Math.Comput.Sci. (2021) 15:877–888
https://doi.org/10.1007/s11786-021-00517-0

Mathematics in Computer Science

Mathematical Multidimensional Modelling and Structural Artificial Intelligence Pipelines Provide Insights for the Designing of Highly Specific AntiSARS-CoV2 Agents

Dimitrios Vlachakis · Panayiotis Vlamos

Received: 15 January 2020 / Revised: 20 December 2020 / Accepted: 26 January 2021 / Published online: 16 June 2021

Abstract COVID19 is the most impactful pandemic of recent times worldwide. It is a highly infectious disease that is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2 virus), To date there is specific drug nor vaccination against COVID19. Therefor the need for novel and pioneering anti-COVID19 is of paramount importance. In this direction, computer-aided drug design constitutes a very promising antiviral approach for the discovery and analysis of drugs and molecules with biological activity against SARS-CoV2. *In silico* modelling takes advantage of the massive amounts of biological and chemical data available on the nature of the interactions between the targeted systems and molecules, as well as the rapid progress of computational tools and software. Herein, we describe the potential of the merging of mathematical modelling, artificial intelligence and learning techniques into seamless computational pipelines for the rapid and efficient discovery and design of potent anti-SARS-CoV-2 modulators.

Keywords SARS-CoV-2 · COVID19 · Mathematical modelling · Artificial intelligence · Drug design · Structural bioinformatics

D. Vlachakis
Laboratory of Genetics, Department of Biotechnology, Genetics and Computational Biology Group, School of Applied Biology and Biotechnology, Agricultural University of Athens, Iera Odos 75 Str. GR11855, Athens, Greece

D. Vlachakis
Laboratory of Molecular Endocrinology, Division of Endocrinology and Metabolism, Center of Clinical, Experimental Surgery and Translational Research, Biomedical Research Foundation of the Academy of Athens, Soranou Ephessiou Str. GR11527, Athens, Greece

D. Vlachakis
University Research Institute of Maternal and Child Health and Precision Medicine, Medical School, National and Kapodistrian University of Athens, Thivon 1 & Papadiamantopoulou Str. GR11527, Athens, Greece
e-mail: dimvl@aua.gr
URL: http://darkdna.gr

P. Vlamos (✉)
Department of Informatics, Ionian University, Plateia Tsirigoti 7, 49100 Corfu, Greece
e-mail: vlamos@ionio.gr
URL: http://www.ionio.gr/~vlamos

1 Introduction

Nowadays, we are going through a pandemic period, which can be considered unprecedented for the modern era. On December 31, 2019, 27 cases of pneumonia with unknown etiology were recorded in Wuhan, China. About a month later, the infection by a new human virus, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), was established. The disease was named COVID-19 by the World Health Organization. As of May 14, 4,248,389 cases and 292,046 confirmed deaths have been reported worldwide [80,81].

Coronaviruses are a family of viruses that are enveloped positive-sense, single-stranded RNA, and their diameter ranges from 60–140 nm. On their surface, they contain spike-like protrusions, making them appear like a crown under the microscope. Four more human coronaviruses have been recorded, HKU1, NL63, 229E, and OC43, which cause mild respiratory disease [64].

It is worth noting that there have been two other cases of human infection with animal beta-coronavirus. The first incident concerns a β genera coronavirus in China, which originates from bats and was transmitted to humans through the palm civet cats, which is its host. This virus caused severe acute respiratory syndrome and had a mortality rate of about 11% before it was reduced [11]. The second case concerns the Middle East respiratory syndrome coronavirus (MERS-CoV), which also originates from bats and had as intermediate host the camels in Saudi Arabia with a mortality rate of about 34% [80].

SARS-CoV-2 has all the characteristics that coronaviruses have, but more specifically, its genetic material has a size of 30 kb, encodes ten proteins, and phylogenetically is closer to viruses bat-SL-CoVZC45 and bat-SL-CoVZXC21 than SARS [49]. Viral infection is transmitted to all ages, mainly through droplets released during coughing or sneezing in symptomatic patients, but is also transmitted before symptoms appear and from asymptomatic patients [56]. The virus has the potential to remain viable on surfaces for several days in favorable environmental conditions. Infection can occur either through inhalation of infected droplets or through contact with infected surfaces. Recent studies have shown that the virus enters through the connection of the angiotensin receptor 2 (ACE2) of the respiratory mucosa and the viral surface glycoprotein Spike [17].

More specifically, coronaviruses are made up of four structural proteins, which include Spike (S), membrane (M), envelop (E), and nucleocapsid (N) [8]. Spike is a trimeric glycoprotein on the surface of the virus that is responsible for the diversity of coronaviruses and their hosts. It includes two functional subunits, S1, which is responsible for binding to the host cell receptor and S2, which is responsible for fusing the two membranes, viral and cellular. Through structural and functional analyzes, it was shown that the ACE2 binds to the spike protein of the virus [62]. ACE2 is expressed in lung cells, mainly epithelial cells, as well as in cells of the heart, ileum, nerves, and bladder [88]. After the spike connects to the host ACE2 receptor, the protease cleavage of the first protein is followed. In the literature, a successive model of two-stage protease cleavage has been proposed as a model, where cleavage is performed at S1 / S2 site for priming and a second cleavage at S'2 site, a position next to a fusion peptide in subunit S2 [6,45]. Afterwards, S1 and S2 are cleaved and remain non-covalently bonded as the distant S1 subunit contributes to the stabilization of S2, which is anchored to the membrane. This event is followed by a cleavage in the S'2 site, which probably, through irreversible, conformational changes, activates the spike for membrane fusion. The coronavirus spike is different from other viruses because it can be cleaved and activated by a number of different proteases SARS-CoV-2 is even more special than the other coronaviruses, as at S1 / S2 site there is a furin cleavage site (RPPA sequence). This omnipresent expression of furin is likely responsible for the increased pathogenesis of the virus [87].

So far, a wide variety of clinical features of COVID-19 have been recorded, with the two extreme cases being the asymptomatic condition on the one hand and the acute respiratory distress syndrome and dysfunction of many organs on the other. The most common clinical features are fever, cough, sore throat, breathlessness, headache, fatigue, and myalgia. Due to these common symptoms, it becomes difficult to distinguish it from other respiratory infections. The development of the infection initially in pneumonia followed by respiratory failure and ultimately leads to death, is associated with the extreme increase in inflammatory cytokines such as IL2, IL7, IL10, GCSF, IP10, MCP1, MIP1A, and TNFα [15].

