect the molecular composition. TIs are 2D descriptors which encode information about molecular size, shape, degree of branching, presence of heteroatoms and multiple bonds [18]. They include Wiener

index [19], Connectivity indices [20], Kier shape [21], Balaban J index [22] and Zagreb indices [23]. Physicochemical descriptors are also usually computed simply from 2D molecular structure. However, 3D geometrical descriptors require information obtainable only from 3D structures of molecules, and thus are conformation dependent and require some form of geometry optimization (e.g. WHIM, MoRSE and GETAWAY, etc.) [24].

Historically, QSAR methodologies used quite simple mathematical methods such as linear and multi-linear regression, to describe and derive that relationship between the descriptors and the property to be predicted. In recent years, Partial Least Squares (PLS) [25] and Multiple Linear Regression (MLR) are the most common techniques applied to examine linear correlations between molecular descriptors and corresponding activity against targets of interest [26-32]. However, in many cases there is no such linear relationship between molecular determinants and bioactivity, and for that reason a range of more sophisticated machine learning methods have been introduced to computational chemistry [33]. While the ultimate advantage of such methods is their ability to generate more accurate predictive models, this is often achieved through their flexibility in representing non-linear relationships and often in handling relatively large descriptor sets. Many machine-learning techniques including Support Vector Machines (SVM), Artificial Neural Network (ANN), k-Nearest Neighbors (k-NN), Random Forest (RF) and genetic algorithms have been applied for better exploration of the bioactive chemical space of examined compounds. For example, Singh et al. [34] used RF to build a QSAR model for EGFR inhibitors. Beyond the QSAR domain, RF is also useful for other tasks in the drug discovery process; Riddick et al. [35] used RF to assess which cancer cell lines were likely to respond to which drugs, while Statnikov et al. [36] classified different types of cancer, and Carlsson et al. [37] used both RF and SVM to assess the mutagenicity of compounds.

One of the first 3D-QSAR methodologies was based on electrostatic interaction energies of superimposed 3D-conformations of a data set of ligands that are effectively included in Comparative Molecular Field Analysis (CoMFA) [38] or Comparative Molecular Similarity Index Analysis (CoMSIA) [39]. The main limitation of the approach is that dynamical properties of the examined compounds could not be included in the calculation [40]. On the other hand, the Molecular Interaction Field (MIF) methodology considers conformations of multiple ligands for determining Grid-Independent Descriptors (GRID) and pharmacophores in 3D-QSAR modeling [41,42]. The GRIDs are used as independent variables for 3D-QSAR modeling, pharmacophore study, and drug design [43-45]. Some very promising novel rational drug discovery methodologies have been developed as combinations of 3D-QSAR modeling and complementary drug target fields [46-49].

Over the past few years, several advanced computer-aided drug design studies have been carried out on a number of epigenetic targets, such as histone arginine methyltransferases, histone deacetylase, and Hsp90 heat shock protein, by use of extensive molecular docking programs combined with 3D-QSAR [50,51] and comparative binding energy [49,52]. These *in silico* approaches were used to understand the activity of known ligands of epigenetic targets and to design novel epigenetic inhibitors. Epigenetic enzymes modulate expression of particular genes and regulate dynamic changes of histone proteins [53], DNA [54], RNA [55] and non-coding RNA [56]. Therefore, the enzymes associated with epigenetics, such as DNA methyltransferases (DNMT) [57], histone deacetylases (HDAC) [58], bromodomains [59], lysine demethylases [60], and lysine methyltransferases [61], are now considered to be very important anticancer drug targets [6]. Over the past few years, a significant increase in inhibitors of histone and non-histone proteins has successfully translated into clinical study [6].

The human HDACs are divided into eleven isoforms (HDAC 1-11) [62,63]. Several pharmacological studies have confirmed that HDAC isoform selective inhibitors could have broader and safer clinical applications than pan-HDAC inhibitors. Recently determined crystal structures of several HDAC isoforms in complex with inhibitors [64,65] will significantly facilitate structure-based drug design of selective HDAC inhibitors and provide important information for understanding the enzymes' functions [5,64]. Both 2D- and 3D-QSAR approaches have been very useful in the design of novel HDAC inhibitors with enhanced potency and isoform selectivity. The developed QSAR models were employed to identify HDAC pharmacophores, optimize HDAC inhibitors and evaluate the inhibitory potency of the novel designed inhibitors [7,66,67]. The CoMFA and CoMSIA 3D-QSAR methodology has been used for development of novel HDAC inhibitors [68-72]. Also, a MIF-based 3D-QSAR approach has been applied on four structurally diverse types of HDAC inhibitors [73]. Combination of the 3D-QSAR modeling and virtual docking has successfully identified compound ZINC70450932 as a novel inhibitor of HDAC1 [7] (**Figure 1**). Developed 3D-QSAR pharmacophore models were used to identify structural requirements for inhibitors binding to the HDAC enzyme [74,75], while Choubey et al. [7] developed a pharmacophore model of histone deacetylase 1 (HDAC1) inhibitors with two hydrogen bond acceptors, one hydrogen bond donor and one aromatic ring, employing 38 structurally diverse inhibitors. HDAC1 has been found to have a crucial role in multiple types of cancers. Also, the electronic structure-activity relationship model, developed by use of an electronic structure-based algorithm, was applied for defining the inhibitory mechanism of the oligodeoxynucleotide DNMT1 inhibitors [76].

Numerous successful cases of 3D-QSAR studies being used in antineoplastic drug discovery have been reported so far. For example, this approach has been applied in the discovery of new Bcl-2 inhibitors [77], Bcl-xl inhibitors [78], mTOR inhibitors [79], and CDK1 inhibitors

[29] (**Table 1**). Actually, most modelling studies for identification of inhibitors of the Bcl-2 family of proteins are based on structure-based methodologies [80-84]**.** Bcl-2 inhibitors exhibit potent anticancer activity against breast cancer cells (MDA-231) and leukemia cells (HL-60) which overexpress Bcl-2 protein [85,86]. Several QSAR models have also been developed for this class of inhibitors [77,78,87] . Almerico and coworkers [78] developed a 3D-QSAR pharmacophore model of Bcl-xl inhibitors from the set of 42 biarylacylsulfonamides. This model was used to identify the structural factors, including an aromatic moiety, negative charge and hydrogen bond acceptor, that govern the activity of these derivatives. Then the model was used as a 3D search query to screen the ZINC database and six hits with new scaffolds were identified (ZINC00784464, ZINC00788197, ZINC03200686, ZINC03212331, ZINC03243504, ZINC03356310) (**Figure 1**) [78]. Recently, Aboalhija et al. performed QSAR analysis to explore the structural features important for Bcl-2 inhibitory activity within a large and potent list of 98 inhibitors. Genetic function algorithm (GFA) coupled with k nearest neighbor (kNN) or multiple linear regression (MLR) analysis was applied to generate the best predictive QSAR models. The resulting QSAR-selected pharmacophores were used as in silico search queries to screen the National Cancer Institute (NCI) database and several hits that exhibit low micromolar cytotoxic activity against MDA-MB-231 were identified [77]. The same procedure and methodology were used for identification of new nanomolar mTOR inhibitors [79].

Nowadays, drug resistance is one of the major problems in cancer therapy. This obstacle could be overcome by using a combination of drugs with different mechanisms or by designing a single chemical entity that simultaneously modulates several targets [88,89]. As a result, development of agents able to interact with more than one biological target for cancer treatment is an interesting new concept in cancer drug design [6]. The first multi-target (mt) approach for the virtual screening has been published for anti-colorectal cancer (anti-CRC)

and anti-breast cancer (anti-BC) agents [90,91]. In this study, its authors designed more efficient anti-CRC and anti-BC drugs against ten and thirteen cell lines, respectively, by using multi-target QSAR and virtual screening techniques. The mt-QSAR models were developed by use of linear discriminant analysis (mt-QSAR-LDA). By analyzing these models, it was possible to identify substructural features responsible for the anticancer activity and thus to design novel potent and versatile agents.

