RECENT ADVANCES IN COMPUTATIONAL DRUG DISCOVERY FOR THERAPY AGAINST CORONAVIRUS SARS-COV-2

Volodymyr Ivanov, Kateryna Lohachova, Yaroslav Kolesnik, Anton Zakharov, Larysa Yevsieieva, Alexander Kyrychenko, Tierry Langer, Sergiy M. Kovalenko, Oleg N. Kalugin

Despite essential experimental efforts focused on studying severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), computational chemistry methods are promising complementary tools in combating coronavirus disease 2019 (COVID-19). The present review aims to provide readers with the recent progress and advances in computational approaches currently used to streamline drug discovery and development in the context of COVID-19. Our review is dual purpose. It is intended (a) to familiarize the readership with the general concept of QSAR, in silico screening, molecular docking and molecular dynamics (MD) simulations and (b) to provide key examples of the recent applications of these computational tools in discovering novel therapeutic agents against COVID-19. We outline how structure- and ligand-based drug design can accelerate the structural elucidation of pharmacological drug targeting and the discovery of preclinical drug candidate molecules. Several examples of MD computational studies demonstrate how atomistic MD simulations can facilitate our understanding of the molecular basis of drug actions and biological mechanisms of virus inhibition in atomic detail. Finally, the short- and long-term perspectives in computational drug discovery are discussed.

The aim of this study is to summarize the last three years' progress and advances in computational approaches currently used to streamline the drug discovery and development process in the context of COVID-19.

Materials and methods. The literature overview of QSAR, in silico screening, machine learning, molecular docking and molecular dynamics (MD) simulations is given in the context of COVID-19. The literature search was performed using online databases, such as Scopus, Web of Science, PDB-protein databank, and PubMed, focusing on the following keywords - human coronavirus, QSAR, molecular docking, virtual screening, machine learning, molecular dynamics, Mpro and PLpro proteases, SARS-CoV-2, respectively.

Results. The review familiarizes the readership with the general concept of QSAR, in silico screening, machine learning, molecular docking and MD simulations and provides key examples of the recent applications of these computational tools in discovering novel therapeutic agents against COVID-19.

Conclusions. New insight into the recent progress and achievements in computer-guided drug discovery for therapeutic agents against SARS-CoV-2 is provided

Keywords: human coronavirus, QSAR, molecular docking, virtual screening, machine learning, molecular dynamics, structure-based drug design, Mpro and PLpro proteases, CADD, SARS-CoV-2

1. Introduction

In the last few years, computational chemistry tools have been extensively employed to guide the development of therapy against Coronavirus Disease 2019 (COVID-19) [1–3]. Several computationally targeting strategies have been suggested, ranging from re-docking of existing FDA-approved drugs aiming for drug repurposing against COVID-19 up to target-specific *in silico* screening of new small-molecules able to block critical viral proteases, such as Mpro and PLpro, avoiding maturation of proteins crucial for the virus life cycle [2, 4–7].

Computer-aided drug discovery (CADD) is a series of computational approaches to guide and streamline the drug discovery and development process to minimize the production cost and time [8, 9]. Two principal strategies of CADD are often encountered: ligand-based and structured-based drug design. Structure-based drug design (SBDD) requires knowledge of the 3D structure of a receptor and utilizes virtual screening\docking and molecular dynamic simulations. In addition, some recent approaches utilized sequence-based [10] and fragment-based [11–13] drug design. A typical workflow of SBDD is summarized in Fig. 1 [6, 14–17]. Ligand-based drug design is based on quantitative structure-activity relationships (QSARs), pharmacophore modelling, artificial intelligence (AI) [18] and machine learning methods [2, 19].