2 Current Drugs, Drug Design and Artificial Intelligence for COVID19

Various laboratory techniques have been reported for diagnosis. The main technique is special molecular testing on throat swabs, nasopharyngeal smear, sputum and secretions of the lower respiratory tract. The main laboratory method of molecular testing is real-time fluorescence (RT-PCR) to detect the positive nucleic acid of the virus. In addition, full genome sequencing and phylogenetic analysis of the above samples can equally detect viral infection [2,68]. However, because in the early stages of the infection the molecular test is likely to be negative, even if the patient has contracted the virus, unusual computed tomography (CT) scans have been used to determine the infection. CT scans generally show infiltrates, ground glass opacities and sub segmental consolidation. Even in the case of asymptomatic patients, CT scans show abnormalities [26]. It is worth noting that the term "confirmed case" is a suspicious case, ie it presents various of the symptoms, with a positive molecular test [28].

At present, there are no specific drugs or vaccines against the virus. However, rapid efforts are being made to test the already broad-spectrum antiviral drugs on the market, such as nucleoside analogues and HIV protease inhibitors [36]. The main drugs used in the tests are oseltamivir, lopinavir, ritonavir, ganciclovir, which were administered to patients for 3–14 days [15]. Even the broad-spectrum antiviral drug remdesivir and chloroquine showed positive results against SARS-CoV-2 infection in vitro [78]. In addition, there are a number of other developing antiviral compounds, such as the candidate compound EIDD-2801 that has shown high therapeutic potential for seasonal and pandemic influenza infections and this represents another possible drug to consider for the treatment of SARS-CoV-2 infection [71]. Finally, several broad-spectrum antiviral therapeutic drugs are to be considered, which provide drug treatment options for COVID-19 infection, and include Lopinavir / Ritonavir, neuraminidase inhibitors, peptide (EK1), RNA synthesis inhibitors [55].

The process of discovering and developing drugs is quite expensive and time consuming. A drug development cycle can last about 14 years and costs about $ 1 billion. Despite rapid advances in chemistry and technology, the process of drug discovery and production remains quite risky as there is a high rate of failure and low efficiency. Thus, computer-aided drug design (CADD) is now used to discover drugs leading to both cost reduction and risk of failure. CADD reflects computational tools for storing, managing, analyzing, and modeling compounds. Therefore, it includes tools for designing compounds, evaluating potential lead candidates, and studying the chemical interactions and physicochemical characteristics of compounds. It is also important that these tools can be used in preclinical tests through which they will be optimized and reprogrammed into new pipelines [44].

The most widely used computational approaches can be divided into three broad categories, structure-based drug design (SBDD), link-based drug design (LBDD), and sequence-based design. SBDD methods are based on knowledge of the structure of the target molecule from crystal structures, NMR data and homology models, and include techniques such as molecular docking and de novo drug design [14]. In cases where knowledge of the structure is not possible the methods LBDD are employed. These tools can provide information on the nature of the interactions between target molecules and ligands and include the quantitative relationship of structure-activity (QSAR), pharmacological modeling, molecular field analysis and similarity 2D and 3D [1]. However, in cases where neither the structure of the target molecule is known nor there is any information about the ligand, methods based on sequence are used. These include tools for multiple analysis and comparison of sequences aimed at identifying potential targets [70,76]. However, the combination of the above methods and tools is important as individual methods are not able to cover the needs of the drug development process [33,44].

The entry of mathematical models into the field of drug discovery and preparation is of paramount importance. Some applications of mathematics in this field are related to the analysis of pharmacokinetics (PK) and pharmacodynamics (PD), which leads to the prediction of the dosage of the drug as well as the duration of potency of the drug in the body [4]. Computational machine learning (ML) and artificial intelligent (AI) methods have recently entered the pharmaceutical industry which provide greater efficiency and economy. The use of artificial intelligence tools in drug detection has been documented, predicting the physical properties, bioactivity and toxicity of a potential drug, and even the structure of the predictions. It is important to note that traditional structural biology techniques usually take several years to resolve a protein structure, whereas AI-based structure predictions only take a few hours to a few days. In addition, the employment of AI has been extended to cell image processing, natural bioac-

tivity and toxicity predictions, quantum mechanics (QM) properties predictions, chemical composition design and computer-aided organic synthesis, aimed at further improving the effectiveness of drug discovery [12].

3 Mathematical Models and Artificial Intelligence in Drug Design

Current –omics technologies produce vast amounts of raw data about intricate biochemical and regulatory processes in living organisms [54]. The accumulation of such large and complex datasets that are difficult to process with traditional data analysis methods epitomizes the term big data [90]. The ability to interpret big data helps to gain insights into molecular mechanisms that modulate different disease processes and access to a number of possible therapeutic agents and their properties. Therefore the use of big data is of high importance in drug design [90].

Drug design is a difficult multidimensional problem in which various compound attributes, including efficacy, pharmacokinetics, and safety, are to be optimized in parallel to provide possible drug candidates [59]. These procedures require a high number of resources, including money, time, personnel, and ideas. The use of computational methods and applied mathematics, although also requiring such resources, are of high value compared to conventional techniques [4]. Specifically, the use of bioinformatics tools can reduce the number of potential drugs and result in savings of money, time, and personnel [37]. A prime example is the FDA's Center for Drug Evaluation and Research (CDER) use of modeling and simulations in drug design [41]. The use of such techniques in drug design continues to rise due to the advancements in data processing power and the formation of new artificial intelligence tools that allow the processing of big data[60]. The term artificial intelligence (AI) applies to the simulation of intelligence in a non-living agent [20]. AI comprises of a number of various technologies that support several tasks and processes [19]. This section will focus on the use of mathematical models and various artificial intelligence bioinformatics tools in different drug design stages.

Mathematical modeling Mathematical modeling is an implementing tool of systems biology and, in the field of drug design, links in a quantitative manner the interactions between a pharmacological molecule or combination of drugs, biological pathways, and intricate disease systems [79]. Mathematical modeling is also heavily linked with pharmacometrics, the scientific field which quantifies drug, disease and trial information in order to aid drug design efficiently [4]. The use of mathematical models in drug design has been heavily encouraged by several scientific agencies [24,32].