Virtual Screening and Virtual Docking methods

Ligand- based virtual screening operates on the hypothesis that molecules with similar values of molecular descriptors should possess similar biological activity [92-95], the so-called similar property principle. This approach analyses the structural and physicochemical similarities between one or more lead compounds and large virtual library of plausible molecular structures, which need not yet actually have been synthesised, while structure-based virtual screening is based on direct modelling of ligands binding to the 3D structure of the drug target. Scores obtained by either or both virtual screening methodologies are then used to rank the examined ligands and to select novel hit compounds for further chemical modifications and testing [9].

Ligand- or structure-based pharmacophore screening on compound databases compares each query molecule with the previously defined spatially located pharmacophore features, such as hydrogen-bond donor, hydrogen-bond acceptor, hydrophobic interactions, steric interactions, aromatic interactions, and positive and negative ionizable regions. Pharmacophore screening algorithms typically use overlay-based scoring functions and root-mean-square deviation for final ranking of hit compounds [96]. Pharmacophore screening has been successfully applied in the search for novel epigenetic inhibitors, such as selective HDAC inhibitors [97,98] and

inhibitors of protein arginine methyltransferases (PRMTs) [99]. Various combinations of ligand and structure-based virtual screening permit very comprehensive study of various aspects of drug action and therefore have higher hit success rates [3,14,100,101]. In particular, parallel [102,103] and hybrid [95,104] virtual screening approaches have been promoted as very precise rational drug design methodologies.

Molecular docking is widely used as a rational drug design tool for protein-ligand complex modelling, indeed nowadays docking studies are considered an essential structure-based drug design strategy [105,106]. Docking into a known binding site of a single target protein can be an affordable and efficient computational technique which allows us to study the ligands' interactions with targets and to estimate the potential energy of the protein–ligand complexes. A relatively simple and quick-to-evaluate objective function is generally used for the many calculations required to search the various possible conformations and positions of ligands in the binding site of the target and to optimize the 3D docked ligand-protein complex structures. However, more elaborate and expensive scoring functions may be used for the relatively few further calculations then required to rank these conformations and to select those corresponding to the strongest binding between the ligand and the target. Using these more sophisticated scoring functions, top ranked conformations of each ligand are compared and used for final ranking of all ligands in the data set [107]. The scoring functions used for the final ranking can be derived by a variety of approaches. One is to design a function which has separate empirical energy terms corresponding to different physicochemical contributions to binding, often including van der Waals, electrostatic, hydrogen bonding, steric, and desolvation terms, each described by a suitable parametric functional form [108]. While the resulting scoring function is thus expressed as a sum of apparently meaningful components, only the total interaction free energy is actually fitted to real data in most cases and hence there is cancellation of error between terms and the interpretability of the individual

components is obscured. A second approach is to base each of the contributing interactions on the corresponding term in a molecular mechanics force field [109]. A third possibility is to derive a statistically based scoring function from analysis of the crystal structures, and in some versions also the binding energies, of known ligand-protein complexes [110]. Thus, the molecular docking approach provides insight into the binding mode of the ligand and target, compares different ligands, and estimates their binding energies. The objective and scoring functions used in docking reflect the fact that binding of a ligand to its target is partly based on their chemical complementarities and physicochemical interactions, and partly on shape complementarities, which may well be conformation-dependent. Flexibility of the target biomolecule can be modelled either by performing virtual docking to set of rigid protein conformations or by examining dynamic ligand-target complexes [111].

Several structure-based virtual screening studies have recently been performed on epigenetic targets using some of the most popular molecular docking programs, such as DOCK [112], AutoDock [113,114], GOLD [115], and Glide [116]. These investigations resulted in the discovery of selective and potent inhibitors of histone arginine methyltransferases [117,118], bromodomain (BRD) type 4 inhibitors [119,120] and DNA methyltransferases type 1 [121,122]. This methodology also found new scaffolds that can be used for experimental optimization or as a starting point for chemical space exploration.

Successful applications such as development of Bcl-2xL agents [80], topoisomerase I and II inhibitors [123], COX-2/5-LOX dual inhibitors [124], Rac1 agents [125], inhibitors of Hsp90 [126], inhibitors of Tip 60 [127], mTOR [128], histone deacetylase inhibitors [129], hTERT inhibitors [130,131] and CDK [29] agents highlight the importance of this virtual docking technique (**Table 2**). These, like many other targets involved in cellular events such as cell growth, differentiation and proliferation, could serve as potential targets for drug discovery directed towards various types of cancer. Since the topoisomerases are essential enzymes

involved in various cell processes, their inhibition is one of the most important mechanisms of anticancer drugs. Ashour et al. [123] designed analogues of oleanolic acid with enhanced activity against topoisomerase I and IIα inhibition based on information obtained from active sites using docking techniques (based on PDB structures 1t8i and 1bgw for topoisomerase I and IIα, respectively) [132,133]. Ten designed compounds with good docking scores and promising binding modes were synthesized, and five of them (known as S2, S3, S5, S7 and S9) showed greater inhibitory activity against topoisomerase I than did the reference molecule camptothecin (CPT). Also, it was found that S2, S3, S5 and S6 showed greater topoisomerase II inhibitory activity than etoposide. Four compounds (S2, S3, S5 and S7) act as dual inhibitors of both enzymes, therefore potentially having an important role in cancer prevention (**Figure 2**) [123]. Based on structure-based drug design, several series of compounds, including *N*-substituted-dihydropyrazoles, dihydropyrazole-coumarin and myricetin, were designed and synthesized as potential human telomerase inhibitors. The binding mode of these compounds was explored in docking studies which provide information that supports rational design of more efficient telomerase inhibitors [130,131].

Apart from providing an insight into the binding mode of the ligands to their target, docking techniques are widely used in virtual screening studies of large databases. An example of this rational design approach was published by Cardama and coworkers [125]. Rac1, a member of Rho family of small GTPases, appears to be a promising and relevant target for the development of novel anticancer drugs since it is overexpressed in breast, colorectal, gastric, testicular, lung and brain cancer [134-137]. Docking-based virtual library screening was conducted on the ZINC database considering the portion of the Rac1 surface area containing a critical Trp56 as the target (PDB ID code 1MH1), and the Rac1 inhibitor ZINC69391 was identified. It was shown that ZINC69391 exerts an antimetastatic effect *in vivo*. A novel analog of this compound (1A-116) with a high docking score was developed by rational

design and it has proved to be more specific and potent at lower doses than its parent compound, both *in vivo* and *in vitro* (**Figure 3**). These results show that these novel Rac1 inhibitors, developed by docking techniques, have a good prospect of being useful as novel agents in anticancer therapy.

Several docking and molecular dynamic studies on epigenetic targets have recently been performed and resulted in the discovery of more selective and potent inhibitors of HDAC [138-141], bromodomain [119,120,142], DNMT [143-145], the histone methyltransferases (HMT) and the histone acetyltransferases (HAT) [146]. The results of these *in silico* studies are very helpful for defining the reaction mechanisms and identifying the key interactions between known active ligands and their epigenetic targets.

Target Prediction

Earlier, we discussed drugs that can interact with more than one target [88,89]. This may on occasion give rise to polypharmacology, beneficial pharmacological effects from a drug inhibiting or modulating multiple protein targets [147]. Another application of multi-target ligands is drug repurposing, where drugs are redirected from one originally intended therapeutic area to another [148]. If computational predictions can identify previously unrecognized drug-target associations, then drug repurposing may progress from serendipity and trial-and-error to being a more mature evidence-based science, and more frequently and reliably deliver significant health benefits to patients.

Too frequently, off-target interactions result in adverse side-effects, which it is important both to understand and to predict [149]. In recent years, such off-target drug interactions with unexpected proteins have led to a significant number of serious adverse drug reactions (ADRs). This has affected both compounds in development, many of which fail to reach the

market due to these unforeseen side-effects, and also marketed drugs whose safety and viability can be seriously compromised in this way [150], and may have to be withdrawn from sale. Both patients and pharmaceutical companies suffer when such withdrawals are required.

The development of multiple target prediction techniques requires the existence of copious molecule-target assay data, incorporating a good range of both ligands and targets. Various experimental technologies, such as target and phenotype based assay, are currently employed to find the underlying macromolecular targets involved in complex biological mechanisms [151].

