Chem Soc Rev

RSC Publishing

REVIEW ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Chem. Soc. Rev.*, 2013, **42**, 2130

Drug repositioning by structure-based virtual screening

Dik-Lung Ma,*[a] Daniel Shiu-Hin Chan[a] and Chung-Hang Leung*[b]

Received 29th August 2012

DOI: 10.1039/c2cs35357a

www.rsc.org/csr

Approved drugs have favourable or validated pharmacokinetic properties and toxicological profiles, and the repositioning of existing drugs for new indications can potentially avoid expensive costs associated with early-stage testing of the hit compounds. In recent years, technological advances in virtual screening methodologies have allowed medicinal chemists to rapidly screen drug libraries for therapeutic activity against new biomolecular targets in a cost-effective manner. This review article outlines the basic principles and recent advances in structure-based virtual screening and highlights the powerful synergy of *in silico* techniques in drug repositioning as demonstrated in several recent reports.

Introduction

The identification of small molecules that interact with specific biological targets and perturb their cellular functions has emerged as a crucial aspect of modern medicinal chemistry and biochemistry over the last few decades. During the early 1990s, extensive development in the field of high-throughput screening (HTS)[1] and combinatorial chemistry[2] spearheaded a revolution in the way that drug discovery was conducted. These techniques enabled medicinal chemists to synthesize large libraries of chemical compounds and to screen them in a fast and efficient manner. However, despite our constantly improving knowledge of molecular biology and the rising sums of money that have been invested into expanding drug pipelines, the rate of new drug approvals has remained painfully stagnant.[3] Several factors have been proposed to explain the failure of combinatorial chemistry and high-throughput screening to yield successful drugs. Firstly, low hit rates in conjunction with a high level of false positives have been observed for particular biological targets. Secondly, promising lead structures identified by HTS *in vitro* often encounter issues with bioavailability and toxicity at a later stage of development. Finally, combinatorial synthesis has

[a] *Department of Chemistry, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China. E-mail: edmondma@hkbu.edu.hk*
[b] *State Key Laboratory of Quality Research in Chinese Medicine, Institute of Chinese Medical Sciences, University of Macau, Macao, China. E-mail: duncanleung@umac.mo*

Dik-Lung Ma

Dik-Lung Ma completed his PhD in 2004 *at the University of Hong Kong under the supervision of Prof. C.-M. Che. He spent the years* 2005–2009 *at the University of Hong Kong, the Hong Kong Polytechnic University, and the Scripps Research Institute with Prof. C.-M. Che, Prof. K.-Y. Wong, and Prof. R. Abagyan. In* 2010 *he was appointed Assistant Professor at the Hong Kong Baptist University. His research mainly focuses on luminescent sensing of biomolecules and metal ions, computer-aided drug discovery, and inorganic medicines.*

Daniel Shiu-Hin Chan

Daniel Shiu-Hin Chan completed his BSc degree in Chemistry at the University of New South Wales, Sydney, Australia. He is currently appointed as a Research Assistant at the Department of Chemistry at the Hong Kong Baptist University under the supervision of Dr Dik-Lung Ma and Dr Chung-Hang Leung. His research interests include drug discovery and the development of luminescent oligonucleotide-based assays for biomolecules and metal ions.

been criticized for generating highly similar chemical structures with limited diversity, thus restricting the chemical space that is sampled in each project.

The pessimistic outlook of present pipelines combined with the impending expiry of blockbuster patents as well as increasingly stringent approval regulations has driven the pharmaceutical industry to seek and develop efficient techniques for drug discovery. Over the last few decades, advances in both computational algorithms and computer processing power have made structure-based virtual screening a popular tool in early drug discovery and development.[4] The use of computational technologies to identify potential drug candidates *in silico* can dramatically reduce the time and cost required for chemical synthesis and *in vitro* testing. Meanwhile, finding new uses for existing drugs has found credence as a cost-effective strategy to complement existing pipelines. The methodology of developing new pharmacological indications for existing drugs is known as drug repositioning (also known as drug redirecting, drug repurposing or drug reprofiling).[5] The number of successful cases of drug repositioning from serendipity or systemic exploration has increased over recent years. However, only a few of them have been identified *via* structure-based virtual screening. This review aims to provide a brief introduction to the application of structure-based virtual screening in the repositioning of existing drugs for new indications. The synergy of these two highly cost-effective methodologies (virtual screening and drug repositioning) may be particularly relevant in the climate of mergers, acquisitions and down-sizing that has engulfed the pharmaceutical industry in recent years.

Drug repositioning: principles, motivation and practical considerations

The dominant paradigm of pharmaceutical research in the past two decades has primarily revolved around target-based drug discovery. In this process, candidate compounds are screened against isolated biomolecules in order to identify molecularly targeted agents with high potency and selectivity towards particular disease targets. However, the results obtained from such efforts have been less than stellar and notable late-stage clinical failures arise frequently due to undesirable pharmacokinetics and/or drug metabolism.[6] This landscape has engendered renewed interest in phenotypic screening and the role of polypharmacology in the search for effective drugs.[7]

A strategy that aims to harness the positive aspects of polypharmacology is drug repositioning, whereby existing drugs are investigated for efficacy against disease targets for other indications. Drug repositioning has gained increasing attention from both academia and industry in recent years.[8,9] The concept behind drug repositioning is based on both the promiscuous nature of small molecules[10] and the interconnectedness of cell signaling networks and signal transduction pathways.[11] Due to the vast numbers of distinct chemical and biochemical entities contained within a living organism, it is practically impossible for a small molecule to achieve 100% selectivity for its cognate target. While such "promiscuity" has traditionally been regarded as undesirable in the drug development process, increasing evidence suggests that certain approved drugs may exert multiple pharmacological effects with little to no explicit connection with their intended indications (polypharmacology).[11] This factor could account for the differential efficacies and side effects typically observed between drugs developed for the same biomolecular target.

A further stimulus for drug repositioning techniques has been the rising costs and difficulties associated with bringing new drugs to the market. The drug discovery and development process can take up to 10 to 17 years from hit identification to clinical approval,[12] with estimated costs of up to 1 billion US dollars. Attempts to accelerate the process by pushing premature drug candidates through the clinic can be risky and can result in severe loss of investment at the late stage. Furthermore, the increasingly stringent approval regulations that have been imposed in recent years have further increased the resource investment required for bringing a new drug to the market. In this context, drug repositioning can significantly decrease the time and cost of development compared to *de novo* discovery due to the availability of pharmacokinetic and toxicological data that have been obtained for existing drugs in previous clinical trials.[12] Therefore, the risks and timelines associated with clinical development of repositioned drugs can be significantly reduced.[13] Furthermore, drug repurposing may represent one of the only few paradigms by which "orphan" or neglected diseases may be profitably pursued by the pharmaceutical industry.

Perhaps the most well-known example of drug repositioning is the serendipitous discovery of sildenafil (trade name: Viagra or Revatio) for the treatment of erectile dysfunction. Sildenafil was originally studied as an anti-anginal compound developed as an inhibitor of the phosphodiesterase family. However, it was later realized that sildenafil most effectively antagonizes the phosphodiesterase 5 (PDE5) isoform, an enzyme not found in heart tissue. It was eventually repositioned for the treatment of male erectile dysfunction in 1998 during Phase I clinical trials.[5] Another famous example is thalidomide, a sedative introduced in the late 1950s that was originally used to treat morning sickness, but was withdrawn in 1961 due to its teratogenic

Chung-Hang Leung

Chung-Hang Leung completed his PhD in 2002 *at City University of Hong Kong. After completing a five year post-doctoral fellowship at the Department of Pharmacology, Yale University, he was appointed Research Assistant Professor at the University of Hong Kong and then at the Hong Kong Baptist University. He is currently appointed as Assistant Professor at University of Macau. His primary research interests are in anti-cancer and anti-inflammatory drug discovery and molecular pharmacology.*

effects in pregnant women. Remarkably, the drug was redeveloped in the 1990s as a treatment for multiple myeloma with strict controls to prevent side effects in pregnant women.[5]

