Computational Methods in Drug Screening and Design

Subjects: Pharmacology & Pharmacy

Contributor: Xubo Lin

Drug development is one of the most significant processes in the pharmaceutical industry. Various computational methods have dramatically reduced the time and cost of drug discovery.

Keywords: multiscale models ; virtual screening ; de novo design ; machine learning

1. Introduction

With the rapid development of both computer hardware, software, and algorithms, drug screening and design have benefited much from various computational methods which greatly reduce the time and cost of drug development. In general, bioinformatics can help reveal the key genes from a massive amount of genomic data [1][2] and thus provide possible target proteins for drug screening and design. As a supplement to experiments, protein structure prediction methods can provide protein structures with reasonable precision [3]. Biomolecular simulations with multiscale models allow for investigations of both structural and thermodynamic features of target proteins on different levels [4], which are useful for identifying drug binding sites and elucidating drug action mechanisms. Virtual screening then searches chemical libraries to provide possible drug candidates based on drug binding sites on target proteins [5][6][7]. With greatly reduced amount of possible drug candidates, in-vitro cell experiments can further evaluate the efficacy of these molecules. In addition to virtual screening, de novo drug design methods [8], which generate synthesizable small molecules with high binding affinity, provide another type of computer-aided drug design direction.

2. Biomolecular Simulations in Drug Screening and Design

The Nobel Prize in Chemistry 2013 was awarded jointly to Martin Karplus, Michael Levitt, and Arieh Warshel for their pioneering contributions in the development of multiscale models for complex biochemical systems, recognizing the important role that theory and computational methods play as a direct and necessary complement to experiments [9]. Further, applications of biomolecular simulations in identifying drug binding sites and elucidating the molecular mechanisms of diseases have been subject to rapid developments [10][11][12].

Depending on the biological problems to be studied, multiscale models will be used in biomolecular simulations. The combination of quantum mechanics (QM) and molecular mechanics (MM) (QM/MM) can be used to study the electronic properties [13], simulate chemical reactions (e.g., the enzyme catalysis mechanism [14]), and calculate spectra [15] in a single simulation, which can be used to elucidate the action mechanism of certain drugs. Another widely used biomolecular method is molecular dynamics (MD) simulation [16][17][18][19][20][21], which applies empirical molecular mechanics (MM) force fields and is based on classical Newtonian mechanics. According to different accuracy requirements, all-atom (AA), united-atom (UA), and coarse-grained (CG) MD simulations as well as explicit/implicit solvent models, which allow simulations of temporal and spatial scales, can be used to facilitate the drug discovery [22][23]. Generally, MD simulations have been used for the identification of potential drug binding sites on target proteins, the calculation of binding free energy between target proteins and drug molecules, the action mechanism of drug molecules, etc. [24][25]. However, it is worth mentioning that MD simulations are also limited to time and length scales. Currently, AAMD simulations can explore time-dependent phenomena of large systems such as viral capsids in atomic detail as long as microseconds or even milliseconds [18][26][27]. Therefore, different types of simulation methods are required for different types of problems. Each of the different simulation methods has advantages, disadvantages, and practical limitations in terms of the size of system that can be simulated, the length of simulation that can be achieved, and the types of phenomena that can be modeled. On the other hand, improved algorithms, fast-growing data sets and computing ability are driving rapid advances in multiscale modeling methods which provide a powerful emerging paradigm for drug discovery.

As mentioned above, multiscale biomolecular simulations are useful in drug design. For example, MD simulations can be used to identify the drug binding sites on the target protein as several simple tools (e.g., fpocket [28]). Moreover, multiscale

biomolecular simulations can be used to reveal the action mechanism of a drug, which happens mainly on a molecular level, but clearly has macroscopic effects [29][30][31]. From the molecular structure level to cellular tissue, the dynamics of drug targets and the surprising complexity of biological systems challenge our scientific understanding. In order to reveal how changes at different levels are linked and interaction network, no single simulation method can solve all these problems involved. Multiscale simulation methods are designed to simulate and analyze cross-scale connections, for example, how one scale change leads to another scale change. An obvious challenge is the integration of data and simulation across length scales and time scales. The current multiscale approach has potential to overcome these limitations by directly combining different levels of descriptions, thus bringing new prospects for drug discovery.

Multiscale simulations play an important role in studying biological processes. Here, we show applications of a combination of different methods to deal with complicated biological processes in this field. In recent years, although MD simulations are widely used, they still cannot consider the change of ionization states in simulated process. For example, MD simulations cannot discuss time dependence in a long time-scale of biological proton transport process, so the time dependence of protein changes with pH also cannot be revealed [32]. Therefore, in calculations and simulation process, it is necessary to combine the Monte Carlo (MC) approach (performed time-dependent simulations) in the proton transport process with the MD approach (performed pH dependent) in a protein model [32]. According to the electrostatic energy of the CG model, which can obtain the free energy of protonation states. In addition, the movement of the MC approach is based on the electrostatic energy of the CG model, and then proton transport time is used to make the scale, so as to correspond to the rate constant predicted by the transition state theory. Hence, they used isomorphism between probabilities obtained from the MC process and probability factors obtained from transition state theory [33], and converted the MC process to a time-dependent simulation with additional simplified modifications [34].