Typically, target prediction methods work via molecular similarity. For a given target, the set of molecules experimentally known to bind to it are encoded using descriptors, and used as training data to build up a profile of what a ligand of this protein typically looks like. In some cases, such binders may form two or more quite separate clusters, possibly corresponding to different ligand scaffolds or to different binding sites. Given that most molecules have not been publicly assayed against most targets, this matrix of training data is sparse and careful interpolation is required. While the problem may be mathematically formulated in different ways [147,152-155], its essence is to use descriptor-based molecular similarity to estimate the probability that a query compound belongs to the set of binders for the given target. Helpfully, Gfeller et al. [156] have made their SwissTargetPrediction server publicly available. Where a model makes explicit use of known binding constants to predict quantitative binding affinities for putative compound-target pairs, the method is termed proteochemometric [147,157].

Another, rather computationally expensive, approach to the same problem is to dock a library of ligands into a panel of protein targets in silico. Glen and Allen [158] discussed the application of multiple dockings to cancer research, whilst when Favia et al. [159] carried

out a study of this kind, their main purpose was protein function prediction as they aimed to identify the native substrates of enzymes. As various authors [152,160] have pointed out, however, the same cross-docking method could be used for multi-target and off-target prediction, provided that the scoring function was accurate enough to compare dockings into diverse protein structures. Meeting this criterion might be problematic, as Warren et al. [107] found no correlation between experimental binding affinities and calculated docking scores for a small but diverse set of targets. Nonetheless, the relative success achieved by Schomburg et al. [161] indicates that target prediction by docking is becoming computationally and methodologically tractable.

Emig et al. [162] have used a target prediction approach to consider both drug repositioning and potential novel targets for cancer along with a number of other diseases. They consider the relationships between the pathological processes involved in different cancers such as hepatocellular carcinoma, colon cancer, ovarian cancer, thyroid cancer, melanoma, and acute myeloid leukemia. They suggest that there is a common core of cancer drug targets relevant to each of these diseases. Nigsch and Mitchell [163] carried out a target prediction study in which they algorithmically grouped together a diverse panel of protein targets according to their associated toxic effects. They found that proteins associated with breast cancer, along with those linked to side effects of breast cancer drugs, clustered together.

Conclusion

Rational drug design has become essential methodology in discovery of antineoplastic agents. The pharmacophore and QSAR approaches employ the chemical structures of experimentally confirmed anticancer compounds to define molecular determinants for activity and use these results to design and evaluate novel molecules. Cheminformatic and virtual screening tools

have been used to rank the designed compounds and to identify new lead structures by searching in various databases.

The 3D-structures of specific drug targets provide essential information for performing structure based virtual screening, molecular docking and dynamic studies. Based on these analyses have been filtered and ranked compounds from large databases, explained binding mode and activity of known ligands, and evaluated the conformational changes and stability of the ligand-target complexes.

Further development of computer-aided drug design approaches, along with improvement in crystallographic and biological methods, will provide deeper knowledge of drug-target interactions and augment discovery of more efficient and safer anticancer drugs.

Conflict of Interest Statement

The Authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

Support was kindly provided by the EU COST Action CM1406 and CA15135. KN and JV kindly acknowledge national project number 172033 supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] Jemal, A.; Bray, F.; Center, M.M.; Ferlay, J.; Ward, E.; Forman, D. Global cancer statistics. *CA Cancer J Clin*. **2011**, 61(2), 69-90.

[2] Stearns, V.; Davidson, N.E.; Flockhart, D.A. Pharmacogenetics in the treatment of breast cancer. *Pharmacogenomics J*. **2004**, 4(3), 143-153.

[3] Wilson, G.I.; Lill, M.A. Integrating structure-based and ligand-based approaches for computational drug design. *Future Med Chem*. **2011**, 3, 735-750.

[4] Alvarez, J.C. High-throughput docking as a source of novel drug leads. *Curr Opin Chem Biol*. **2004**, 8, 365–370.

[5] Roche, J.; Bertrand, P. Inside HDACs with more selective HDAC inhibitors. *Eur J Med Chem*. **2016**, 121, 451-483.

[6] Ganesan, A. Multitarget drugs: an epigenetic epiphany. *Chem Med Chem*. **2016**, 11, 1227-1241.

[7] Choubey, S.K.; Jeyaraman, J. A mechanistic approach to explore novel HDAC1 inhibitor using pharmacophore modeling, 3D-QSARanalysis, molecular docking, density functional and molecular dynamics simulation study. *J Mol Graph Model*. **2016**, 70, 54-69.

[8] Kufareva, I.; Abagyan, R. Type-II kinase inhibitor docking, screening, and profiling using modified structures of active kinase states. *J Med Chem*. **2008**, 51 (24), 7921-7932.

[9] Ripphausen, P.; Nisius, B.; Peltason, L.; Bajorath, J. Quo vadis, virtual screening? A comprehensive survey of prospective applications. *J Med Chem*. **2010**, 53, 8461-8467.

[10] Hu, B.; Lill, M.A. PharmDock: a pharmacophore-based docking program. *J Cheminform*. **2014**, 6, 14.

[11] Wong, Y.H.; Lin, CL.; Chen, T.S.; Chen, C.A.; Jiang, P.S.; Lai, Y.H.; Chu, L.; Li, C.W.; Chen, J.J.; Chen, B.S. Multiple target drug cocktail design for attacking the core network markers of four cancers using ligand-based and structure-based virtual screening methods. *BMC Med Genomics*. **2015**, 8 Suppl 4, S4.

[12] Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.;

Gramatica, P.; Martin, Y.C.; Todeschini R.; Consonni, V.; Kuzmin, V.E.; Cramer, C.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. QSAR modeling: Where have you been? Where are you going to? *J Med Chem*. **2014**, 57, 4977-5010.

[13] Cramer, R.D. The Inevitable QSAR Renaissance. *J Comp Aided Mol Des*. **2012**, 26, 35-38.

[14] Vucicevic, J.; Srdic-Rajic, T.; Pieroni, M.; Laurila, J.M.; Perovic, V.; Tassini, S.; Azzali, E.; Costantino, G.; Glisic, S.; Agbaba, D.; Scheinin, M.; Nikolic, K.; Radi, M.; Veljkovic, N. A combined ligand- and structure-based approach for the identification of rilmenidine-derived compounds which synergize the antitumor effects of doxorubicin. *Bioorg Med Chem*. **2016**, 24(14), 3174-3183.

[15] Gagic, Z.; Nikolic, K.; Ivkovic, B.; Filipic, S.; Agbaba, D. QSAR studies and design of new analogs of vitamin E with enhanced antiproliferative activity on MCF-7 breast cancer cells. *J Taiwan Inst Chem Eng*. **2016**, 59, 33–44.

[16] Ivkovic, B.M.; Nikolic, K.; Ilic, B.B.; Zizak, Z.S.; Novakovic, R.B.; Cudina, O.A.; Vladimirov, S.M. Phenylpropiophenone derivatives as potential anticancer agents: Synthesis, biological evaluation and quantitative structure activity relationship study. *Eur J Med Chem*. **2013**, 63, 239-255.

[17] Shahlaei, M. Descriptor selection methods in quantitative structure-activity relationship studies: a review study. *Chem Rev*. **2013**, 113(10), 8093-8103.

[18] Harary, F. Recent results in topological graph theory. *Acta Mathem Academ Scient Hung.* **1964**, 15, 405-412.

[19] Wiener, H. Structural determination of paraffin boiling points. *J Am Chem Soc*. **1941**, 69, 17-20.

[20] Randic, M. On characterization of molecular branching. *J Am Chem Soc*. **1975**, 97, 6609-6614.

[21] Kier, L.B. Indexes of molecular shape from chemical graphs. *Acta Pharm Jugosl*. **1986**, 36, 171-188.

[22] Balaban, A.T. Highly discriminating distance-based topological index. *Chem Phys Lett*. **1982**, 89, 399-404.

[23] Kier, L.B.; Hall, L.H. Molecular Connectivity in structure activity analysis. Wiley&Sons: New York **1986**.

[24] Todeschini, R.; Consonni, V. Handbook of molecular descriptors. Wlley-VCH, New York, **2000**.

[25] Wold, S.; Ruhe, A.; Wold, H.; Dunn, W.J. The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. *SIAM J Sci Stat Comput*. **1984**, 5, 735-743.

