Article:
Lolli, M, Narramore, S, Fishwick, CWG et al. (1 more author) (2015) Refining the chemical toolbox to be fit for educational and practical purpose for drug discovery in the 21st Century. Drug Discovery Today, 20 (8). pp. 1018-1026. ISSN 1359-6446

https://doi.org/10.1016/j.drudis.2015.04.010

Short Title: Refining the chemical toolbox for drug discovery in the 21st Century

Authors:

Dr Marco Lolli
Assistant Professor of Medicinal Chemistry
Department of Science and Drug Technology
University of Turin
Via Pietro Giuria,
9 - 10125 Turin
ITALY
Phone: +39-0116 707993

Ms Sarah Narramore
PhD student
School of Chemistry
University of Leeds
Leeds LS2 9JT, UK
Phone: +44-(0)113 3436510

Professor Colin W.G. Fishwick
Head of Organic Chemistry
School of Chemistry
University of Leeds
Leeds LS2 9JT, UK
Phone: 44-(0)113 3436510

Dr Klaus Pors
Senior Lecturer in Chemical Biology
Institute of Cancer Therapeutics
University of Bradford
Bradford, BD7 1DP, U.K.
Phone: 44-(0)127 4236482

Corresponding Author: Pors, K. (k.pors1@bradford.ac.uk)

Keywords: Drug discovery, chemical space, diversity-oriented synthesis (DOS), bioisosterism, scaffold hopping, medicinal chemistry.

Teaser: Modernisation of theoretical and practical chemistry taught in academia is required for (i) expanding the chemical toolbox and (ii) educating the next generation of drug-hunters.

Refining the chemical toolbox to be fit for educational and practical purpose for drug discovery in the 21st Century

Marco Lolli, Sarah Narramore, Colin W.G. Fishwick and Klaus Pors

Abstract

We live in a time where exploration and generation of new knowledge is occurring on a colossal scale. Medicinal chemists have traditionally played a key role in drug discovery, however the many unmet medical demands in the healthcare sector emphasises the need to evolve the medicinal chemistry discipline. To rise to the challenges in the 21st Century there is a necessity to refine the chemical toolbox for both educational and practical reasons. This article provides suggestions of modern strategies that are beneficial to teaching in academia, but are also important reminders of strategies that potentially can lead to better drugs.

Drug discovery is a complex, time-consuming and a very costly process. In recent years the pharmaceutical industry has been facing increases in the cost of pharmaceutical R&D, driven by the high failure rate of projects and development candidates. Between 2005 and 2010 nine of the largest pharmaceutical companies in the world achieved an average output less than 1 new molecular entity (NME) per company per year despite their combined research expenditure climbing to $60 billion dollars during this period [1]. Failure in phase 2 clinical trials has been highlighted as a key industry challenge, as only 25% of compounds that currently enter phase 2 successfully progress to full phase 3 clinical studies. It is well-known that a large number of drug candidates fail during the development process due to problems with absorption, distribution, metabolism, excretion or toxicity (ADMET) [2]. For this reason, the evaluation of compounds for their ADMET profile is being carried out at increasingly early stages of the development process. At the same time, strategies have been developed to remove compound structures with undesirable properties when choosing compound sets [3], in order to limit the use of a significant amount of resources to develop candidates which are ultimately unsuccessful [4]. Medicinal chemists play a key role in the drug discovery process [5], however the many unmet medical demands in the healthcare sector emphasise the need to evolve the medicinal chemistry discipline. To rise to the challenges in the 21st Century [6], there is a necessity to refine the chemical toolbox for both educational and practical reasons.

Education and training of undergraduate students is, at large, still rooted in the "classical" way of carrying out medicinal chemistry with emphasis on structure-activity relationship (SAR), physicochemical properties, hit-lead optimisation, solubility and formulation. Paradigms such as 'Rule of 5', Ligand Efficiency (LipE),

‘leadlike’ property considerations and perhaps newer additions such as rotatable bonds in the ligand and polar surface area are taught as fundamental concepts of good medicinal chemistry. However, the question is whether the medicinal chemistry taught in academia globally is truly fit for purpose to solve the grand challenges of treating diseases such as cancer, Parkinson’s, and Alzheimer’s. The successful completion of the Human Genome Project has opened up exciting new opportunities not only for the medicinal chemist but for other researchers who sees themselves as drug-hunters. If we are going to successfully tap into the wealth of this limitless ‘omic’ data it will require iterative and synergistic interactions between scientists working at the chemistry/biology interface and the broader research community for years to come. Decoding the information stored in the functional genome, including thousands of predicted gene products, will only be possible with new scientific tools and methods that are fit for purpose [6,7]. Nonetheless, due to the increasingly poor productivity of the pharmaceutical industry as measured by the number of FDA-approved NCEs, there is a need to constantly re-evaluate the drug discovery models we employ.

The purpose of this paper is to contribute to the ongoing discussion about modernisation of medicinal chemistry [5,8-15]. As discussed in the following sections, modern approaches including computational chemistry, bioisosterism, scaffold hopping and diversity-oriented synthesis to search for valuable hit compounds in new chemical space are important tools for the medicinal chemists of the future, which can be incorporated into undergraduate and postgraduate teaching with relative ease.

Computer aided hit identification

In the early stages of a drug discovery project computational assistance can be particularly useful in the identification of hits which can form the basis of a medicinal chemistry program. Any target that has structural data available, or even just known ligands, can be assessed computationally to identify potential hit compounds either using de novo design in which novel molecules are designed from scratch or by using virtual screening of databases (VS) [16,17]. Design approaches that seek to identify hits based on their ability to complement features of a protein structure are called structure-based and approaches that look for hits based on similarity to known inhibitors or pharmacophore models are called ligand-based (for a simplistic overview, see Figure 1) [18].

Insert Figure 1

Virtual screening

Back in the nineties, virtual screening (VS), which is a means of efficiently docking sets of molecular structures to a target protein, emerged as a valuable tool able to guide and focus experimental efforts on smaller, filtered sets of compounds with increased probability of showing the desired biological activity. Chemical space is a term widely used in drug discovery to refer to the vast number of theoretically possible molecules. Sets of molecules can be mapped to illustrate the distribution of their properties [19]. VS allows the interrogation of large portions of chemical space without the need to physically possess the molecules in question. In the last 3 years, over 2000 papers have been published which mention the use of VS in applications as diverse as antimalarials [20], cognitive impairment [21] and cancer [22]. VS is

becoming increasingly popular due to its ease of use relative to the cost, time and expertise needed to carry out HTS. There are many software packages available for carrying out small molecule docking, several of which are free for academic use, and as a result, the literature is full of papers describing the results of VS campaigns against a wide variety of targets.