In recent years, a wide variety of techniques have been investigated in conjunction with drug repositioning in an attempt to accelerate or streamline the early phases of pharmaceutical research, such as phenotypic screening,[14,15] high-throughput experimental screening[16,17] and database mining.[18,19] These strategies aim to discover potential candidates that can be directly repositioned or alternatively to identify critical functionalities that can be modified in order to influence the pharmacology of existing drugs. In a large-scale study, Keiser and co-workers utilized a statistics-based chemo-informatics approach to compare 3665 FDA-approved and investigational drugs against 65 241 ligands organized into 246 targets drawn from the MDL Drug Data Report (MDDR) database,[20] which yielded 23 new drug–target associations.[21] Very recently, Pollastri and co-workers have designed analogues of the human Aurora kinase inhibitor, danusertib, in order to generate potent and selective compounds against the corresponding trypanosomal Aurora kinase 1 (TbAUK1) for the treatment of human African trypanosomiasis, a neglected tropical disease.[22]

The last several years have witnessed a significant surge of interest in drug repositioning strategies and applications. For example, recent summaries have reviewed the applications of drug repurposing for conditions ranging from stroke,[23] malaria and TB,[24] pediatric diseases,[25] hematological malignancies,[26] schistosomiasis.[27] Readers interested in the latest developments in the general field of drug repositioning may further peruse a number of comprehensive review articles or perspectives published recently.[28–32]

It should be stressed that while drug repositioning offers potentially tremendous savings in time and cost, it does not guarantee a smooth journey to the market. Repositioned leads may suffer from non-optimized interactions with the new biomolecular target, and may require further structural modification in order to enhance specificity towards the new target while, optionally, reducing activity against the original target. Furthermore, although available pharmacokinetic, pharmacodynamic and toxicological data for the drug candidate may allow the compound to bypass routine preliminary screening assays, the dosage and efficacy of the repositioned drug will still have to be rigorously tested in a suitable animal disease model for the secondary indication.[8]

On the clinical front, additional factors have to be considered to ensure that the polypharmacological nature of the repositioned drug does not pose undue safety concerns.[33] For example, drug repositioning may potentially uncover additional side effects or risks unobserved in patients of the primary indication. Furthermore, it is likely that the dosage required for repositioned indication would vary significantly from that of the original indication, and thus the optimum dosage and associated pharmacokinetic and metabolic parameters would have to be fully re-evaluated. In addition, alternative formulations or delivery methods for the repositioned compound will necessitate additional safety trials to identify potential adverse effects. Finally, the ever-increasing stringency of drug approval regulations means that discontinued drug compounds used in the past may have to undergo complete safety profiling as part of the drug repositioning process.

Drug repositioning campaigns may sometimes encounter patent-related issues.[8] The costly and arduous nature of drug development compels pharmaceutical companies to take extensive measures to guarantee the return of their investment. However, such protection policies may present additional roadblocks that hinder the commercialization of repositioned drugs. Other important aspects such as the market share, intellectual property and formulation of new development plans of the repositioned drugs should also be considered carefully during the repositioning process.[34]

Structure-based virtual screening

Virtual screening, including target- and ligand-based approaches, has emerged as an effective methodology for prioritizing compounds for high-throughput screening and in the structure-guided optimization of lead compounds.[35–37] Structure-based virtual screening (also known as molecular docking) allows the interactions between large libraries of compounds and a biomolecular target to be evaluated in a rapid and cost-effective manner. Besides *in silico* screening, molecular docking can also be used to study the binding mode of small molecules to the target binding site.[38]

The combination of structure-based virtual screening and drug repositioning potentially represents a highly efficient methodology for the development of new medicines. Existing drugs constitute a class of privileged structures with verified bioavailability and compatibility. Consequently, there is reduced likelihood that a highly potent hit compound identified through virtual screening would disappoint in cellular or animal studies. Furthermore, virtual screening allows academic groups or smaller companies to conduct exploratory repositioning of existing drugs against novel disease targets without the expense of purchasing thousands of compounds.

Drug repositioning by molecular docking can be operated *via* a ''single target'' approach which aims to identify potential interactions between the drug candidates and a particular target of interest. Alternatively, ''inverse docking'' can be used to investigate the binding of an existing drug against a panel of known therapeutic targets.[39] This approach could potentially be used to uncover novel drug–target interactions as well as evaluate the selectivity of the drug compounds against the desired target. Furthermore, *in silico* techniques can be used to predict the absorption, distribution, metabolism, elimination and toxicity (ADMET) characteristics of modified drug analogues.[40–42]

The molecular docking process generally involves three main steps: (i) generation of the molecular model of the target and the pre-treatment of the ligands, (ii) conformational sampling of the ligands within the binding site of the target, and (iii) a score assignment reflecting the binding energy of the ligand–target complex. This is also acceptable to molecular docking for drug repositioning setting. For *in silico* drug repositioning, specialized

virtual libraries are available that contain information about both existing drugs as well as compounds that have made it to late stage clinical trials but were un-marketed. Some freely accessible drug databases include DrugBank,[43] e-Drug 3D,[44] KEGG DRUG[45] and the superdrug database.[46]

Docking

In the docking process, three methods are commonly employed to tackle the issue of ligand flexibility, namely (i) ligand incremental construction, (ii) generation of multiple conformers (rotamer library) before docking and (iii) stochastic methods. In ligand incremental construction, the ligands are first partitioned into small fragments which are individually docked into the receptor pocket based on the geometric fit. The entire ligand is then incrementally assembled from the ligand fragments within the binding pocket.[47] In the second strategy, multiple low-energy conformations of the same ligand are generated, which are then individually docked against the receptor pocket.[48]

However, the most widely used strategies to account for ligand flexibility are stochastic methods such as Monte Carlo (MC)[49] or genetic algorithms (GA).[50] A Monte Carlo algorithm operates by stochastically varying one parameter at a time in order to generate new conformations which are accepted or rejected based on Boltzmann considerations. The modeling is initiated at a high enough temperature such that there is a significant chance of accepting the next conformation sampled. The temperature is then progressively decreased during docking in order to trap the ligand–target complex in a low energy state as a consequence of the reduced conformational freedom during cooling. On the other hand, genetic algorithms adopt a totally different approach inspired by Darwin's theory of evolution. The conformations of the ligands begin as a random population of states modeled as a set of chromosomes. The ligand is allowed to perform random crossovers and mutations in order to produce another set of ligands with different conformations. The "fittest" compounds, which possess the lowest binding energies with the target, are accepted and used to produce a new generation. This cycle is iteratively repeated a number of times until the local energy minimum of the target–ligand complex has been reached.

Traditional scoring functions

Scoring functions are used to predict the binding free energy of the target–ligand complex, which is a measure of the binding potency of the small molecule for the biomolecular target. Scoring functions can be grouped into three categories: force-field-based, empirical-based and knowledge-based scoring.[4,51] Force-field-based scoring functions compute the potential energy of the entire ligand–target complex by summing together contributions from different interactions, such as van der Waals' or electrostatic energies between the atoms of the ligand and the biomolecule. In empirical-based scoring functions, a set of weighted parameters corresponding to various types of interactions between the target and ligand are calibrated with a training set of experimental structures. The coefficients of the different terms, which usually include polar interactions such as hydrogen bonding and electrostatic interactions, and apolar interactions such as aromatic and lipophilic interactions, are optimized by fitting the structures and binding affinities of the training set into a multiple linear regression model. In knowledge-based scoring, the structural information of known target–ligand complexes is extracted and used to derive atomic interaction potentials describing the interactions between the atoms of the ligand and the target. It should be noted that each type of scoring function has its inherent drawbacks. For example, empirical scoring relies strongly on the similarity between the predicted binding modes of the ligand–target complex with the compounds of the training set. Thus, the predictive power of empirical scoring functions is influenced by both the quality of the structural data of the biomolecular target as well as that of the binding affinity data of the training set. Meanwhile, force-field scoring functions may neglect to account for entropic change, while exaggerating the contribution of intermolecular energy values in the binding free energy calculation. Furthermore, current docking and scoring algorithms are not yet able to accurately model the contributions of protein flexibility[52] or water-mediated ligand binding.[53] Considerable effort is continually being invested into improving the performance of scoring algorithms, such as including additional terms into the scoring functions to better estimate the solvation effect of the ligands or the change in entropic energy during target–ligand binding. Furthermore, the involvement of higher level quantum mechanical calculations to devise a more accurate scoring function has also been reported.[54] Finally, the use of multiple scoring functions in concert, termed consensus scoring, has been proposed in order to provide a better estimation of the binding affinity of the ligand–target complex.[55,56]

