Drug Design Progress of In silico, In vitro and In vivo Researches

Qifeng Bai\textsuperscript{1,*,#}, Lanlan Li\textsuperscript{2,*,#}, Shanhui Liu\textsuperscript{2}, Shujun Xiao\textsuperscript{1,2}, Yu Guo\textsuperscript{1,2}

\textsuperscript{1}Key Lab of Preclinical Study for New Drugs of Gansu Province, School of Basic Medical Sciences, Lanzhou University, Lanzhou, Gansu 730000, P. R. China

\textsuperscript{2}Institute of Urology, Lanzhou University Second Hospital; Key Laboratory of Urological Diseases in Gansu Province, Lanzhou University; Gansu Nephro-Urological Clinical Center, Lanzhou, Gansu, China, 730030

#These authors contributed equally to this work.

Abstract

Drug design, referred to the fields of pharmacology, biotechnology and medicine, is \textit{in silico}, \textit{in vitro} and \textit{in vivo} assay processes of finding new candidate medications based on the biological targets. The \textit{in silico} experiments of drug discovery are involved in the macromolecular structure databases, small molecule databases, molecular docking, de novo drug design and molecular dynamics simulations. The \textit{in vitro} experiments of drug discovery need evaluate the direct interaction information between ligands and targets as well as the function of ligands on signaling pathway in the cell. The \textit{in vivo} experiments of drug discovery give the convincing evidence for preclinical trial at the physiological level. In this review, we outline the drug design components of databases, virtual screening tools, biochemical assays, cell-based system and animal models.

Corresponding Author: Qifeng Bai, Key Lab of Preclinical Study for New Drugs of Gansu Province, School of Basic Medical Sciences, Lanzhou University, Lanzhou, Gansu 730000, P. R. China, Email: baiqf@lzu.edu.cn; Lanlan Li, Institute of Urology, Lanzhou University Second Hospital; Key Laboratory of Urological Diseases in Gansu Province, Lanzhou University; Gansu Nephro-Urological Clinical Center, Lanzhou, Gansu, China, 730030, Email: llii12@lzu.edu.cn

Keywords: virtual screening tools, biochemical assays, cell-based system and animal models

Received: May 05, 2018 \hspace{1cm} Accepted: Aug 03, 2018 \hspace{1cm} Published: Aug 16, 2018

Editor: George Kordas, Twice ERC laureate (Nanotherapy Advanced Grant and PoC Grant) ERC LS7 Panel Member, Member of the ASF 15 Panel
Introduction

Modern new drug design is the integrated and long-term processes which will cost tens to hundreds of millions dollars from candidate compounds trial to Food and Drug Administration (FDA) approval. The preclinical drug design trial can be made of in silico, in vitro, and in vivo experiments. The development of information technology and big data accelerate the speed of drug discovery because the high effective and targeted databases are constructed under these circumstances. The famous databases designed for drug discovery are introduced in detail following review part. Based on the accurate structural models of biological targets and small molecules, the molecular docking software can be used to estimate the affinity of ligand in the pocket of biological macromolecules. In addition, the de novo method is another computer-aided design for new drug generation on basis of 3D-structural targets and pharmacophore model. To study the dynamical interaction between receptors and ligands at the atomic level, the molecular dynamics (MD) simulations supply a reliable and accurate way to explore the binding mechanism between ligands and targets. Our review describes the popular molecular docking, de novo drug and MD simulation software in drug discovery field. The candidate compounds screened in silico still need be validated to make sure the compound has pharmaceutical activity. The direct binding experiment in vitro between ligands and targets should be trial by using the methods of X-ray crystallography structural analysis, Surface Plasmon Resonance (SPR), etc. Because the direct binding assay cannot guarantee the activity of screened ligands, the cell signaling pathway response experiment should be performed to check the activity of ligands in cell. Due to the complex physiological environment, the active ligands in vitro trial may not show any response to the targeting disease in vivo. Hence, it need choose the suitable animal models for in vivo experiments. In this review, we introduce the database, in silico drug design software, in vitro experiment, and in vivo animal models (see Figure 1). Of course, the sequence of in silico, in vitro and in vivo experiments can be changed according to specific conditions. For instance, if the animal models are easier to be got and cheaper to be bought than biochemical assays, the in vivo experiments can be placed before the in vitro trial. Generally, our review gives the researchers an easy understanding contour for drug design.

