ELSEVIER

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

Marco Lolli[1], Sarah Narramore[2], Colin W.G. Fishwick[2] and Klaus Pors[3]

[1] Department of Science and Drug Technology, University of Turin, Via Pietro Giuria, 9-10125 Turin, Italy
[2] School of Chemistry, University of Leeds, Leeds LS2 9JT, UK
[3] Institute of Cancer Therapeutics, University of Bradford, Bradford BD7 1DP, UK

We live in a time where exploration and generation of new knowledge is occurring on a colossal scale. Medicinal chemists have traditionally taken key roles in drug discovery; 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 there is a necessity to refine the chemical toolbox for educational and practical reasons. This review proposes modern strategies that are beneficial to teaching in academia but are also important reminders of strategies that can potentially lead to better drugs.

Introduction

Drug discovery is a complex, time-consuming and 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 one new molecular entity (NME) per company per year despite their combined research expenditure climbing to US$60 billion during this period [1]. Failure in Phase II clinical trials has been highlighted as a key industry challenge, because only 25% of compounds that currently enter Phase II successfully progress to full Phase III clinical studies. It is well known that a large number of drug candidates fail during the development process owing to problems with 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] to limit the use of a significant amount of resources to develop candidates that are ultimately unsuccessful [4]. Medicinal chemists play a key part 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 educational and practical reasons.

Education and training of undergraduate students is, usually, still rooted in the 'classical' way of carrying out medicinal chemistry with emphasis on SAR, physicochemical properties, hit-to-lead optimisation, solubility and formulation. Paradigms such as 'Rule of Five', ligand efficiency (LE), '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 or not 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 also for other researchers who see themselves as drug hunters. If we are going to tap into the wealth of limitless omic data successfully 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]. Nevertheless, owing to the increasingly poor productivity of the

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

1359-6446/
http://dx.doi.org/10.1016/j.drudis.2015.04.010

Please cite this article in press as: Lolli, M. et al. Refining the chemical toolbox to be fit for educational and practical purpose for drug discovery in the 21st Century, Drug Discov Today (2015), http://dx.doi.org/10.1016/j.drudis.2015.04.010

pharmaceutical industry as measured by the number of FDA-approved new chemical entities (NCEs), there is a constant need to re-evaluate the drug discovery models we employ.

The purpose of this review 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, and these techniques 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 that 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 (VS) of databases [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 (Fig. 1 provides a simplistic overview) [18].

Virtual screening

Back in the 1990s, 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 enables the interrogation of large portions of chemical space without the need to possess the molecules in question. In the past three years, over 2000 papers have been published that mention the use of VS in applications as diverse as antimalarials [20], cognitive impairment [21] and cancer [22]. VS is becoming increasingly popular because of 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 serum- and glucocorticoid-regulated kinase (SGK)1 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 two 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 seven were active including compound **1** (Fig. 2). Library

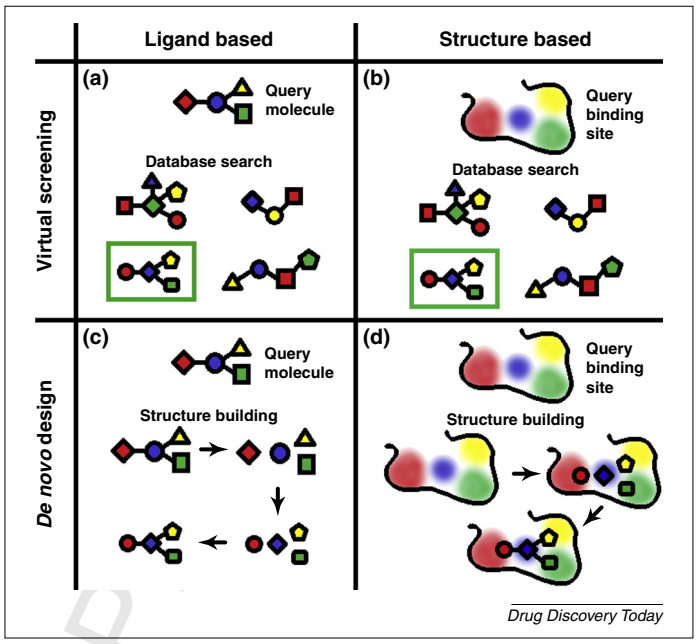

FIGURE 1

Different approaches to computer-aided hit identification. **(a)** Ligand-based virtual screening (VS), **(b)** structure-based VS, **(c)** ligand-based *de novo* design and **(d)** structure-based *de novo* design.

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, ADMET and pharmacokinetic (PK) properties was developed.

Recently, Li *et al.* have reported the use of structure-based VS to identify inhibitors that 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 that 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 micromolar range, and further shape similarity searching within the ZINC database identified even more active compounds that were used as the starting points for a full medicinal chemistry campaign resulting in compound **2** (Fig. 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, because 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

Please cite this article in press as: Lolli, M. et al. Refining the chemical toolbox to be fit for educational and practical purpose for drug discovery in the 21st Century, Drug Discov Today (2015), http://dx.doi.org/10.1016/j.drudis.2015.04.010