Machine learning methods and scoring functions

Machine learning methods have attracted increasing attention for estimation of ligand–target interactions. These methods avoid the error-prone calculation of intractable terms such as ligand solvation and entropic changes that occur during ligand–target binding. In recent studies, machine learning-based methods have been demonstrated to exhibit comparable or even superior performance in estimating ligand binding affinity compared to traditional scoring functions. For example, Durrant and McCammon used neural networks (NN), which are computer models designed to mimic the microscopic organization of the human brain, to characterize the binding affinities of protein–ligand complexes.[57] This neural-network-based scoring function is able to "learn" the favourable binding interactions and characteristics contributing to the ligand–target complex. Ballester and Mitchell have reported a novel scoring function termed RF-score that utilizes the random forest machine learning technique to implicitly capture binding effects that are difficult to model explicitly.[58] In the context of drug repositioning, Bourne and co-workers have used support vector machines (SVM), trained by associating individual energy terms from molecular docking with the known binding affinity of the

ligands, to identify phosphodiesterase inhibitors as direct inhibitors of *Mycobacterium tuberculosis*.[59]

Protein flexibility

Shape complementarity between protein and ligand is a fundamental concept in cellular biology. In the 1950s, Koshland proposed that the binding pocket of a protein undergoes significant structural changes upon interaction with its substrate.[60] The movement of various amino acid residues within the receptor pocket forms a "true receptor pocket" for the binding of the ligand. It is now well-known that most proteins possess varying degrees of flexibility, which can be largely attributed to the flexible nature of amino acid side chains. This flexibility can range from a slight perturbation of the original binding pocket to a complete reconstitution of the receptor pocket if very significant shifts are involved.[61] Thus, an inadequate treatment of protein flexibility could result in the rate of both false positives and false negatives in a virtual screening experiment.

Several approaches have been investigated to tackle the issue of protein (receptor) flexibility in recent years. One strategy is to utilize multiple high-quality static receptor conformations in docking runs and to select the highest-scoring conformation for further investigation.[62] For example, Miteva and co-workers have used normal mode analysis to choose relevant conformations appropriate for subsequent ligand docking using cyclic dependent receptor kinase 2 (CDK2) as a model structure.[63]

A similar strategy is to utilize multiple receptor conformations (MRC) of the biomolecule obtained from different X-ray or NMR structures, or generated from molecular dynamics simulations *in silico*.[64,65] For instance, Okamoto and co-workers have developed a MultiCopyMD method to generate a target ensemble of the ligand in the receptor binding site by molecular dynamic (MD) simulations.[66] This target ensemble was subsequently used in structure-based virtual screening to identify a potent and selective inhibitor of death-associated protein kinase 1 (DAPK1).[66] In another approach, Abagyan and co-workers have developed an efficient "four-dimensional" docking methodology that allows ligands to be fitted against multiple receptor conformations in a single docking simulation.[67,68] This strategy generates a single 4D grid map by merging the 3D grids of the most stable structures of multiple target conformers, and the success rate was dramatically enhanced from 50% for a single docking run to 80–90% for the 4D ensemble docking algorithm. In another work, Abagyan and co-workers have investigated strategies for the selection of experimental protein conformations for virtual ligand screening and have found that the use of ensemble conformations of receptors co-crystallized with larger ligands provided the best results.[69] However, it has been noted that the use of excessively large numbers of conformers in ensemble docking can lead to an increased number of false positive results.[67,70]

An alternative approach to account for receptor flexibility is to employ "soft docking" where the interaction of the protein target and the ligand is continuously changed to allow partial clashing between the atoms of the ligand and target.[71] Soft docking has generally given superior results compared to rigid docking protocols but has also been associated with an increased number of false positive hits in high-throughput screening campaigns.[72]

Finally, some docking algorithms are able to explicitly model receptor flexibility using predictive molecular dynamics (MD) or exhaustive sampling of the protein–ligand conformational space.[73,74] However, this is usually constrained to the ligand binding domain as the explicit inclusion of receptor flexibility for the whole protein in the docking calculations would be too computationally demanding.[75] An alternative option is to only model side chain flexibility whereby the receptor backbone is fixed while the side chain conformations are sampled.

Post-docking analysis

Scoring functions play a central role in the identification of potential drug hits in a virtual screening campaign. However, the inability of current scoring functions to accurately predict binding energy values places a major limitation on the quality of the docking results. While manual visual inspection has and still finds use in post-docking analysis, the technique becomes proportionately less efficient as the sampling size is increased. A number of methods have been devised in recent years to eliminate false positive hits obtained from the initial docking experiment in order to improve the hit-rate in subsequent *in vitro* assays.

Marcou and Rognan have proposed the use of molecular interaction fingerprints (IFPs), which are simple bit strings that convert the 3D information of protein–ligand interactions into a 1D vector representation that can be compared quickly using traditional metrics for prioritizing the most relevant poses of low-molecular-weight fragments or molecular scaffolds in order to increase the accuracy of a docking run.[76] In their study, the authors examined the utility of IFPs as post-docking filters for the screening of CDK2 inhibitors using four different docking algorithms (FlexX, Gold, glide, and Surflex) and the Tanimoto metric (Tc-IFP) to evaluate the similarity of the predicted X-ray IFP to the re-docked pose. The authors demonstrated that scoring by the Tanimoto similarity of IFPs to a given reference was statistically more accurate compared to conventional scoring functions.

Based on the assumption that active compounds should have certain specific interactions or contacts with their corresponding target, Bertho and co-workers have reported a strategy to analyze the binding pose of ligands termed the automatic analysis of poses using self-organizing maps (AuPosSOM).[77] This approach aims to rank docked poses by examining the interatomic contacts between the ligand and the binding site of the biomolecular target. The authors showed that this strategy was able to discriminate active compounds from decoys using contact footprint clustering techniques.

Lai and co-workers have reported an approach using binding energy landscape analysis to discriminate true positive hits from decoy compounds.[78] This strategy focuses on two parameters

(the energy gap and the number of local binding wells in the landscape) that reflect the thermodynamic stability and the kinetic accessibility of the binding energy landscape. The authors successfully used this strategy to discriminate inhibitors from high-scoring decoys of neuraminidase and cyclooxygenase-2.

Case studies

Over the last decade, there have been numerous reports of the application of structure-based virtual screening for drug repositioning. In this section, we describe interesting case studies of existing drugs that have been re-discovered for new indicators with the aim of highlighting the efficient combination of structure-based virtual screening and drug repurposing for the streamlining of early-phase drug discovery. These examples complement those described in a recent review by Ekins and co-workers that describe the application of *in silico* repositioning of existing drugs for rare and neglected diseases.[79]

Repositioned drug as a G-protein coupled receptor inhibitor

G-protein coupled receptors (GPCRs) are a large family of transmembrane receptors sharing a common seven α-helix structural motif and they have historically represented an important class of drug targets.[80,81] GPCRs control a variety of signal transduction pathways including the release of hormones and neurotransmitters as well as visual and olfactory sensory functions.[82] All 5-hydroxytryptamine (5-HT) receptors (5-HT_1 to 5-HT_7), except the 5-HT_3 receptor which is a ligand-gated ion channel, are GPCRs that control intracellular signaling cascades through the production or inhibition of secondary messengers.[83] In 2012, Huang and co-workers conducted a structure-based virtual screening campaign to identify new 5-HT_{2A} inhibitors.[84] 5-HT_{2A} is one of the most well-studied 5-HT receptors and the inhibition of the 5-HT_{2A} signaling pathway leads to anti-psychotic and anti-depressive effects.[85] A total of 1430 drugs from DrugBank and the ZINC FDA drug database were screened against two receptor models using in-house automatic docking and the MM-GB/SA rescoring protocol.[86] The best-scoring pose of each compound was rescored and the 200 highest-scoring compounds were retained for visual inspection and further analysis. The authors chose the top-scoring compounds identified using the cyproheptadine model for further investigation due to the higher enrichment and superior pose prediction of this model. These compounds were further filtered by discarding compounds that lack the ability to form favourable hydrogen bonding interactions with the conserved Asp residue of the receptor pocket, which resulted in 99 remaining compounds. Among these 99 existing drugs, 73 compounds were further eliminated due to their reported activities against other GPCRs. Finally, 6 of the remaining 26 compounds were chosen for *in vitro* investigation based on their commercial availability. The well-known multiple kinase inhibitor sorafenib (Fig. 1), used for the treatment of renal cell carcinoma (RCC) and hepatocellular carcinoma (HCC), was identified to inhibit 5-HT_{2A} activity in a competitive binding assay. Pharmacological profiling results indicated that this drug bound strongly to 5-HT_{2B} and 5-HT_{2C} over 5-HT_{2A}. The authors then synthesized two related sorafenib analogues and evaluated their binding affinity against the 5-HT_{2B} receptor. Interestingly, the replacement of the aromatic nitrogen atom in sorafenib by the carbon atom in compound **1** lead to a slight improvement in binding affinity, suggesting that the nitrogen atom does not form significant interactions with the receptor. On the other hand, the addition of a methyl group on the amide nitrogen afforded compound **2** which displayed dramatically reduced binding affinity towards 5-HT_{2B}. The lowered binding affinity of compound **2** was rationalized by the removal of hydrogen bond capability between the amide nitrogen of the ligand and the carbonyl group of a lysine residue in the receptor binding pocket.