A typical use of mathematical models in drug design is found in cancer research [32,79]. Some of the most characteristic mathematical models used in cancer drug development are associated with network-based medicine. Network medicine is an area of research that focuses on molecular and genetic interactions and has a number of therapeutic applications. The main three biological networks are protein interaction networks, gene regulatory networks, and expression-based networks [69]. Network-based mathematical models, as an example, can quantify the relationship between proteins in the protein-protein interactome. The information received can elucidate particular network 'nodes', in this case proteins, that are of high importance in a specific disease and therefore, appealing drug targets. Using network-based mathematical models in several diseases, a research team may find better targets for drug development, disease biomarkers, or drug combinations that can enhance therapeutic action [16,69].

Another example of mathematical model use in drug development is infectious diseases such as malaria [47]. In this case, mathematical models can be used to assess the candidate drug's ability to kill the parasite during specific stages of infection. Such models make use of two factors, compound pharmacokinetics and compound pharmacodynamics [65]. Pharmacokinetics focuses on how the drug's concentration increases and drops over time as determined by a number of set parameters such as absorption, distribution, and metabolism. Pharmacokinetics is described through the use of differential equations [65]. On the other hand, pharmacodynamics focuses on the relationship between the drug concentration and its killing efficacy. Pharmacodynamics makes use of the dose-response curve [65]. By combining both pharmacokinetics and pharmacodynamics, predictions can be made on a potential drug's efficacy and decay regarding different dosage schedules [65].

Machine Learning Machine Learning (ML) is a subset of AI which refers to several algorithms that execute intelligent predictions based on a –often large- data set [42]. In ML, instead of following a specific set of predetermined instructions by the programmer, the computer is trained using large amounts of data and has the ability to learn how to perform the tasks and improve its performance [73]. Machine learning tasks fall broadly into three main categories, which include supervised learning, unsupervised learning, and sequential learning [53]. In supervised learning, a computer aims to predict a new observation label based on labeled examples in a large training dataset. In unsupervised learning, the computer's goal is to detect underlying relationships or patterns in unlabeled data. Lastly, in sequential learning, algorithms work through trial-and-error, and iteratively use external observations to find the ideal decision proportionate to the environment they interact with [53].

Deep Learning (DL) is a prime example of machine learning algorithms currently being used in drug design. DL uses artificial neural networks (ANNs) with multiple layers of nonlinear processing units for learning data representations [13]. An ability of deep learning is that it can have supervised or unsupervised learning of feature presentations on each layer [10]. DL has been applied to various stages in drug design. Deep learning has been extensively used in predicting the properties of a candidate drug. Properties such as bioactivity, toxicity, or solubility are quite important in selecting a drug. Predicting such properties through deep neural networks is considered a case of supervised learning. Specifically, the input includes a candidate drug, while the output is expected to produce a label that showcases if a drug has the desired properties. In a characteristic study, two deep neural networks are integrated, the generative and the predictive. In the generative deep neural network, a candidate drug molecule's structure is encoded as text (encoded string), the deep neural network receives such strings and is trained to provide chemically feasible encoded strings. One of the most popular standards in encoding molecule structure as text is simplified molecular-input line-entry system (SMILES), which this study uses. The predictive deep neural network, on the other hand, is expected to calculate the compound's desired properties. In the first stage, both networks are trained separately through a supervised learning algorithm, while in the second stage, they are jointly trained with a reinforced learning approach in order to create a bias towards the desired drug properties [50].

Another example of machine learning algorithms used in drug design is support vector machines (SVM) [9]. Support vector machines are supervised learning models that in the field of drug design are used for topics tasks as classification, regression, and ranking/virtual screening [38]. SVM machines have been used in drug development H1N1 by focusing on the inhibition of a viral enzyme called neuraminidase. In this specific case, SVM models were built to single out active, and weakly active neuraminidase inhibitors where compounds were encoded by MARCSS fingerprints and their molecular descriptors were calculated by ADRIANA.Code [35]. The models displayed high accuracy showcasing SVMs' potential in drug design [35].

Drug Repurposing and the Case for Natural Language Processing Drug repurposing or drug reposition is a strategy for finding new uses for approved or investigational drugs that are different from their original therapeutic purpose [51]. Drug repurposing has specific advantages in regards to developing an entirely new drug. Repurposed drugs have been already tested and found sufficiently safe in humans if early trials have been completed. Moreover, the time needed for drug development is greatly shortened since several drug development steps have already been completed. Lastly, drug repurposing cost is expected to be lower than developing a drug de novo [51]. Various artificial intelligence technologies have been used in search of repositionable drugs, with a prime example being natural language processing.

Natural language processing (NLP) is a field of artificial intelligence that includes applications related to language such as text analysis, translation, and speech recognition [19]. NLP main applications involve the creation, understanding and classification of published research and clinical documentation. As mentioned, NLP has also been used to identify possible repositionable drugs [85]. In a specific study, natural language processing was used to mine text from electronic health records and lead to the identification of metformin as a drug that can be repurposed for cancer treatment [84].

Personalized Medicine as the Future of Drug Development Indeed, a brief mention should be made for personalized medicine. Personalized medicine or precision medicine aims to guide health care decisions toward the most effective treatment for a given individual or subpopulation [23]. Researchers can use mathematical models or

artificial intelligence tools on large –omics datasets to identify characteristics that are unique to specific individuals or subpopulations and develop drugs that are tailored for them by taking into account said attributes [18].

4 Mathematics in Drug Design

Modern drug development is characterized by low productivity and high costs. Moreover, less than 10% of new compounds entering clinical trials make it to the market [31]. Some of these hurdles can be addressed through the use of mathematical models. A mathematical model can be defined as the description of a biological system using mathematical concepts and language to create a proper explanation of the mechanisms underlying the system's function or to predict the effect of various parameters on the aforementioned system [43]. The process of constructing a mathematical model is called mathematical modeling, which is also known as predictive modeling, simulation, or decision analysis [43]. In silico studies can use mathematical models to elucidate various pharmacological attributes of potential therapeutic compounds. It is important to note that the FDA's Center for Drug Evaluation and Research (CDER) is using modeling and simulations in numerous stages of drug development [4].