In silico Database for Drug Design

With the development of bioinformatics, big data, biology, chemistry and medicine, more and more databases are design to service for the drug discovery. The databases are divided into macromolecular and small molecular databases. The macromolecular structure databases contain the crystal structures of proteins, nucleic acids, or other biopolymers. The small molecular databases are design to service for the drug discovery data, biology, chemistry and medicine, more and more databases are constructed under these circumstances. The databases of these biological macromolecules. In addition, the de novo method is another computer-aided design for new drug generation on basis of 3D-structural targets and pharmacophore model. To study the dynamical interaction between receptors and ligands at the atomic level, the molecular dynamics (MD) simulations supply a reliable and accurate way to explore the binding mechanism between ligands and targets. Our review describes the popular molecular docking, de novo drug and MD simulation software in drug discovery field. The candidate compounds screened in silico still need be validated to make sure the compound has pharmaceutical activity. The direct binding experiment in vitro between ligands and targets should be trial by using the methods of X-ray crystallography structural analysis, Surface Plasmon Resonance (SPR), etc. Because the direct binding assay cannot guarantee the activity of screened ligands, the cell signaling pathway response experiment should be performed to check the activity of ligands in cell. Due to the complex physiological environment, the active ligands in vitro trial may not show any response to the targeting disease in vivo. Hence, it need choose the suitable animal models for in vivo experiments. In this review, we introduce the database, in silico drug design software, in vitro experiment, and in vivo animal models (see Figure 1). Of course, the sequence of in silico, in vitro and in vivo experiments can be changed according to specific conditions. For instance, if the animal models are easier to be got and cheaper to be bought than biochemical assays, the in vivo experiments can be placed before the in vitro trial. Generally, our review gives the researchers an easy understanding contour for drug design.
Table 1. Macromolecular structure databases

<table>
<thead>
<tr>
<th>Structure databases</th>
<th>Description</th>
<th>Weblink</th>
</tr>
</thead>
<tbody>
<tr>
<td>wwPDB</td>
<td>Worldwide macromolecular structures</td>
<td><a href="http://www.wwpdb.org">www.wwpdb.org</a></td>
</tr>
<tr>
<td>RCSB PDB</td>
<td>Macromolecular structures</td>
<td><a href="http://www.rcsb.org">www.rcsb.org</a></td>
</tr>
<tr>
<td>PDBe</td>
<td>Macromolecular structures in Europe</td>
<td><a href="http://www.ebi.ac.uk/pdbe">www.ebi.ac.uk/pdbe</a></td>
</tr>
<tr>
<td>PDBj</td>
<td>Macromolecular structures in Japan</td>
<td>pdbj.org</td>
</tr>
<tr>
<td>BMRB</td>
<td>Macromolecular structures from NMR</td>
<td><a href="http://www.bmrbr.wisc.edu">www.bmrbr.wisc.edu</a></td>
</tr>
<tr>
<td>NDB</td>
<td>Nucleic acid structure database</td>
<td>ndbserver.rutgers.edu</td>
</tr>
<tr>
<td>JenaLib</td>
<td>3D biopolymer structures</td>
<td>jenalib.leibniz-fli.de</td>
</tr>
<tr>
<td>PDBbind</td>
<td>3D macromolecular structures in complex with ligands</td>
<td><a href="http://www.pdbbind.org.cn">www.pdbbind.org.cn</a></td>
</tr>
<tr>
<td>IUPHAR/BPS Guide to PHARMACOLOGY</td>
<td>Pharmacological targets and substances</td>
<td><a href="http://www.guidetopharmacology.org">www.guidetopharmacology.org</a></td>
</tr>
<tr>
<td>ModBase</td>
<td>Theoretical 3D macromolecular structures</td>
<td>modbase.compbio.ucsf.edu/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>modbase-cgi/index.cgi</td>
</tr>
<tr>
<td>PMP</td>
<td>Theoretical 3D macromolecular structures</td>
<td><a href="http://www.proteinmodelportal.org">www.proteinmodelportal.org</a></td>
</tr>
<tr>
<td>SWISS-MODEL</td>
<td>Theoretical 3D macromolecular structures</td>
<td>swissmodel.expasy.org/repository</td>
</tr>
</tbody>
</table>

Figure 1. The diagram of *in silico*, *in vitro* and *in vivo* for drug design.
G protein-coupled receptors (GPCRs)\textsuperscript{21-23}, transporter\textsuperscript{24}, ion channels receptors\textsuperscript{25}, and so on. IUPHAR/BPS Guide to PHARMACOLOGY\textsuperscript{26, 27} provides the information of clinical, approved drugs and candidate compounds for the popular targets such as GPCRs, ion channels receptors, kinases, transporters, and so on. It is very simple for users to find the drug comprehensive resources which contain 3D crystal structure targets, relative database links, bioactive ligands from literature, antibodies, functional assays, physiological functions, disease models and so on.

Until now, the number of proteins with crystal structures is limited because there are still some intractably unresolved crystal proteins and nucleic acids structures. The targets without the crystal structures cannot supply the accurate 3D models for drug discovery. In order to screen the candidate compounds based on the active sties of targets without crystal structure, it need predict the 3D structures from the sequence of proteins or nucleic acids. The ModBase\textsuperscript{28}, PMP\textsuperscript{29}, and SWISS-MODEL\textsuperscript{30} can help the researchers build the 3D theoretical models of targets by comparative modeling methods. The theoretical 3D macromolecular structures also give the reliable 3D models to screen the potential drugs from the large of small molecule database. Some studies have shown homology model can be considered as the valid target to perform the virtual screening\textsuperscript{31, 32}. The crystal and predicted models have been widely applied into the research field of drug discovery.