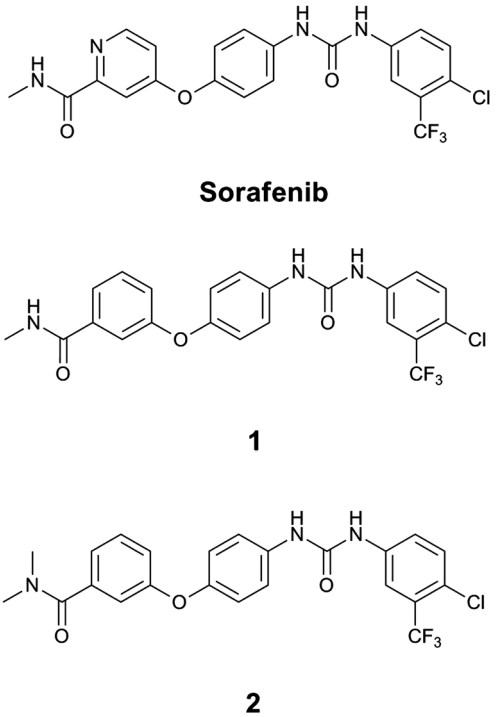


Fig. 1 Chemical structures of the anti-cancer drug sorafenib and its derivatives **1** and **2**.

Repositioned drug as a nuclear receptor antagonist

Nuclear receptors are ligand-inducible transcription factors that mediate gene expression in response to steroid hormones or other molecules. Some nuclear receptors are termed ''orphan receptors'' as their cognate ligands have not yet been identified or may not exist.[87] These nuclear receptors commonly share three structural domains: the N-terminal transactivation domain (TAD), the DNA-binding domain (DBD), and the C-terminal ligand-binding domain (LBD) which is connected to the DBD through a hinge region. The LBDs of nuclear receptors share

high sequence and structural similarity which makes the design and discovery of selective nuclear receptor antagonists difficult. In 2007, Bisson, Abagyan and co-workers applied structure-based high throughput docking techniques in order to identify androgen receptor antagonists from a database of existing drugs.[88] The human androgen receptor (AR) belongs to the superfamily of nuclear receptors and its dysregulation has been associated with the development and progression of prostate cancer. The authors first constructed two initial models of the AR LBD domains: the bicalutamide-biased model (B-model) and the flutamide-biased model (F-model). The models were then calibrated using a training set of 24 known antagonists in order to identify a suitable conformer that could effectively distinguish AR antagonists from decoy compounds. Two representative conformers were selected for subsequent virtual screening experiments. Ligands from a database of marketed oral drugs were first prepared by energy minimization in the absence of the receptor. The compounds were then docked against the receptor pocket using the ICM method. The highest-scoring binders were visually inspected, and 11 compounds scoring highest in both structural model variants were selected for further biological investigation. In the *in vitro* assays, three phenothiazine derivatives (acetophenazine, fluphenazine and pericyazine, Fig. 2) exhibited anti-AR transactivation efficacies in an *in vitro* chloramphenicol acetyltransferase (CAT) reporter assay. These three phenothiazines are marketed as antipsychotics that interact strongly with serotonergic and dopaminergic GPCRs. A subsequent similarity search based on the phenothiazine core structure was performed which yielded the phenothiazine derivative **3** as a potent anti-androgenic agent with sub-micromolar potency in the CAT reporter assay (Fig. 2). Further experiments demonstrated that **3** showed almost no cross-reactivity with other receptors and behaved as an anti-androgenic agent with diminished binding affinity towards the original biological targets of the phenothiazines, such as serotonin and dopamine GPCRs. This report from Bisson, Abagyan and co-workers suggested that through modification of the functional group on the drug, a markedly reduced affinity to the original target, along with enhancement of the secondary pharmacological effect, could be achieved and this can lead to the repositioning of drugs with limited ''off-target'' and undesirable side effects.

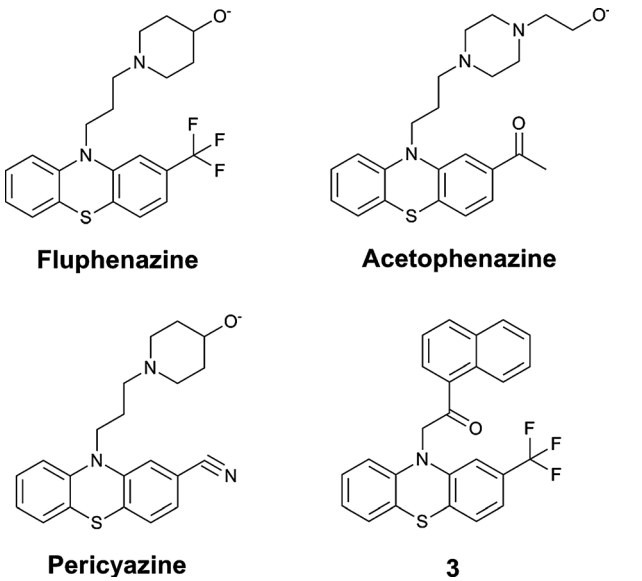


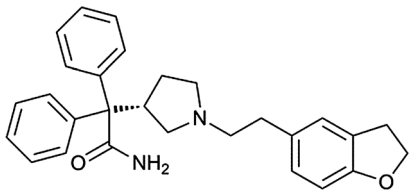
Fig. 2 Chemical structures of anti-psychotic phenothiazine derivatives repositioned as antagonists of the androgen receptor (AR).

Repositioned drugs as modulators of protein–protein interactions

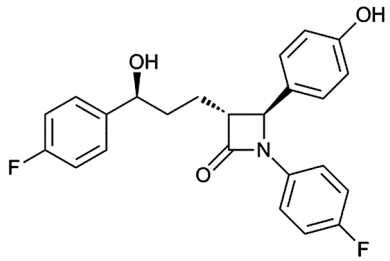
Historically, the pharmaceutical industry has largely focused on the discovery and development of small molecule inhibitors that target the active/binding sites of enzymes or protein receptors, which are typically small, well-defined and solvent-shielded.[89] As the low-lying fruits from those areas have dried up, recent efforts have turned towards protein–protein interactions (PPI) that play key roles in controlling cell proliferation, differentiation and apoptosis. However, targeting protein–protein interfaces is considered a particularly challenging task due to their typically large size ($\sim$1500–3000 Å) compared to protein-small-molecule binding sites ($\sim$300–1000 Å) as well as their amorphous nature that lack well-defined cavities for recognition by small molecules. Moreover, another problem associated with some protein–protein targets is their inherent flexibility. The movements of side chains and perturbations of loops under dynamic equilibrium or the influence of small molecule modulators may affect the conformation of the protein surface, leading to the formation of ''transient'' binding pockets that may be absent in the static structure of the free protein target or the protein–protein complex.[90] These factors make the discovery of small molecule PPI modulators generally more difficult than for traditional enzyme or receptor inhibitors.

Recent advances in the biophysical characterization of PPIs have provided insight into how these surfaces might be targeted. Some studies have shown that the stability of these macromolecular complexes is dependent on particular ''hot spots'' rather than the size of the protein–protein interface.[91] These hot spots are small subsets of amino acid residues between the protein–protein interfaces which are determined to provide a major contribution towards the binding affinity of the surface.[92] Such developments have stimulated the discovery of both small-molecule[93] and proteomimetic[94] PPI modulators over the recent years, with some of them identified through structure-based virtual screening. We highlight here two examples of existing drugs identified through structure-based virtual screening that have been repositioned as an inhibitor of the protein–protein interface.