Impressive examples of VS include a strategy by Halland et. al. who, very recently, carried out ligand based VS using ROCS (a shape comparison tool produced by OpenEye [23]) to great effect in the search for SGK1 inhibitors [24]. SGK1 has been linked to several diseases including cancer, diabetes and neurological degeneration, so selective and potent inhibitors could have a wide range of applications. A library of 2 million commercially available structures was screened to find structures similar to known inhibitors of the protein and 78 compounds were selected for biological evaluation of which 7 were active including compound **1**, Figure 2. Library synthesis was then employed to explore SAR around the two most promising hits and a series of highly potent and selective compounds with attractive physicochemical, ADME and pharmacokinetic (PK) properties was developed.

Recently, Li et. al. have reported the use of structure-based VS to identify inhibitors which target the human androgen receptor, which is known to regulate the progression of prostate cancer. They identified a binding site in the DNA binding region of the receptor which had no previously known ligands, but was identified by the group as a promising target for inhibition. By performing VS using Glide [25] they were able to identify some compounds with activities in the low micro molar range, and further shape similarity searching within the ZINC database identified even more active compounds which were used as the starting points for a full medicinal

chemistry campaign resulting in compound **2** (Figure 2) which exhibited submicromolar activity against the enzyme [26].

De novo design

De novo design can be used to generate entirely novel starting points for synthesis and has the potential to design compounds with very high ligand efficiency. De novo design is attractive because it provides the opportunity to sample a much wider area of chemical space than that offered by the contents of a commercial compound library. However this approach does require a greater investment of chemical resource early in a drug discovery project, as hits generated by the software must be synthesised to establish their activity rather than simply being purchased.

Recently Yule et al. reported a series of potent, antibiotic inhibitors of the ATP binding site of DNA gyrase [27]. The fragment connection approach used within SPROUT [28] produced an initial hit molecule, which was synthesised and shown to be active against the enzyme. This hit was then developed into a series of promising inhibitors including compound **3** (Figure 2) that are now being further explored for drug-like potential.

A different approach to structure-based de novo design has recently been used by the Baell group at the University of Melbourne in the design of a series of peptidomimetic inhibitors of BCL-X_L using the program SYBYL [29]. BCL-X_L is part of the BCL-2 family of proteins which are involved in the apoptotic cascade and thus implicated in tumorigenesis. An unsubstituted benzoyl urea scaffold was identified and then decorated with groups, which were oriented in such a way as to make desired interactions in the enzyme active sites. A thorough SAR investigation

produced several inhibitors with low micromolar IC_{50}s including compound **4** (Figure 2).

Ligand-based de novo design was recently employed by Rodrigues et. al. to generate promising anti-angiogenesis compounds, which inhibit the kinase VGFR-2 [30]. They used the program DOGS (design of genuine structures) to identify analogues of a known VGFR-2 inhibitor and synthetic routes to the designed compounds, including compound **5** (Figure 2), which proved to be effective inhibitors of cell proliferation.

Insert Figure 2

Bioisosterism: new ways to access diverse chemical space.

When compound libraries for use in VS and HTS were first being constructed, the focus was mostly on synthesising as many compounds as possible. However, it was soon realised that chemical diversity was a more important aim and recently quality has become the guiding principle [5]. The meaning of '*quality'* in this case involves a balance of multiple criteria including structural diversity, bioavailability, permeability, drug-likeness and target focus, among others [31]. In 1997, Lipinski's "Rule of Five" began a trend of focussing drug discovery efforts on confined areas of chemical space likely to contain compounds with pharmaceutically relevant properties. Since then, many related modified or related rules have been proposed [32].

In order to enhance molecular complexity organic synthesis was called upon to

support VS and HTS to accomplish the role which natural chemodiversity has played traditionally [33] in covering large and novel areas of chemical space. In particular, assistance has come from a traditional, powerful tool used by the medicinal chemist: bioisosteric replacement [34]. This approach is often used not only to improve ADME and physicochemical properties (water solubility, logP, pKa, etc.), but also to increase potency, enhance selectivity and decrease toxicity, or to gain a novel intellectual property (IP) position.

The concept of bioisosterism is based on the notion that single atoms, groups or whole molecules that exhibit similar volume, shape, and/or physicochemical properties can produce broadly similar biological effects. Two functional groups should be called isosteres if they share similar physicochemical properties. When isosteres are able to share also the same biological profile they are called bioisosteres. Bioisosteres cannot behave in exactly the same manner as the functional group they are intended to replace and, as a consequence, some bioisosteres might be successful at some targets but not at others. It could be said that while chemistry defines the isosteric similarity between groups, only the biological target can determine the similarity of bioisosteres.

Traditionally, this powerful strategy has been applied to single functional groups. Employed since the 70´s by Wermut [35], the proof of the importance of this technology in the modern drug discovery setting is substantiated by the fact that a comprehensive review published by Meanwell in the Journal of Medicinal Chemistry (JMC) in 2011 [36] became the most downloaded JMC article in following years. Quite recently, the replacement of carboxylic acids with suitable bioisosteric surrogates has been reviewed [37]. The carboxylic acid has been much investigated in terms of bioisosteric replacements; this interest can be explained from its frequent

presence in pharmacophores resulting in the large number (>450) of carboxylic acid-containing drugs marketed worldwide. Bioisosteres of the α-amino acidic moiety have also been reviewed recently [38]. In this case, bioisosterism of the whole α-amino acid moiety is accomplished with heterocyclic bioisosteres that often display an acidic function.

A new strategy referred to as scaffold hopping has matured as a subset of bioisosteric replacement, for discovering structurally novel compounds to explore chemical space [39] . This process, also known as lead hopping, typically starts with known active compounds and modifies the central core structure of the molecule to give a novel chemotype [40]. The core may be of direct functional importance in interacting with the protein target or it may provide the necessary scaffolding to allow substitution with functional groups in the appropriate geometric configuration to interact with the target. In other words, scaffold hopping is a useful strategy in drug design to 'jump' into different areas of chemical space, for example to identify easily synthesisable small molecules that mimic complex natural products. Thus, scaffold hopping represents an attractive strategy to address novelty within confined chemical space, and by consequence achieve favourable intellectual property positions [41]. A classification of scaffold hopping into four major categories, namely heterocycle replacements, ring opening or closure, peptidomimetics, and topology-based hopping, has been recently proposed [40]. Some examples of recently published applications of this technique to oncology related targets are presented below.