Drug discovery not only needs the reliable models, but also is relied on huge of small molecules with reasonable conformations. Table 2 shows the popular and special purpose small molecule databases for drug discovery. ZINC database\textsuperscript{33, 34} is considered as an open-access commercially-available screening library which contains over 35 million purchasable small molecules for virtual screening. ZINC database, which supplies Lead-Like, Drug-Like, Fragment-Like for different research needs, has been reported to be used to screen the candidate compounds by molecular docking methods\textsuperscript{35, 36}. PubChem database currently consists of compounds, substances and bioassay databases which recruit the 93.9 million, 249 million and 1.25 million entries, respectively. Therefore, PubChem\textsuperscript{37} database can be used to dig out the potential bioactivated compounds through molecular docking and deep learning studies. DrugBank\textsuperscript{38} is a comprehensive and open-access database containing detail information of drug data and drug targets. In combination with the ZINC database, it has been used to find the ligands via molecular docking based screening\textsuperscript{39, 40}. The databases of ChemBridge, Specs and ChemDiv can supply with the commercially customized screening libraries for the drug development targeting to research receptors such as cyclophilin D\textsuperscript{41}, \(p\)-Hydroxyphenylpyruvate dioxygenase (HPPD)\textsuperscript{42}, SETDB1/ESET\textsuperscript{43}, and so on. The e-Drug3D\textsuperscript{44} database recruits the 1852 molecular structures which approved by Food and Drug Administration (FDA) between 1939 and 2017 with a molecular weight \(\leq 2000\). The e-Drug3D can be used as the benchmark for finding the effective candidate compounds. Super Natural II\textsuperscript{45} is a database of natural compounds with physicochemical properties, predicted toxicity information, 2d structures, and vendors. In all, \(\sim 75\%\) of FDA approved small molecular drugs are from natural compounds or its derivatives\textsuperscript{46}. Hence, Super Natural II database has a good promising for drug development by \textit{in silico} method. The Cambridge Structural Database (CSD)\textsuperscript{47} and Crystallography Open Database (COD)\textsuperscript{48} are mainly interested in the collection of small molecule organic and inorganic crystal structures. It cannot only be used for drug virtual screening, but also give the accurate structural information for other computing such as quantum chemical calculation. The KEGG Ligand database\textsuperscript{49} comprises COMPOUND, REACTION, and ENZYME which is responsible for collecting the chemical compounds, chemical reactions, and enzyme molecules, respectively. The nucleic acid ligand database (NALDB)\textsuperscript{50} and small molecule pathway database (SMPDB)\textsuperscript{51} are designed for the special purpose of drug discovery. The NALDB provides the detail experimental data of small molecules which target to nucleic acid structures. The SMPDB gives the ligands interactive network on signal pathway found in human which is used to elucidate the drug-action signaling pathways. The MarinChem3D (mc3d.qnlm.ac), which locates at National Laboratory for Marine Science and Technology (Qingdao) in China, publishes more than 30,000 well-defined 3D structures of marine natural products. It can be used to perform the virtual screening studies directly. The MarinChem3D
<table>
<thead>
<tr>
<th>Ligand databases</th>
<th>Description</th>
<th>Weblink</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ZINC</strong></td>
<td>Over 35 million purchasable compounds for virtual screening</td>
<td>zinc.docking.org</td>
</tr>
<tr>
<td><strong>PubChem</strong></td>
<td>Over 90 million compounds</td>
<td>pubchem.ncbi.nlm.nih.gov</td>
</tr>
<tr>
<td><strong>DrugBank</strong></td>
<td>More than 11,000 drug entries</td>
<td><a href="http://www.drugbank.ca">www.drugbank.ca</a></td>
</tr>
<tr>
<td><strong>Specs</strong></td>
<td>Providing high throughput screening compounds</td>
<td><a href="http://www.specs.net">www.specs.net</a></td>
</tr>
<tr>
<td><strong>ChemBridge</strong></td>
<td>Over 1.1 million druglike and leadlike compounds</td>
<td><a href="http://www.chembridge.com/screening_libraries">www.chembridge.com/screening_libraries</a></td>
</tr>
<tr>
<td><strong>ChemDiv</strong></td>
<td>Over 1,5 M individual solid screening compounds</td>
<td><a href="http://www.chemdiv.com/services-menu/screening-libraries">www.chemdiv.com/services-menu/screening-libraries</a></td>
</tr>
<tr>
<td><strong>e-Drug3D</strong></td>
<td>1852 FDA approved drugs between 1939 and 2017</td>
<td>cheminfo.ipmc.cnrs.fr/MOLDB/index.html</td>
</tr>
<tr>
<td><strong>Super Natural II</strong></td>
<td>325,508 natural compounds</td>
<td>bioinf-applied.charite.de/supernatural_new/index.php</td>
</tr>
<tr>
<td><strong>CSD</strong></td>
<td>Over 900,000 small-molecule organic crystal structures</td>
<td><a href="http://www.ccdc.cam.ac.uk/solutions/csd-system/components/csd">www.ccdc.cam.ac.uk/solutions/csd-system/components/csd</a></td>
</tr>
<tr>
<td><strong>COD</strong></td>
<td>Over 390,000 inorganic crystals and small organic compounds</td>
<td><a href="http://www.crystallography.net/cod">www.crystallography.net/cod</a></td>
</tr>
<tr>
<td><strong>KEGG Ligand</strong></td>
<td>Universe of chemical substances and reactions</td>
<td><a href="http://www.genome.jp/kegg/ligand.html">www.genome.jp/kegg/ligand.html</a></td>
</tr>
<tr>
<td><strong>NALDB</strong></td>
<td>Nucleic acid ligand database</td>
<td>bsbe.iiti.ac.in/bsbe/naldb/HOME.php</td>
</tr>
<tr>
<td><strong>SMPDB</strong></td>
<td>Over 30,000 small molecule pathways found in humans</td>
<td>smpdb.ca</td>
</tr>
<tr>
<td><strong>MarinChem3D</strong></td>
<td>Over 30,000 kinds of marine compounds</td>
<td>mc3d.qnlm.ac</td>
</tr>
</tbody>
</table>
gives a promising way to find candidate ligands targeting to receptors from the ocean.