Mathematical approaches provide great value to drug discovery based on their potential impact and relatively low cost in comparison with running clinical trials [4]. The use of mathematical models and associated computer simulations showcase several benefits. Firstly, discrepancies between systems behavior predicted by a mathematical model and behaviors observed in actual experiments can indicate components that are missing from the mathematical model, therefore assisting in developing a more concise picture of a biological process. Even if it is uncertain which components are missing from the system being investigated, the mathematical model results could potentially guide the design of additional experiments to clarify the issue [22]. Moreover, mathematical models allow for a systematic approach for investigating system perturbations induced by drug administration. Lastly, due to the rapid technological advancements of the past years, mathematical models are not as limited by experimental constraints as lab experiments [22].

The earliest and some of the most heavily used applications of mathematics to drug design, though, are related to pharmacokinetic and pharmacodynamic analyses [4]. Pharmacokinetics describes the time course of drug concentration, while pharmacodynamics describes how drug effects change with concentration [25]. Pharmacokinetic models portray the concentration of a potential drug in the relevant organism compartment (e.g., circulating blood). On the other hand, pharmacodynamic models link this concentration to a biomarker that is hypothesized to be associated with a disease state, oftentimes taking directly into account the pharmacological target's modulation [4].

The pharmacokinetic characteristics of a drug include parameters such as concentration related to time (C(t)), the elimination rate constant (K), half-life ($t_{1/2}$), apparent volume of distribution (V_d), and total clearance rate (CL) [86]. The simplest pharmacokinetic model considers clearance as a function of both volume of distribution and elimination, in which

$CL = V_d K$ [39].

Even though pharmacokinetic analysis can be done without specifying any mathematical models, such models are beneficial in therapeutic decision making. The simplest linear pharmacokinetic model, which assumes that a drug is administered as an instantaneous bolus and that complete drug distribution is also instantaneous, is

$C(t) = \frac{Dose}{V_d}(e^{-kt})$ [39]

The assumptions made on the model above are often non-valid. In cases in which a potential drug is administered as a slow bolus or infusion, during drug administration, the concentration is increasing, and infusion time (T) should be taken into consideration.

$$C(t) = \frac{Dose}{V_d K T}(1 - e^{-kt})$$

The post-infusion drug concentration can be represented as

$C(t) = C(T)e^{-k(1-T)}$ [39]

The pharmacokinetic data though can be even more complex, and more thorough mathematical models may be needed. A prime example is nonlinear pharmacokinetic models, which consider that the pharmacokinetic behavior of a drug is saturable, and are very relevant to anticancer agents [39].

The pharmacodynamic characteristics of a drug include the archetypal effect of a drug (E) as a function of drug concentration (C) and an unknown pharmacodynamic parameter (θ_{pd}). These models can predict drug effect and the mathematical model describing one is

$E(C) = f(C, \theta_{pd})$ [57]

One of the most used pharmacodynamic models is the Emax model (4). The Emax model is a nonlinear mathematical model derived from the classic drug-receptor theory (9). The Emax model's main parameters are Emax and EC_{50}, with an Emax model being either stimulating or inhibiting (4). Emax is a pharmacodynamic parameter that describes the maximum effect of a potential drug, and EC50 describes the concentration producing 50% of the maximum effect (6). The Emax model can be described as

$Emax = \frac{Emax \times C}{EC_{50}+C}$ or $Emax = \frac{Emax \times Dose}{ED_{50}+Dose}$,

where ED_{50} is the dose that produces 50% of the maximum effect.

The Emax model can be improved by including E which accounts for the baseline physiological conditions, such as blood pressure, in the absence of a drug. The Emax can then be described as

$Emax = E_0 + \frac{Emax \times C}{EC_{50}+C}$ or $Emax = E_0 + \frac{Emax \times Dose}{ED_{50}+Dose}$ [57]

Elucidating the relationship between pharmacokinetics and pharmacodynamics allows researchers to understand a drug's action mechanism and select the optimal compounds for further profiling. Furthermore, pharmacokinetics and pharmacodynamic models can showcase if a potential drug's attributes are in agreement with clinical expectations.

Lastly, drug development in complex diseases such as cancer makes use of more advanced mathematical models. In this case, network-based mathematical models can be used to identify drug combinations for better treatment. Moreover, network-based mathematical models can quantify the relationship between 'nodes' in the protein interactome. The information received can elucidate particular network 'nodes', proteins in this case, that are of high importance in a specific disease and, therefore, appealing drug targets.

All of the above reflects the importance of mathematics in drug design. The use of mathematical models is not a new phenomenon since pharmacokinetics and pharmacodynamics have been using them for many years. What is new, though, is the impressive leaps in computational technology characterizing the past decade, which allow the use of more complex models that provide more detailed information to drug designers. The incorporation of such models can allow for better dose predictions, better drug response predictions, and even more individualized therapies (Moore, Allen et al. 2019).

5 Specifics of Computational Drug Design

Recapping the previous statements, the process of developing new drugs that can be effective, highly specific, and non-toxic is a major challenge for the pharmaceutical industry. Standard development methods, from the synthesis and evaluation of candidate compounds to extensive clinical trials, demand a great investment in terms of both time and cost [48]. In the face of these challenges, computational methods for drug design have presented themselves as an attractive alternative and have had many successful applications [58,77]. Computational drug design can be broadly split into two major categories, structure-based and ligand-based drug design. Structure-based drug design is the preferred method when the structure of the biological target is well documented through various experimental methods. When such structural data is unavailable, ligand-based drug design is applied [72].

Structure-based Drug Design For structure-based drug design, an accurate view of the biomolecular target's structure is crucial for the employment of the method. This prerequisite is usually achieved through NMR, X-ray crystallography, and cryo-EM [40]. High throughput methods have led to the generation of an incredible amount of solved structures that can be obtained through public databases such as Protein Data Bank [7], which houses more than 163.000 structural entries. In the case where the structure cannot be retrieved, *in silico* methods

such as homology modeling can lead to the prediction of the unknown three-dimensional structure. Modeling, by comparison, takes into account the fact that structure is more associated with the protein's function and biological properties compared to its sequence [27]. Homology modeling, a subset of comparative modeling, includes the element of an evolutionary relationship between the proteins in question. Towards the execution of homology modeling, the sequence of interest is aligned, using widely available tools such as Blast . Altschul et al. [5], and one or several proteins with an experimentally defined structure are chosen as templates. During the building of the model, well-aligned regions between the target sequence and the template allow for the copying of the coordinates. Missing regions are modeled via a variety of available loop-modeling methods, and the conformations of side-chains are subsequently predicted [66]. Refinement of the generated models can be carried out with the use of various methods, such as genetic algorithms [83], or molecular dynamics [52], which have become a staple for homology modeling in the drug design setting [34]. The final step of the process is the evaluation of the generated models, which often includes the comparison of protein structures with known structure and the model of the target protein [66].