In this review, we outline the progress and recent advances of CADD in anti-COVID-19 therapy. The different strategies for the rapid identification of promising drug candidate molecules against various drug targets

implicated in the pathogenesis of the coronavirus SARS-CoV-2 are overviewed. Section 2 outlines how the structural elucidation of pharmacological drug targeting and the discovery of preclinical drug candidate molecules can be accelerated using structure- and ligand-based drug design. Section 3 provides fundamentals and some promising examples of using molecular dynamics (MD) simulations for critical SARS-CoV-2 proteins and their complexes with non-covalent inhibitors as essential tools for drug design and development [20]. Several examples of MD computational studies will demonstrate how atomistic MD simulations can facilitate our understanding of the molecular basis of drug actions and biological mechanisms of virus inhibition in atomic detail.
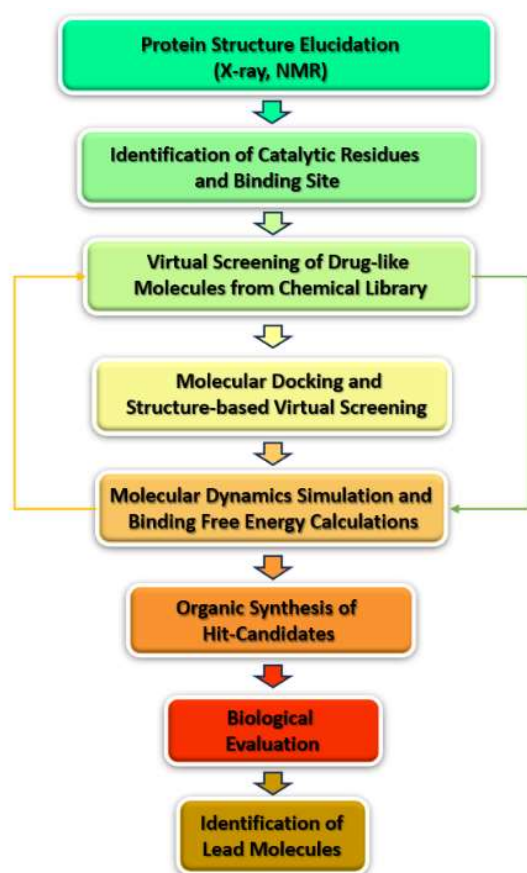

Fig. 1. A typical workflow of structure-based drug design

2. Materials and methods

The literature overview of QSAR, *in silico* screening, machine learning, molecular docking and molecular dynamics (MD) simulations is given in the context of COVID-19. The literature search was performed using online databases, such as Scopus, Web of Science, PDB-protein databank, and PubMed focusing on the following keywords – human coronavirus, QSAR, molecular docking, virtual screening, machine learning, molecular dynamics, Mpro and PLpro proteases, SARS-CoV-2, respectively.

3. Result and discussion
3. 1. Computational methods in drug discovery

Rapid progress in developing new anti-COVID-19 drugs has been achieved due to the critical role of com-

putational methods and approaches, which become an essential complementary tool in drug discovery. One of popular workflows of computer-aided drug discovery for therapy against COVID-19 is summarized in Fig. 2.

Most available *in silico* screening tools can be divided into two large groups. The first one is based on statistical Quantitative Structure – Activity Relationships (QSAR) methodology. The second group includes molecular modelling approaches.
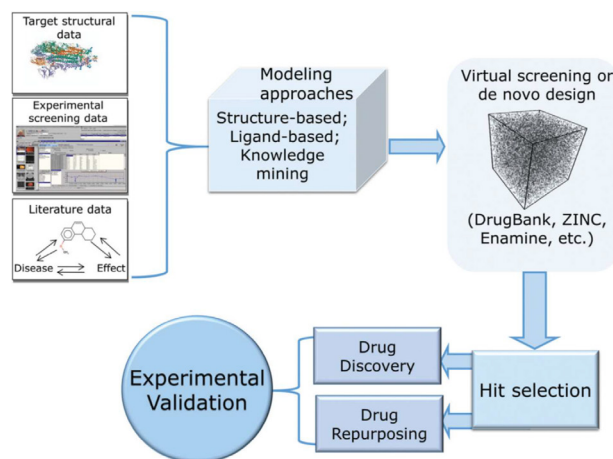

Fig. 2. Common workflow and key approaches in computer-aided drug discovery for therapy against COVID-19. Reproduced from [2] with permission from the Royal Society of Chemistry

3. 1. 1. QSAR methods

Various implementations of QSAR methods in the context of COVID-19 studies have been reported [6, 14, 21, 22].