Ge et al. recently [42] demonstrated scaffold hopping could be applied to the identification of potential anti-tumour agents based on known inhibitors. The

approach started by collecting sixteen anti-cancer compounds previously isolated from natural sources (Clausena vestita). The sixteen hits, besides sharing the same carbazole alkaloid scaffolds, were found to inhibit cell growth with a pIC_{50} 3.80-5.37 range. Starting from these structures, the application of a scaffold hopping approach enabled the identification of eight hits, including compound **6** (Figure 2), showing similar anti-tumour activities at the cellular level to the starting compounds but characterised by completely different scaffolds. These represent a valuable starting point for the generation of new lead candidates.

Human dihydroorotate dehydrogenase (hDHODH) is a class-2 dihydroorotate dehydrogenase extensively used by proliferating cells. Its inhibition in autoimmune and inflammatory diseases, cancers, and multiple sclerosis is therefore of substantial clinical importance. The strategies employed in the development of DHODH inhibitors have been recently reviewed [43]. Scaffold hopping approaches have recently been applied to generate new lead candidates from known DHODH inhibitors. In the first example [44] a computational method able to design novel hDHODH inhibitors was developed. After a validation step, 13 novel hDHODH inhibitors were designed using a scaffold hopping strategy (e.g. compound **7**, Figure 2). In the second example [45], structural analogies present in two well-known DHODH inhibitors (Leflunomide and Brequinar) were identified and resulted in the design of a new class of inhibitors based on the 4-hydroxy-1,2,5-oxadiazol-3-yl scaffold (e.g. compound **8**, Figure 2). In Figure 3, the binding mode of compound **8** is compared with the lead compound Brequinar.

Insert Figure 3

Fragment-based drug design

Fragment-based drug design (FBDD) has over the past decade become routinely used to identify lead molecules which bind to troublesome biochemical targets for which standard techniques, such as high-throughput screening (HTS), had failed. [46]. FBDD begins with identification of small chemical fragments which bind weakly to a target, these fragments are then extended or connected to give potent lead molecules. The main benefit of this approach is its ability to explore chemical space very efficiently, with even modestly sized fragment libraries comprising molecules with less than 12 heavy atoms [47]. The successful application of FBDD has depended on a number of recent advances including better biochemical and biophysical detection techniques which are sensitive enough to identify the weak interactions seen in fragment binding, and higher throughput X-ray crystallography [48]. Indeed a critical stage in many FBDD projects is the detection of bound fragments in an appropriate ligand-binding site in crystal structures. Different from the traditional X-ray crystallography which is initially focused on acquiring the ligand-protein complex, the primary screening in FDBB employs a cocktail of ligands for each soaking experiment in order to minimise the experimental effort of screening an entire library. Effective use of FBDD also takes into account both detection method and fragment chemodiversity [49] while fragment solubility is a key design criterion. For recent examples of successful FBDD applications, see e.g. a recent review focussed on HIV [50].

Diversity-oriented synthesis

The strength of combinatorial chemistry (combichem) is based on the creation of large populations of molecules, or libraries that can be screened efficiently en masse

in a short period of time. Whilst the vast amounts of money spent on development of combichem libraries has not yet resulted in many drug successes, combichem has spun out many exciting technologies that now occupy a central place in the biotech industry. The mapping of the human genome may have provided a new area of application for combichem in combination with other HTS methodologies including techniques and instruments developed for DNA microarrays. Indeed, high-density chemical microarrays can now be synthesised in situ on glass slides or be printed through covalent linkage or non-specific adsorption to the surface of the solid-support with fully automatic arrays. In conjunction with the “one-bead one-compound” combinatorial library method, chemical microarrays have proven to be very valuable in hit identification and lead optimisation [6,51]. One limitation of combichem is that the small molecules developed using this technique do not cover a broad area of chemical space and in recognition of this, diversity-oriented synthesis (DOS) as a relatively new technique for exploring chemical space, is gaining momentum. One of the aims of DOS is to populate undeveloped chemical space by using dependable reactions to generate novel chemical scaffolds. Such scaffolds are designed to be structurally diverse in order to navigate new chemical space, enhancing the possibility of uncovering new biologically active molecules. Multicomponent reaction (MCR) chemistry has noted as a useful method for generating staring points for DOS with a high degree of molecular diversity. Interestingly, as with other chemical disciplines mentioned in this review, MCR strategies often utilise routine chemical transformations comprising three or more reactants to form a product that contains all the starting components. Although MCR has existed for more than a century, the recent excitement in this field has roots in discoveries related to generation of new reactions, asymmetric catalysts and natural

products and other targets of biological interest [52]. Important in the context of this review is that MCR, either as a part of DOS strategies or as an independent tool, can provide great opportunities for drug hunters and researchers interested in small molecules with biological activity [53]. This includes the development of highly specific tools for chemical genetics, but also to provide a new chemical pool for drug discovery [54-58]. Chemical genetics is increasingly providing scientists with a new set of complementary chemical tools to investigate biological systems [59]. Given the emerging limitations of using siRNA to completely suppress target expression as well as off-target effects, there have been calls for highly specific small molecules to be employed as chemical tools to interrogate new biological targets [60]. The constant evolution of chemical techniques and their application to biological systems will inevitably speed up the process of discovering NCEs with improved therapeutic efficacy.

Computational prediction of ADMET

Computational tools for the prediction of ADMET are becoming increasingly useful and accurate, and are now used extensively by medicinal and computational chemists [61]. However, currently available models are still some way away from being able to perfectly predict the complex and interconnected ADMET properties of a novel chemical entity. Compounds often fail in late development due to unforeseen problems with metabolism and toxicity (less so absorption, distribution and excretion as good methods for determining these properties are already well established). The earlier these problems are detected, the less money will be spent developing ultimately unsuccessful compounds and the greater the chance of designing a modified version which avoids the liabilities of its parent compound. The

pharmaceutical industry is therefore keen to implement newer and better methods, both in vitro and in silico, for predicting ADMET liabilities in compounds before significant resources are invested in them. [62]. A variety of approaches have been used to generate models for the prediction of ADMET. Three of these approaches: quantitative SAR (QSAR), human knowledge based and structure-based models are described below.

