Medicinal chemistry approaches to neglected diseases drug discovery

http://www.producao.usp.br/handle/BDPI/50275

Downloaded from Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo
Medicinal Chemistry Approaches to Neglected Diseases Drug Discovery

Leonardo G. Ferreira and Adriano D. Andricopulo*

Abstract: The prevalence of a variety of neglected diseases is an increasing serious public health problem in developing countries, particularly in the poorest and most remote areas with very little or no access to medical care. The consequences in terms of morbidity and mortality due to these infections are devastating and have a major social and economic impact in several relevant aspects. According to the World Health Organization, these diseases are one of the most important scientific and technological challenges that face humankind in the 21st century. Although they affect more than a billion people around the world, there are only a few safe and effective drugs currently available. The urgent need for new drugs has led pharmaceutical and academic R&D centers to employ more knowledge-based platforms, as an unprecedented opportunity to make a significant impact on the lives of disadvantaged people through the discovery of novel therapeutic options. In this perspective we discuss the successful application of modern medicinal chemistry approaches to neglected diseases.

Keywords: Neglected diseases, Drug design, Virtual screening, Molecular dynamics, QSAR, Molecular modeling, Pharmacophore, LBDD, SBDD, Inhibitor.

NEGLECTED DISEASES

According to the World Health Organization (WHO), neglected tropical diseases (NTDs) are a heterogeneous group of 17 infectious conditions, which have as common attributes the prevalence in tropical areas of the globe, primarily striking developing countries, where the poorest population is the main target [1]. Approximately 150 countries are endemic in one of them, and a considerable fraction of 70% is plagued by two or more NTDs [1, 2]. In this group are included diseases such as Chagas’ disease, dengue, human African trypanosomiasis (HAT), schistosomiasis, leishmaniasis, leprosy and lymphatic filariasis. Although NTDs affect approximately one and a half billion people around the world, the burden they cause are mostly concentrated in rural areas and degraded urban zones. Precarious habitations, water pollution, poor sanitary infrastructure and the consequent proliferation of vector insects are among the top causes of high concentration of NTDs in these environments [3]. In recent years, this general tendency has been changing because of the migration of an increasing number of cases to developed countries, mainly in North America, Europe and Asia. In addition to the group of NTDs defined by WHO, a broader terminology can be employed to define the array of neglected diseases (NDs), which includes other deadly infectious conditions that generally receive relatively greater pharmaceutical interest and research funding, such as tuberculosis and malaria [4]. A report from the International Federation of Pharmaceutical Manufacturers & Associations (IFPMA), published in 2012, revealed a 40% increase of R&D projects on NDs, when compared with 2011 [1, 5]. However, considering the 336 new chemical entities (NCEs) approved during the last decade, only four are indicated to NDs, three for malaria and one for diarrhoeal disease [4].

The treatments for NDs are extremely limited and often ineffective. Several drugs currently available were developed to treat other conditions, such as cancer, bacterial and viral infections [6]. Examples include the compounds eflornithine and miltefosine, initially developed for cancer therapy, but presently used in the treatment of HAT and leishmaniasis, respectively (Figure 1). Another example is nifurtimox, a drug used to treat Chagas’ disease. Combination therapy with eflornithine and nifurtimox was found to be more safe and effective than treatment with eflornithine alone, and it has been recommended as first-line treatment for second-stage HAT. Melarsoprol, a drug formerly used in the treatment of late-stage HAT (Figure 1), is highly toxic to humans, a life-threatening characteristic of many antiprotozoal drugs. Another example is nifurtimox, whose gastrointestinal and neurological side effects have limited its therapeutic indication. In the past, nifurtimox was also used to treat HAT in combination with melarsoprol. Based on the present
situation and the serious existent problems, there is an urgent need for the development of new, effective and safe drugs for NDs [7, 8].