*In silico Software for Drug Design*

Although the macromolecular structure and small molecule databases are an important factor for drug discovery *in silico*, it still needs the effective software for performing virtual screening on targets and small molecule databases. In the past decades, various molecular docking software emerges based on different algorithms and molecular formats. Generally, the computational methods for drug discovery can be divided into ligand-based (indirect) and structure-based (direct) techniques\(^5\). The ligand-based drug design methods contain quantitative structure-activity relationship (QSAR)\(^5\), pharmacophore\(^5\), etc. The structure-based drug design contains molecular docking and *de novo* methods. With the development of genomics and the accumulation of pharmacological information, the big data and deep learning have permeated into the drug discovery fields. For instance, the tensorflow\(^5\), which is a deep learning software library, has been used for the drug discovery and molecular dynamics simulations\(^5\). These popular computational methods have been integrated into different software for drug design. The QSAR and pharmacophore model can be constructed by Schrödinger or Discovery Studio software. The molecular docking software can be divided into free academic and commercial programs. Besides, molecular dynamics simulations\(^5\) are considered as the accurate and dynamical way to study the interaction between targets and ligands. Table 3 shows the free academic, commercial molecular docking programs and molecular dynamics simulations software. LeDock\(^5\) is designed based on CHARMM force field parameters by using simulated annealing search algorithm. LeDock shows the very high accuracy in pose prediction and is free for the purpose of academic use. rDock is an open source molecular docking program which can be used to dock ligands into the active sites of proteins and nucleic acids\(^5\). AutoDock Vina and AutoDock are two free academic programs for molecular docking. AutoDock can use the flexibility algorithm to dock the ligands into the proteins by Lamarckian genetic algorithm\(^5\),. AutoDock Vina, which is considered as the new generation of AutoDock, has faster run speed and more accurate binding mode predictions than AutoDock\(^5\). UCSF DOCK\(^5\) is the first molecular docking program which contains rigid and flexible ligand docking based on the geometric algorithms\(^5\). UCSF DOCK can be used to screen the small molecules subset of ZINC database directly. The LigandFit\(^6\), Glide\(^6\)-\(^9\), GOLD\(^8\), MOE Dock and Surflex-Dock\(^7\) use the commercial licenses to service for drug virtual screening\(^71\)-\(^76\). They show the powerful ability to screen the drugs from small molecule database through checking their citations on Google scholar. Moreover, wang et al. systematically summarize the advantage on the accuracies of binding pose and binding affinity of molecular docking software by comparing with five free academic license programs (LeDock, rDock, AutoDock, AutoDock Vina, and UCSF DOCK) and five commercial license programs (GOLD, LigandFit, MOE Dock, Glide, and Surflex-Dock)\(^7\). Furthermore, *de novo* drug design, which is another computer-aided method for drug discovery, can create the new ligands. LigBuilder\(^7\) and MOE Fragment-Based Design (www.chemcomp.com) are the representative *de novo* drug design software based on fragment linking and growing in the active pocket of targets. The computer-aided software has successfully applied into virtual screening studies and accelerated the process of drug discovery\(^79\)-\(^81\). Besides, although the ligands can locate at the active pocket of receptor very well, they may not become the medicine due to the poor ADME (absorption, distribution, metabolism, and excretion)\(^82\). ADME can be used to build the computer modeling for the prediction of structure-property relationships and reduce the trial failure of drugs in the clinical phases\(^83\). The ADME of drugs can be predicted based on the supported molecular format files by Schrödinger or Discovery Studio software.

Molecular docking can perform the high-throughput screening on the huge of small molecule database, while molecular dynamics (MD) simulation is the low-throughput method to evaluate ligand binding pathways\(^84\). The software Amber\(^85\), Gromacs\(^86\), NAMD\(^87\), and CHARMM\(^88\) are four popular molecular dynamics package mainly designed for simulations of lipids, nucleic acids and proteins (see Table 3). MD software can be used to study the dynamical interaction between targets and ligands at the atomic level. It can profile more detail and accurate interaction information for
### Table 3. Software of molecular docking and molecular dynamics simulations