QSAR models were first developed in the 1960s as a way to correlate chemical descriptors with the experimental characteristics of a compound, so as to make predictions about the properties of new molecules. QSAR models have been developed for many applications including prediction of drug binding to biological targets as well as prediction of the various aspects of ADMET. However, these models are limited by the training sets used to generate them and so the accuracy of the predictions they give depends upon the similarity of the molecules being investigated to the molecules used in the training set. There is also a tendency for more complex models of this type to be “black box” models in that the reasoning behind a given prediction is hard to interpret [61].

Human-knowledge based models make predictions based on a thorough understanding of metabolic processes and mechanisms of toxicity. These tools have an advantage over QSAR in that they provide a rationale for their predictions which can be acted on by researchers. However, they are limited by the availability of empirical knowledge about the metabolism and toxicity of given chemical structures [62]. Available software packages include Lhasa’s Derek Nexus [63,64] and EPA’s OncoLogic [65].

More recently, docking based protocols have been investigated for the prediction of

metabolism and toxicity by docking the drug molecules in question into crystal structures of the active sites of metabolism and toxicity related targets[66]. These targets include enzymes from the cytochrome P450 (CYP) family which are involved in phase 1 metabolism; nuclear receptors, such as the androgen receptor, which are involved in important signalling processes; and ion channels including the human ether-a-go-go (hERG) channel which is linked to irregularities in heart rhythm. However, predictions of this kind are made more challenging by the tendency for proteins key to ADMET to be highly promiscuous with flexible or even multiple binding sites [67].

There is a great need for more accurate and generally applicable computational tools for both medicinal chemistry, and the wider fields of toxicology and environmental protection and the rapid rate of progress in the field seems determined to meet those needs[67].

Merging innovation and teaching

So how do we make ends meet? The speed with which we have generated new knowledge, especially at the molecular biology level, in the last part of the 20th Century and in the beginning of the 21st Century has been breath-taking and yet the methodologies we employ in medicinal chemistry remain, in comparison, somewhat pedestrian. Despite over half a century of medicinal chemistry practice that has matured a step at a time into a reliable discipline, there is no doubt that medicinal chemists need to take to their heart innovations in VS, DOS, chem- and bioinformatics, chemical genetics and chemical biology. This is especially the case in light of estimates that indicate that known chemical space, including public databases and corporate collections contains on the order of 100 million molecules

and there may be as many as 10^{60} compounds which would comply with Lipinski's rule of 5 [68]. A more modest estimate of 10^{20}–10^{24} molecules can be reached if a combination of known fragments are considered [69], but this number is still staggeringly large and in fact it has been suggested that almost all small molecules (>99.9%) have never been synthesised and are still available to be prepared and tested [19]. However, just as much as astronomical "space" is a void, much of chemical space contains nothing of biological interest [70]. Therefore, the drug-hunter needs to be innovative in order to overcome current challenges and boundaries to deliver smarter small molecules that can modernise our healthcare. However, the success in reducing attrition rates and sustaining productivity in an increasingly challenging environment with limited funding in both academia and industry is a challenge of immense magnitude [7]. Accordingly, medicinal chemists and other drug-hunters must have a willingness to constantly refine their knowledge and skill-sets as well as being willing to explore new strategies and paradigms.

Ironically, modernising medicinal chemistry in some ways takes us back to the early stages of the last century, when the great Paul Erlich and colleagues were stepping into new territory, by carrying out pharmacological studies using dyes from the textile industry. Clearly we understand much better where we are heading and we have a lot more experience under our belt, yet we are nowhere near finding a cure for some of the major diseases such as cancer. We are pretty good at using new tools and embracing technologies that allow us to visualise molecular interactions, draw upon previous information and knowledge to predict properties, avoid inefficiencies and redundancies and also explore new ways in which the experience and expertise can be used to explore scientific boundaries. Paradoxically, however, to meet the

healthcare challenges of the 21st Century we may have to, once again, learn how to "work in the dark" to search new areas of chemical space in the hunt for novel drugs. To achieve this, academia and industry must work together and share technical advances, but also share the burden of developing the next generations of drug-hunters. Efforts are already being made to advance education at the undergraduate level. An example of this was reported by Fray et al who reported a study carried out at the University of Nottingham in the UK and GSK [71], which was focussed on the development of a practical medicinal chemistry course, suitable for third-year undergraduates. In the course, the students learned about successfully designing more potent and soluble analogues of a weakly potent, insoluble phosphatidylinositide 3-kinase delta (PI3Kd) inhibitor through compound array design, molecular modelling, screening, data-analysis and the synthesis of target compounds in the laboratory. While this project benefitted from significant industrial support (including lectures, student mentoring and provision of a consumables budget) and is therefore more difficult to implement universally it was a great example of how academia and industry can work together to create a learning environment that embraced conceptual design and innovation in medicinal chemistry.

At the University of Leeds, fourth year undergraduates studying medicinal chemistry and postgraduate master's students studying chemical biology and drug design can attend a computational drug design course, which allows them to gain an insight into structure-based drug design and the use of computational methods. Students explore the active site of abl kinase and use computational docking to investigate possible variations of a known ligand (Imatinib) [72] in a simple virtual SAR investigation. The final part of the course involves use of SPROUT to carry out de

novo design looking for novel structures expected to inhibit the enzyme. Students also study medicinal chemistry strategy, biological chemistry and the use of important chemical biology tools.

One of Europe's strongest life science clusters, Medicon Valley, is based in Copenhagen but spans into eastern Denmark and the south-western part of Sweden. Students at the University of Copenhagen, which is located in the heart of the Medicon Valley, are major benefactors of this tight network of collaborations between academics and scientists at biotech companies. As an example, students studying on the Master's degree in Medicinal Chemistry at the University of Copenhagen are offered a state-of-the-art educational approach, which is focussed on the combination of understanding theoretical aspects of rational drug design with an underpinning of practical work that provides insight into the use of modern computerised chemical methods in drug design. All the approaches are translational and organised in way to optimise and prepare the students to meet the demands for immediate and future challenges of drug discovery.