[26] Taha, M.O.; Bustanji, Y.; Al-Ghussein, M.A.; Mohammad, M.; Zalloum, H.; Al-Masri, I.M.; Atallah, N. Pharmacophore Modeling, Quantitative Structure-Activity Relationship Analysis, and in Silico Screening Reveal Potent Glycogen Synthase Kinase-3beta Inhibitory Activities for Cimetidine, Hydroxychloroquine, and Gemifloxacin. *J Med Chem*. **2008**, 51, 2062−2077.

[27] Al-Nadaf, A.; Abu Sheikha, G.; Taha, M.O. Elaborate LigandBased Pharmacophore Exploration and QSAR Analysis Guide the Synthesis of Novel Pyridinium-Based Potent Beta-Secretase Inhibitory Leads. *Bioorg Med Chem*. **2010**, 18, 3088−3115.

[28] Al-Sha'er, M.A.; Taha, M.O. Discovery of novel CDK1 inhibitors by combining pharmacophore modeling, QSAR analysis and in silico screening followed by in vitro bioassay. *Eur J Med Chem*. **2010**, 45, 4316-4330.

[29] Abdula, A.M.; Khalaf, R.A.; Mubarak, M.S.; Taha, M.O. Discovery of New Beta-D-Galactosidase Inhibitors via Pharmacophore Modeling and QSAR Analysis Followed by in silico Screening. *J Comput Chem*. **2011**, 32, 463−482.

[30] Habash, M.; Taha, M. O. Ligand-Based Modelling Followed by Synthetic Exploration Unveil Novel Glycogen Phosphorylase Inhibitory Leads. *Bioorg Med Chem*. **2011**, 19, 4746−4771.

[31] Shahin, R.; Alqtaishat, S.; Taha, M.O. Elaborate Ligand-Based Modeling Reveal New Submicromolar Rho Kinase Inhibitors. *J Comput-Aided Mol Des*. **2012**, 26, 249−266.

[32] Suaifan, G.A.; Shehadehh, M.; Al-Ijel, H.; Taha, M.O. Extensive Ligand-Based Modeling and in Silico Screening Reveal Nanomolar Inducible Nitric Oxide Synthase (iNOS) Inhibitors. *J Mol Graphics Modell*. **2012**, 37, 1−26.

[33] Mitchell, J.B.O. Machine learning methods in chemoinformatics. *WIREs Comput Mol Sci*. **2014**, 4, 468-481.

[34] Singh, H.; Singh, S.; Singla, D.; Agarwal, S.M.; Raghava, G.P.S. QSAR based model for discriminating EGFR inhibitors and non-inhibitors using Random forest. *Biology Direct*. **2015,** 10, 10.

[35] Riddick, G.; Song, H.; Ahn, S.; Walling, J.; Borges-Rivera, D.; Zhang, W.; Fine, H.A. Predicting in vitro drug sensitivity using Random Forests. *Bioinformatics* **2011**, 27(2), 220-224.

[36] Statnikov, A.; Wang, L.; Aliferis, C. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. *BMC Bioinformatics* **2008**, 9, 319.

[37] Carlsson, L.; Helgee, E.A.; Boyer, S. Interpretation of Nonlinear QSAR Models Applied to Ames Mutagenicity Data. *J Chem Inf Model*. **2009**, 49, 2551-2558.

[38] Cramer, R.D.; Patterson, D.E.; Bunce, J.D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J Am Chem Soc*. **1988**, 110, 5959-5967.

[39] Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indexes in a Comparative-Analysis (CoMSIA) of Drug Molecules To Correlate and Predict Their Biological-Activity. *J Med Chem*. **1994**, 37, 4130−4146.

[40] Acharya, C.; Coop, A.; Polli, J.E.; Mackerell, A.D. Jr. Recent advances in ligand based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr Comput Aided Drug Des*. **2011**, 7, 10–22.

[41] Goodford, P.J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J Med Chem*. **1985**, 28, 849-857.

[42] Dixon, S.L.; Smondyrev, A.M.; Rao, S.N. PHASE: A Novel Approach to Pharmacophore Modeling and 3D Database Searching. *Chem Biol Drug Des*. **2006**, 67, 370-372.

[43] Pastor, M.; Cruciani, G.; McLay, I.; Pickett, S.; Clementi, S. GRid-INdependent Descriptors (GRIND): A Novel Class of Alignment- Independent Three-Dimensional Molecular Descriptors. *J Med Chem*. **2000**, 43, 3233-3243.

[44] Durán, A.; Zamora, I.; Pastor, M. Suitability of GRIND-Based Principal Properties for the Description of Molecular Similarity and Ligand-Based Virtual Screening. *J Chem Inf Model*. **2009**, 49, 2129-2138.

[45] Ballante, F.; Ragno, R. 3-D QSAutogrid/R: an alternative procedure to build 3-D QSAR models. Methodologies and applications. *J Chem Inf Model*. **2012**, 52(6), 1674-1685.

[46] Ortiz, A.R.; Pisabarro, M.T.; Gago, F.; Wade, R.C. Prediction of Drug Binding Affinities by Comparative Binding Energy Analysis. *J Med Chem*. **1995**, 38, 2681-2691.

[47] Gohlke, H.; Klebe, G. DrugScore Meets CoMFA: Adaptation of Fields for Molecular Comparison (AFMoC) or How To Tailor Knowledge-Based Pair-Potentials to a Particular Protein. *J Med Chem*. **2002**, 45, 4153−4170.

[48] Varela, R.; Walters, W.P.; Goldman, B.B.; Jain, A.N. Iterative Refinement of a Binding Pocket Model: Active Computational Steering of Lead Optimization. *J Med Chem*. **2012**, 55, 8926−8942.

[49] Silvestri, L.; Ballante, F.; Mai, A.; Marshall, G. R.; Ragno, R., Histone deacetylase inhibitors: structure-based modeling and isoform-selectivity prediction. *J Chem Inf Model*. **2012**, 52, 2215-2235.

[50] Ragno, R.; Simeoni, S.; Castellano, S.; Vicidomini, C.; Mai, A.; Caroli, A.; Tramontano, A.; Bonaccini, C.; Trojer, P.; Bauer, I.; Brosch, G.; Sbardella, G., Small molecule inhibitors of histone arginine methyltransferases: homology modeling, molecular docking, binding mode analysis, and biological evaluations. *J Med Chem*. **2007**, 50, 1241-1253.

[51] Ballante, F, Caroli, A.; Wickersham, R.B., 3rd; Ragno, R. Hsp90 inhibitors, part 1: definition of 3-D QSAutogrid/R models as a tool for virtual screening. *J Chem Inf Model*. 54(3), 956-969.

[52] Caroli, A.; Ballante, F.; Wickersham, R.B., 3rd; Corelli, F.; Ragno, R., Hsp90 inhibitors, part 2: combining ligand-based and structure-based approaches for virtual screening application. *J Chem Inf Model*. **2014**, 54, 970-977.

[53] Zentner, G.E.; Henikoff, S. Regulation of nucleosome dynamics by histone modifications. *Nat Struct Mol Biol*. **2013**, 20, 259-266.

[54] Heyn, H.; Esteller, M. DNA methylation profiling in the clinic: applications and challenges. *Nat Rev Genet*. **2012**, 13, 679-692.

[55] Meyer, K.D.; Saletore, Y.; Zumbo, P.; Elemento, O.; Mason, C.E.; Jaffrey, S.R. Comprehensive analysis of mRNA methylation reveals enrichment in 30 UTRs and near stop codons. *Cell*. **2012**, 149, 1635-1646.

[56] Mercer, T.R.; Mattick, J.S. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol*. **2013**, 20, 300-307.

[57] Fahy, J.; Jeltsch, A.; Arimondo, P.B. DNA methyltransferase inhibitors in cancer: a chemical and therapeutic patent overview and selected clinical studies. *Expert Opin Ther Pat*. **2012**, 22, 1427-1442.

[58] Huang, H.; Lin, S.; Garcia, B.A.; Zhao, Y. Quantitative Proteomic Analysis of Histone Modifications. *Chem Rev*. **2015**, 115, 2376–2418.

[59] Filippakopoulos, P.; Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat Rev Drug Discov*. 13(5), 337-356.