In brief, a QSAR study can be proceeded through the following steps:

(*i*) collection of experimental data: biological activity parameters IC_{50}, EC_{50}, RBA (Relative Binding Affinity), or binary information – active/inactive, etc., identification of a set of systems with known biological activity;

(*ii*) separation of molecular systems with known activity into two groups: training and test sets;

(*iii*) computation of different molecular descriptors for each molecule from the training set;

(*iv*) formulation of a possible mathematical (statistical) model of biological activity: multilinear and nonlinear regression models, classification rules (discrimination function, logistic regression), artificial neural network;

(*v*) validation of the obtained model using a test set: comparison with experimental data, verification of the predictive ability of the model.

Each step in the above list can include many different approaches. The first step (*i*) involves collecting experimental data on the activity of certain molecules, detailing the molecular structure, and creating an appropriate database. Task (*ii*) divides the sample into test and training samples. Because of such division, statistically representative samples should be obtained. In practice, this stage determines the success of the subsequent validation procedure of the accepted statistical models. The

rational selection of train and test sets has been discussed elsewhere [23, 24].

Task (*iii*) requires the computation of a wide set of descriptors, outlining the most diverse aspects of molecular structure. To date, there are several thousands of such molecular descriptors implemented in computer programs. Among the popular programs for computing descriptors are DRAGON [25] and PaDEL-Descriptor [26]. A contemporary version of DRAGON-7 can calculate up to 5270 molecular descriptors, organized in different logical blocks, which correspond to different aspects of molecular structure. Among these descriptors, in addition to 1-, 2-, and 3-dimensional molecular parameters, such as constitutional, topological and electro-topological, this set also includes some drug-like indices, physicochemical molecular properties, etc. Comprehensive overviews of available descriptors are given elsewhere [27, 28].

The next part (*iv*) is the most complex and multi-faceted. The main problem associated with using a wide descriptor set is likely multicollinearity and redundancy of the data. Therefore, it requires reducing the descriptor set and selecting the most important descriptors relevant to the activity under consideration. To date, many approaches have been considered to accomplish this goal. Among them, classical approaches based on the analysis of the factor structure of the task should be mentioned. However, factor analysis (or its simplified version – principal component analysis) [29] does not correspond to such a purely pragmatic goal as descriptor selection. Nevertheless, such information provides the deepest understanding of mutual correlations of descriptors.

Several methods exist for reducing the descriptor set and constructing compact regression equations. Popular in recent years, the genetic algorithm method allows the generate of equations with a given number of descriptors [30]. In addition to full search and genetic algorithms, methods based on stochastic predictor search strategies, such as the "ant colony" method [31] and the "random forest" method [32], should be mentioned. It should be noted that, unlike regularization approaches, these search methods are prone to overfitting. Some competition to the genetic algorithms approach is a relatively new application to QSAR problem – method LASSO (Least Absolute Shrinkage and Selection Operator), which allows to unambiguously rank the descriptors and construct the corresponding regression equations, including a given number of parameters [33, 34]. Ideologically, the LASSO approach (L1-regularization of Least Squares method) is close to stepwise selection [35] and ridge regression [36] (L2-regularization of Least Squares method). As a result, the approaches mentioned above give the possibility to obtain a function that describes biological activity (BA) as a function of linear combination (1).

$$BA = a_0 + \sum_{i \ll N} a_i d_i, \qquad (1)$$

where the number of descriptors (d_i) must be significantly less than the number of objects/molecules (N), a_i – regression coefficients.