On the other side of the Oresund strait, in Sweden, the University of Lund offers an undergraduate course in Medicinal Chemistry. This course also emphasises the importance of understanding the practical methods and techniques used for discovering novel candidate drugs. The course frequently make use of interconnected mini-projects where student teams are asked to mimic every aspect of the drug design process, from the beginning (the experimental design) until the end (evaluation of the biological activity of the target compound). The students carry out computer modelling of receptor-ligand complexes, organic synthesis of ligands, examination of the biochemical activity and physicochemical properties of the

ligands, and studies of ligand metabolism along with debates and presentations about their self-designed inhibitors.

In France, the University Paris Diderot offers an international master's degree entitled "In silico drug design" (ISDD), which encompasses both modern theory via lectures and practical teaching in virtual screening and data analysis. The Master's ISDD includes two international learning paths "bioactive molecules" and "biological macromolecules" which each encompasses at least one semester abroad. Up to half of the course is taught in English and the students are taught by international experts from Bulgaria, Canada, Denmark, Finland, Italy, Spain, Sweden and the UK amongst others. This means that the students are exposed to modern approaches to drug discovery using a variety of in silico approaches including virtual screening of therapeutic targets and drug safety parameters. In addition to in silico-based approaches, the students also receive training in chemical biology, a more recent scientific discipline, based on the fundamental principles of mature disciplines such as chemistry, biochemistry, biophysics, pharmacology, medicine, bioinformatics and biostatistics. The master's students are also exposed to experts from industry including Cambridge MedChem Consulting, Sanofi-Aventis and Schrödinger, which gives the students a sense of which tools are employed in the drug discovery process.

As a part of the students gaining practical experience, the students have an obligation to conduct an internship of minimum 5 months in a research center in academia or in industry. While on placements, the students are able to sharpen their understanding of research innovation by being directly involved in research projects asking challenging questions to current problems. Typical projects involving

participation by the internship students include (i) the design of novel competitive peptide-based compounds against key histone lysine demethylases based on cheminformatics and docking tools, (ii) evaluation of drug adverse effects due to genetic variation using biostatistics and structural bioinformatics, (iii) development of methods combining molecular dynamics simulations with free energy estimations or (iv) design of peptides and small molecules targeting protein-protein interactions based on computational approaches.

While constant refinement of computational and practical techniques is vital to teaching the next generation of medicinal chemists and chemical biologists, it is important that the current change from small molecules to biologics is also being taught in academia. The Drug Innovation Master's programme at the University of Utrecht in the Netherlands is addressing current demands by pharmaceutical and biotechnology companies for interdisciplinary research focussed on vaccines, monoclonal antibodies and gene therapeutics in addition to small molecule discovery. This 2-year Drug Innovation programme is a vital part of the largest Life Sciences cluster of the Netherlands, comprised of academic, clinical and industrial research groups. Much of the teaching can be tailored by the students and consists almost entirely of training-on-the-job with one-on-one supervision. The teaching is also based on the collaborative initiatives between different instituions. For example the Proteomics Centre collaborate in oncology research with proteomics, medicinal chemistry, drug delivery and clinical pharmacology while master's students can also be exposed to innovation in tissue engineering and stem cell biology. The fact that about 70% of the master's students choose to follow a PhD programme in the Netherlands or elsewhere is testament to teaching being carried out that addresses the current and future demands in drug innovation.

Although the above-mentioned institutions provide educational platforms for modern drug discovery that prepare students for meeting the current and future demands by pharmaceutical and biotech companies, it is worth pointing out that we also need to implement educational tools that help to address late-stage obstacles such as finding unanticipated toxicities in patients. Toxicities are a major problem in drug development and understanding mechanisms of toxicity could yield ways to circumvent them and/or enable exclusion of patients who might be prone to the most severe ones [73]. Off-target toxicities may be caused by the class of agent, immune reactions or toxic metabolites while individual genetic predisposition may impact treatment outcome in a number of ways [73,74].

Although there are incremental improvements in bringing undergraduate and postgraduate degrees in drug discovery up to speed with the demands of the health care sector, true progress in developing efficient innovative medicines is halted by the reduction in funding streams, key patent expirations, fierce price competition from generics and high regulatory hurdles. As discussed by Gharbia and Childers in two recent accounts [75,76], new business models are being developed which will help to guide companies through economic uncertainty.

Concluding remarks

A recent study [77] argued that medicinal chemists ought to pursue innovative goals and balance efficiency and innovation. The entrepreneurial medicinal chemist ought to broaden the discipline to chemical biology while maintaining close feedback from clinical development. The training in innovative thinking and entrepreneurship should be implemented already at the undergraduate level so one of the challenges is to make sure that science we teach at undergraduate and postgraduate level is

updated and not outdated. Exploring new regions of chemical space has reinvigorated the hunt for highly specific chemical tools and innovative drugs, expanding many more potential avenues that can be pursued in order to find success. We have, in this article, focussed on computational methods, bioisosterism/scaffold hopping and DOS as these are methodologies often used to develop NCEs while at the same time topics that can easily be implemented in medicinal chemistry and chemical biology courses without a great need for financial support. We appreciate that our contribution of moving drug discovery into the 21st Century is at large based on existing techniques and strategies that is already employed in Pharma and the more established academic institutions with a financial platform that enables state-of-the-art infrastructure. Nonetheless, we believe that better education of the drug hunter on a global scale will enable a gradual evolution of student teaching and ultimately lead to greater exposure of drug hunters to modern technologies and strategies.

In whatever shape or form the drug-hunter comes in, we live in a time where exploration and generation of new knowledge is occurring on a colossal scale; medicinal chemists can draw on an impressive volume of distilled knowledge in order to power through chemical space, but the journey will only be successful if we understand how to make the best use of both traditional and innovative tools. We are at the beginning of an exciting journey into unknown chemical space hoping to find new inhibitors and druggable targets, which will put us on the path towards cures for the major diseases we have not been able to treat so far.

Acknowledgements

We thank Anne-Claude Camproux and Olivier Taboureau at the University Paris Diderot in France for their help in structuring information about the international master's degree in "In silico drug design" (for more information, see http://isddteach.sdv.univ-paris-diderot.fr/en/accueil.html). Ulf Nilsson at the Department of Chemistry, University of Lund (Sweden) assisted with information about Medicinal Chemistry undergraduate course. Bente Frolund at the Department of Drug Design and Pharmacology, University of Copenhagen (DK) and Monique Slijper at the University of Utrecht (The Netherlands) is thanked for constructive dialogues about their Master's degrees. More information can be found at http://studies.ku.dk/masters/medicinal-chemistry/ and http://www.uu.nl/masters/en/drug-innovation/.