3. Drug Design and Virtual Screening

The design, discovery, and development of drugs are complex processes involving many different fields of knowledge and are considered a time-consuming and laborious inter-disciplinary work [35][36][37][38]. Different drug design methods and virtual screening will be very useful to design and find rational drug molecules based on the target macromolecule that interacts with the drug and thus speed up the whole drug discovery process. Here, we will discuss structure-based drug design, ligand-based drug design, and virtual screening.

3.1. Structure-Based Drug Design

Structure-based drug design must be performed with available structural models of the target proteins, which are provided by X-ray diffraction, nuclear magnetic resonance (NMR) or molecular simulation (homologous protein modeling, etc.) [39][40][41][42][43]. Keeping in mind the complexity of cancers which show diverse phenotypes and multiple etiologies, a one-size-fits-all drug design strategy for the development of cancer chemotherapeutics does not yield successful results. Lately, Arjmand et al. [44] adopted a series of methods, such as the combination of X-ray crystal structures and molecular docking, to design, synthesize, and characterize novel chromone based-copper(II) antitumor inhibitors. In general, after obtaining the structure of the receptor macromolecule by x-crystal single-crystal diffraction technique or multi-dimensional NMR, molecular modeling software can be used to analyze the physicochemical properties of drug binding sites on the receptor, especially including electrostatic field, hydrophobic field, hydrogen bond, and key residues. Then, the small molecule database is searched, or the drug design technique is used to identify the suitable molecules whose molecular shapes match the binding sites of the receptor and binding affinity is high. Then, these molecules are synthesized and their biological activities will be tested for further drug development. In short, structure-based drug design plays an extremely important role in drug design.

3.2. Ligand-Based Drug Design

Unlike structure-based drug design, ligand-based drug design doesn’t search small molecule libraries. Instead, it relies on knowledge of known molecules binding to the target macromolecule of interest. Using these known molecules, a pharmacophore model that defines the minimum necessary structural characteristics a molecule must possess in order to bind to the target can be derived [45][46]. Then, this model can be further used to design new molecular entities that interact with the target. On the other hand, ligand-based drug design can also use quantitative structure–activity relationships (QSAR) [47][48] in which a correlation between calculated properties of molecules and their experimentally determined biological activity is derived, to predict the activity of new analogs. Both the pharmacophore model and QSAR model will be discussed in detail in the following sessions.

3.3. Virtual Screening

In recent years, the rapid development of computational resources and small molecule databases have led to major breakthroughs in the development of lead compounds. As the number of new drug targets increases exponentially, computational methods are increasingly being used to accelerate the drug discovery process. This has led to the increased use of computer-assisted drug design and chemical bioinformatics techniques such as high-throughput docking, homology search and pharmacophore search in databases for virtual screening (VS) technology [48]. Virtual screening is an important part of computer-aided drug design methods. It may be the cheapest way to identify potential lead compounds, and many successful cases have proven successful using this technology.

The primary technique for identifying new lead compounds in drug discovery is to physically screen large chemical libraries for biological targets. In experiments, high-throughput screening identifies active molecules by performing separate biochemical analysis of more than one million compounds. However, this technology involves significant costs and time. Therefore, a cheaper and more efficient calculation method came into being, namely, virtual high-throughput screening. The method has been widely used in the early development of new drug. The main purpose is to determine the novel active small molecule structure from the large compound libraries. It is consistent with the purpose of high-throughput screening. The difference is that virtual screening can save a lot of experimental costs by significantly reducing the number of compounds for the measurement of the pharmacological activity, while high-throughput screening needs to perform experiments with all compounds in the database. Here, we will discuss common methods of virtual screening.

4. Multiscale De Novo Drug Design toward Personalized Medicine

Computer-based de novo design methods of drug-like molecules are mainly for generating small molecule compounds with ideal physicochemical and pharmacological properties. In the past decades, fragment-based drug discovery had appeared as a novel concept that has proved a good prospect for improving lead optimization, in order to decrease the clinical attrition rates in drug design. It is an approach that uses small molecular fragments to deduce the biomolecular targets [49]. Fragment-based de novo design has obtained the long-term clinical success [50].

Despite the fact that modern drug discovery has made some successes in offering effective drugs, drug design has been affected by several factors, such as the tremendous chemical space for exploring drug molecules [51]. Further, as a large number of data increase in biological, chemical, and clinical medicine, it is obvious that the drug design should be solved with multiscale optimization methods, and concentrate on the data beyond molecular levels [52]. Thus, it is essential to discuss the function of multiscale models in drug discovery, and how they have predicted multiple biological properties in different biological targets.