The tumor necrosis factor-α (TNF-α) trimer is an important human cytokine that is involved in inflammatory response through regulation of diverse signaling pathways. The aberration of cellular levels of TNF-α has been implicated in a variety of inflammatory disorders.[95] The clinically improved biopharmaceutical infliximab targets TNF-α trimerization and is routinely used to treat inflammatory disorders such as rheumatoid



Darifenacin



Ezetimibe

Fig. 3 Chemical structures of the TNF-α PPI inhibitors: darifenacin (overactive bladder syndrome) and ezetimibe (hypercholesterolemia).

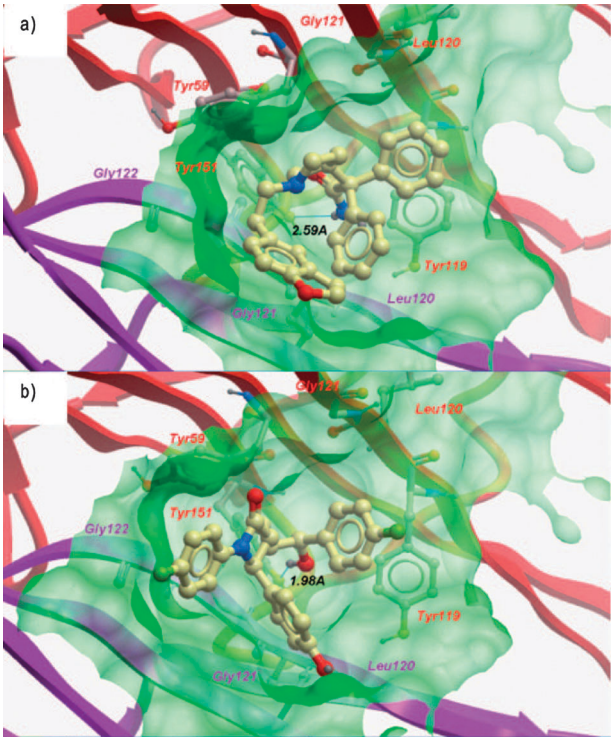
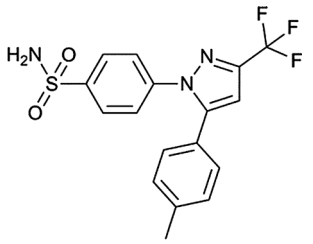


Fig. 4 Low-energy binding conformations of (a) darifenacin and (b) ezetimibe bound to the TNF-α dimer generated by molecular docking. The TNF-α dimer is depicted as ribbon form displaying subunit A (purple) and subunit B (red). The small molecules are depicted as ball-and-stick models showing carbon (yellow), hydrogen (grey), oxygen (red), nitrogen (blue), and fluorine (green) atoms. The binding pocket of the TNF-α dimer is represented as a translucent green surface. Reproduced from ref. 99 by permission of John Wiley & Sons, Ltd.

arthritis, psoriatic arthritis, and Crohn's disease. However, the use of TNF-α antibodies such as infliximab can elicit an autoimmune anti-antibody response or weaken the body's immune system to opportunistic infections. This has stimulated the development of few small molecule inhibitors of TNF-α trimerization identified *via* experimental screening[96] or structure-based virtual screening.[97,98] In 2011, Ma, Leung and co-workers applied structure-based, high-throughput virtual screening methods to identify small-molecule inhibitors of tumor necrosis factor-α (TNF-α) from a library of marketed drugs.[99] The high scoring structures were visually inspected and seven of these compounds were experimentally tested for TNF-α inhibition using an enzyme-linked immunosorbent assay (ELISA). Two of these compounds (darifenacin and ezetimibe, Fig. 3) were subsequently demonstrated to disrupt the TNF-α–TNF-α receptor interaction *in vitro* and down-regulate TNF-α-driven gene expression in human cells. Darifenacin (trade name: Enablex) is currently used in the treatment of overactive bladder (OAB) syndrome by antagonizing the action of the muscarinic acetylcholine receptor, while ezetimibe (trade name: Zedia) is a potent inhibitor of cholesterol absorption in intestine and is usually co-prescribed with statins for the control of cholesterol levels (Fig. 4).

The second example of an existing drug repositioned as a PPI inhibitor through virtual screening relates to the signal transduction and activator of transcription 3 (STAT3). The STAT3 dimer is a key transcription factor through which receptors of multiple cytokines and growth factors exert their effects. The phosphorylation of STAT3 at the tyrosine 705 residue in the SH2 domain induces the formation of STAT3 homodimers or STAT3–STAT1 heterodimers that subsequently translocate into the nucleus and activate target genes through binding to specific DNA-response elements. STAT3 is constitutively activated in many types of cancers and has been linked to tumor progression through enhancing angiogenesis, metastasis, growth and survival of cancer cells.[100] In 2011, Li and co-workers employed a multiple ligand simultaneous docking (MLSD) approach to identify a small molecule inhibitor of STAT3 dimerization.[101] They first constructed a fragment library based upon reported small molecule inhibitors of the STAT3 SH2 domain. The fragments were then filtered using a similarity search of chemically or structurally similar entities from a drug scaffold database in order to weed out any fragments with undesirable ADMET properties. In the second stage of the screening, three drug fragments from the library, including one polar fragment and two non-polar fragments, were simultaneously docked against a pre-treated model of the STAT3 SH2 binding pocket and a docking score was assigned to each fragment. Finally, the high-scoring fragments were linked together to generate fifteen virtual templates using various chemical linkers such as amide, amine, ether or alkene linkers. The virtual templates were used as pharmacophores for screening an FDA-approved drug database. Thirteen of the fifteen pharmacophore models identified celecoxib (Fig. 5) as a top hit for STAT3 inhibition. *In vitro* experiments demonstrated that celecoxib could down-regulate STAT3 phosphorylation in a dose-dependent manner, selectively antagonize interleukin-6 (IL-6)-induced STAT3 phosphorylation and inhibit cancer cell growth with micromolar potency.



Celecoxib

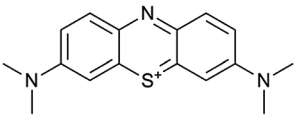
Fig. 5 Chemical structure of the STAT3 transcription factor dimerization inhibitor celecoxib, originally prescribed as an nonsteroidal anti-inflammatory drugs (NSAID).

Celecoxib (trade name: Celebrex) is a non-steroidal anti-inflammatory sulfa drug and a selective COX-2 inhibitor that is mainly used for the treatment of various conditions such as rheumatoid arthritis, acute pain, and colorectal polyps.

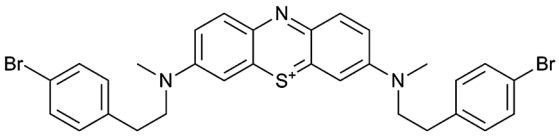
Repositioned drug as a stabilizer of G-quadruplex DNA

DNA-interacting anti-cancer agents are typically designed to target double-helical DNA through covalent cross-linking of DNA strands, thereby preventing subsequent replication and transcription processes. Recently, non-canonical forms of DNA such as the G-quadruplex have emerged as an alternative drug target due to their putative involvement in cell division and transcription as well as their low abundance in cells, which may make them easier to selectively target.[102–104] G-quadruplexes are four-stranded nucleic acid structures formed from co-planar stacked guanine tetrads stabilized by Hoogsteen hydrogen bonding and monovalent cations such as potassium or sodium ions located in the central ion channel.[105–112] In 2011, Ma, Leung and co-workers utilized structure-based virtual screening techniques to identify existing drugs as potential c-*myc* G-quadruplex stabilizers.[113] The c-*myc* oncogene encodes a transcription factor that is believed to regulate a diverse range of genes related to cell proliferation and apoptosis, and the stabilization of the c-*myc* promoter G-quadruplex by small molecules has been reported to repress c-*myc* oncogenic expression and inhibit cancer cell growth.[103]

From the preliminary docking results, methylene blue (MB) was identified as a promising scaffold for further optimizations (Fig. 6). MB is a well-known DNA intercalating dye with numerous pharmacological properties and has been investigated for the treatment of malaria, cancer, and nitrate poisoning. Based on the phenothiazinium template of MB, the authors designed 50 MB derivatives with side chains of various lengths and functionalities using ICM-chemist-pro software. This small focused library was docked to the c-*myc* G-quadruplex and the three high-scoring compounds were synthesized and evaluated for biological activities. Compound **4** (Fig. 6) was found to stabilize the c-*myc* Pu27 G-quadruplex in a dose-dependent fashion in a polymerase chain reaction (PCR)-stop assay, and



Methylene Blue



4

Fig. 6 Chemical structures of methylene blue and its derivative compound **4** as c-*myc* G-quadruplex stabilizers.

could also interfere with c-*myc* promoter activity in cancer cells and inhibit the cancer cell growth with micromolar potency. According to molecular modelling analysis, the increased activity of the MB analogue compared to the parent compound MB was attributed to the ability of the pendant side chains of **4** to interact with the groove region of the c-*myc* G-quadruplex (Fig. 7). This study demonstrated that structure-based virtual screening could also be applied for the repositioning of existing drugs against biomolecular targets other than proteins. In addition, *in silico* techniques were utilized to perform preliminary structure-based optimization in order to improve the potency of the MB derivative against the G-quadruplex motif.