<table>
<thead>
<tr>
<th>Software</th>
<th>Description</th>
<th>Weblink</th>
</tr>
</thead>
<tbody>
<tr>
<td>LeDock</td>
<td>Protein-ligand docking</td>
<td><a href="http://www.lephar.com/software.htm">www.lephar.com/software.htm</a></td>
</tr>
<tr>
<td>rDock</td>
<td>Ligands against proteins and nucleic acids</td>
<td>rdock.sourceforge.net</td>
</tr>
<tr>
<td>AutoDock</td>
<td>Protein-ligand docking</td>
<td>autodock.scripps.edu</td>
</tr>
<tr>
<td>AutoDock Vina</td>
<td>Protein-ligand docking</td>
<td>vina.scripps.edu</td>
</tr>
<tr>
<td>UCSF DOCK</td>
<td>Protein-ligand docking</td>
<td>dock.compbio.ucsf.edu</td>
</tr>
<tr>
<td>LigandFit</td>
<td>Protein-ligand docking</td>
<td>accelrys.com</td>
</tr>
<tr>
<td>Glide</td>
<td>Protein-ligand docking</td>
<td><a href="http://www.schrodinger.com">www.schrodinger.com</a></td>
</tr>
<tr>
<td>GOLD</td>
<td>Protein-ligand docking</td>
<td><a href="http://www.ccdc.cam.ac.uk/solutions/csddiscovery/components/gold">www.ccdc.cam.ac.uk/solutions/csddiscovery/components/gold</a></td>
</tr>
<tr>
<td>MOE Dock</td>
<td>Protein-ligand docking</td>
<td><a href="http://www.chemcomp.com">www.chemcomp.com</a></td>
</tr>
<tr>
<td>Surflex-Dock</td>
<td>Protein-ligand docking</td>
<td><a href="http://www.jainlab.org">www.jainlab.org</a></td>
</tr>
<tr>
<td>Amber</td>
<td>Molecular dynamics simulations</td>
<td>ambermd.org</td>
</tr>
<tr>
<td>Gromacs</td>
<td>Molecular dynamics simulations</td>
<td><a href="http://www.gromacs.org">www.gromacs.org</a></td>
</tr>
<tr>
<td>NAMD</td>
<td>Molecular dynamics simulations</td>
<td><a href="http://www.ks.uiuc.edu/Research/namd">www.ks.uiuc.edu/Research/namd</a></td>
</tr>
<tr>
<td>CHARMM</td>
<td>Molecular dynamics simulations</td>
<td><a href="http://www.charmm.org">www.charmm.org</a></td>
</tr>
</tbody>
</table>

### Table 4. Biological assays measuring the binding between macromolecules

<table>
<thead>
<tr>
<th>Assays</th>
<th>Binding Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface Plasmon Resonance</td>
<td>Protein-protein, protein-DNA, protein-RNA, protein-drug, antibody-antigen, DNA-DNA</td>
</tr>
<tr>
<td>Enzyme-Linked Immunosorbent Assay</td>
<td>Antibody-antigen, protein-ligand</td>
</tr>
<tr>
<td>Isothermal Titration Calorimetry</td>
<td>Protein-protein, protein-drug, drug-DNA, protein-DNA, enzyme-substrate</td>
</tr>
<tr>
<td>Electrophoretic Mobility Shift Assay</td>
<td>Protein-DNA, protein-RNA</td>
</tr>
<tr>
<td>Thermal shift assay</td>
<td>Protein-drug, enzyme-substrate</td>
</tr>
<tr>
<td>Protein Fluorescence Quenching</td>
<td>Protein-drug</td>
</tr>
<tr>
<td>Differential Scanning Calorimetry</td>
<td>Protein-drug</td>
</tr>
<tr>
<td>Nuclear Magnetic Resonance</td>
<td>Protein-drug</td>
</tr>
<tr>
<td>Affinity Chromatography</td>
<td>Protein-drug</td>
</tr>
<tr>
<td>GST Pull-down</td>
<td>Protein-protein</td>
</tr>
<tr>
<td>Footprinting</td>
<td>Protein-DNA</td>
</tr>
<tr>
<td>Chromatin Immunoprecipitation</td>
<td>Protein-DNA</td>
</tr>
</tbody>
</table>
ligands in the pocket of receptor than molecular docking. Especially, the Poisson–Boltzmann or generalized Born and surface area continuum solvation (MM/PBSA and MM/GBSA) give the good approaches to compute the binding free energy between ligands and biological macromolecules. As the experiment reported, MM/GBSA shows the faster and better prediction of binding affinities between ligands and targets than MM/PBSA in the absence of metal. MD simulations have become the popular method to study the mechanism of activated, inactivated states and ligand interaction on different targets, such as GPCRs, ion channel receptors, etc.

In vitro Biochemical Assays for Compound Screening

One advantage of drug design using high throughput assays and computational tools is that it can largely reduce the use of animals in activity testing. Furthermore, in vitro experiments complemented with computational methods have been extensively used in early drug discovery to select compounds with more favorable ADME and toxicological profiles. Most commonly, drugs are organic small molecules produced through chemical synthesis, but biopolymer-based drugs (also known as biopharmaceuticals) produced through biological processes are becoming increasingly more common. In addition, mRNA-based gene silencing technologies may have therapeutic applications.