5. Machine Learning Methods Accelerate Drug Development

In the process of drug discovery, machine intelligence methods have mostly been used in the above-mentioned computational methods over the past few decades [53]. With the booming era of “big” data, machine learning methods have developed into deep learning approaches, which are a more efficient way for drug designers to deal with important biological properties from large amount of compound databases.

6. Applications of Multiscale Methods in Drug Discovery

6.1. Molecular Dynamics of Cardiac Modelling

Multiscale modeling of the drugs in an excitable system is critical because experiments on a single system scale cannot reveal the underlying effects of multiple drug interactions. A computationally based approach to predict the emergency effects of drugs on excitatory rhythms may form an interactive technology-driven process for the drug and disease screening industry, research and development academia, and patient-oriented medical clinic. There are potentially far-reaching implications because millions of people affected by arrhythmia each year will benefit from improved risk stratification of drug-based interventions.

Much progress has been made in developing multiscale computational modeling and simulation approaches for predicting effects of cardiac ion channel blocking drugs. Structural modeling of ion channel interactions with drugs is a critical approach for current and future drug discovery efforts. Modeling of drug receptor sites within an ion channel structure can be useful to identify key drug-channel interaction sites. Drug interactions with cardiac ion channels have been modeled at the atomic scale in simulated docking and MD simulations, as well as at the level of channel function to simulate drug effects on channel behavior [29][54][55][56][57][58][59][60]. Structural modeling of drug-channel interactions at the atomic scale

may ultimately allow for the design of novel high-affinity and subtype selective drugs for specific targeting of receptors for cardiac and neurological disorders.

6.2. Cancer Modeling and Network Biology

The World Health Organization (WHO) stated that cancer remains one of the most dangerous diseases today. Considering that cancer is a multifactorial disease, there is increasing interest in multi-target compounds that can target multiple intracellular pathways. However, the study of large data sets for the analysis of anticancer compounds is difficult, with a large amount of data and high data complexity. For example, the ChEMBL database [61] compiles big datasets of very heterogeneous preclinical assays. Bediaga et al. [62] have reported a PTML-LDA model of the ChEMBL dataset for the preclinical determination of anticancer compounds. PTML is a model that combines perturbation theory (PT) ideas and ML methods to solve similar problems. They compared this model with other PTML models which was reported by Speck-Planche et al. [63][64][65][66] and then concluded that this is the only one that can predict activity against multiple cancers. Speck-Planche et al. also derived a multi-task (mtk) chemical information model combining Broto Moreau autocorrelation with ANN from a dataset containing 1933 peptide cases. This model is used to virtually design and screen peptides with potential anti-cancer activity against different cancer cell lines and low cytotoxicity to a variety of healthy mammalian cells, and the model shows greater than 92% in both training and prediction (test)accuracy.

In addition, due to the inherent complexity of tumors, it is necessary to analyze their growth at different scales. It includes many phenomena that occur at various spatial scales from tissue to molecular length. The complexity of cancer development is manifested in at least three scales that can be distinguished and described by mathematical models, namely microscale, mesoscale, and macroscale. Wang et al. conducted a number of studies on how to use multiscale models for the identification and combination therapy of drug targets [67][68][69][70][71][72][73]. This method is based on quantification of relationship between intracellular epidermal growth factor receptor (EGFR) signaling kinetics, lung cancer extracellular epidermal growth factor (EGF) stimulation and multicellular growth. The multiscale modeling of tumors combined with systemic pharmacology will contribute to the development of practical smart drugs. It will produce a comprehensive system-level approach to determine the dynamics and effects of existing and new drugs in preclinical trials, model organisms and individual patients. In addition, mathematical and computational studies will provide a better way to understand many factors that influence the effects of drugs, thus helping to uncover better ways to therapeutically interfere with disease.

6.3. Multiscale Modeling for Drug Discovery in Brain Disease

Multiscale models can be also used to identify pathophysiological processes to allow disease staging. In many cases, like cancer, treatments vary depending on the stage of disease. The model can help determine prognosis, which is an important clinical determination that can help determine the right type of medication to be administered or discovered. Several models focus on the neuronal network levels, including Cutsuridis and Moustafa for Alzheimer's disease [74], and Lytton for epilepsy [75]. ANN is a class of ML techniques that can be used for clinical analysis of big data including that related to drug testing, which is critical for drug discovery. In addition, Anastasio [76] introduced process algebra, a computer technology widely used to analyze complex computing systems, used here to calculate neurology. Sirci et al. [77] described how network (map) theory is used to identify similarities and differences between different pharmacological agents. In this type of study, each drug is a node, and the edges between drugs represent chemical and transcription-based interactions that characterize the drug.