Unfortunately, in a number of cases, it is impossible to obtain a compact regression (1). In this case, the arsenal of chemoinformatics includes regression models based on the use of the factor structure of the descriptor set together with the response of the system (biological activity data). Principal Component Regression (PCR) and Partial Least Squares (PLS) allow to build a regression model without explicitly selecting "important" descriptors [37]. PCR and PLS models can lead to a linear equation in which the number of descriptors can be much larger than the number of objects (molecules) (2).

$$BA = \sum_{j=1}^{n_s} PC_j = a_0 + \sum_{\substack{i=1,M \\ M \geq N}} a_i d_i, \qquad (2)$$

where PC_j – principal components of the correlation matrix, n_s – number of factors to be taken into account.

One possible 3D QSAR regression model is the CoMFA (Comparative Molecular Field Analysis) approach. The CoMFA method is based on the observation that primarily intermolecular effects, which are fundamentally noncovalent and depend mainly on molecular features and shape, determine ligand-target interaction. CoMFA aims is to study the correlations between 3D characteristics of molecules and their biological activity. The 3D molecular descriptors are steric (Lennard-Jones), electrostatic (Coulomb) potential and possibly other parameters (e.g., lipophilicity). A feature of the CoMFA model is that such parameters are calculated as a three-dimensional map describing a given property in space. To describe such fields, the molecule is assumed to be in a three-dimensional spatial lattice defined by nodes. A given property (biological activity) is expressed through the superposition of field values in the nodes of the spatial lattice (3).

$$\log \frac{1}{C} = \sum_i a_i E(r_i) + \sum_i b_i S(r_i). \qquad (3)$$

The CoMFA approach allows not only to describe the relevant molecular property, but also to obtain a visual map giving a view of the "localization" of the property on molecular fragments. Examples of the CoMFA calculations on the COVID-19 problem have recently been reported [38, 39].

Another helpful group of approaches is related to the construction of classification functions. For typical binary classification (e.g. active/inactive, carcinogen/non-carcinogen, etc.), the discriminant function, which is traditional for such purposes, can be used [40]. The linear variant of which (linear discriminant analysis, LDA) is connected with the construction of a hyperplane in the multidimensional space of descriptors, separating active objects from inactive ones. The results of the LDA analysis can be represented as an inequality (4):

$$b_0 + \sum_{i=1,M} b_i d_i > 0, \qquad (4)$$

where b_i is the optimal LDA coefficients that guarantee the separation of active and inactive molecules. A system

with a certain activity (or inactivity, depending on the calibration of the equation) must satisfy (4).

Logistic regression (sigmoid function) is an alternative approach to solving the classification problem [41] (5):

$$p = 1/\left(1 + \exp(-z)\right), \quad z = b_0 + \sum_{i=1,M} b_i d_i. \quad (5)$$

In (5), the parameter $0 \leq p \leq 1$ defines the classification of the object into two classes. The problem of descriptor selection is also essential for LDA and logistic regression. In connection with the classification problem, a relatively new approach should be mentioned, such as support vector machine (SVM) [42]. The main idea of SVM is transforming high-dimensional descriptor space to make the active/inactive groups linearly separable. In the transformed space, the boundaries are constructed in such a way as to maximize the difference between groups.

Moreover, in the context of the classification problem, the kernel approach [43, 44] and the classification tree method should also be mentioned [45]. However, an artificial neural network (ANN) is perhaps the most popular machine learning technique that can be used for QSAR description of a wide range of molecular properties [19]. The recent update on artificial intelligence methods for repurposing and discovering new drugs against COVID-19 is given in [46, 47]. In the case of the classification problem, the formal mathematical description of a single neuron can be represented as (5). The number of neurons in ANN determines the flexibility of the nonlinear predictive model. An example of ANN discrimination of organic molecules by antibacterial activity has been given [48].

Schematically, a neural network for classification or quantitative prediction of molecular properties can be represented as shown in Fig. 3. Neurons are usually divided according to their functional features into input, hidden and output neurons. The topology of a neural network (neural network architecture) can be incredibly diverse [46, 49]. Deep generative models are powerful tools for exploring chemical space, enabling the on-demand generation of molecules with desired physical, chemical or biological properties. To build chemical space with up to 67 million molecules, some chemical language models have been suggested [50–52]. Generative modelling approaches for drug discovery have been reviewed in [53].