Conclusions

Drug repositioning has emerged as a cost-effective alternative to traditional *de novo* drug discovery and development. In today's economic climate, the ability to accelerate and streamline the drug development process by bypassing most early-stage pharmacokinetic and toxicological barriers represents an attractive strategy that could supplement existing pipelines. Encouragingly, major drug companies are embracing a pilot program launched by the US National Institutes of Health (NIH) National Center for Advancing Translational Sciences (NCATS) which allows academic researchers to investigate the repurposing of safety-tested but unapproved drugs for new therapeutic indications.[9,114,115] Meanwhile, computational methodologies have found increasing use in the early stages of drug discovery and can significantly reduce the time and cost of biological screening by weeding out non-binders *in silico*. The combination of virtual screening and repurposing thus potentially represents a very efficient and cost-effective strategy in drug discovery.

The successful examples in the literature described in this review have shown that structure-based virtual screening techniques can be applied for the cost-effective repositioning of existing drugs against a range of pharmacological targets, including enzymes, protein–protein interactions, and non-canonical

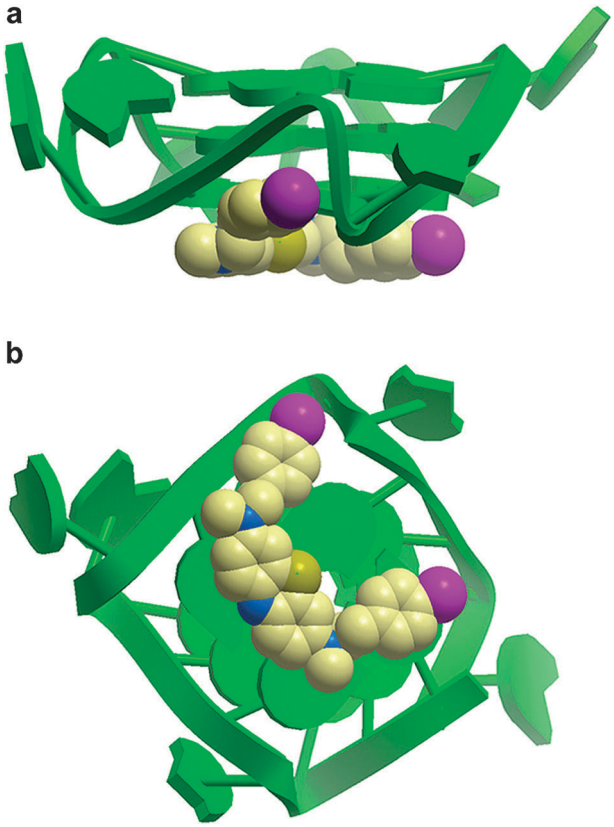


Fig. 7 Hypothetical molecular models showing the (a) side view; (b) top view of the interactions of compound **4** with the c-*myc* G-quadruplex structure. The G-quadruplex is displayed as a ribbon representation (green), while compound **4** is depicted as a space-filling representation showing carbon (beige), nitrogen (blue), sulphur (green-yellow) and bromine (purple atoms). Reproduced from ref. 113 by permission of Elsevier Inc.

nucleic acid structures. Ma, Leung and co-workers successfully identified two FDA-approved drugs as inhibitors of TNF-α trimerization using a direct high-throughput molecular docking approach. Using a fragment-based multiple ligand simultaneous docking approach, Li and co-workers constructed a virtual molecule which was used to identify celecoxib as an antagonist of STAT3 transcription factor dimerization. Moreover, Ma, Leung and co-workers applied high-throughput virtual screening to discover methylene blue as a promising scaffold for c-*myc* G-quadruplex stabilization, and employed *in silico* optimization to generate a more potent stabilizer of this non-canonical DNA motif.

Most of the studies highlighted in this review have centered around the single-target approach rather than on the network pharmacology of repositioned drugs.[116] In our view, structure-based virtual screening deserves further attention for the inverse docking of existing drug candidates in order to harness the polypharmacology of known drugs in a valuable manner. Besides the studies highlighted here, other successful cases in drug repositioning have been reported using alternative computational techniques that are outside the scope of this review.[117,118]

Although no perfect computational software is available, we envisage the continued refinement of docking and scoring algorithms along with the concomitant advancement of computational processing power will continue to empower structure-based molecular docking as an efficient technique in the early stages of drug discovery and drug repositioning.

Acknowledgements

This work is supported by Hong Kong Baptist University (FRG2/11-12/009), Centre for Cancer and Inflammation Research, School of Chinese Medicine (CCIR-SCM, HKBU), the Health and Medical Research Fund (HMRF/11101212), the Research Grants Council (HKBU/201811 and HKBU/204612), the Science and Technology Development Fund, Macao SAR (001/2012/A) and the University of Macau (SRG013-ICMS12-LCH, MYRG091(Y1-L2)-ICMS12-LCH and MYRG121(Y1-L2)-ICMS12-LCH).

Notes and references

1 K. H. Bleicher, H.-J. Bohm, K. Muller and A. I. Alanine, *Nat. Rev. Drug Discovery*, 2003, **2**, 369–378.
2 H. Mario Geysen, F. Schoenen, D. Wagner and R. Wagner, *Nat. Rev. Drug Discovery*, 2003, **2**, 222–230.
3 K. I. Kaitin, *Clin. Pharmacol. Ther.*, 2010, **87**, 356–361.
4 D. B. Kitchen, H. Decornez, J. R. Furr and J. Bajorath, *Nat. Rev. Drug Discovery*, 2004, **3**, 935–949.
5 T. T. Ashburn and K. B. Thor, *Nat. Rev. Drug Discovery*, 2004, **3**, 673–683.
6 G. R. Zimmermann, J. Lehár and C. T. Keith, *Drug Discovery Today*, 2007, **12**, 34–42.
7 J. T. Metz and P. J. Hajduk, *Curr. Opin. Chem. Biol.*, 2010, **14**, 498–504.
8 T. T. Ashburn and K. B. Thor, *Nat. Rev. Drug Discovery*, 2004, **3**, 673–683.
9 A. Mullard, *Nat. Rev. Drug Discovery*, 2012, **11**, 505–506.
10 A. L. Hopkins, *Nature*, 2009, **462**, 167–168.
11 A. L. Hopkins, *Nat. Chem. Biol.*, 2008, **4**, 682–690.
12 E. L. Tobinick, *Drug News Perspect.*, 2009, **22**, 119–125.
13 S. H. Sleigh and C. L. Barton, *Pharm. Med.*, 2010, **24**, 151–159.
14 K. A. O'Connor and B. L. Roth, *Nat. Rev. Drug Discovery*, 2005, **4**, 1005–1014.
15 A. G. Reaume, *Drug Discovery Today: Ther. Strategies*, 2011, **8**, 85–88.
16 H. Zhang, D. Z. Qian, Y. S. Tan, K. Lee, P. Gao, Y. R. Ren, S. Rey, H. Hammers, D. Chang, R. Pili, C. V. Dang, J. O. Liu and G. L. Semenza, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 19579–19586.
17 Y. Zhang, Y. Byun, Y. R. Ren, J. O. Liu, J. Laterra and M. G. Pomper, *Cancer Res.*, 2009, **69**, 5867–5875.
18 C. Andronis, A. Sharma, V. Virvilis, S. Deftereos and A. Persidis, *Briefings Bioinf.*, 2011, **12**, 357–368.
19 F. Moriaud, S. B. Richard, S. A. Adcock, L. Chanas-Martin, J.-S. Surgand, M. Ben Jelloul and F. Delfaud, *Briefings Bioinf.*, 2011, **12**, 336–340.
20 A. Schuffenhauer, J. Zimmermann, R. Stoop, J.-J. van der Vyver, S. Lecchini and E. Jacoby, *J. Chem. Inf. Comput. Sci.*, 2002, **42**, 947–955.