The research and development (R&D) trends for NDs have been changing because of the growing awareness on the subject in all levels of society, including governments, general public and the scientific community. As a consequence, the investments on basic and applied research have increased, mainly through public-private partnerships [9]. These multilateral efforts are generally coordinated and implemented by non-profit organizations in association with private companies and public research institutions. The Drugs for Neglected Diseases Initiative (DNDi); the Special Programme for Research and Training in Tropical Diseases (TDR) of WHO, and the Medicines for Malaria Venture (MMV) are involved in several projects in partnership with major pharmaceutical companies, such as Pfizer, GlaxoSmithKline, Novartis, AstraZeneca, Johnson & Johnson, Sanofi-Aventis, and others [9, 10]. These collaborations connecting private, public and government partners, include countries in all continents of world [6].

MEDICINAL CHEMISTRY AND DRUG DISCOVERY

Historically, the drug R&D process was based on empirical strategies, such as the systematic screening of large collections of compounds on different experimental models, and the synthesis of derivatives of known active compounds [11]. More recently, a new paradigm has emerged in academic and industrial environments because of the lack of innovation and the limited investigation of the chemical and biological space imposed by the traditional models [12]. For this reason, an increasing use of highly integrated drug discovery approaches has been noted, involving a combination of experimental and computational methods [13].

Among the core technologies employed in the early identification of novel small molecule hits are the high throughput screening (HTS) and virtual screening (VS) methods (Figure 2) [14, 15]. HTS relies on the use of miniaturized and automated assays – binding, competition and enzymatic assays – upon a collection of generally hundreds of thousands of structurally diverse molecules. The output of the experiments consists of a large set of molecules that holds low to moderate binding affinity for the target, and that may be suitable to structural modification. Consequently, hit-to-lead optimization efforts are undertaken to produce molecules with improved properties (e.g. potency, affinity, and selectivity) [16].

After the selection of lead compounds, the next step comprises the process of lead optimization (LO). At the optimization stage, structure- (SBDD) and ligand-based drug design (LBDD) methods are performed to improve the pharmacodynamics, pharmacokinetics and safety of the leads [17, 18]. The understanding of such a broad range of parameters is essential to provide the right balance in the characterization of the most promising compounds for further investigation [19].
The drug discovery process has been boosted, among other technological developments, by advances in combinatorial chemistry [20]. This approach provided the synthetic chemists the ability to generate large collections of compounds by exploring a small number of starting points in all allowed combinations through a given reaction sequence. From a medicinal chemist's point of view, combinatorial chemistry became an important tool in the organization of libraries of compounds for biological screening and hit identification [21].

On one hand, SBDD approaches rely on the use of 3D structural data derived from a variety of molecular targets [22]. VS, molecular docking and molecular dynamics (MD) are among some of the most important methods [23, 24], where the main goal is to simulate the interaction between small-molecule ligands and biological targets to gather insights into critical intermolecular events [25]. On the other hand, LBDD methods focus on the investigation of known biologically active ligands to produce knowledge that can be applied in multiple ways in the design of new compounds [26, 27]. Pharmacophores, quantitative structure-activity relationships (QSAR) and ligand-based VS (LBVS) are strategies usually employed in LBDD studies, both in academia and industry [28-30]. In addition, current data point out that the integration of SBDD and LBDD provides invaluable insights in lead identification and optimization, which are essential aspects of the drug discovery process [30-32].

STRUCTURE-BASED DRUG DESIGN STRATEGIES

The phase of target identification and validation is an important basis of successful drug development in key areas of human health [33]. Technological advances in protein crystallography and nuclear magnetic resonance (NMR) have allowed the resolution of more than 80,000 proteins [34]. The vast availability of structural information for an increasing number of proteins has strongly contributed to the field of SBDD in modern drug discovery. This knowledge is broadly used in the area of NTDs research for the design of high-affinity small-molecule ligands [23, 28, 35].

Figure 3 illustrates four crystal structures of extensively explored molecular targets for NTDs. These enzymes from the pathogens *Schistosoma mansoni* (3A), *Trypanosoma cruzi* (3B), *Trypanosoma brucei* (3C) and *Leishmania* (3D) are vital for the survival and proliferation of the parasites. The binding modes depicted in Figure 3 reveal the essential intermolecular interactions present in the complexes between proteins and small-molecule inhibitors.