[60] Zheng, Y.; Ma, J.; Wang, Z.; Li, J.; Jiang, B.; Zhou, W.; Shi, X.; Wang, X.; Zhao, W.; Liu, H. systematic review of histone lysine-specific demethylase 1 and its inhibitors. *Med Res Rev*. **2015**, 35, 1032–1071.

[61] Kaniskan H.Ü.; Konze K.D.; Jin J. Selective Inhibitors of Protein Methyltransferases. *J Med Chem*. **2015**, 58, 1596–1629.

[62] Marmorstein, R. Structure of histone deacetylases: review insights into substrate recognition and catalysis. *Structure* **2001**, 9, 1127-1133.

[63] Khan, N.; Jeffers, M.; Kumar, S.; Hackett, C.; Boldog, F.; Khramtsov, N.; Qian, X.; Mills, E.; Berghs, S.C.; Carey, N.; Finn, P.W.; Collins, L.S.; Tumber, A.; Ritchie, J.W.; Jensen, P.B.; Lichenstein, H.S.; Sehested, M. Determination of the class and isoform selectivity of small-molecule histone deacetylase inhibitors. *Biochem J*. **2008**, 409, 581-589.

[64] Micelli, C.; Rastelli, G. Histone deacetylases: structural determinants of inhibitor selectivity. *Drug Discov Today* **2015**, 20, 718–735.

[65] Yang, H.; Christianson, D.V. Histone deacetylase 6 structure and molecular basis of catalysis and inhibition. *Nat Chem Biol.* **2016**, 12, 741–747.

[66] Xie, A.; Liao, C.; Li, Z.; Ning, Z.; Hu, W.; Lu, X.; Shi, L.; Zhou, J. Quantitative structure-activity relationship study of histone deacetylase inhibitors. *Curr Med Chem Anticancer Agents* **2004**, 4, 273-299.

[67] Dessalew, N. QSAR study on aminophenylbenzamides and acrylamides as histone deacetylase inhibitors: an insight into the structural basis of antiproliferative activity. *Med Chem Res*. 2007, 16, 449-460.

[68] Chen, H.-F., Kang, J.-H., Li, Q., Zeng, B.-S., Yao, X.-J., Fan, B.-T., Yuan, S.G.; Panay, A.; Doucet, J.P. 3D-QSAR study on apicidin inhibit histone deacetylase. *Chin J Chem*. **2003**, 21, 1596-1607.

[69] Guo, Y.; Xiao, J.; Guo, Z.; Chu, F.; Cheng, Y.; Wu, S. Exploration of a binding mode of indole amide analogues as potent histone deacetylase inhibitors and 3D-QSAR analyses. *Bioorg Med Chem*. **2005**, 13, 5424-5434.

[70] Juvale, D.C.; Kulkarni, V.V.; Deokar, H.S.; Wagh, N.K.; Padhye, S.B.; Kulkarni, V.M. 3D-QSAR of histone deacetylase inhibitors: hydroxamate analogues. *Org Biomol Chem*. **2006**, 4, 2858-2868.

[71] Liu, B.; Lu, A.-J.; Liao, C.-Z.; Liu, H.-B.; Zhou, J.-J.; 3D-QSAR of sulfonamide hydroxamic acid HDAC inhibitors. *Acta Phys Chim Sin*. **2005**, 21, 333-337.

[72] Agarwal, N.; Bajpai, A.; Srivastava, V.; Gupta, S.P. A quantitative structure-activity relationship and molecular modeling study on a series of biaryl imidazole derivatives acting as H+/K+-ATPase inhibitors. *Biochem Res Int*. **2012**, 2013(2013), Article ID 810691.

[73] Ragno, R.; Simeoni, S.; Valente, S.; Massa, S.; Mai, A. 3-D QSAR studies on histone deacetylase inhibitors. A GOLPE/GRID approach on different series of compounds. *J Chem Inf Model* **2006**, 46, 1420-1430.

[74] Chen, Y.; Li, H.; Tang, W.; Zhu, C.; Jiang, Y.; Zou, J.; Yu, Q.; You, Q. 3D-QSAR studies of HDACs inhibitors using pharmacophore-based alignment. *Eur J Med Chem*. **2009**, 44, 2868-2876.

[75] Mahipal, A.; Tanwar, O.P.; Karthikeyan, C.; Moorthy, N.S.H.N.; Trivedi, P. 3D QSAR of aminophenyl benzamide derivatives as histone deacetylase inhibitors. *Med Chem*. **2010**, 6, 277-285.

[76] Clark, J., Shevchuk, T., Kho, M.R., Smith, S.S.Methods for the design and analysis of oligodeoxynucleotide- based DNA (cytosine-5) methyltransferase inhibitors. *Anal Biochem*. **2003**, 321, 50-64.

[77] Aboalhaija, N.H.; Zihlif, M.A.; Taha, M.O. Discovery of new selective cytotoxic agents against Bcl-2 expressing cancer cells using ligand-based modeling. *Chem Biol Interact*. **2016**, 250, 12-26.

[78] Almerico, A.M.; Tutone, M.; Lauria, A. 3D-QSAR pharmacophore modeling and in silico screening of new Bcl-xl inhibitors. *Eur J Med Chem*. **2010**, 45(11), 4774-4782.

[79] Khanfar, M.A.; Taha, M.O. Elaborate Ligand-Based Modeling Coupled with Multiple Linear Regression and k Nearest Neighbor QSAR Analyses Unveiled New Nanomolar mTOR Inhibitors. *J Mol Graph Model*. **2013**, 53, 2587-2612.

[80] Azam, S.S.; Abro, A.; Tanvir, F.; Parvaiz, N. Identification of unique binding site and molecular docking studies for structurally diverse Bcl-xL inhibitors. *Med Chem Res*. **2014**, 23, 3765-3783.

[81] Mukherjee, P.; Desai, P.; Zhou, Y.D.; Avery, M. Targeting the BH3 domain mediated protein-protein interaction of Bcl-xL through virtual screening. *J Chem Inf Model*. **2010**, 50(5), 906-923.

[82] Zheng, C.H.; Zhou, Y.J.; Zhu, J.; Ji, H.T.; Chen, J.; Li, Y.W.; Sheng, C.Q.; Lu, J.G.; Jiang, J.H.; Tang, H.; Song, Y.L. Construction of a three-dimensional pharmacophore for Bcl-2 inhibitors by flexible docking and the multiple copy simultaneous search method. *Bioorg Med Chem*. **2007**, 15(19), 6407-6417.

[83] Pinto, M.; Orzaez Mdel, M.; Delgado-Soler, L.; Perez, J.J.; Rubio-Martinez, J. Rational design of new class of BH3-mimetics as inhibitors of the Bcl-xL protein. *J Chem Inf Model*. **2011**, 51(6), 1249-1258.

[84] Sivakumar, D.; Gorai, B.; Sivaraman, T. Screening efficient BH3-mimetics to hBcl-B by means of peptidodynmimetic method. *Mol Biosyst*. **2013**, 9(4), 700-712.

[85] Enyedy, I.J.; Ling, Y.; Nacro, K.; Tomita, Y.; Wu, X.; Cao, Y.; Guo, R.; Li, B.; Zhu, X.; Huang, Y.; Long, Y.Q.; Roller, P.P.; Yang, D.; Wang, S. Discovery of small-molecule inhibitors of Bcl-2 through structure-based computer screening. *J Med Chem*. **2001**, 44(25), 4313-4324.

[86] Wang, J.L.; Liu, D.; Zhang, Z.J.; Shan, S.; Han, X.; Srinivasula, S.M.; Croce, C.M.; Alnemri, E.S.; Huang, Z. Structure-based discovery of an organic compound that binds Bcl-2 protein and induces apoptosis of tumor cells. *Proc Natl Acad Sci U S A*. **2000**, 97(13), 7124-7129.

[87] Levoin, N.; Vo, D.D.; Gautier, F.; Barillé-Nion, S.; Juin, P.; Tasseau, O.; Grée, R. A combination of in silico and SAR studies to identify binding hot spots of Bcl-xL inhibitors. *Bioorg Med Chem*. **2015**, 23(8), 1747-1757.

[88] Morphy, R.; Rankovic, Z. Designed multiple ligands. An emerging drug discovery paradigm. *J Med Chem*. **2005**, 48(21), 6523-6543.