The predominant strategy used over the last decades consists, first of all, in clarifying the biochemical processes underlying a disease, then identifying an appropriate drug target and finally developing a suitable assay that allows the screening of chemical libraries for small molecules interfering with the target. Based on the various protein targets of diseases, multiple approaches including biological or in silico had been designed to screen new drugs towards the diseases treatment. These approaches are, however, very expensive and demand a great deal of background knowledge. For classical pharmacology, many chemical libraries of synthetic small molecules, natural products or extracts were screened in vitro or in vivo, such as intact cells, whole organisms or cell-free systems to identify substances that have a desirable therapeutic effect.

The early experimental process to approach drug discovery involves several well defined biochemical assays to screen those compounds that can interact or bind with certain binding partners, such as receptor/ligand binding analysis, enzyme-activity evaluation. In addition, techniques such as X-ray crystallography structural analysis, NMR, calorimetry, affinity chromatography, ELISA and protein mass spectrometry are common strategic tools in protein-binding studies and play an important role to enhance the structural basis of rational drug design. These techniques aim to detect the separation of compounds from the studied proteins and monitor changes in intrinsic parameters of the targets upon forming a complex with tested drugs. Table 4 shows the biological assays used to detect the binding event between two binding partners, such as protein-protein, protein-nucleic acid, protein-small molecules. In the next, the applications of isothermal titration calorimetry (ITC) and Surface Plasmon Resonance (SPR) are reviewed in detail.

The rapid development of science technology has prompted the emergence of several new approaches. Isothermal titration calorimetry (ITC) is one of the products of the rapid science technology development. It is most often used to investigate the binding of small molecules to larger macromolecules, such as proteins or DNA. In the measurement using ITC, several important parameters involving the binding process can be calculated, including binding affinity, enthalpy changes, and the binding stoichiometry. According to the obtained parameters, the final Gibbs energy changes and entropy changes can be specifically determined. As ITC gives not only the binding affinity, but also the thermodynamics of the binding interaction, it is typically used as a secondary screening technique in high throughput drug discovery to eliminate false positive hits after primary screening. Characterization of the binding thermodynamics allows further hit selection and lead optimization as ITC can provide insights into the structure-activity relationship (SAR) for ligand interaction with the target. Comparing to other techniques such as fluorescence assays and NMR for studying the complex formation, ITC does not need any fluorescent probes or radioactive tags for data analysis. In addition, proteins used in the measurement do not require chemical modification that is ease of use and cost. In spite of various advantages,
low throughput, low sensitivity, and large sample requirement are major concerns, which may hamper its application\textsuperscript{126-128}.

Another new technique is Surface Plasmon Resonance (SPR), which is the product of nano-science development. The emergence of SPR had greatly reduced the detection limit of biological analysis and it is widely used for the study of ligand binding interactions\textsuperscript{7, 129, 130}. SPR is label-free in that a label molecule is not required for detection of the analytes and capable of measuring real-time quantitative binding affinities and kinetics in sequential binding events\textsuperscript{131, 132}. Moreover, SPR is especially interesting as it can present kinetic information according to affinity data and can be used for thermodynamic studies\textsuperscript{133-136}. At the same time, SPR biosensor assays can be applied in a wide range of proteins, including membrane proteins, such as G-protein-coupled receptors (GPCRs)\textsuperscript{137-141}. Another application of SPR technology is early ADME (absorption, distribution, metabolism, and excretion) profile prediction for lead compounds in drug discovery trial\textsuperscript{142}. SPR have the merits of real-time measurement, label-free and widespread biomolecules, it is emerging as an essential tool for drug development and has been widely used as a primary screening methodology for drug discovery\textsuperscript{125, 143, 144}. During a SPR measurement, the information includes data about concentration of a binding partner in a mixture as well as kinetic rate constants (association, dissociation rate constants and the equilibrium dissociation constant) for the binding interactions\textsuperscript{145}. Thus, SPR provides insights into the efficacy, safety, duration of action, indication, and patient tolerability of a drug.

Typically, in a compound screening campaign, the selection of evaluation means is highly depended on properties of the target to be studied\textsuperscript{146}. In studies of prion diseases, researchers have proposed several methods to screen new prion inhibitors that would benefit prion-related patients. As prion protein is prone to convert from a dominantly soluble α-helix structure to β-rich insoluble pathogenic aggregates, efforts have been extensively made toward the aggregation dynamic process and discovery new compounds interrupting the aggregation\textsuperscript{147-150}. Thioflavin T (ThT) can specifically bind to β-rich aggregates accompanying a red shift of the fluorescence emission spectra. Hence, it is frequently used as a dye to monitor the aggregation of prion protein\textsuperscript{151-154}. Based on this knowledge, Li et al. evaluated the inhibitory effects of several compounds on prion aggregation using ThT as a detector\textsuperscript{155, 156}. Once the aggregation process is interrupted or interfered by a compound, the increase and intensity of ThT fluorescence are delayed or weakened obviously indicating the efficiency of tested compound. Moreover, by combining multiple approaches in drug screening, it would provide more and much accurate information of the drug-target interaction profile. For example, ITC combined with chromatography has been used to identify and isolate unknown target proteins such as receptors or cell/tissue lysates\textsuperscript{127, 157}. Overall, to conduct a successful drug discovery, the evaluation assays should be carefully selected according to different investigating targets.