In order to predict the web of intermolecular interactions between proteins or between a protein and a small molecule (ligand), molecular docking studies are carried out. Docking takes into account features such as shapes, electrostatic interactions, hydrogen bonds, van der Waals and Coulombic interactions [46]. Various manners of fitting of the ligand in the binding site are explored, and a docking score represents binding potentiality. A key component of the docking methodology is a sampling algorithm that filters the different conformations for one that best matches the receptor structure, ensuring accuracy and time efficiency [66]. There is a broad spectrum of sampling algorithms with varying degrees of freedom, like Monte Carlo simulation, incremental construction, genetic algorithms and fast shape matching [61], implemented in various docking software such as DOCK [29], GLIDE [21], GOLD 3.1 [74] and others. Another component of the docking methodology is the scoring function, which complements the algorithms that are presented above by evaluating and ranking the conformations of the ligand that are predicted. These scoring functions can be broadly classified into empirical, force field, and knowledge-based [61]. During the initial or later stages of a docking approach, molecular dynamics (MD) simulations can be used as an additional tool. MD runs can be run before docking or as a last filter, in order to evaluate and rank the predicted modes of binding as they are generated by the docking approach [89].

The features of electronic and steric nature which are necessary for the best level of interaction between a ligand and a target help to shape a pharmacophore model of the target binding site. During structure-based drug design, pharmacophore models can be incorporated into algorithms for de novo design, or can be used in order to screen databases of compounds [66].

One last component of structure-based drug design is the de novo design of ligands. This approach encompasses a ligand-growing or a ligand-linking methodology. In the ligand-growing methodology, the ligand is fragmented and one of the fragments undergoes docking into the binding site. Through the addition of functional groups to the original fragment, the ligand is elongated. In the second methodology, several fragments undergo docking into adjoining pockets of the target site. Once all the fragments have been connected, a complete compound is created [66].

Ligand-Based Drug Design Ligand-based drug design is characterized as an indirect approach, given that the structure of the biomolecular target is unknown or cannot be predicted with the use of methods such as homology modeling. Ligand-based drug design involves the study of ligands that are known to interact with the target of interest, with the goal of understanding the correlation between the pharmacological profile of the ligands and their respective chemical and physical properties [30]. These properties that are the most important for the interaction are kept, while the noise of irrelevant information is removed.

Quantitative Structure-Activity Relationships (QSAR) is an elemental approach to ligand-based drug design, based on the notion that similar activity is a result of similar physiochemical or structural properties [3]. In broad strokes, QSAR involves the following steps. Ligands whose biological activity of interest has been quantified experimentally are identified. Descriptors that exhibit a level of association with the physiochemical and structural properties of interest are subsequently chosen, and then used in order to identify correlations between them and

the biological activity. These correlations must account for the variation in terms of activity observed in the set of ligands. The generated QSAR model is used for the optimization of the active ligands. In the last step, the activity of the predicted molecules is verified experimentally [1].

The classic QSAR approach, named 2D-QSAR, cannot provide an adequate level of accuracy when it comes to the relationship between biological activity and the arrangements of physiochemical properties in the three-dimensional space. This has led to the development of 3D-QSAR, where the generated QSAR model relies on descriptors that account for the three-dimensional characteristics of the active molecules. A widely used 3D-QSAR method, Comparative Molecular Field Analysis (CoMFA), computes the fields that surround the molecule, operating under the notion that these molecular fields are the basis of the molecules' biological activity. An extension of this approach, Comparative Molecular Similarity Indices Analysis (CoMISA), adds fields such as hydrogen bond donors and acceptors, and hydrophobic interactions, to the mix [63]. Efforts have been made, in parallel, to incorporate the element of conformational flexibility in the 3D-QSAR approach. A widely-used software package that follows this approach uses a poling algorithm [67] to sample the space of conformations in terms of the ligands. Finally, multidimensional QSAR (mQSAR) attempts to address the phenomena that add to the energy of ligand binding. This includes adaptation to the binding pocket and conformational entropy, which are addressed by 5D-QSAR and 4D-QSAR, respectively [66].

Lastly, a crucial component of ligand-based drug design is pharmacophore mapping. In order for a drug to elicit a specific biological response, it must possess functional groups with a specific arrangement. This arrangement, in the drug design approach, is the definition of the pharmacophore. For the creation of a ligand-based pharmacophore, active molecules are stacked in a way that ensures their chemical features are overlaid to the maximum. The alignment of the molecules can be achieved through a manner of ways, from rigid methods, which demand information of the active conformations of ligands to semi-flexible and flexible methods. The latter require significant computational power to perform a conformational search while the alignment is taking place. After the alignment, pharmacophore features are extracted, usually involving elements such as hydrophobic properties or hydrogen bond receptors and donors [1]. The generation of the pharmacophore can be executed through a variety of available packages, like LigandScout [82] or MOE [75], all of which handle the steps of the methodology in a different manner.

6 Conclusion

In conclusion, we are going through a period that requires the rapid design of drugs and vaccines in order for humanity to be able to reverse the pandemic from COVID-19. Traditional experimental studies need plenty of time and money to provide the right solutions in order to test and launch the new drug on the world market. At this juncture, computational methods would be able to provide solutions in less time and with less money. Through the methods of machine learning and artificial intelligence it is possible to predict the characteristics, both chemical and structural, of the molecules thus leading to a reduction in the risk of complications and failure.