[89] Hopkins, A.L. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol*. **2008**, 4(11), 682-90.

[90] Speck-Planche, A., Kleandrova, V.V.; Luan, F.; Cordeiro, M.N. Chemoinformatics in anti-cancer chemotherapy: multi-target QSAR model for the in silico discovery of anti-breast cancer agents. *Eur J Pharm Sci*. **2012**, 47(1), 273-279.

[91] Speck-Planche, A.; Kleandrova, V.V.; Luan, F.; Cordeiro, M.N. Rational drug design for anti-cancer chemotherapy: multi-target QSAR models for the in silico discovery of anti-colorectal cancer agents. *Bioorg Med Chem*. **2012**, 20(15), 4848-4855.

[92] Venkatraman, V.; Pérez-Nueno, V.I.; Mavridis, L.; Ritchie, D.W. Comprehensive Comparison of Ligand-Based Virtual Screening Tools Against the DUD Data set Reveals Limitations of Current 3D Methods. *J Chem Inf Model*. **2010**, 50, 2079–2093.

[93] Koeppen, H., Kriegl, J., Lessel, U., Tautermann, C.S., Wellenzohn, B. Ligand-Based Virtual Screening. In Virtual Screening. Wiley-VCH Verlag GmbH & Co. KGaA, **2011**, 61−85.

[94] Schuster, D. 3D pharmacophores as tools for activity profiling. *Drug Discov Today Technol*. **2010**, 7, e203-270.

[95] Spitzer, G.M.; Heiss, M.; Mangold, M.; Markt, P.; Kirchmair, J.; Wolber, G., Liedl, K.R. One concept, three implementations of 3D pharmacophore-based virtual screening: distinct coverage of chemical search space. *J Chem Inf Model*. **2010**, 50, 1241-1247.

[96] Sanders, M.P.A.; Barbosa, A.J.M.; Zarzycka, B.; Nicolaes, G.A.F.; Klomp, J.P.G.; de Vlieg, J.; Rio, A.D. Comparative analysis of pharmacophore screening tools. *J Chem Inf Model* **2012**, 52, 1607–1620.

[97] Vadivelan, S.; Sinha, B.N.; Rambabu, G.; Boppana, K.; Jagarlapudi, S.A.R.P. Pharmacophore modeling and virtual screening studies to design some potential histone deacetylase inhibitors as new leads. *J Mol Graph Model*. **2008**, 26, 935–946.

[98] Kandakatla, N.; Ramakrishnan, G. Ligand based pharmacophore modeling and virtual screening studies to design novel HDAC2 inhibitors. *Adv Bioinform*. **2014**, 2014 (2014), Article ID 812148.

[99] Wang, J.; Chen, L.; Sinha, S.H.; Liang, Z.; Chai, H.; Muniyan, S.; Chou, Y.W.; Yang, C.; Li, K.K.; Lin, M.F.; Jiang, H.; Zheng, Y.G.; Luo, C. Pharmacophore-based virtual screening and biological evaluation of small molecule inhibitors for protein arginine methylation. *J Med Chem*. **2012**, 55, 7978–7987.

[100] Drwal, M.N.; Griffith, R. Combination of ligand- and structure-based methods in virtual screening. *Drug Discov Today Technol*. **2013**, 10, e395–401.

[101] Drwal, M.N.; Agama, K.; Wakelin, L.P.; Pommier, Y.; Griffith, R. Exploring DNA topoisomerase I ligand space in search of novel anticancer agents. *PLoS One* **2011**, 6 (9), e25150.

[102] Svensson, F.; Karlén, A.; Sköld, C. Virtual screening data fusion using both structure- and ligand-based methods. *J Chem Inf Model*. **2012**, 52, 225-232.

[103] Swann, S.L.; Brown, S.P.; Muchmore, S.W.; Patel, H.; Merta, P.; Locklear, J. A unified, probabilistic framework for structure- and ligand-based virtual screening. *J Med Chem*. **2011**, 54, 1223-1232.

[104] Planesas, J.M.; Claramunt, R.M.; Teixidó, J.; Borrell, J.I.; Pérez-Nueno, V.I. Improving VEGFR-2 docking-based screening by pharmacophore postfiltering and similarity search postprocessing. *J Chem Inf Model*. **2011**, 51, 777-787.

[105] Kroemer, R.T. Structure-based drug design: docking and scoring. *Curr Protein Pept Sci*. **2007**, 8, 312-328.

[106] Meng, X.Y.; Zhang, H.X.; Mezei, M.; Cui, M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des*. **2001**, 7, 146-157.

[107] Warren, G.L.; Andrews, C.W.; Capelli, A.M.; Clarke, B.; LaLonde, J.; Lambert, M.H.; Lindvall, M.; Nevins, N.; Semus, S.F.; Senger, S.; Tedesco, G.; Wall, I.D.; Woolven, J.M.; Peishoff, C.E.; Head, M.S. A Critical Assessment of Docking Programs and Scoring Functions. *J Med Chem*. **2006**, 49, 5912-5931.

[108] Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J Comp-Aid Mol Des*. **2002**, 16, 11-26.

[109] Huang, N.; Kalyanaraman, C.; Bernacki, K.; Jacobson, M.P. Molecular mechanics methods for predicting protein-ligand binding. *Phys Chem Chem Phys*. **2006**, 8, 5166-5177.

[110] Ballester, P.J.; Mitchell, J.B.O. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. *Bioinformatics* **2010**, 26:1169-1175.

[111] Yuriev, E.; Ramsland, P.A. Latest developments in molecular docking: 2010–2011 in review. *J Mol Recognit*. **2013**, 26, 215–239.

[112] Ewing, T.J.; Makino, S.; Skillmam, G.; Kuntz, I.D. DOCK 4.0: Search strategies for automated molecular docking of flexibile molecule databases. *J Comput Aided Mol Des*. **2001**, 15, 411-428.

[113] Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. *J Comput Chem*. **1998**, 19, 1639-1662.

[114] Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S., Olson, A.J. Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. *J Comput Chem*. **2009**, 30, 2785-2791.

[115] Jones, G.; Willett, P.; Glen, R.C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J Mol Biol*. **1995**, 245, 43–53.

[116] Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K.; Shaw, D.E.; Francis, P.; Shenkin, P.S.; Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem*. **2004**, 47, 1739-1749.

[117] Spannhoff, A.; Heinke, R.; Bauer, I.; Trojer, P.; Metzger, E.; Gust, R.; Schüle, R.; Brosch, G.; Sippl, W.; Jung, M. Target-based approach to inhibitors of histone arginine methyltransferases. *J Med Chem*. **2007**, 50, 2319-2325.

[118] Spannhoff, A.; Machmur, R.; Heinke, R.; Trojer, P.; Bauer, I.; Brosch, G.; Schüle, R.; Hanefeld W.; Sippl, W.; Jung, M. A novel arginine methyltransferase inhibitor with cellular activity. *Bioorg Med Chem Lett*. **2007**, 17, 4150-4153.

[119] Vidler, L.R.; Filippakopoulos, P.; Fedorov, O.; Picaud, S.; Martin, S.; Tomsett, M.; Woodward, H.; Brown, N.; Knapp, S.; Hoelder, S. Discovery of novel small-molecule inhibitors of BRD4 using structure-based virtual screening. *J Med Chem*. **2013**, 56, 8073–8088.

[120] Zhao, H.; Gartenmann, L.; Dong, J.; Spiliotopoulos, D.; Caflisch, A. Discovery of BRD4 bromodomain inhibitors by fragment-based high-throughput docking. *Bioorg Med Chem Lett*. **2014**, 24, 2493-2496.

[121] Chen, S., Wang, Y., Zhou, W., Li, S., Peng, J., Shi, Z., Hu, J.; Liu, Y.C.; Ding, H.; Lin.; Y, Li, L.; Cheng, S.; Liu, J.; Lu, T.; Jiang, H.; Liu, B.; Zheng, M.; Luo, C. Identifying novel selective non-

nucleoside DNA methyltransferase 1 inhibitors through docking-based virtual screening. *J Med Chem*. **2014**, 57, 9028-9041.

[122] Kuck, D., Singh, N., Lyko, F., Medina-Franco, J.L. Novel and selective DNA methyltransferase inhibitors: docking-based virtual screening and experimental evaluation. *Bioorg Med Chem*. **2010**, 18, 822–829.