In addition, Ferreira da Costa et al. [78] report the first PTML (PT + ML) study of a large number of ChEMBL datasets for preclinical determinations of compounds for dopamine pathway proteins. Molecular docking or ML models can be used to solve a specific protein, but these models cannot explain the large and complex large data sets of preclinical assays reported in public databases. PT models, on the other hand, allow us to predict the properties of a query compound or molecular system in an experimental analysis with multiple boundary conditions based on previously known reference cases. In their work, the best PTML model found in the training and external validation series has an accuracy of 70–91%. Hansch's model is a classic method for solving quantitative structural binding relationships (QSBR) in pharmacology and medicinal chemistry. Abeijon et al. [79] developed a new PT-QSBR Hansch model based on PT and QSBR methods for a large number of drugs reported in ChEMBL, focusing on a protein expressed in the hippocampus of the brain of Alzheimer's disease (AD) patients. Now, by decomposing how risks and causes are combined in complex systems to produce disease, and how to prevent or improve these diseases through multi-stage, multi-target, multi-drug techniques, multiscale modeling is gradually being grasped.

6.4. Infectious Diseases

From AIDS, hepatitis C, influenza, and other disease-related viruses to the current 2019-nCoV, we have been working hard to develop antiviral drugs targeting them. However, the unique structure and proliferation of the virus pose a natural challenge for drug development. Viruses do not have their own cellular structure and metabolic system, and must replicate and proliferate in host cells. Therefore, it is difficult to find compounds that target only viral targets without affecting the normal function of host cells. At present, the main way that some antiviral drugs work is to inhibit viral replication. However, many of the tools used for virus replication come from human cells, such as ribosomes, and the corresponding antiviral drugs will also bring great side effects to the human body. Therefore, the discovery of drugs requires the introduction of a multi-scale model to screen out drugs that can inhibit viral replication while reducing the damage to the human body.

So far, retroviral infections, such as HIV, are incurable diseases. ChEMBL manages big data capabilities through complex datasets, which make the information difficult to analyze because these datasets describe numerous features for predicting new drugs for retroviral infections. Without proper model, it is impossible to make full use of these features. Hence, Vásquez-Domínguez et al. [80] proposed a PTML model for the ChEMBL dataset, which can be efficiently used for preclinical experimental analysis of antiretroviral compounds. The PT operator is based on a multi-conditional moving average, which combines different functions and simplifies the management of all data. The PTML model they proposed was the first to consider multiple features combined with preclinical experimental antiretroviral tests. In order to simultaneously explore antibacterial activity against Gram-negative pathogens and in vitro safety related to absorption, distribution, metabolism, elimination, and toxicity (ADMET), the Speck-Planchee et al. [81] further proposed the first mtk-QSBER model. The accuracy of this model in both the training and prediction (test) sets is higher than 97%. They also have developed a chemoinformatic model for simultaneous prediction of anti-cocci activities and in vitro safety [82]. The best model displayed accuracies around 93% in both training and prediction (test) sets. Additionally, focusing on exploring anti-hepatitis C virus (HCV), the accuracy shown in the training and prediction (test) sets is higher than 95% using this model [83]. Cytotoxicity is one of the main concerns in the early development of peptide-based drugs. Kleandrova et al. [84] introduce the first multi-task processing (mtk) computational model focused on predicting both antibacterial activity and peptide cytotoxicity. Gonzalez-Diaz et al. [85] developed a model called LNN-ALMA to generate complex networks of the AIDS prevalence with respect to the preclinical activity of anti-HIV drugs.

Multiscale models are also imperfect and have their limitations. Models are expressions and simplifications of real life. No model can represent everything that can happen in the system. All models contain specific assumptions, and models vary widely in their comprehensiveness, quality, and utility. In other words, each model can only solve limited problems. Hence, we need to integrate different computational models and data in order to make full use of these models.

References

1. Yamanishi, Y.; Araki, M.; Gutteridge, A.; Honda, W.; Kanehisa, M. Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. Bioinformatics 2008, 24, i232–i240.
2. Bakheet, T.M.; Doig, A.J. Properties and identification of human protein drug targets. Bioinformatics 2009, 25, 451–457.
3. Moult, J.; Fidelis, K.; Kryshtafovych, A.; Schwede, T.; Tramontano, A. Critical assessment of methods of protein structure prediction (CASP)—Round XII. Proteins Struct. Funct. Bioinf. 2018, 86, 7–15.
4. Ayton, G.S.; Noid, W.G.; Voth, G.A. Multiscale modeling of biomolecular systems: In serial and in parallel. Curr. Opin. Struct. Biol. 2007, 17, 192–198.
5. Shoichet, B.K. Virtual screening of chemical libraries. Nature 2004, 432, 862–865.
6. Forli, S.; Huey, R.; Pique, M.E.; Sanner, M.F.; Goodsell, D.S.; Olson, A.J. Computational protein–ligand docking and virtual drug screening with the AutoDock suite. Nat. Protoc. 2016, 11, 905–919.
7. Rosales, A.R.; Wahlers, J.; Limé, E.; Meadows, R.E.; Leslie, K.W.; Savin, R.; Bell, F.; Hansen, E.; Helquist, P.; Munday, R.H. Rapid virtual screening of enantioselective catalysts using CatVS. Nat. Catal. 2019, 2, 41.
8. Schneider, G.; Clark, D.E. Automated de novo drug design: Are we nearly there yet? Angew. Chem. 2019, 131, 10906–10917.
9. Karplus, M. Development of multiscale models for complex chemical systems: From H+ H2 to biomolecules (Nobel lecture). Angew. Chem. 2014, 53, 9992–10005.
10. Jorgensen, W.L. The many roles of computation in drug discovery. Science 2004, 303, 1813–1818.
11. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of molecular dynamics and related methods in drug discovery. J. Med. Chem. 2016, 59, 4035–4061.