SBDD methods can be applied to the solution of several drug design problems, generally starting with the investigation of a 3D structure of a given biological target. Next, the appropriate binding cavity of the protein may be considered in the processes of ligand optimization or *in silico* screening of new ligands. Subsequently, a privileged set of compounds is either synthesized or purchased, and then experimentally tested *in vitro* against the target protein (Figure 4) [31].

Virtual Screening and Molecular Docking

The *in silico* screening of libraries of compounds is one of most applied strategies to identify hits in the early phases of the drug discovery process. VS is the
Figure 3: Crystallographic structures of enzyme-inhibitor complexes. (A) Schistosoma mansoni purine nucleoside phosphorylase (PNP) in complex with a nucleoside analog (PDB ID: 3DJF); (B) Cruzain from Trypanosoma cruzi in complex with a bromophenoxy-acetamide inhibitor (PDB ID: 3KUU); (C) Trypanosoma brucei dihydrofolate reductase (DHFR) in complex with the inhibitor pyrimethamine (PDB ID: 3QFX); (D) Leishmania major pteridine reductase 1 with a 2,4-diaminopteridine inhibitor (PDB ID: 3H4V). Protein structures in cartoon, inhibitors in stick and water molecules in red spheres.

Figure 4: SBDD in drug design.
Among the several phases comprising the VS process, the filtering step is important to reduce the size of the database, and to restrict the chemical space being explored in terms of lead- or drug-like properties. ADME: absorption, distribution, metabolism and excretion.

SBVS methods involve the computational docking of large libraries of small-molecule compounds into the binding site of the target protein [37]. Molecular docking is a well-known method used to predict the binding
conformation of the ligands within the binding pocket of the biological receptor [38]. It has been largely employed to evaluate the bound conformation preference and the free binding energy [39]. Binding conformations predicted by molecular docking are in general close to the corresponding binding conformations observed in crystallographic structures [38]. However, the difficulty to find correct binding solutions increases proportionally with the degree of flexibility of the docked molecules. In addition, the docking algorithms are applied to classify the compounds using a variety of scoring functions. The majority of the currently employed scoring functions is based on force fields, empirical data and knowledge of the protein-ligand interactions in order to predict the free energy of binding between the docked ligands and the macromolecular target [38, 39]. The ranking list is ultimately used to select a subset of promising compounds for experimental (in vitro) evaluation (Figure 5) [40].

**Molecular Dynamics**

The selection of promising lead compounds requires the integration of different strategies. In this framework, SBVS can be combined with molecular dynamics (MD) simulations, considering that numerous molecular targets may undergo conformational changes of varied extensions. MD simulations play an essential role in the characterization of molecular motion [41]. In some cases, the receptor’s changes are small and the ligand fits into a well-defined cavity. In contrast, some proteins may undergo major alterations in their structure during the process of molecular recognition [42]. In the latter case, it is useful to generate an ensemble of structures, and then consider the most representative conformational states to produce meaningful SBVS results [43, 44].

MD simulations have been successfully employed in the elucidation of the molecular basis underlying experimental data, for instance, in the case of enzyme inhibition measurements [43, 45]. Regardless of some important limitations, such as the high computational cost and system size, MD has contributed in several ways in drug design projects, mainly when combined with other medicinal chemistry approaches [46].

**LIGAND-BASED DRUG DESIGN STRATEGIES**

In a considerable number of cases, the discovery of new compounds with pharmacological activity can be done without the use of the target’s structural information. LBDD strategies are widely applied to the design and optimization of ligands based on the properties extracted from previously known active molecules (Figure 6) [47].

Pharmacophore modeling is one of the most important LBDD strategies employed in drug design [48]. This technology captures common structural features from a specific set of ligands, which are thought to be essential in the process of molecular binding to the target protein. Firstly, the conformational space for each molecule is delineated to characterize the conformations of the entire set of ligands [49]. Next, the several ligands are aligned and the key common features are identified. This results in the generation of a so called 3D pharmacophore hypothesis, which can then be used to search large databases of compounds with unknown activities [48, 49].