References

1. Acharya, C., Coop, A., Polli, J.E., Mackerell Jr., A.D.: Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. Curr. Comput. Aided Drug Des. **7**(1), 10–22 (2011)
2. Adhikari, S.P., Meng, S., Wu, Y.J., Mao, Y.P., Ye, R.X., Wang, Q.Z., Sun, C., Sylvia, S., Rozelle, S., Raat, H., Zhou, H.: Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review. Infect Dis. Poverty **9**(1), 29 (2020)
3. Akamatsu, M.: Current state and perspectives of 3D-QSAR. Curr. Top. Med. Chem. **2**(12), 1381–1394 (2002)
4. Allen, R., Moore, H.: Perspectives on the role of mathematics in drug discovery and development. Bull. Math. Biol. **81**(9), 3425–3435 (2019)
5. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J.: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**(17), 3389–3402 (1997)
6. Belouzard, S., Chu, V.C., Whittaker, G.R.: Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. Proc. Natl. Acad. Sci. U S A **106**(14), 5871–5876 (2009)

7. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E.: The protein data bank. Nucleic Acids Res. **28**(1), 235–242 (2000)
8. Bosch, B.J., van der Zee, R., de Haan, C.A., Rottier, P.J.: The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. J. Virol. **77**(16), 8801–8811 (2003)
9. Burbidge, R., Trotter, M., Buxton, B., Holden, S.: Drug design by machine learning: support vector machines for pharmaceutical data analysis. Comput. Chem. **26**(1), 5–14 (2001)
10. Cao, C., Liu, F., Tan, H., Song, D., Shu, W., Li, W., Zhou, Y., Bo, X., Xie, Z.: Deep learning and its applications in biomedicine. Genomics Proteomics Bioinf. **16**(1), 17–32 (2018)
11. Chan-Yeung, M., Xu, R.H.: SARS: epidemiology. Respirology **8**(Suppl), S9–14 (2003)
12. Chan, H.C.S., Shan, H., Dahoun, T., Vogel, H., Yuan, S.: Advancing drug discovery via artificial intelligence. Trends Pharmacol. Sci. **40**(8), 592–604 (2019)
13. Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., Blaschke, T.: The rise of deep learning in drug discovery. Drug Discov. Today **23**(6), 1241–1250 (2018)
14. Chen, L., Morrow, J.K., Tran, H.T., Phatak, S.S., Du-Cuny, L., Zhang, S.: From laptop to benchtop to bedside: structure-based drug design on protein targets. Curr. Pharm. Des. **18**(9), 1217–1239 (2012)
15. Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J.a, Yu, T., Zhang, X., Zhang, L.: Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet **395**(10223), 507–513 (2020)
16. Cheng, F., Kovács, I.A., Barabási, A.-L.: Network-based prediction of drug combinations. Nat. Commun. **10**(1), 1197 (2019)
17. Cheng, Z.J., Shan, J.: 2019 Novel coronavirus: where we are and what we know. Infection **48**(2), 155–163 (2020)
18. Colijn, C., Jones, N., Johnston, I.G., Yaliraki, S., Barahona, M.: Toward precision healthcare: context and mathematical challenges. Front. Physiol. **8**, 136 (2017)
19. Davenport, T., Kalakota, R.: The potential for artificial intelligence in healthcare. Future Healthc. J. **6**(2), 94–98 (2019)
20. Dias, R., Torkamani, A.: Artificial intelligence in clinical and genomic diagnostics. Genome Med. **11**(1), 70 (2019)
21. Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C., Mainz, D.T.: Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. J. Med. Chem. **49**(21), 6177–6196 (2006)
22. Fischer, H.P.: Mathematical modeling of complex biological systems: from parts lists to understanding systems behavior. Alcohol Res. Health : J. Natl. Inst. Alcohol Abuse Alcohol. **31**(1), 49–59 (2008)
23. Ginsburg, G.S., Phillips, K.A.: Precision medicine: from science to value. Health Aff. (Project Hope) **37**(5), 694–701 (2018)
24. Gobburu, J.V.S., Lesko, L.J.: Quantitative disease, drug, and trial models. Annu. Rev. Pharmacol. Toxicol. **49**(1), 291–301 (2009)
25. Holford, N.: Pharmacodynamic principles and the time course of immediate drug effects. Transl. Clin. Pharmacol. **25**(4), 157–61 (2017)
26. Huang, P., Liu, T., Huang, L., Liu, H., Lei, M., Xu, W., Hu, X., Chen, J., Liu, B.: Use of chest CT in combination with negative RT-PCR assay for the 2019 novel coronavirus but high clinical suspicion. Radiology **295**(1), 22–23 (2020)
27. Illergard, K., Ardell, D.H., Elofsson, A.: Structure is three to ten times more conserved than sequence-a study of structural response in protein cores. Proteins **77**(3), 499–508 (2009)
28. Jin, Y.H., Cai, L., Cheng, Z.S., Cheng, H., Deng, T., Fan, Y.P., Fang, C., Huang, D., Huang, L.Q., Huang, Q., Han, Y., Hu, B., Hu, F., Li, B.H., Li, Y.R., Liang, K., Lin, L.K., Luo, L.S., Ma, J., Ma, L.L., Peng, Z.Y., Pan, Y.B., Pan, Z.Y., Ren, X.Q., Sun, H.M., Wang, Y., Wang, Y.Y., Weng, H., Wei, C.J., Wu, D.F., Xia, J., Xiong, Y., Xu, H.B., Yao, X.M., Yuan, Y.F., Ye, T.S., Zhang, X.C., Zhang, Y.W., Zhang, Y.G., Zhang, H.M., Zhao, Y., Zhao, M.J., Zi, H., Zeng, X.T., Wang, Y.Y., Wang, X.H., FTZHOWUNC Management, E-BMCOCIE Research Team: M. Promotive association for and C. health: a rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). Mil. Med. Res. **7**(1), 4 (2020)
29. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R., Ferrin, T.E.: A geometric approach to macromolecule-ligand interactions. J. Mol. Biol. **161**(2), 269–288 (1982)
30. Kurogi, Y., Guner, O.F.: Pharmacophore modeling and three-dimensional database searching for drug design using catalyst. Curr. Med. Chem. **8**(9), 1035–1055 (2001)
31. Lalonde, R.L., Kowalski, K.G., Hutmacher, M.M., Ewy, W., Nichols, D.J., Milligan, P.A., et al.: Model-based drug development. Clin. Pharmaco. Ther. **82**(1), 21–32 (2007)
32. Lavezzi, S.M., Borella, E., Carrara, L., De Nicolao, G., Magni, P., Poggesi, I.: Mathematical modeling of efficacy and safety for anticancer drugs clinical development. Expert Opin. Drug Discov. **13**(1), 5–21 (2018)
33. Leelananda, S.P., Lindert, S.: Computational methods in drug discovery. Beilstein J. Org. Chem. **12**, 2694–2718 (2016)
34. Li, W., Tang, Y., Liu, H., Cheng, J., Zhu, W., Jiang, H.: Probing ligand binding modes of human cytochrome P450 2J2 by homology modeling, molecular dynamics simulation, and flexible molecular docking. Proteins **71**(2), 938–949 (2008)
35. Li, Y., Kong, Y., Zhang, M., Yan, A., Liu, Z.: Using support vector machine (SVM) for classification of selectivity of H1N1 neuraminidase inhibitors. Mol. Inform. **35**(3–4), 116–124 (2016)
36. Lu, H.: Drug treatment options for the 2019-new coronavirus (2019-nCoV). Biosci. Trends **14**(1), 69–71 (2020)
37. Malathi, K., Ramaiah, S.: Bioinformatics approaches for new drug discovery: a review. Biotechnol. Genet. Eng. Rev. **34**(2), 243–260 (2018)
38. Maltarollo, V.G., Kronenberger, T., Espinoza, G.Z., Oliveira, P.R., Honorio, K.M.: Advances with support vector machines for novel drug discovery. Expert Opin. Drug Discov. **14**(1), 23–33 (2019)