[123] Ashour, A.; El-Sharkawy, S.; Amer, M.; Abdel Bar, F.; Katakura, Y.; Miyamoto, T.; Toyota, N.; Bang, T.H.; Kondo, R.; Shimizu, K. Rational design and synthesis of topoisomerase I and II inhibitors based on oleanolic acid moiety for new anti-cancer drugs. *Bioorg Med Chem*. **2014**, 22(1), 211-220.

[124] Cai, H.; Huang, X.; Xu, S.; Shen, H.; Zhang, P.; Huang, Y.; Jiang, J.; Sun, Y.; Jiang, B.; Wu, X.; Yao, H.; Xu, J. Discovery of novel hybrids of diaryl-1,2,4-triazoles and caffeic acid as dual inhibitors of cyclooxygenase-2 and 5-lipoxygenase for cancer therapy. *Eur J Med Chem*. **2016**, 108, 89-103.

[125] Cardama, G.A.; Comin, M.J.; Hornos, L.; Gonzalez, N.; Defelipe, L.; Turjanski, A.G.; Alonso, D.F.; Gomez, D.E.; Menna, P.L . Preclinical development of novel Rac1-GEF signaling inhibitors using a rational design approach in highly aggressive breast cancer cell lines. *Anticancer Agents Med Chem*. **2014**, 14(6), 840-851.

[126] Dutta Gupta, S.; Revathi, B.; Mazaira, G.I.; Galigniana, M.D.; Subrahmanyam, C.V.; Gowrishankar, N.L.; Raghavendra, N.M. 2,4 dihidroxy benzaldehyde derived Schiff bases as small molecule Hsp90 inhibitors: rational identification of a new anticancer lead. *Bioorg Chem*. **2015**, 59, 97-105.

[127] Gao, C.; Bourke, E.; Scobie, M.; Famme, M.A.; Koolmeister, T.; Helleday, T.; Eriksson, L.A.; Lowndes, N.F.; Brown, J.A. Rational design and validation of a Tip60 histone acetyltransferase inhibitor. *Sci Rep*. 2014, 4, Article No. 5372.

[128] Khanfar, M.A.; AbuKhader, M.M.; Alqtaishat, S.; Taha, M.O. Pharmacophore modeling, homology modeling, and in silico screening reveal mammalian target of rapamycin inhibitory activities for sotalol, glyburide, metipranolol, sulfamethizole, glipizide, and pioglitazone. *J Mol Graph Model*. **2013**, 42, 39-49.

[129] Parker, J.P.; Nimir, H.; Griffith, D.M.; Duff, B.; Chubb, A.J.; Brennan, M.P.; Morgan, M.P.; Egan, D.A.; Marmion, C.J. A novel platinum complex of the histone deacetylase inhibitor belinostat: rational design, development and in vitro cytotoxicity. *J Inorg Biochem*. **2013**, 124, 70-77.

[130] Xue, W.; Song, B.A.; Zhao, H.J.; Qi, X.B.; Huang, Y.J.; Liu, X.H. Novel myricetin derivatives: Design, synthesis and anticancer activity. *Eur J Med Chem*. **2015**, 97, 155-163.

[131] Wang, Y.; Cheng, F.X.; Yuan, X.L.; Tang, W.J.; Shi, J.B.; Liao, C.Z.; Liu, X.H. Dihydropyrazole derivatives as telomerase inhibitors: Structure-based design, synthesis, SAR and anticancer evaluation in vitro and in vivo. *Eur J Med Chem*. **2016**, 112, 231-251.

[132] Staker, B.L.; Feese, M.D.; Cushman, M.; Pommier, Y.; Zembower, D.; Stewart, L.; Burgin, A.B. Structures of three classes of anticancer agents bound to the human topoisomerase I-DNA covalent complex. *J Med Chem*. **2005**, 48(7), 2336-2345.

[133] Berger, J.M.; Gamblin, S.J.; Harrison, S.C.; Wang, J.C. Structure and mechanism of DNA topoisomerase II. *Nature* **1996**, 379(6562), 225-232.

[134] Ellenbroek, S.I.; Collard, J.G. Rho GTPases: Functions and association with cancer. *Clin Exp Metastasis* **2007**, 24(8), 657-672.

[135] Fritz, G.; Just, I.; Kaina, B. Rho GTPases are over-expressed in human tumors. *Int J Cancer* **1999**, 81(5), 682-687.

[136] Fritz, G.; Brachetti, C.; Bahlmann, F.; Schmidt, M.; Kaina, B. Rho GTPases in human breast tumors: Expression and mutation analyses and correlation with clinical parameters. *Br J Cancer* **2002**, 87(6), 635-644.

[137] Kamai, T.; Yamanishi, T.; Shirataki, H.; Takagi, K.; Asami, H.; Ito, Y.; Yoshida, K. Overexpression of RhoA, Rac1, and Cdc42 GTPases is associated with progression in testicular cancer. *Clin Cancer Res*. **2004**, 10, 4799-4805.

[138] Ragno, R.; Mai, A.; Massa, S.; Cerbara, I.; Valente, S.; Bottoni, P.; Scatena, R.; Jesacher, F.; Loidl, P.; Brosch, G. 3-(4-Aroyl-1-methyl-1H-pyrrol-2-yl)-N-hydroxy-2-propenamides as a new class of synthetic histone deacetylase inhibitors. 3. Discovery of novel lead compounds through structure-based drug design and docking studies. *J Med Chem*. **2004**, 47, 1351-1359.

[139] Subha, K.; Kumar, G.R. Assessment for the identification of better HDAC inhibitor class through binding energy calculations and descriptor analysis. *Bioinformation* **2008**, 3, 218-222.

[140] Wang, D.F.; Helquist, P.; Wiech, N.L.; Wiest, O. Toward selective histone deacetylase inhibitor design: homology modeling, docking studies, and molecular dynamics simulations of human class I histone deacetylases. *J Med Chem*. **2005**, 48, 6936-6947.

[141] Butler, K.V.; Kalin, J.; Brochier, C.; Vistoli, G.; Langley, B.; Kozikowski, A.P. Rational design and simple chemistry yield a superior, neuroprotective HDAC6 inhibitor, tubastatin A. *J Am Chem Soc*. **2010**, 132, 10842–10846.

[142] Ran, T.; Zhang, Z.; Liu, K.; Lu, Y.; Li, H.; Xu, J.; Lu, S.; Liu, H.; Lu, T.; Chen, Y. Insight into the key interactions of bromodomain inhibitors based on molecular docking, interaction fingerprinting, molecular dynamics and binding free energy calculation. *Mol Biosyst*. **2015**, 11, 1295-1304.

[143] Singh, N.; Dueñas-González, A.; Lyko, F.; Medina-Franco, J.L. Molecular modeling and molecular dynamics studies of hydralazine with human DNA methyltransferase 1. *Curr Med Chem*. **2009**, 4(5), 792-799.

[144] Aldawsari, F.S.; Aguayo-Ortiz, R.; Kapilashrami, K.; Yoo, J.; Minkui, L.; Medina-Franco, J.L.; Velázquez-Martínez C.A. Resveratrol-salicylate derivatives as selective DNMT3 inhibitors and anticancer agents. *J Enzyme Inhib Med Chem*. **2016**, 31, 695-703.

[145] Medina-Franco, J.L.; Yoo, J. Docking of a novel DNA methyltransferase inhibitor identified from high-throughput screening: insights to unveil inhibitors in chemical databases. *Mol Divers*. **2013**, 17, 337–344.

[146] Angeles, E., Vázquez-Valadéz, V.H., Vázquez-Valadéz, O., Velázquez-Sánchez, A.M., Ramírez, A., Martínez, L.; Diaz-Barriga, S.; Romero-Rojas, A.; Cabrera, G.; Lopez-Castanares R.; Duenas-Gonzalez, A. Computational studies of 1-hydrazinophthalazine (hydralazine) as antineoplasic agent. Docking studies on methyltransferase. *Lett Drug Des Discov*. **2005**, 2, 282–286.

[147] Paricharak, S.; Cortes-Ciriano, I.; IJzerman, A.P.; Malliavin, T.E.; Bender, A. Proteochemometric modelling coupled to in silico target prediction: an integrated approach for the

simultaneous prediction of polypharmacology and binding affinity/potency of small molecules. *J Chem Info*. **2015**, 7, 15.