Bibliography

1 Bunnage, M.E. (2011) Getting pharmaceutical R&D back on target. *Nat Chem Biol* 7 (6), 335-339

2 Hann, M.M. and Keseru, G.M. (2012) Finding the sweet spot: the role of nature and nurture in medicinal chemistry. *Nat Rev Drug Discov* 11 (5), 355-365

3 Baell, J.B. and Holloway, G.A. (2010) New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem* 53 (7), 2719-2740

4 Lavecchia, A. and Di Giovanni, C. (2013) Virtual screening strategies in drug discovery: a critical review. *Curr Med Chem* 20 (23), 2839-2860

5 Johnstone, C. (2012) Medicinal chemistry matters - a call for discipline in our discipline. *Drug Discov Today* 17 (11-12), 538-543

6 Pors, K. (2011) Drug Discovery into the 21st Century. *in Drug Discovery and Development - Present and Future, edited by Izet M. Kapetanovic (Ed.), ISBN: 978-953-307-615-7, InTech, Available from: http://www.intechopen.com/articles/show/title/drug-discovery-into-the-21st-century.*

7 Pors, K. et al. (2009) The changing landscape of cancer drug discovery: a challenge to the medicinal chemist of tomorrow. *Drug Discov Today* 14 (21-22), 1045-1050

8 Satyanarayanajois, S.D. and Hill, R.A. (2011) Medicinal chemistry for 2020. *Future Med Chem* 3 (14), 1765-1786

9 Hoffmann, T. and Bishop, C. (2010) The future of discovery chemistry: quo vadis? Academic to industrial-the maturation of medicinal chemistry to chemical biology. *Drug Discov Today* 15 (7-8), 260-264

10 Doyle, M.P. (2010) The evolving nature of chemical education: challenges and opportunities. *Future Med Chem* 2 (2), 247-249

11 Pitman, S. et al. (2014) Spotlight on medicinal chemistry education. *Future Med Chem* 6 (8), 865-869

12 Roughley, S.D. and Jordan, A.M. (2011) The medicinal chemist's toolbox: an analysis of reactions used in the pursuit of drug candidates. *J Med Chem* 54 (10), 3451-3479

13 Lombardino, J.G. and Lowe, J.A., 3rd. (2004) The role of the medicinal chemist in drug discovery--then and now. *Nat Rev Drug Discov* 3 (10), 853-862

14 Erhardt, P.W. (2002) Medicinal chemistry in the new millennium. A glance into the future. *Pure and Applied Chemistry* 74 (5), 703-785

15 MacCoss, M. and Baillie, T.A. (2004) Organic chemistry in drug discovery. *Science* 303 (5665), 1810-1813

16 Kutchukian, P.S. and Shakhnovich, E.I. (2010) De novo design: balancing novelty and confined chemical space. *Expert Opinion on Drug Discovery* 5 (8), 789-812

17 Schneider, G. (2010) Virtual screening: an endless staircase? *Nat Rev Drug Discov* 9 (4), 273-276

18 Ripphausen, P. et al. (2011) State-of-the-art in ligand-based virtual screening. *Drug Discovery Today* 16 (9–10), 372-376

19 Reymond, J.L. and Awale, M. (2012) Exploring chemical space for drug discovery using the chemical universe database. *ACS Chem Neurosci* 3 (9), 649-657

20 Shah, F. et al. (2012) Computer-Aided Drug Design of Falcipain Inhibitors: Virtual Screening, Structure–Activity Relationships, Hydration Site Thermodynamics, and Reactivity Analysis. *Journal of Chemical Information and Modeling* 52 (3), 696-710

21 Mountford, S.J. et al. (2014) Synthesis, Structure–Activity Relationships and Brain Uptake of a Novel Series of Benzopyran Inhibitors of Insulin-Regulated Aminopeptidase. *Journal of Medicinal Chemistry* 57 (4), 1368-1377

22 Zheng, C. et al. (2014) Synthesis and Biological Evaluation of Novel Tetrahydro-β-carboline Derivatives as Antitumor Growth and Metastasis Agents through Inhibiting the Transforming Growth Factor-β Signaling Pathway. *Journal of Medicinal Chemistry* 57 (3), 600-612

23 Hawkins, P.C. et al. (2007) Comparison of shape-matching and docking as virtual screening tools. *J Med Chem* 50 (1), 74-82

24 Halland, N. et al. (2014) Discovery of N-[4-(1H-Pyrazolo[3,4-b]pyrazin-6-yl)-phenyl]-sulfonamides as Highly Active and Selective SGK1 Inhibitors. *ACS Medicinal Chemistry Letters* Epub 23 Oct. DOI: 10.1021/ml5003376

25 Halgren, T.A. et al. (2004) Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry* 47 (7), 1750-1759

26 Li, H. et al. (2014) Discovery of Small-Molecule Inhibitors Selectively Targeting the DNA-Binding Domain of the Human Androgen Receptor. *Journal of Medicinal Chemistry* 57 (15), 6458-6467

27 Yule, I.A. et al. (2014) Pyridine-3-carboxamide-6-yl-ureas as novel inhibitors of bacterial DNA gyrase: Structure based design, synthesis, SAR and antimicrobial activity. *European Journal of Medicinal Chemistry* 86 (0), 31-38

28 Gillet, V. et al. (1993) SPROUT - A PROGRAM FOR STRUCTURE GENERATION. *Journal of Computer-Aided Molecular Design* 7 (2), 127-153

29 Brady, R.M. et al. (2014) De-Novo Designed Library of Benzoylureas as Inhibitors of BCL-XL: Synthesis, Structural and Biochemical Characterization. *Journal of Medicinal Chemistry* 57 (4), 1323-1343

30 Rodrigues, T. et al. (2013) Steering Target Selectivity and Potency by Fragment-Based De Novo Drug Design. *Angewandte Chemie International Edition* 52 (38), 10006-10009

31 Deng, Z.L. et al. (2013) Exploring the biologically relevant chemical space for drug discovery. *J Chem Inf Model* 53 (11), 2820-2828