\textit{In vitro Cell-based Approaches}

As a common method used in drug discovery, cell-free systems have various advantages, such as fast, microscale and high throughput and the screening results are accurate and stable\textsuperscript{146, 158, 159}. However, the molecular testing models are designed to specific target, which can only provide limited information about the target interactions. Moreover, the effects of certain drugs on an organism are complex and the interactions between two partners are involved in multiple levels that cannot be predicted using biochemical assays. Hence, single molecular high-throughput drug screening technology can no longer meet the needs of new drug discovery today. Then more biologically relevant cell-based screening assays have been developed and are widely used to predict responses of an organism to drugs\textsuperscript{160, 161}. In addition, the cell culture is selected as a model system to predict cellular toxicity, which plays an important role in drug discovery process\textsuperscript{162}.

Tumor cell lines are common cell models used in diseases investigation and drug discovery. Current cell-based models rely heavily on immortalized cell lines, usually derived from human tumors. These models have advantages, such as cost-effective scale up and well consistency. Table 5 displays common human-derived cell lines used in cancer-related investigations. For normal cell lines, they are usually used as controls and...
Additional toxicology evaluations. Additionally, these cell lines are amenable to genetic engineering, permitting gain and loss of function analysis. While these models demonstrate advantages, they offer limited biological relevance when compared to the intact organ and primary cell types. Currently, primary cells and tissue slices are the gold standards for drug discovery, as they exhibit greater resemblance to the organ of interest [163-165].

Human pluripotent stem cells hold great promise in research and medicine for their unique ability of self-renewing and differentiating to various cell lineages in the body. For different studies, human pluripotent stem cells can be controlled to differentiate to desired cell types to fulfill the investigation purpose [166-169]. As general cell lines used in drug screening present limited relevance to the organ of interest, stem cells provide exciting new models and bring new changes for drug discovery and development as well as drug toxicity testing to treat different human diseases [170]. As we know, it is still a great challenge for drug discovery to neurodegenerative diseases because the biological mechanisms are complex and poorly understood. The lack of models that accurately characterize these dysfunctions blocks further investigations. Fortunately, recent advances in stem cell technology offer researchers available tools to generate human neurons to develop disease resemble assays for small molecules screening. The emergence of adult tissues or cells derived induced pluripotent stem cells (iPSC), which bypass the need for embryos, promotes new investigations of stem cells. Bright et al. [171] developed a specific antibody BMS-986168 for the Tau fragment based on human-induced pluripotent stem cells from patients with sporadic Alzheimer’s disease (AD). In 2017, this antibody was licensed by Biogen and entered Phase II clinical trials for AD treatment. Retigabine, another drug under clinical Phase II trial was also derived from iPSC models generated from amyotrophic lateral sclerosis (ALS) patients [172]. In addition, the recent advances in the production of stem cell-derived hepatocytes and cardiomyocytes combined with cutting-edge engineering technologies supplement the application of stem cells as an attractive alternative model for current drug discovery, which will deliver safer and more efficacious medicines for the patient [173-175].

Moreover, the use of stem cell-derived in vitro systems could reduce animal use and facilitate mechanisms investigation of the toxicants at the same time in toxicity evaluations [170]. Advancements in pluripotent stem cells and 3D culturing techniques promote the creation of organoids that can accurately recapitulate the properties of various specific subregions of many human organs. Tumor organoids resemble the original tumors much
better than cell lines, having a 3D structure, a variety of cells, and similar growth characteristics, they can be applied as a pre-clinical cancer model for drug discovery\textsuperscript{176, 177}. In addition, organoids provide another opportunity to construct cellular models of human diseases that can be used to deeply study the causes of diseases and further identify possible treatment in laboratory\textsuperscript{178, 179}. We believe that advances in stem cell biology would provide more accurate human tissue and disease models for drug development.

In vivo Animal Models

Compared to in vitro screening, in vivo testing is better suited for observing the overall effects of an experiment on a living subject. As in vitro assays can sometimes yield misleading results with drug candidate molecules that are irrelevant in vivo, efficacy verification in vivo is especially crucial in drug discovery process. In addition, whole-organism in vivo screening holds several advantages to small molecule discovery for its target agonistic and holistic.

One of the most important and widely used model organisms in scientific research is zebrafish, which possesses numerous advantages and is the pioneer model for drug screening. It has been used in various research fields, such as gene expression and sequencing\textsuperscript{180, 181}, cancer models\textsuperscript{182, 183}, immune system\textsuperscript{184} and infectious diseases\textsuperscript{185, 186}. Ongoing research programs have promoted zebrafish model to develop novel therapeutic agents in drug discovery. Drug screens based on zebrafish can not only identify novel classes of compounds with biological effects, but also discover novels uses or targets of existing drugs\textsuperscript{187}. Using a zebrafish screen, a bioactivity from an extract of Jasminum gilgianum plant was discovered to induce the formation of ectopic tailbuds in larvae\textsuperscript{188}. Similarly, some new targets of old medications were identified from zebrafish chemical screening, such as the phenothiazine antipsychotics, which was demonstrated to be toxic to MYC overexpressing thymocytes\textsuperscript{189}. For the commercially available antiangiogenic statin rosuvastatin, a new function was discovered to suppress the growth of prostate cancer in a zebrafish screening of known bioactive compounds\textsuperscript{190}. Other new compound classes were also identified based on zebrafish screens, such as GS4012\textsuperscript{191} and lenaldekar\textsuperscript{192}.