12. Abel, R.; Wang, L.; Harder, E.D.; Berne, B.; Friesner, R.A. Advancing drug discovery through enhanced free energy calculations. Acc. Chem. Res. 2017, 50, 1625–1632.
13. Fan, J.; Lin, L.; Wang, C.-K. Excited state properties of non-doped thermally activated delayed fluorescence emitters with aggregation-induced emission: A QM/MM study. J. Mater. Chem. C 2017, 5, 8390–8399.
14. Jindal, G.; Warshel, A. Exploring the Dependence of QM/MM Calculations of Enzyme Catalysis on the Size of the QM Region. J. Phys. Chem. B 2016, 120, 9913–9921.
15. Morzan, U.N.; Alonso de Armino, D.J.; Foglia, N.O.; Ramirez, F.; Gonzalez Lebrero, M.C.; Scherlis, D.A.; Estrin, D.A. Spectroscopy in complex environments from QM–MM simulations. Chem. Rev. 2018, 118, 4071–4113.
16. Buchete, N.-V.; Hummer, G. Peptide folding kinetics from replica exchange molecular dynamics. Phys. Rev. E 2008, 77, 030902.
17. Liu, Y.; Strümpfer, J.; Freddolino, P.L.; Gruebele, M.; Schulten, K. Structural characterization of λ-repressor folding from all-atom molecular dynamics simulations. J. Phys. Chem. Lett. 2012, 3, 1117–1123.
18. Sothiselvam, S.; Liu, B.; Han, W.; Ramu, H.; Klepacki, D.; Atkinson, G.C.; Brauer, A.; Remm, M.; Tenson, T.; Schulten, K. Macrolide antibiotics allosterically predispose the ribosome for translation arrest. Proc. Natl. Acad. Sci. USA 2014, 111, 9804–9809.
19. Hernández-Rodríguez, M.; C Rosales-Hernández, M.; E Mendieta-Wejebe, J.; Martínez-Archundia, M.; Correa Basurto, J. Current tools and methods in Molecular Dynamics (MD) simulations for drug design. Curr. Med. Chem. 2016, 23, 3909–3924.
20. Takada, S.; Kanada, R.; Tan, C.; Terakawa, T.; Li, W.; Kenzaki, H. Modeling structural dynamics of biomolecular complexes by coarse-grained molecular simulations. Acc. Chem. Res. 2015, 48, 3026–3035.
21. Mortier, J.; Rakers, C.; Bermudez, M.; Murgueitio, M.S.; Riniker, S.; Wolber, G. The impact of molecular dynamics on drug design: Applications for the characterization of ligand–macromolecule complexes. Drug Discov. Today Technol. 2015, 20, 686–702.
22. Durrant, J.D.; McCammon, J.A. Molecular dynamics simulations and drug discovery. BMC Biol. 2011, 9, 71.
23. Borhani, D.W.; Shaw, D.E. The future of molecular dynamics simulations in drug discovery. J. Comput. Aided Mol. Des. 2012, 26, 15–26.
24. Wang, Y.; Lupala, C.S.; Liu, H.; Lin, X. Identification of Drug Binding Sites and Action Mechanisms with Molecular Dynamics Simulations. Curr. Top. Med. Chem. 2018, 18, 2268–2277.
25. Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J. Chem. Inf. Model. 2011, 51, 69–82.
26. Perilla, J.R.; Schulten, K. Physical properties of the HIV-1 capsid from all-atom molecular dynamics simulations. Nat. Commun. 2017, 8, 1–10.
27. Yu, I.; Mori, T.; Ando, T.; Harada, R.; Jung, J.; Sugita, Y.; Feig, M. Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. eLife 2016, 5, e19274.
28. Schmidtke, P.; Le Guilloux, V.; Maupetit, J.; Tuffíry, P. Fpocket: Online tools for protein ensemble pocket detection and tracking. Nucleic Acids Res. 2010, 38, W582–W589.
29. Clancy, C.E.; An, G.; Cannon, W.R.; Liu, Y.; May, E.E.; Ortoleva, P.; Popel, A.S.; Sluka, J.P.; Su, J.; Vicini, P. Multiscale modeling in the clinic: Drug design and development. Ann. BioMed. Eng. 2016, 44, 2591–2610.
30. Amaro, R.E.; Mulholland, A.J. Multiscale methods in drug design bridge chemical and biological complexity in the search for cures. Nat. Rev. Chem. 2018, 2, 0148.
31. Speck-Planche, A. Recent advances in fragment-based computational drug design: Tackling simultaneous targets/biological effects. Future Med. Chem. 2018, 10, 2021–2024.
32. Olsson, M.H.; Warshel, A. Monte Carlo simulations of proton pumps: On the working principles of the biological valve that controls proton pumping in cytochrome c oxidase. Proc. Natl. Acad. Sci. USA 2006, 103, 6500–6505.
33. Messer, B.M.; Roca, M.; Chu, Z.T.; Vicatos, S.; Kilshtain, A.V.; Warshel, A. Multiscale simulations of protein landscapes: Using coarse-grained models as reference potentials to full explicit models. Proteins Struct. Funct. Bioinf. 2010, 78, 1212–1227.
34. Braun-Sand, S.; Burykin, A.; Chu, Z.T.; Warshel, A. Realistic simulations of proton transport along the gramicidin channel: Demonstrating the importance of solvation effects. J. Phys. Chem. B 2005, 109, 583–592.
35. Veerareddy, P.R. Diverse Strategies in Drug Discovery and Development. EC Pharm. Toxicol. 2018, 6, 601–603.