[148] Kinnings, S.L.; Liu, N.; Tonge, P.J.; Jackson, R.M.; Xie, L.; Bourne, P.E. A machine learning-based method to improve docking scoring functions and its application to drug repurposing. *J Chem Inf Model*. **2011**, 51, 408-419.

[149] Lounkine, E.; Keiser, M.J.; Whitebread, S.; Mikhailov, D.; Hamon, J.; Jenkins, J.L.; Lavan, P.; Weber, E.; Doak, A.K.; Côté, S.; Shoichet, B.K.; Urban, L. Large-scale prediction and testing of drug activity on side-effect targets. *Nature* **2012** 486:361-367.

[150] Kubinyi H. Drug research: myths, hype and reality. *Nat Rev Drug Discov.* **2003**, 2, 665-668.

[151] Terstappen, G.C.; Schlüpen, C.; Raggiaschi, R.; Gaviraghi, G. Target deconvolution strategies in drug discovery. *Nat Rev Drug Discov.* **2007**, 6:891-903.

[152] Jenkins, J.L.; Bender, A.; Davies, J.W. In silico target fishing: Predicting biological targets from chemical structure. *Drug Discovery Today: Technologies* **2006**, 3:413-421.

[153] Mavridis, L.; Mitchell, J.B. Predicting the protein targets for athletic performance-enhancing substances. *J Chemoinform*. **2013**, 5(1), 31.

[154] Nikolic, K.; Mavridis, L.; Bautista-Aguilera, O.M.; Marco-Contelles, J.; Stark, H.; do Carmo Carreiras, M.; Rossi, I.; Massarelli, P.; Agbaba, D.; Ramsay, R.R.; Mitchell, J.B. Predicting targets of compounds against neurological diseases using cheminformatic methodology. *J Comput Aided Mol Des*. **2015**, 29(2), 183-198.

[155] Martinez-Jimenez, F.; Papadatos, G.; Yang, L.; Wallace, I.M.; Kumar, V.; Pieper, U.; Sali, A.; Brown, J.R.; Overington, J.P.; Marti-Renom, M.A. Target Prediction for an Open Access Set of Compounds Active against Mycobacterium tuberculosis. *PLoS Comput Biol.*, **2013**, 9, 1003253.

[156] Gfeller, D.; Grosdidier, A.; Wirth, M.; Daina, A.; Michielin, O.; Zoete, V. SwissTargetPrediction: a web server for target prediction of bioactive small molecules. *Nucleic Acids Res.* 2014, 42:W32-W38.

[157] Cortes-Ciriano, I.; Ain, Q.U.; Subramanian, V.; Lenselink, E.B.; Mendez-Lucio, O.; IJzerman, A.P.; Wohlfahrt, G.; Prusis, P.; Malliavin, T.E.; van Westen, G.J.P.; Bender, A. Polypharmacology

modelling using proteochemometrics (PCM): recent methodological developments, applications to target families, and future prospects. *Med Chem Commun*. **2015**, 6:24-50.

[158] Glen, R.C.; Allen, S.C. Ligand-protein docking: Cancer Research at the interface between biology and chemistry. *Curr Med Chem*. **2003**, 10:767-782.

[159] Favia, A.D.; Nobeli, I.; Glaser, F.; Thornton, J.M. Molecular Docking for Substrate Identification: The Short-Chain Dehydrogenases/Reductases. *J Mol Biol.* **2008**, 375, 855-874.

[160] Schomburg, K.T., Rarey, M. What is the potential of structure-based target prediction methods? *Future Med Chem.* **2014**, 6, 1987-1989.

[161] Schomburg, K.T.; Bietz, S.; Briem, H.; Henzler, A.M.; Urbaczek, S.; Rarey, M. Facing the Challenges of Structure-Based Target Prediction by Inverse Virtual Screening. *J Chem Inf Model*. **2014**, 54, 1676-1686.

[162] Emig, D.; Ivliev, A.; Pustovalova, O.; Lancashire, L.; Bureeva, S.; Nikolsky, Y, et al. Drug Target Prediction and Repositioning Using an Integrated Network-Based Approach. *PLoS ONE* **2013**, 8, e60618.

[163] Nigsch, F.; Mitchell, J.B.O. Toxicological Relationships Between Proteins Obtained from Protein Target Predictions of Large Toxicity Databases. *Toxicol Appl Pharmacol.* **2008**, 231, 225-234.

Figures

Table 1. Reported QSAR studies used in rational drug design of antineoplastic agents

Drug target	Methodology	Software package	References
B-cell lymphoma 2 (Bcl-2) family of proteins	3D-QSAR Virtual Screening	DiscoveryStudio, version 2.2.5, Biovia Inc. (www.biovia.com) MATLAB, version 7.4.0.287- R2007a, MathWorks Inc. (www. mathworks.com)	[77]
B-cell lymphoma 2 extra-large (Bcl-2xl) family of proteins	3D-QSAR Virtual Screening Molecular Docking	PHASE, version 2.5, Schrödinger, LLC (www.schrodinger.com) Glide, version 4.5., Schrödinger LLC (www.schrodinger.com)	[78]
Mammalian target of rapamycin (mTOR)	Pharmacophore modelling 3D-QSAR	CATALYST, Accelrys Software Inc. (www.accelrys.com) Discovery Studio, version 2.55, Biovia Inc. (www.biovia.com)	[79]
Cyclin-dependent kinase 1 (CDK1)	Pharmacophore modelling 3D-QSAR Virtual Screening	CATALYST, version 4.11, Accelrys Software Inc. (www.accelrys.com) CERIUS2, version 4.10, Accelrys Software Inc. (www.accelrys.com)	[29]
13 breast cancer cell lines	mt-QSAR	LDA modules of StatSoft. STATISTICA, version 6.0 (www.statsoft.com)	[90]
10 colorectal cancer cell lines	mt-QSAR	LDA modules of StatSoft. STATISTICA, version 6.0 (www.statsoft.com)	[91]

Table 2. Reported structure-based studies used in rational design of antineoplastic agents

Drug target	Methodology	Software package	References
B-cell lymphoma extra- large (Bcl-xL) family of proteins	Molecular Docking	GOLD, version 5.1, CCDC (www.ccdc.cam.ac.uk) AutoDock Vina (vina.scripps.edu)	[80]
Topoisomerase I and II	Molecular Docking Virtual Screening	Molegro Virtual Docker (MVD) (molegro-virtual-docker.software.informer.com) GOLD, version 4.1 and 5.0, CCDC (www.ccdc.cam.ac.uk)	[123]
Cyclooxygenase-2 and 5-lipoxygenase	Molecular Docking	AutoDock Vina, (vina.scripps.edu)	[124]
Rac1-GEF	Molecular Docking Virtual Screening	Autodock4 program eHITS, SimBioSys Inc. (www.simbiosys.ca)	[125]
Hsp90	Molecular Docking	Surflex Geom X, Sybyl X-1.2 version (sybyl-x.software.informer.com)	[126]
Tip60 histone acetyltransferases	Molecular Docking Molecular Dinamics	Molecular Operating Environment (MOE), version 2010.10, Chemical computing group Inc. (www.chemcomp.com) YASARA, version 10.7.20 (www.yasara.org)	[127]
Mammalian target of rapamycin (mTOR)	Pharmacophore modeling Molecular Docking/Virtual Screening	CATALYST, Accelrys Software Inc. (www.accelrys.com) LigandFit docking engine	[128]
Histone deacetylase (HDAC)	Molecular Docking	Molecular Operating Environment (MOE), Chemical computing group Inc. (www.chemcomp.com)	[129]
Human telomerase reverse transcriptase (hTERT) and tetrahymena telomerase p65	Molecular Docking Molecular Dynamics	Glide, version 5.9, Schrödinger, LLC (www.schrodinger.com) Desmond, Schrödinger, LLC (www.schrodinger.com)	[130]
Human telomerase reverse transcriptase (hTERT)	Molecular Docking Molecular Dynamics	Glide, Schrödinger, LLC (www.schrodinger.com) Desmond, Schrödinger, LLC (www.schrodinger.com)	[131]