21 M. J. Keiser, V. Setola, J. J. Irwin, C. Laggner, A. I. Abbas, S. J. Hufeisen, N. H. Jensen, M. B. Kuijer, R. C. Matos, T. B. Tran, R. Whaley, R. A. Glennon, J. Hert, K. L. H. Thomas, D. D. Edwards, B. K. Shoichet and B. L. Roth, *Nature*, 2009, **462**, 175–181.
22 S. O. Ochiana, V. Pandarinath, Z. Wang, R. Kapoor, M. J. Ondrechen, L. Ruben and M. P. Pollastri, *Eur. J. Med. Chem.*, 2012, DOI: 10.1016/j.ejmech.2012.1007.1038.
23 S. C. Fagan, *Pharmacotherapy*, 2010, **30**, 51S–54S.
24 A. Nzila, Z. Ma and K. Chibale, *Future Med. Chem.*, 2011, **3**, 1413–1426.
25 J. Blatt and S. J. Corey, *Drug Discovery Today*, 2012, DOI: 10.1016/j.drudis.2012.1007.1009.
26 M. A. Sukhai, P. A. Spagnuolo, S. Weir, J. Kasper, L. Patton and A. D. Schimmer, *Blood*, 2011, **117**, 6747–6755.
27 C. R. Caffrey and W. E. Secor, *Curr. Opin. Infect. Dis.*, 2011, **24**, 410–417.
28 T. I. Oprea, J. E. Bauman, C. G. Bologa, T. Buranda, A. Chigaev, B. S. Edwards, J. W. Jarvik, H. D. Gresham, M. K. Haynes, B. Hjelle, R. Hromas, L. Hudson, D. A. Mackenzie, C. Y. Muller, J. C. Reed, P. C. Simons, Y. Smagley, J. Strouse, Z. Surviladze, T. Thompson, O. Ursu, A. Waller, A. Wandinger-Ness, S. S. Winter, Y. Wu, S. M. Young, R. S. Larson, C. Willman and L. A. Sklar, *Drug Discovery Today: Ther. Strategies*, 2011, **8**, 61–69.
29 S. J. Weir, L. J. DeGennaro and C. P. Austin, *Cancer Res.*, 2012, **72**, 1055–1058.
30 J. Aubé, *ACS Med. Chem. Lett.*, 2012, **3**, 442–444.
31 Y. A. Lussier and J. L. Chen, *Sci. Transl. Med.*, 2011, **3**, 96ps35.
32 D. Cavalla and C. Singal, *Drug Discovery Today*, 2012, **17**, 104–109.
33 S. H. Sleigh and C. L. Barton, *Pharm. Med.*, 2010, **24**, 151–159.
34 T. X. Witkowski, *Drug Discovery Today: Ther. Strategies*, 2011, **8**, 139–143.
35 C. N. Cavasotto and A. J. Orry, *Curr. Top. Med. Chem.*, 2007, **7**, 1006–1014.
36 P. Ripphausen, B. Nisius, L. Peltason and J. r. Bajorath, *J. Med. Chem.*, 2010, **53**, 8461–8467.
37 S. Kalyaanamoorthy and Y.-P. P. Chen, *Drug Discovery Today*, 2011, **16**, 831–839.
38 P. D. Lyne, *Drug Discovery Today*, 2002, **7**, 1047–1055.
39 S. L. Kinnings, N. Liu, N. Buchmeier, P. J. Tonge, L. Xie and P. E. Bourne, *PLoS Comput. Biol.*, 2009, **5**, e1000423.
40 M. P. Gleeson, A. Hersey and S. Hannonqbua, *Curr. Top. Med. Chem.*, 2010, **11**, 358–381.
41 S. Modi, M. Hughes, A. Garrow and A. White, *Drug Discovery Today*, 2012, **17**, 135–142.
42 G. Moroy, V. Y. Martiny, P. Vayer, B. O. Villoutreix and M. A. Miteva, *Drug Discovery Today*, 2012, **17**, 44–55.
43 D. S. Wishart, C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang and J. Woolsey, *Nucleic Acids Res.*, 2006, **34**, D668–D672.
44 E. Pihan, L. Colliandre, J.-F. Guichou and D. Douguet, *Bioinformatics*, 2012, **28**, 1540–1541.
45 M. Kanehisa, S. Goto, M. Furumichi, M. Tanabe and M. Hirakawa, *Nucleic Acids Res.*, 2010, **38**, D355–D360.
46 A. Goede, M. Dunkel, N. Mester, C. Frommel and R. Preissner, *Bioinformatics*, 2005, **21**, 1751–1753.
47 B. Kramer, M. Rarey and T. Lengauer, *Proteins*, 1999, **37**, 228–241.
48 S. K. Kearsley, D. J. Underwood, R. P. Sheridan and M. D. Miller, *J. Comput. Aided Mol. Des.*, 1994, **8**, 565–582.
49 T. N. Hart and R. J. Read, *Proteins*, 1992, **13**, 206–222.
50 G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, *J. Comput. Chem.*, 1998, **19**, 1639–1662.
51 H. Gohlke and G. Klebe, *Angew. Chem., Int. Ed.*, 2002, **41**, 2644–2676.
52 H. A. Carlson, *Curr. Opin. Chem. Biol.*, 2002, **6**, 447–452.
53 C. R. Corbeil and N. Moitessier, *J. Chem. Inf. Model.*, 2009, **49**, 997–1009.
54 S. Raub, A. Steffen, A. Kämper and C. M. Marian, *J. Chem. Inf. Model.*, 2008, **48**, 1492–1510.
55 R. D. Clark, A. Strizhev, J. M. Leonard, J. F. Blake and J. B. Matthew, *J. Mol. Graphics Modell*, 2002, **20**, 281–295.
56 R. Teramoto and H. Fukunishi, *J. Chem. Inf. Model.*, 2008, **48**, 288–295.
57 J. D. Durrant and J. A. McCammon, *J. Chem. Inf. Model.*, 2010, **50**, 1865–1871.
58 P. J. Ballester and J. B. O. Mitchell, *Bioinformatics*, 2010, **26**, 1169–1175.
59 S. L. Kinnings, N. Liu, P. J. Tonge, R. M. Jackson, L. Xie and P. E. Bourne, *J. Chem. Inf. Model.*, 2011, **51**, 408–419.
60 D. E. Koshland, *Angew. Chem., Int. Ed.*, 1995, **33**, 2375–2378.
61 R. Najmanovich, J. Kuttner, V. Sobolev and M. Edelman, *Proteins*, 2000, **39**, 261–268.
62 M. Totrov and R. Abagyan, *Curr. Opin. Struct. Biol.*, 2008, **18**, 178–184.
63 O. Sperandio, L. Mouawad, E. Pinto, B. Villoutreix, D. Perahia and M. Miteva, *Eur. Biophys. J.*, 2010, **39**, 1365–1372.
64 M. Totrov and R. Abagyan, *Curr. Opin. Chem. Biol.*, 2008, **18**, 178–184.
65 M. Rueda, G. Bottegoni and R. Abagyan, *J. Chem. Inf. Model.*, 2009, **50**, 186–193.
66 M. Okamoto, K. Takayama, T. Shimizu, K. Ishida, O. Takahashi and T. Furuya, *J. Med. Chem.*, 2009, **52**, 7323–7327.
67 G. Bottegoni, I. Kufareva, M. Totrov and R. Abagyan, *J. Med. Chem.*, 2008, **52**, 397–406.
68 M. Rueda and R. Abagyan, in *Protein-Ligand Interactions*, ed. H. Gohlke, Wiley-VCH Verlag GmbH & Co. KGaA, 2012, pp. 283–294.
69 M. Rueda, G. Bottegoni and R. Abagyan, *J. Chem. Inf. Model.*, 2009, **50**, 186–193.
70 G. Bottegoni, I. Kufareva, M. Totrov and R. Abagyan, *J. Comput. Aided Mol. Des.*, 2008, **22**, 311–325.
71 D. B. Kokh and W. Wenzel, *J. Med. Chem.*, 2008, **51**, 5919–5931.
72 A. M. Ferrari, B. Q. Wei, L. Costantino and B. K. Shoichet, *J. Med. Chem.*, 2004, **47**, 5076–5084.