36. EFPIA, M.; Marshall, S.; Burghaus, R.; Cosson, V.; Cheung, S.; Chenel, M.; Dellapasqua, O.; Frey, N.; Hamrén, B.; Harnisch, L. Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation. CPT Pharm. Syst. Pharmacol. 2016, 5, 93–122.
37. Wang, T.; Wu, M.-B.; Zhang, R.-H.; Chen, Z.-J.; Hua, C.; Lin, J.-P.; Yang, L.-R. Advances in computational structure-based drug design and application in drug discovery. Curr. Top. Med. Chem. 2016, 16, 901–916.
38. Basith, S.; Cui, M.; Macalino, S.J.; Choi, S. Expediting the design, discovery and development of anticancer drugs using computational approaches. Curr. Med. Chem. 2017, 24, 4753–4778.
39. Bhuvaneshwari, S.; Sankaranarayanan, K. Identification of potential CRAC channel inhibitors: Pharmacophore mapping, 3D-QSAR modelling, and molecular docking approach. SAR QSAR Environ. Res. 2019, 30, 81–108.
40. Levoin, N.; Calmels, T.; Krief, S.; Danvy, D.; Berrebi-Bertrand, I.; Lecomte, J.-M.; Schwartz, J.-C.; Capet, M. Homology model versus x-ray structure in receptor-based drug design: A retrospective analysis with the dopamine D3 receptor. ACS Med. Chem. Lett. 2011, 2, 293–297.
41. Jacobson, K.A.; Costanzi, S. New insights for drug design from the X-ray crystallographic structures of G-protein-coupled receptors. Mol. Pharmacol. 2012, 82, 361–371.
42. He, G.; Gong, B.; Li, J.; Song, Y.; Li, S.; Lu, X. An improved receptor-based pharmacophore generation algorithm guided by atomic chemical characteristics and hybridization types. Front. Pharmacol. 2018, 9, 1463.
43. Yang, H.; Du Bois, D.R.; Ziller, J.W.; Nowick, J.S. X-ray crystallographic structure of a teixobactin analogue reveals key interactions of the teixobactin pharmacophore. Chem. Commun. 2017, 53, 2772–2775.
44. Arjmand, F.; Afsan, Z.; Roisnel, T. Design, synthesis and characterization of novel chromone based-copper (ii) antitumor agents with N, N-donor ligands: Comparative DNA/RNA binding profile and cytotoxicity. RSC Adv. 2018, 8, 37375–37390.
45. Yang, S.-Y. Pharmacophore modeling and applications in drug discovery: Challenges and recent advances. Drug Discov. Today 2010, 15, 444–450.
46. Kist, R.; Timmers, L.F.S.M.; Caceres, R.A. Searching for potential mTOR inhibitors: Ligand-based drug design, docking and molecular dynamics studies of rapamycin binding site. J. Mol. Graph. Model. 2018, 80, 251–263.
47. Tropsha, A. Best practices for QSAR model development, validation, and exploitation. Mol. Inf. 2010, 29, 476–488.
48. Vucicevic, J.; Nikolic, K.; Mitchell, J.B. Rational drug design of antineoplastic agents using 3D-QSAR, cheminformatic, and virtual screening approaches. Curr. Med. Chem. 2019, 26, 3874–3889.
49. Reker, D.; Perna, A.M.; Rodrigues, T.; Schneider, P.; Reutlinger, M.; Mönch, B.; Koeberle, A.; Lamers, C.; Gabler, M.; Steinmetz, H. Revealing the macromolecular targets of complex natural products. Nat. Chem. 2014, 6, 1072.
50. Reutlinger, M.; Rodrigues, T.; Schneider, P.; Schneider, G. Multi-objective molecular de novo design by adaptive fragment prioritization. Angew. Chem. Int. Ed. 2014, 53, 4244–4248.
51. Röper, S.; Kolb, H.C.; Jahnke, W.; Erlanson, D. Click chemistry for drug discovery. Rrag. Bas. Appr. Drug Discov. 2006, 34, 313–339.
52. Sanz, F.; Pognan, F.; Steger-Hartmann, T.; Díaz, C.; Cases, M.; Pastor, M.; Marc, P.; Wichard, J.; Briggs, K.; Watson, D.K. Legacy data sharing to improve drug safety assessment: The eTOX project. Nat. Rev. Drug Discov. 2017, 16, 811.
53. Wang, T.; Wu, M.-B.; Lin, J.-P.; Yang, L.-R. Quantitative structure–activity relationship: Promising advances in drug discovery platforms. Expert Opin. Drug Discov. 2015, 10, 1283–1300.
54. Kernik, D.C.; Morotti, S.; Wu, H.; Garg, P.; Duff, H.J.; Kurokawa, J.; Jalife, J.; Wu, J.C.; Grandi, E.; Clancy, C.E. A computational model of induced pluripotent stem-cell derived cardiomyocytes incorporating experimental variability from multiple data sources. J. Physol. 2019, 597, 4533–4564.
55. Zhang, Y.; Barocas, V.H.; Berceli, S.A.; Clancy, C.E.; Eckmann, D.M.; Garbey, M.; Kassab, G.S.; Lochner, D.R.; McCulloch, A.D.; Tran-Son-Tay, R. Multi-scale modeling of the cardiovascular system: Disease development, progression, and clinical intervention. Ann. BioMed. Eng. 2016, 44, 2642–2660.
56. Malisi, C.; Schumann, M.; Toussaint, N.C.; Kageyama, J.; Kohlbacher, O.; Höcker, B. Binding pocket optimization by computational protein design. PLoS ONE 2012, 7, e52505.
57. Clegg, L.E.; Mac Gabhann, F. Molecular mechanism matters: Benefits of mechanistic computational models for drug development. Pharmacol. Res. 2015, 99, 149–154.
58. Moreno, J.D.; Lewis, T.J.; Clancy, C.E. Parameterization for in-silico modeling of ion channel interactions with drugs. PLoS ONE 2016, 11, e0150761.