32 Vistoli, G. et al. (2008) Assessing drug-likeness--what are we missing? *Drug Discov Today* 13 (7-8), 285-294

33 Hilker, M. (2014) New synthesis: parallels between biodiversity and chemodiversity. *J Chem Ecol* 40 (3), 225-226

34 Wermuth, C.G. et al. Bioisosterism. . *In Comprehensive Medicinal Chemistry II Editors-in-Chief: John B. Taylor and David J. Triggle ISBN: 978-0-08-045044-5.*

35 Wermut, C.G. (2008) The Practice of Medicinal Chemistry (3rd Ed).

36 Meanwell, N.A. (2011) Synopsis of some recent tactical application of bioisosteres in drug design. *J Med Chem* 54 (8), 2529-2591

37 Ballatore, C. et al. (2013) Carboxylic acid (bio)isosteres in drug design. *ChemMedChem* 8 (3), 385-395

38 Poulie, C.B. and Bunch, L. (2013) Heterocycles as nonclassical bioisosteres of alpha-amino acids. *ChemMedChem* 8 (2), 205-215

39 Brown, N. (2014) Bioisosteres and Scaffold Hopping in Medicinal Chemistry *Mol. Inf.* 33, 458 – 462

40 Sun, H. et al. (2012) Classification of scaffold-hopping approaches. *Drug Discov Today* 17 (7-8), 310-324

41 Brown, N. (2014) Scaffold Hopping in Medicinal Chemistry. *First Edition. Wiley-VCH Verlag GmbH & Co. KGaA*

42 Ge, H. et al. (2014) Scaffold hopping of potential anti-tumor agents by WEGA: a shape-based approach Med. Chem. Commun. *Med. Chem. Commun* 5, 737.

43 Munier-Lehmann, H. et al. (2013) On dihydroorotate dehydrogenases and their inhibitors and uses. *J Med Chem* 56 (8), 3148-3167

44 Shih, K.-C. et al. (2014) Development of a Human Dihydroorotate Dehydrogenase (hDHODH) Pharma-Similarity Index Approach with Scaffold-Hopping Strategy for the Design of Novel Potential Inhibitors. *PLOS ONE* 9 (2), 1-11

45 Lolli, M.L. et al. (2012) New inhibitors of dihydroorotate dehydrogenase (DHODH) based on the 4-hydroxy-1,2,5-oxadiazol-3-yl (hydroxyfurazanyl) scaffold. *Eur J Med Chem* 49, 102-109

46 Davies, T.G. and Tickle, I.J. (2012) Fragment Screening Using X-Ray Crystallography in Fragment-Based Drug Discovery and X-Ray Crystallography. *Topics in Current Chemistry, Springer-Verlag Berlin Heidelberg.* 317, 33-59
47 Joseph-McCarthy, D. et al. (2014) Fragment-based lead discovery and design. *J Chem Inf Model* 54 (3), 693-704
48 Larsson, A. et al. (2011) Efficiency of hit generation and structural characterization in fragment-based ligand discovery. *Curr Opin Chem Biol* 15 (4), 482-488
49 Turnbull, A.P. and Boyd, S.M. (2012) Targeting cancer using fragment based drug discovery. *Anticancer Agents Med Chem* 12 (1), 40-48
50 Tiefendbrunn, T. and Stout, C.D. (2014) Towards novel therapeutics for HIV through fragment-based screening and drug design. *Prog Biophys Mol Biol* 116 (2-3), 124-140
51 Lam, K.S. and Renil, M. (2002) From combinatorial chemistry to chemical microarray. *Curr Opin Chem Biol* 6 (3), 353-358
52 Biggs-Houck, J.E. et al. (2010) Recent advances in multicomponent reactions for diversity-oriented synthesis. *Curr Opin Chem Biol* 14 (3), 371-382
53 Domling, A. et al. (2012) Chemistry and biology of multicomponent reactions. *Chem Rev* 112 (6), 3083-3135
54 Collins, I. and Jones, A.M. (2014) Diversity-Oriented Synthetic Strategies Applied to Cancer Chemical Biology and Drug Discovery. *Molecules* 19 (11), 17221-17255
55 Comer, E. et al. (2014) Diversity-oriented synthesis-facilitated medicinal chemistry: toward the development of novel antimalarial agents. *J Med Chem* 57 (20), 8496-8502
56 Koehler, A.N. et al. (2003) Discovery of an inhibitor of a transcription factor using small molecule microarrays and diversity-oriented synthesis. *J Am Chem Soc* 125 (28), 8420-8421
57 Kuruvilla, F.G. et al. (2002) Dissecting glucose signalling with diversity-oriented synthesis and small-molecule microarrays. *Nature* 416 (6881), 653-657
58 Schreiber, S.L. (2000) Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* 287 (5460), 1964-1969
59 Lehar, J. et al. (2008) Combination chemical genetics. *Nat Chem Biol* 4 (11), 674-681
60 Hoelder, S. et al. (2012) Discovery of small molecule cancer drugs: successes, challenges and opportunities. *Mol Oncol* 6 (2), 155-176
61 Cherkasov, A. et al. (2014) QSAR modeling: where have you been? Where are you going to? *J Med Chem* 57 (12), 4977-5010
62 Valerio, L.G., Jr. (2009) In silico toxicology for the pharmaceutical sciences. *Toxicol Appl Pharmacol* 241 (3), 356-370
63 Marchant, C.A. et al. (2008) In silico tools for sharing data and knowledge on toxicity and metabolism: derek for windows, meteor, and vitic. *Toxicol Mech Methods* 18 (2-3), 177-187
64 *http://www.lhasalimited.org/products/derek-nexus.htm* accessed 27/1/15
65 *http://www.epa.gov/oppt/sf/pubs/oncologic.htm* accessed 27/1/15
66 Kirchmair, J. et al. (2012) Computational prediction of metabolism: sites, products, SAR, P450 enzyme dynamics, and mechanisms. *J Chem Inf Model* 52 (3), 617-648
67 Moroy, G. et al. (2012) Toward in silico structure-based ADMET prediction in drug discovery. *Drug Discov Today* 17 (1-2), 44-55
68 Bohacek, R.S. et al. (1996) The art and practice of structure-based drug design: a molecular modeling perspective. *Med Res Rev* 16 (1), 3-50
69 Ertl, P. (2003) Cheminformatics analysis of organic substituents: identification of the most common substituents, calculation of substituent properties, and automatic identification of drug-like bioisosteric groups. *J Chem Inf Comput Sci* 43 (2), 374-380
70 Kirkpatrick, P. and Ellis, E. (2004) Chemical space. *Nature* 432, 823
71 Fray, M.J. et al. (2013) A practical drug discovery project at the undergraduate level. *Drug Discov Today* 18 (23-24), 1158-1172