Another well-known LBDD method is quantitative structure-activity relationships (QSARs) [50]. Over the last decades QSAR models have been successfully used to describe and quantify the underlying correlations between molecular properties and biological activity [29, 50]. These relationships are converted to mathematical models that can be used to predict the biological activity of novel compounds not yet synthesized. According to the type of descriptor, QSAR methods can be broadly classified as 1D, 2D or 3D. Several types of molecular descriptors can be employed: physicochemical (e.g. logP and pKa; 1D QSAR), topological and fingerprints (2D QSAR), and molecular interaction fields (3D QSAR) [51]. QSAR studies are usually applied to a specific chemical space to be explored in the design of new molecules with improved properties, particularly in the hit-to-lead and lead optimization phases [52, 53].

LBVS methods do not consider the 3D structure of the macromolecular target. In this case, the molecular properties of active compounds are used as the search criteria to identify molecules with similar structure. The search can be based on 1D, 2D or 3D molecular descriptors. The selection of adequate descriptors is critical, and depends on the nature of the molecules used as search templates. Typically used descriptors include 1D bit strings, 2D common substructures, and 3D pharmacophore hypothesis and molecular shape [54, 55]. LBVS approaches are broadly divided in filtering or activity-based methods (Figure 7). By using filtering methods, large databases can be rapidly screened for molecules containing specific properties (e.g. lead-like, drug-like) or substructures, or absence
of toxic or undesired groups. In activity-based methods, QSAR models, pharmacophore hypothesis and classification models can be used to screen libraries of compounds, as well as to classify them as actives or inactives, and predict their activity [56].

The integration between LBDD and SBDD approaches is present in a vast number of cases described in the literature [57]. The use of these strategies requires a careful design to ensure maximum benefit [57]. In the following section we briefly discuss some examples of studies in the area of NDs.

**CURRENT MEDICINAL CHEMISTRY APPROACHES IN NEGLECTED DISEASES**

Valuable strategies in medicinal chemistry have been employed to the discovery and optimization of lead compounds for NDs. A few successful examples of the integration of VS, QSAR, molecular docking, pharmacophore modeling and MD are described below.

Recently, a series of compounds exhibiting potent activity against *Plasmodium falciparum* was discovered [58]. The authors initially developed several QSAR models based on the *in vitro* activity of nearly 3,000
compounds against *P. falciparum* cells. Subsequently, the ZINC database was filtered for drug-likeness properties. Next, the compounds selected were submitted to a procedure (applicability domain – AD – analysis) that delimitates the chemical space in which a QSAR model provides reliable predictions for both internal and external validation sets. Predictions were provided by QSAR models for the compounds within the AD, and a subset of promising compounds was experimentally tested. Several hits were found to inhibit the parasite’s growth in cell cultures. The interesting chemical diversity of the series can be further explored in the design of novel antimalarial agents (Figure 8A).

The enzyme 2-trans-enoyl-ACP-CoA-reductase (InhA) is a key target of the fatty acid synthesis pathway of *Mycobacterium tuberculosis*, the causative agent of tuberculosis. A subset of the ZINC database was used in a strategy combining VS and 3D pharmacophore modeling [59]. Firstly, a 3D pharmacophore model was generated based on available InhA crystal structures. The information derived from SBDD and LBDD investigations allowed the generation of a pharmacophore consisting of four key points, which guided the selection of molecules with complementary properties to the InhA binding site. Subsequently, four docking programs were employed to compare the binding mode of the previously selected molecules. Three molecules possessing IC\(_{50}\) values ranging from 24 to 83 μM were identified. The study resulted in the discovery of promising novel scaffolds – oxadiazols and thiadiazols – as inhibitors of *M. tuberculosis* InhA (Figure 8B).