39. Ratain MJ, Plunkett WK Jr. Principles of Pharmacokinetics. In: Kufe DW, Pollock RE, Weichselbaum RR, et al., (eds) Holland-Frei Cancer Medicine. 6th edition. Hamilton (ON): BC Decker; 2003. Available from: https://www.ncbi.nlm.nih.gov/books/NBK12815/
40. Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F., Sali, A.: Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. **29**, 291–325 (2000)
41. Moore, H., Allen, R.: What can mathematics do for drug development? Bull. Math. Biol. **81**(9), 3421–3424 (2019)
42. Nichols, J.A., Herbert Chan, H.W., Baker, M.A.B.: Machine learning: applications of artificial intelligence to imaging and diagnosis. Biophys. Rev. **11**(1), 111–118 (2019)
43. Ningthoujam, S.S., Talukdar, A.D., Sarker, S.D., Nahar, L., Choudhury, M.D.: Prediction of medicinal properties using mathematical models and computation, and selection of plant materials. Comput. Phytochem. (2018). https://doi.org/10.1016/B978-0-12-812364-5.00002-X
44. Ou-Yang, S.S., Lu, J.Y., Kong, X.Q., Liang, Z.J., Luo, C., Jiang, H.: Computational drug discovery. Acta Pharmacol. Sin **33**(9), 1131–1140 (2012)
45. Ou, X., Liu, Y., Lei, X., Li, P., Mi, D., Ren, L., Guo, L., Guo, R., Chen, T., Hu, J., Xiang, Z., Mu, Z., Chen, X., Chen, J., Hu, K., Jin, Q., Wang, J., Qian, Z.: Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. Nat. Commun. **11**(1), 1620 (2020)
46. Pagadala, N.S., Syed, K., Tuszynski, J.: Software for molecular docking: a review. Biophys. Rev. **9**(2), 91–102 (2017)
47. Patel, K., Simpson, J.A., Batty, K.T., Zaloumis, S., Kirkpatrick, C.M.: Modelling the time course of antimalarial parasite killing: a tour of animal and human models, translation and challenges. Br. J. Clin. Pharmacol. **79**(1), 97–107 (2015)
48. Paul, S.M., Mytelka, D.S., Dunwiddie, C.T., Persinger, C.C., Munos, B.H., Lindborg, S.R., Schacht, A.L.: How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat. Rev. Drug Discov. **9**(3), 203–214 (2010)
49. Pickett, B.E., Sadat, E.L., Zhang, Y., Noronha, J.M., Squires, R.B., Hunt, V., Liu, M., Kumar, S., Zaremba, S., Gu, Z., Zhou, L., Larson, C.N., Dietrich, J., Klem, E.B., Scheuermann, R.H.: ViPR: an open bioinformatics database and analysis resource for virology research. Nucleic Acids Res. **40**, D593–598 (2012)
50. Popova, M., Isayev, O., Tropsha, A.: Deep reinforcement learning for de novo drug design. Sci. Adv. **4**(7), eaap7885 (2018)
51. Pushpakom, S., Iorio, F., Eyers, P.A., Escott, K.J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., Norris, A., Sanseau, P., Cavalla, D., Pirmohamed, M.: Drug repurposing: progress, challenges and recommendations. Nat. Rev. Drug Discov. **18**(1), 41–58 (2019)
52. Raval, A., Piana, S., Eastwood, M.P., Dror, R.O., Shaw, D.E.: Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. Proteins **80**(8), 2071–2079 (2012)
53. Réda, C., Kaufmann, E., Delahaye-Duriez, A.: Machine learning applications in drug development. Comput. Struct. Biotechnol. J. **18**, 241–252 (2020)
54. Ristevski, B., Chen, M.: Big data analytics in medicine and healthcare. J. Integr. Bioinform. **15**(3), 20170030 (2018)
55. Rothan, H.A., Byrareddy, S.N.: The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. J. Autoimmun. **109**, 102433 (2020)
56. Rothe, C., Schunk, M., Sothmann, P., Bretzel, G., Froeschl, G., Wallrauch, C., Zimmer, T., Thiel, V., Janke, C., Guggemos, W., Seilmaier, M., Drosten, C., Vollmar, P., Zwirglmaier, K., Zange, S., Wolfel, R., Hoelscher, M.: Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. N. Engl. J. Med. **382**(10), 970–971 (2020)
57. Salahudeen, M.S., Nishtala, P.S.: An overview of pharmacodynamic modelling, ligand-binding approach and its application in clinical practice. Saudi Pharm. J. **25**(2), 165–75 (2017)
58. Schames, J.R., Henchman, R.H., Siegel, J.S., Sotriffer, C.A., Ni, H., McCammon, J.A.: Discovery of a novel binding trench in HIV integrase. J. Med. Chem. **47**(8), 1879–1881 (2004)
59. Schneider, G.: Automating drug discovery. Nat. Rev. Drug Discov. **17**(2), 97–113 (2018)
60. Schneider, P., Walters, W.P., Plowright, A.T., Sieroka, N., Listgarten, J., Goodnow, R.A., Fisher, J., Jansen, J.M., Duca, J.S., Rush, T.S., Zentgraf, M., Hill, J.E., Krutoholow, E., Kohler, M., Blaney, J., Funatsu, K., Luebkemann, C., Schneider, G.: Rethinking drug design in the artificial intelligence era. Nat. Rev. Drug Discov. **19**(5), 353–364 (2020)
61. Sethi, A., Joshi, K., Sasikala, K., Alvala, M.: Molecular docking in modern drug discovery: principles and recent applications. In: Gaitonde, V., Karmakar, P., Trivedi, A. (eds.) Drug Discovery and Development - New Advances. IntechOpen, London (2020)
62. Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., Li, F.: Structural basis of receptor recognition by SARS-CoV-2. Nature **581**(7807), 221–224 (2020)
63. Singh, B., Mal, G., Gautam, S.K., Mukesh M.: Computer-aided drug discovery. Advances in Animal Biotechnology (2019)
64. Singhal, T.: A review of coronavirus disease-2019 (COVID-19). Indian J. Pediatr. **87**(4), 281–286 (2020)
65. Slater, H.C., Okell, L.C., Ghani, A.C.: Mathematical modelling to guide drug development for malaria elimination. Trends Parasitol. **33**(3), 175–184 (2017)
66. Sliwoski, G., Kothiwale, S., Meiler, J., Lowe Jr., E.W.: Computational methods in drug discovery. Pharmacol. Rev. **66**(1), 334–395 (2013)
67. Smellie, A., Teig, S.L., Towbin, P.: Poling: promoting conformational variation. J. Comput. Chem. **16**(2), 171–187 (1995)
68. Sohrabi, C., Alsafi, Z., O'Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, R.: World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). Int. J. Surg. **76**, 71–76 (2020)
69. Sonawane, A.R., Weiss, S.T., Glass, K., Sharma, A.: Network medicine in the age of biomedical big data. Front. Genet. **10**, 294 (2019)