72 Deininger, M.W. and Druker, B.J. (2003) Specific targeted therapy of chronic myelogenous leukemia with imatinib. *Pharmacol Rev* 55 (3), 401-423

73 Liu, S. and Kurzrock, R. (2014) Toxicity of targeted therapy: Implications for response and impact of genetic polymorphisms. *Cancer Treat Rev* 40 (7), 883-891

74 Liebler, D.C. and Guengerich, F.P. (2005) Elucidating mechanisms of drug-induced toxicity. *Nat Rev Drug Discov* 4 (5), 410-420

75 Abou-Gharbia, M. and Childers, W.E. (2014) Discovery of innovative therapeutics: today's realities and tomorrow's vision. 2. Pharma's challenges and their commitment to innovation. *J Med Chem* 57 (13), 5525-5553

76 Abou-Gharbia, M. and Childers, W.E. (2013) Discovery of innovative therapeutics: today's realities and tomorrow's vision. 1. Criticisms faced by the pharmaceutical industry. *J Med Chem* 56 (14), 5659-5672

77 Hoffmann, T. and Bishop, C. (2010) The future of discovery chemistry: quo vadis? Academic to industrial--the maturation of medicinal chemistry to chemical biology. *Drug Discov Today* 15 (7-8), 260-264

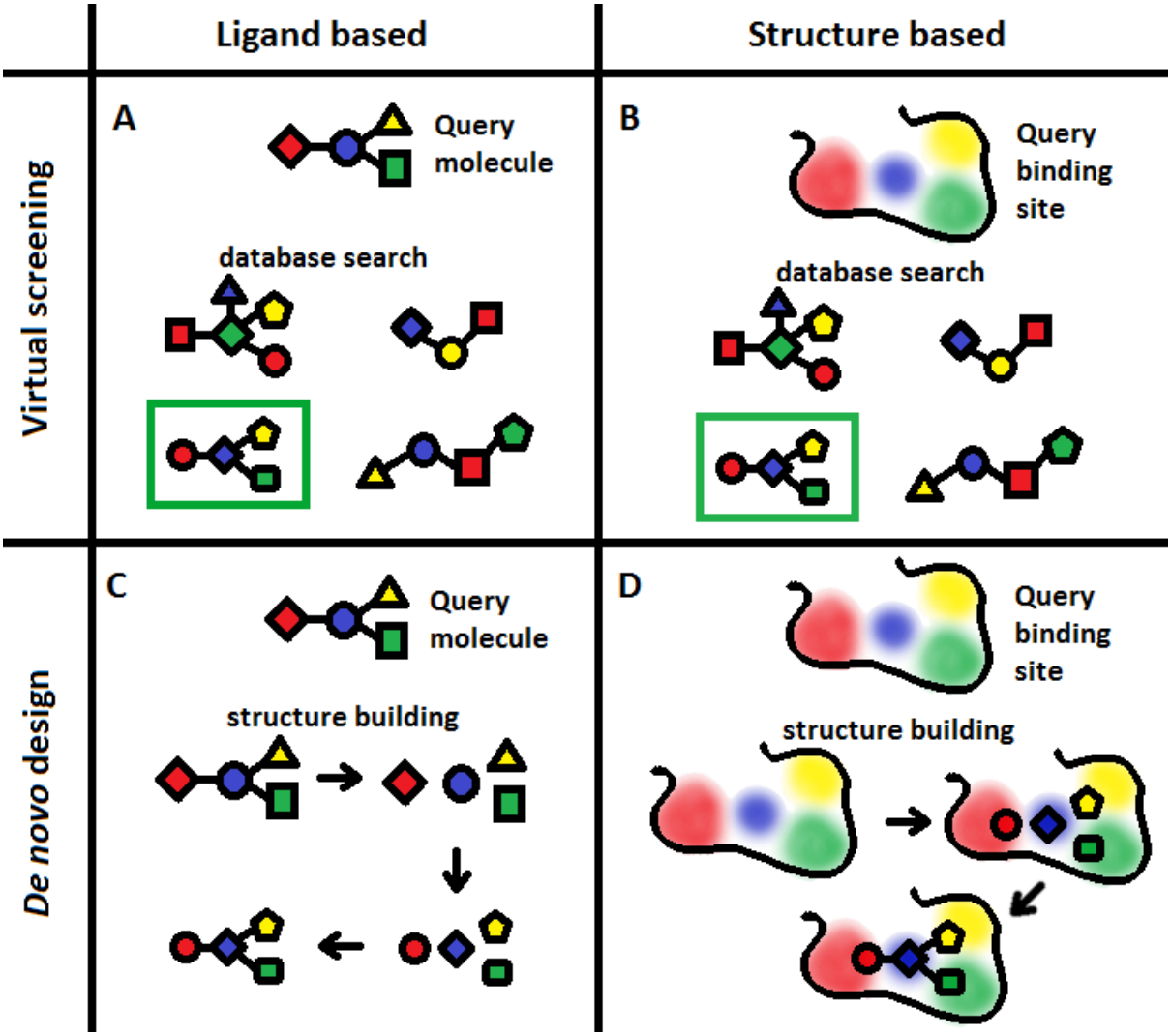
Figure Legends

Figure 1 Different approaches to computer-aided hit identification. A) ligand-based VS, B) structure-based VS C) ligand-based de novo design, D) structure-based de novo design.

Figure 2 Inhibitors designed using the chemical tools described in this paper.

Figure 3 Figure 3: Starting from of well-known DHODH inhibitor Brequinar (A, B), the application of a scaffold hopping strategy resulted in the design of a new class of inhibitors based on the 4-hydroxy-1,2,5-oxadiazol-3-yl scaffold. A comparable lead potency (low nM on rDHODH) by compound **8** (A, C) was obtained by lower binding contacts and, by reflex, with higher ligand efficiency (LE) [45].

Figures


Ligand based
Structure based
Virtual screening
De novo design
A
Query molecule
database search
B
Query binding site
database search
C
Query molecule
structure building
D
Query binding site
structure building