There are a significant number of practical applications of ANN to describe the physicochemical properties of molecules (boiling and melting points, vapour pressure, viscosity, heat capacity, flash point, enthalpy of formation, etc.) and biological activity (toxicity, carcinogen activity, antibacterial properties, various medication effects, etc.) [54]. Some comprehensive reviews and a general description of the above-mentioned methods are given elsewhere [46, 55, 56].

Part (v) is the most problematic step of QSAR investigations. The importance of validating QSPR/QSAR models has been pointed out [23, 24, 57]. However, only recently several papers have been published that pro-

posed some coefficients characterizing the quality of the obtained equations [58, 59]. The optimal partitioning of datasets into training and test samples is also closely related to the validation problem. There are a number of publications discussing approaches to such partitioning [60, 61]. Historically, the first validation criteria used to assess the effectiveness of models were a number of indices that are now known as "internal validation" approaches. Among them are the well-known determination coefficient (R^2) and root means squares deviation (RMSD) calculated for a train set. Moreover, corresponding values R^2 and RMSD were obtained by means of Leave–One (or Many) – Out cross-validation and bootstrap methods [62]. The "external validation" mainly uses the calculation of different parameters for a test set.
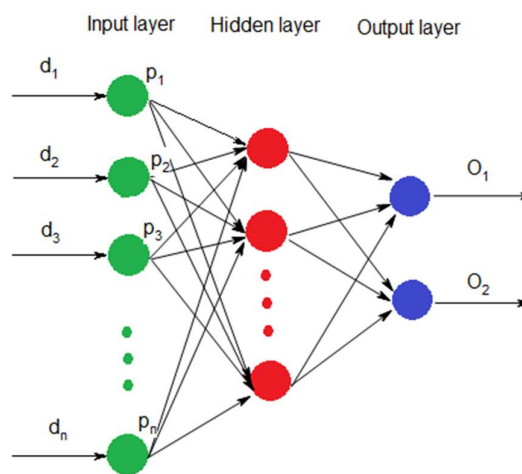

Fig. 3. Schematic representation of neural network with hidden layer

The second group of *in silico* methods in drug discovery is based on molecular modelling (MM). MM includes computational approaches that focus on direct ligand-protein modelling and the screening for the identification of promising molecular structures. In general, the following steps can outline a procedure for such modelling:

(*a*) identification of target (proteins, binding sites), collection of experimental screening data, collection of drug activity data, database building or database selection;

(*b*) virtual screening (VS);

(*c*) structure-based modelling, docking;

(*d*) hit identification;

(*e*) lead generation (optimization);

(*f*) validation *in vitro*, *in vero*, *in vivo*.

For step (*a*), there are many important databases containing information on the biological activity (especially with respect to COVID-19) and theoretical studies of known compounds. The most important databases are given in Table 1.

In step (*b*), these data from public databases as well as local databases can be used for subsequent screening against selected proteins. The goal of VS is to choose subsets of chemical libraries so that they are enriched with compounds that have the desired affinity for a given target.

Table 1

Online databases and chemical libraries.

Database	Description	Reference
ChEMBL	Database of bioactive molecules with drug-like properties containing chemical, bioactivity and genomic data	https://www.ebi.ac.uk/chembl
PubChem	Collection of freely accessible chemical information, such as chemicals by name, molecular formula, and structure	https://pubchem.ncbi.nlm.nih.gov/
DrugBank	Database for drug interactions, pharmacology, chemical structures, targets, metabolism	https://go.drugbank.com/
KEGG	Database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem	https://www.genome.jp/kegg/
Kaggle	Data science community with powerful tools for machine learning	https://www.kaggle.com/
U.S. Food and Drug Administration (FDA)	The FDA is responsible for protecting the public health by ensuring the safety, efficacy, and security of human and veterinary drugs, and biological products	https://www.fda.gov
Therapeutic Target Database	Database providing information about the known and explored therapeutic protein and nucleic acid targets, the targeted disease	http://db.idrblab.net/ttd
PharmGKB	Comprehensive online resource for curating knowledge about the impact of genetic variation on drug response for clinicians	https://www.pharmgkb.org
BindingDB	Public molecular recognition database supporting research, education and practice in drug discovery, pharmacology and related fields	https://www.bindingdb.org/bind/index.jsp
STITCH	Database of proteins from 630 organisms to over 74,000 different chemicals, including 2,200 drugs	http://stitch.embl.de
Protein Data Bank	On-line public database of X-ray structure of proteins and their complexes	https://www.rcsb.org https://pdb101.rcsb.org