59. Moreno, J.D.; Zhu, W.; Mangold, K.; Chung, W.; Silva, J.R. A molecularly detailed Nav1. 5 model reveals a new class I antiarrhythmic target. JACC Basic Trans. Sci. 2019, 4, 736–751.

60. Moreno, J.D.; Zhu, Z.I.; Yang, P.-C.; Bankston, J.R.; Jeng, M.-T.; Kang, C.; Wang, L.; Bayer, J.D.; Christini, D.J.; Trayanova, N.A. A computational model to predict the effects of class I anti-arrhythmic drugs on ventricular rhythms. Sci. Transl. Med. 2011, 3, 98ra83.

61. Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A.P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L.J.; Cibrián-Uhalte, E. The ChEMBL database in 2017. Nucleic Acids Res. 2017, 45, D945–D954.

62. Bediaga, H.; Arrasate, S.; Gonzalez-Diaz, H. PTML combinatorial model of ChEMBL compounds assays for multiple types of cancer. ACS Comb. Sci. 2018, 20, 621–632.

63. Speck-Planche, A.; Cordeiro, M.N.D. Fragment-based in silico modeling of multi-target inhibitors against breast cancer-related proteins. Mol. Divers. 2017, 21, 511–523.

64. Speck-Planche, A.; Kleandrova, V.V.; Luan, F.; Cordeiro, M.N.D. Chemoinformatics in anti-cancer chemotherapy: Multi-target QSAR model for the in silico discovery of anti-breast cancer agents. Eur. J. Pharm. Sci. 2012, 47, 273–279.

65. Speck-Planche, A.; Kleandrova, V.V.; Luan, F.; Cordeiro, M.N.D. Rational drug design for anti-cancer chemotherapy: Multi-target QSAR models for the in silico discovery of anti-colorectal cancer agents. Bioorg. Med. Chem. 2012, 20, 4848–4855.

66. Planche, A.S.; Kleandrova, V.V.; Luan, F.; Cordeiro, M. Unified multi-target approach for the rational in silico design of anti-bladder cancer agents. Anti-Cancer Agent. Med. Chem. 2013, 13, 791–800.

67. Butner, J.D.; Cristini, V.; Wang, Z. Multiscale Modeling of Ductal Carcinoma In Situ. Biophys. J. 2019, 116, 322–323.