In an interesting work, a combined strategy comprising docking studies and a HTS assay of approximately 20,000 compounds was developed employing the enzyme cruzain from *T. cruzi*, a validated target for Chagas' disease [60]. An initial set with nearly 1,000 hits was identified from the HTS campaign. Molecular docking studies were performed upon these hits, and the top scoring compounds were evaluated against the target enzyme. This procedure allowed the selection of a subset of potent inhibitors with \( K_i \) (the dissociation constant for the enzyme-inhibitor complex) values in the low micromolar and nanomolar range (Figure 8C). These inhibitors, belonging to five different structural classes, are attractive starting points for future development.

The association of protein crystallography and molecular docking led to the study of a series of potent inhibitors of the enzyme purine nucleoside phosphorylase (PNP) from *Schistosoma mansoni*, the causative agent of schistosomiasis [61]. Crystallographic data of the enzyme were used to assist the SBDD investigation of a series of deazaguanine derivatives and other purine bases. The compounds showed excellent inhibitory profile, with IC\(_{50}\) values in the nanomolar range (Figure 9A). The
authors also performed kinetic evaluations upon the human PNP to assess the selectivity of the inhibitors. Three compounds showed good selectivity towards the parasite’s enzyme. The obtained crystallographic complex for one of the inhibitors (Figure 3A) revealed several key structural features for selectivity that can be explored in the optimization of this series of inhibitors of \( S.\ mansonii \) PNP.

The RNA-editing ligase (REL1) is an essential enzyme of the RNA editing pathway, a unique process of protozoa, which include the parasite \( T.\ cruzi \).

**Figure 8:** Chemical structures of compounds identified through the integration of medicinal chemistry approaches. (A) Two active compounds against \( P.\ falciparum \) cells. (B) Inhibitors of the InhA enzyme from \( M.\ tuberculosis \); (C) Inhibitors of the enzyme cruzain from \( T.\ cruzi \).

**Figure 9:** (A) Structures of \( S.\ mansonii \) PNP deazaguanine inhibitors; (B) Structures of \( T.\ brucei \) REL1 inhibitors.
**brucei**, the causative agent of HAT. Drug-like inhibitors were discovered through the combination of MD and SBVS approaches [62]. A subset of the National Cancer Institute (NCI) database was virtually screened using a crystal structure of the enzyme (PDB ID: 1XDN). The top scoring compounds were redocked into the binding pocket using structures of the enzyme obtained through MD – a process called relaxed complex scheme (RCS). Following this, the most promising compounds were filtered for drug-likeness, leading to the identification of potent inhibitors. The structure of the most potent inhibitor was used in a similarity search for new REL1 inhibitors, employing the entire NCI database (Figure 9B). This series of disulfonic acid derivatives are lead candidates for the development of new anti-HAT agents.

**CONCLUSIONS**

Although NDs account for 12% of the global health burden, only about 1% of the new NCEs launched over the last decade were targeted to these conditions [4]. To address this critical gap, pharmaceutical companies, not-for-profit organizations, academic and research institutions have been progressively more engaged in several modalities of partnerships to raise the levels of investments and research activities in the area of NDs. As a consequence, good opportunities are becoming increasingly available for these alarming public health problems. Current research has taken advantage of the existing open databases of small-molecules and macromolecular targets, whose availability is likely to increase over the next years. This aspect is of great importance, particularly for the application of SBDD and LBDD methods. Successful examples of the use of these approaches can be found for a variety of different diseases and drug discovery programs. It seems clear that, from a careful examination of these cases, important information can emerge to support knowledge-based decisions. In this context, we believe that the use of a diversity of strategies in medicinal chemistry will continue to play a significant role in the area of NDs for the foreseeable future.

**ACKNOWLEDGEMENTS**

We gratefully acknowledge financial support from FAPESP (The State of Sao Paulo Research Foundation, grant #2013/07600-3) and CNPq (The National Council for Scientific and Technological Development), Brazil.


Ferreira and Andricopulo


Ferreira and Andricopulo


Received on 17-12-2013 Accepted on 13-01-2014 Published on 06-03-2014

DOI: http://dx.doi.org/10.12970/2308-8044.2014.02.01.4

© 2014 Ferreira and Andricopulo; Licensee Synergy Publishers.
This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.