73 C. B-Rao, J. Subramanian and S. D. Sharma, *Drug Discovery Today*, 2009, **14**, 394–400.
74 O. Guvench and A. D. MacKerell Jr, *Curr. Opin. Struct. Biol.*, 2009, **19**, 56–61.
75 F. Spyrakis, A. BidonChanal, X. Barril and F. J. Luque, *Curr. Top. Med. Chem.*, 2011, **11**, 192–210.
76 G. Marcou and D. Rognan, *J. Chem. Inf. Model.*, 2006, **47**, 195–207.
77 G. Bouvier, N. Evrard-Todeschi, J.-P. Girault and G. Bertho, *Bioinformatics*, 2010, **26**, 53–60.
78 D. Wei, H. Zheng, N. Su, M. Deng and L. Lai, *J. Chem. Inf. Model.*, 2010, **50**, 1855–1864.
79 S. Ekins, A. J. Williams, M. D. Krasowski and J. S. Freundlich, *Drug Discovery Today*, 2011, **16**, 298–310.
80 W. K. Kroeze, D. J. Sheffler and B. L. Roth, *J. Cell Sci.*, 2003, **116**, 4867–4869.
81 M. C. Lagerstrom and H. B. Schioth, *Nat. Rev. Drug Discovery*, 2008, **7**, 339–357.
82 D. M. Rosenbaum, S. G. F. Rasmussen and B. K. Kobilka, *Nature*, 2009, **459**, 356–363.
83 A. B. Pithadia and S. N. Jain, *J. Clin. Med. Res.*, 2009, **1**, 72–80.
84 X. Lin, X.-P. Huang, G. Chen, R. Whaley, S. Peng, Y. Wang, G. Zhang, S. X. Wang, S. Wang, B. L. Roth and N. Huang, *J. Med. Chem.*, 2012, **55**, 5749–5759.
85 D. E. Nichols and C. D. Nichols, *Chem. Rev.*, 2008, **108**, 1614–1641.
86 C. R. W. Guimarães and M. Cardozo, *J. Chem. Inf. Model.*, 2008, **48**, 958–970.
87 Y. Shi, *Drug Discovery Today*, 2007, **12**, 440–445.
88 W. H. Bisson, A. V. Cheltsov, N. Bruey-Sedano, B. Lin, J. Chen, N. Goldberger, L. T. May, A. Christopoulos, J. T. Dalton, P. M. Sexton, X.-K. Zhang and R. Abagyan, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 11927–11932.
89 S. Fletcher and A. D. Hamilton, *Curr. Opin. Chem. Biol.*, 2005, **9**, 632–638.
90 D. Kozakov, D. R. Hall, G.-Y. Chuang, R. Cencic, R. Brenke, L. E. Grove, D. Beglov, J. Pelletier, A. Whitty and S. Vajda, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 13528–13533.
91 I. S. Moreira, P. A. Fernandes and M. J. Ramos, *Proteins*, 2007, **68**, 803–812.
92 J. A. Wells and C. L. McClendon, *Nature*, 2007, **450**, 1001–1009.
93 T. Berg, in *Protein-protein complexes Analysis, Modeling and Drug Design*, ed. M. Zacharias, Imperial College Press, 2010, pp. 318–339.
94 L. K. Tsou, Y. Cheng and Y.-C. Cheng, *Curr. Opin. Pharmacol.*, 2012, **12**, 403–407.
95 J. R. Bradley, *J. Pathol.*, 2008, **214**, 149–160.
96 C.-H. Leung, H.-J. Zhong, H. Yang, Z. Cheng, D. S.-H. Chan, V. P.-Y. Ma, R. Abagyan, C.-Y. Wong and D.-L. Ma, *Angew. Chem., Int. Ed.*, 2012, **51**, 9010–9014.
97 M. M. He, A. S. Smith, J. D. Oslob, W. M. Flanagan, A. C. Braisted, A. Whitty, M. T. Cancilla, J. Wang, A. A. Lugovskoy, J. C. Yoburn, A. D. Fung, G. Farrington, J. K. Eldredge, E. S. Day, L. A. Cruz, T. G. Cachero, S. K. Miller, J. E. Friedman, I. C. Choong and B. C. Cunningham, *Science*, 2005, **310**, 1022–1025.
98 D. S.-H. Chan, H.-M. Lee, F. Yang, C.-M. Che, C. C. L. Wong, R. Abagyan, C.-H. Leung and D.-L. Ma, *Angew. Chem., Int. Ed.*, 2010, **49**, 2860–2864.
99 C.-H. Leung, D. S.-H. Chan, M. H.-T. Kwan, Z. Cheng, C.-Y. Wong, G.-Y. Zhu, W.-F. Fong and D.-L. Ma, *ChemMedChem*, 2011, **6**, 765–768.
100 Z.-l. Yuan, Y.-j. Guan, L. Wang, W. Wei, A. B. Kane and Y. E. Chin, *Mol. Cell. Biochem.*, 2004, **24**, 9390–9400.
101 H. Li, A. Liu, Z. Zhao, Y. Xu, J. Lin, D. Jou and C. Li, *J. Med. Chem.*, 2011, **54**, 5592–5596.
102 J.-L. Mergny and C. Helene, *Nat. Med.*, 1998, **4**, 1366–1367.
103 S. Balasubramanian, L. H. Hurley and S. Neidle, *Nat. Rev. Drug Discovery*, 2011, **10**, 261–275.
104 A. De Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J.-F. Riou and J.-L. Mergny, *Biochimie*, 2008, **90**, 131–155.
105 G. N. Parkinson, M. P. H. Lee and S. Neidle, *Nature*, 2002, **417**, 876–880.
106 A. T. Phan and J. L. Mergny, *Nucleic Acids Res.*, 2002, **30**, 4618–4625.
107 O. Stegle, L. Payet, J.-L. Mergny, D. J. C. MacKay and J. L. Huppert, *Bioinformatics*, 2009, **25**, i374–i382.
108 A. Risitano and K. R. Fox, *Biochemistry*, 2003, **42**, 6507–6513.
109 H.-M. Lee, D. S.-H. Chan, F. Yang, H.-Y. Lam, S.-C. Yan, C.-M. Che, D.-L. Ma and C.-H. Leung, *Chem. Commun.*, 2010, **46**, 4680–4682.
110 D.-L. Ma, V. P.-Y. Ma, D. S.-H. Chan, K.-H. Leung, H.-J. Zhong and C.-H. Leung, *Methods*, 2012, **57**, 106–114.
111 D.-L. Ma, D. S.-H. Chan, W.-C. Fu, H.-Z. He, H. Yang, S.-C. Yan and C.-H. Leung, *PLoS One*, 2012, **7**, e43278.
112 H. Yang, H.-J. Zhong, K.-H. Leung, D. S.-H. Chan, V. P.-Y. Ma, W.-C. Fu, R. Nanjunda, W. D. Wilson, D.-L. Ma and C.-H. Leung, *Eur. J. Pharm. Sci.*, 2013, **48**, 130–141.
113 D. S.-H. Chan, H. Yang, M. H.-T. Kwan, Z. Cheng, P. Lee, L.-P. Bai, Z.-H. Jiang, C.-Y. Wong, W.-F. Fong, C.-H. Leung and D.-L. Ma, *Biochimie*, 2011, **93**, 1055–1064.
114 M. Allison, *Nat. Biotechnol.*, 2012, **30**, 571–572.
115 F. S. Collins, *Nat. Rev. Drug Discovery*, 2011, **10**, 397.
116 M. A. Yildirim, K.-I. Goh, M. E. Cusick, A.-L. Barabasi and M. Vidal, *Nat. Biotechnol.*, 2007, **25**, 1119–1126.
117 S. L. Kinnings, N. Liu, N. Buchmeier, P. J. Tonge, L. Xie and P. E. Bourne, *PLoS Comput. Biol.*, 2009, **5**, e1000423.
118 S. R. Vasudevan, J. B. Moore, Y. Schymura and G. C. Churchill, *J. Med. Chem.*, 2012, **55**, 7054–7060.