68. Butner, J.D.; Fuentes, D.; Ozpolat, B.; Calin, G.A.; Zhou, X.; Lowengrub, J.; Cristini, V.; Wang, Z. A multiscale agent-based model of ductal carcinoma in situ. IEEE Trans. BioMed. Eng. 2019.

69. Dogra, P.; Butner, J.D.; Chuang, Y.-l.; Caserta, S.; Goel, S.; Brinker, C.J.; Cristini, V.; Wang, Z. Mathematical modeling in cancer nanomedicine: A review. BioMed. Microdevices 2019, 21, 40.

70. Wang, Z.; Deisboeck, T.S. Mathematical modeling in cancer drug discovery. Drug Discov. Today 2014, 19, 145–150.

71. Karolak, A.; Rejniak, K.A. Mathematical modeling of tumor organoids: Toward personalized medicine. In Tumor Organoids; Springer: Berlin/Heidelberg, Germany, 2018; pp. 193–213.

72. Karolak, A.; Rejniak, K.A. Micropharmacology: An in silico approach for assessing drug efficacy within a tumor tissue. Bull. Math. Biol. 2019, 81, 3623–3641.

73. Kim, M.; Gillies, R.J.; Rejniak, K.A. Current advances in mathematical modeling of anti-cancer drug penetration into tumor tissues. Front. Oncol. 2013, 3, 278.

74. Cutsuridis, V.; Moustafa, A.A. Multiscale models of pharmacological, immunological and neurostimulation treatments in Alzheimer's disease. Drug Discov. Today Dis. Models 2016, 19, 85–91.

75. Lytton, W.W. Computer modeling of epilepsy: Opportunities for drug discovery. Drug Discov. Today Dis. Models 2016, 19, 27–30.

76. Anastasio, T.J. Modeling neurological disease processes using process algebra. Drug Discov. Today Dis. Models 2016, 19, 43–49.

77. Sirci, F.; Napolitano, F.; di Bernardo, D. Computational Drug Networks: A computational approach to elucidate drug mode of action and to facilitate drug repositioning for neurodegenerative diseases. Drug Discov. Today Dis. Models 2016, 19, 11–17.

78. Ferreira da Costa, J.; Silva, D.; Caamaño, O.; Brea, J.M.; Loza, M.I.; Munteanu, C.R.; Pazos, A.; García-Mera, X.; González-Díaz, H. Perturbation theory/machine learning model of ChEMBL data for dopamine targets: Docking, synthesis, and assay of new l-prolyl-l-leucyl-glycinamide peptidomimetics. ACS Chem. Neurosci. 2018, 9, 2572–2587.

79. Abeijon, P.; Garcia-Mera, X.; Caamano, O.; Yanez, M.; Lopez-Castro, E.; J Romero-Duran, F.; Gonzalez-Diaz, H. Multi-target mining of Alzheimer disease proteome with Hansch's QSBR-perturbation theory and experimental-theoretic study of new thiophene isosters of rasagiline. Curr. Drug Targets 2017, 18, 511–521.

80. Vásquez-Domínguez, E.; Armijos-Jaramillo, V.D.; Tejera, E.; Gonzalez-Diaz, H. Multioutput Perturbation-Theory Machine Learning (PTML) Model of ChEMBL Data for Antiretroviral Compounds. Mol. Pharm. 2019, 16, 4200–4212.

81. Speck-Planche, A.; Cordeiro, M.N.D. De novo computational design of compounds virtually displaying potent antibacterial activity and desirable in vitro ADMET profiles. Med. Chem. Res. 2017, 26, 2345–2356.

82. Speck-Planche, A.; Cordeiro, M.N.D.S. Chemoinformatics for medicinal chemistry: In silico model to enable the discovery of potent and safer anti-cocci agents. Future Med. Chem. 2014, 6, 2013–2028.

83. Speck-Planche, A.; Dias Soeiro Cordeiro, M.N. Speeding up early drug discovery in antiviral research: A fragment-based in silico approach for the design of virtual anti-hepatitis C leads. ACS Comb. Sci. 2017, 19, 501–512.
84. Kleandrova, V.V.; Ruso, J.M.; Speck-Planche, A.; Dias Soeiro Cordeiro, M.N. Enabling the discovery and virtual screening of potent and safe antimicrobial peptides. simultaneous prediction of antibacterial activity and cytotoxicity. ACS Comb. Sci. 2016, 18, 490–498.
85. González-Díaz, H.; Herrera-Ibatá, D.M.; Duardo-Sánchez, A.; Munteanu, C.R.; Orbegozo-Medina, R.A.; Pazos, A. ANN multiscale model of anti-HIV drugs activity vs. AIDS prevalence in the US at county level based on information indices of molecular graphs and social networks. J. Chem. Inf. Model. 2014, 54, 744–755.