Another vertebrate animal model frequently used in drug discovery are Xenopus frogs, which belong to the amphibians. As Xenopus frogs share a long evolutionary history with mammals, they are the excellent models to predict human biology. Xenopus have been extensively adopted as a convenient first-line animal model at various stages of drug discovery and development. Since the early 1980s, Xenopus embryogenesis has been much explored and a protocol termed frog embryo teratogenesis assay was applied to identify drugs that pose potential teratogenic hazards, including mortality and malformation\textsuperscript{193}. Embryos and tadpoles were severed as versatile animal models to investigate blood vascular development and angiogenesis. Subsequently, Roland et al.\textsuperscript{194} uncovered pathways involved in the development of the lymphatic and blood vascular system in Xenopus tadpoles and discovered new compounds and pathways that were not previously known to mediate lymphatic or vascular development. Due to large number and size of the eggs, the rapid development of the embryos and the fact they are amenable to pharmacological, surgical and genetic techniques, Xenopus laevis has been successfully used in searching for embryonic signaling pathways targeting compounds\textsuperscript{195}, particularly the Wnt/\(\beta\)-catenin pathway\textsuperscript{196, 197}.

Despite successful application in drug discovery, the zebrafish and Xenopus models are largely different from the mammals in various aspects including genetics, immune system, and metabolism. Hence, more advanced mammal models were developed to decrease the gap between animal and human diseases, such as dog, rabbit, rat, mouse and the non-human primate Rhesus and Orangutan Monkey. In a research to determine the isolates of pancreatic secretion could be used to treat dogs with diabetes on an animal model dog, promoted the discovery of insulin and then the use in diabetes treatment\textsuperscript{198}. For rabbits used as organism models, they are frequently used to produce antibodies in immunology\textsuperscript{199}. In addition, rabbits are important models to study cardiovascular disease\textsuperscript{200}. Guinea pigs are vertebrate models extensively used by early bacteriologists as hosts for bacterial infections and infectious diseases including viral and parasitic infections\textsuperscript{201}. The mouse is one of the classical model vertebrates and has become the popular choice for
developing various in vivo mammalian models as it shares about 85% genome identity to humans and has many physiological systems that are similar to those in humans. In addition, mouse has characteristics such as short life-cycle, techniques for genetic manipulation (inbred strains, stem cell lines, and methods of transformation) and non-specialist living requirements that are predominant and convenient to use. It is commonly used for scientific research in medicine, psychology and genetics. Comparing to the mouse, the rat has larger size of organs and suborganellar structures and it is particularly used as toxicology models and neurological models. While for the non-human primate Rhesus and Orangutan Monkeys, they are conventional animal models used in hepatitis, HIV, Parkinson's disease, cognition, and vaccines investigations. For different scientific research purposes, numerous animal models can be constructed. For example, human tumor xenograft, orthotopic/intratibial tumor models, murine tumor xenograft and patient-derived tumor grafts were built in nude mice or rats to investigate oncology development or screen new cancer drugs. Using the constructed models, development of many human tumors including colon cancer, breast cancer, lung cancer, prostate cancer, ovarian cancer, renal cancer, cervical cancer, pancreatic cancer and melanoma has been investigated at different degrees. At the same time, therapeutic compounds targeting these tumors were also discovered either from the existing drugs or from new synthetic compounds.

Despite numerous models developed for drug discovery, most therapeutic drugs still fail in clinical trials. One of the reasons is attributed to sufficient clinical predictive power of our current model systems. Despite the high genetic similarities between human and mice, physiological differences affect the course of diseases in mice models when some genetic disorders in human do not have the same symptoms in mice. In addition, the cell lines and xenografts commonly used are inadequate models that can not highly mimic and accurately predict human diseases. Generally, for a drug discovery research in the laboratory, in combination with the chemical assays, cell-based and in vivo testing would perform more efficiently to obtain effective lead compounds for further drug development. At last, new models are still needed to be developed for scientific researches in the future.

Conclusions
As above review, the entire drug design process can be profiled from in silico, in vitro and in vivo experiments (see Figure 1). Our review summarizes the functions of macromolecular structure databases, small molecule databases, molecular docking software, de novo drug design software, MD simulations software, biochemical assays, cell-based system and animal models. This review shows the detailed individual component of drug design, and gives the comprehensive understanding for the progress of drug design.

Acknowledgments
The work is supported by the National Natural Science Foundation of China (Grant No. 21605066) and Fundamental Research Funds for the Central Universities (Grant No. Izujbk-2018-92).

Additional information
Competing financial interests: The authors declare no competing financial interests.

References
resonance for biological and chemical sensing applications. Chem. Soc. Rev. 43, 3426-3452.


106. Acker, M. G. & Auld, D. S. (2014) Considerations for the design and reporting of enzyme assays in
high-throughput screening applications. Perspectives in Science 1, 56-73.