70. Tang, Y., Zhu, W., Chen, K., Jiang, H.: New technologies in computer-aided drug design: toward target identification and new chemical entity discovery. Drug Discov. Today Technol. **3**(3), 307–313 (2006)
71. Toots, M., Yoon, J.J., Cox, R.M., Hart, M., Sticher, Z.M., Makhsous, N., Plesker, R., Barrena, A.H., Reddy, P.G., Mitchell, D.G., Shean, R.C., Bluemling, G.R., Kolykhalov, A.A., Greninger, A.L., Natchus, M.G., Painter, G.R., Plemper, R.K.: Characterization of orally efficacious influenza drug with high resistance barrier in ferrets and human airway epithelia. Sci. Transl. Med. **11**, 515 (2019)
72. Usha, T., Shanmugarajan, D., Goyal, A.K., Kumar, C.S., Middha, S.K.: Recent updates on computer-aided drug discovery: time for a paradigm shift. Curr. Top Med. Chem. **17**(30), 3296–3307 (2017)
73. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., Zhao, S.: Applications of machine learning in drug discovery and development. Nat. Rev. Drug Discov. **18**(6), 463–477 (2019)
74. Verdonk, M.L., Chessari, G., Cole, J.C., Hartshorn, M.J., Murray, C.W., Nissink, J.W., Taylor, R.D., Taylor, R.: Modeling water molecules in protein-ligand docking using GOLD. J. Med. Chem. **48**(20), 6504–6515 (2005)
75. Vilar, S., Cozza, G., Moro, S.: Medicinal chemistry and the molecular operating environment (MOE): application of QSAR and molecular docking to drug discovery. Curr. Top Med. Chem. **8**(18), 1555–1572 (2008)
76. Wang, F., Liu, D., Wang, H., Luo, C., Zheng, M., Liu, H., Zhu, W., Luo, X., Zhang, J., Jiang, H.: Computational screening for active compounds targeting protein sequences: methodology and experimental validation. J. Chem. Inf. Model **51**(11), 2821–2828 (2011)
77. Wang, L., Gu, Q., Zheng, X., Ye, J., Liu, Z., Li, J., Hu, X., Hagler, A., Xu, J.: Discovery of new selective human aldose reductase inhibitors through virtual screening multiple binding pocket conformations. J. Chem. Inf. Model **53**(9), 2409–2422 (2013)
78. Wang, M., Cao, R., Zhang, L., Yang, X., Liu, J., Xu, M., Shi, Z., Hu, Z., Zhong, W., Xiao, G.: Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res. **30**(3), 269–271 (2020)
79. Wang, Z., Deisboeck, T.S.: Mathematical modeling in cancer drug discovery. Drug Discov. Today **19**(2), 145–150 (2014)
80. WHO: Coronavirus disease (COVID-2019) situation reports. (2020) Retrieved May 15th 2020, from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports
81. WHO: Middle East Respiratory Syndrome Coronavirus. (2020) Retrieved 16 Feb 2020, 2020, from https://www.who.int/emergencies/mers-cov/en/
82. Wolber, G., Langer, T.: LigandScout: 3-D Pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J. Chem. Inf. Model. **45**(1), 160–169 (2005)
83. Xiang, Z.: Advances in homology protein structure modeling. Curr. Protein Pept. Sci. **7**(3), 217–227 (2006)
84. Xu, H., Aldrich, M.C., Chen, Q., Liu, H., Peterson, N.B., Dai, Q., Levy, M., Shah, A., Han, X., Ruan, X., Jiang, M., Li, Y., Julien, J.S., Warner, J., Friedman, C., Roden, D.M., Denny, J.C.: Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality. J. Am. Med. Inf. Assoc. : JAMIA **22**(1), 179–191 (2015)
85. Xue, H., Li, J., Xie, H., Wang, Y.: Review of drug repositioning approaches and resources. Int. J. Biol. Sci. **14**(10), 1232–1244 (2018)
86. Yu, R.H., Cao, Y.X.: A method to determine pharmacokinetic parameters based on andante constant-rate intravenous infusion. Sci. Rep. **7**(1), 13279 (2017)
87. Yuki, K., Fujiogi, M., Koutsogiannaki, S.: COVID-19 pathophysiology: a review. Clin. Immunol. (Orlando, Fla.) **215**, 108427 (2020)
88. Zhang, Y., Geng, X., Tan, Y., Li, Q., Xu, C., Xu, J., Hao, L., Zeng, Z., Luo, X., Liu, F., Wang, H.: New understanding of the damage of SARS-CoV-2 infection outside the respiratory system. Biomed. Pharmacother. = Biomed. Pharmacother. **127**, 110195 (2020)
89. Zhao, H., Caflisch, A.: Molecular dynamics in drug design. Eur. J. Med. Chem. **91**, 4–14 (2015)
90. Zhu, H.: Big data and artificial intelligence modeling for drug discovery. Annu. Rev. Pharmacol. Toxicol. **60**(1), 573–589 (2020)