In this review, we separate docking-based approaches from those based on molecular similarity assessment, which we consider virtual screening methods. Usually, VS is based on available quantitative information about 3D protein structure, corresponding binding site and the reference (core) ligand. There currently are a large number of approaches to VS. They can be categorized into several groups; however, conceptually, VS is based on similarity/diversity concept [63].

3. 1. 2. Similarity-based virtual screening

According to the similarity principle, structurally related molecules should exhibit similar biological activity. As the similarity parameter, one can use the value S_{AB} (6) [64, 65]:

$$S_{AB} = \int F_A(r) F_B(r) \, dr \left/ \sqrt{\int F_A(r) F_A(r) \, dr \int F_B(r) F_B(r) \, dr}, \right. \quad (6)$$

where $F_A(r)$ and $F_B(r)$ are the molecular fields connected to the tested (A) and the target (B) molecules. The functions $F(r)$ represent atom-centered steric ($S(r)$) and electrostatic fields ($E(r)$), so (6) can be viewed as a "distance" or correlation coefficient between molecular fields. It is evident that the similarity parameter S_{AB} depends on the mutual superposition of molecules and their conformations and can take the value parameter within a range of $0 < S_{AB} < 1$ for positively defined fields. To account for both steric- and electrostatic-fields, the similarity index can be defined as a weighted value of the two components $F(r)$ by (7).

$$F(r) = w_1 S(r) + w_2 E(r). \quad (7)$$

A method based on the Tanimoto coefficient is similar to the above approach. Chemical structures can be characterized using binary data collected on the corresponding

vectors $|x_A\rangle = |0101001100101.......\rangle$ (for molecule A) and for molecule B. These vectors (fingerprints) describe the presence (or absence) of certain elements of the chemical structure [66, 67]. The fingerprints are popular elements for VS [68, 69]. The fingerprints can be obtained using various descriptor calculation programs (PaDEL-Descriptor [26], DRAGON [25]) and RDKit, which is an important contemporary tool for chemoinformatics [70].

The Tanimoto index can then be calculated from a particular fingerprint as (8):

$$T_{AB} = \frac{c}{a+b-c} = \frac{\langle x_A | x_B \rangle}{\langle x_A | x_A \rangle + \langle x_B | x_B \rangle - \langle x_A | x_B \rangle}, \quad (8)$$

where a is the total number of yes (or "1") bits in molecule A, b is the number of "yes" bits in molecule B, and c is the number of "yes" bits in both structures A and B. A larger value corresponds to a greater similarity of the compounds. An example of a direct similarity search for SARS-CoV-2 Mpro inhibitors is given in [71], where the Tanimoto index was calculated using Morgan fingerprints [72], which is probably the most popular in contemporary investigations. A Tanimoto similarity threshold of 75 % was used to select several compounds from DrugBank as potentially active substances. Different fingerprint similarity-based approaches for VS are described elsewhere [73–75].

There are also a number of different approaches to VS based on graph theory indices and other methods. Some instructive examples have been described in [76, 77].

3. 1 .3. Pharmacophore-based virtual screening

Virtual screening based on pharmacophores is an important approach in modern drug design [78, 79].

Pharmacophores are usually defined as important chemical features corresponding to efficient protein-ligand interactions [80]. In other words, a pharmacophore can be understood as regions in three-dimensional space where certain functional groups can lead to effective intermolecular interactions. These regions are as follows:

– hydrogen bond donor regions. Recognize the structural elements (and spatial regions) of promising hydrogen bonding donor ligands located in the corresponding sites favourable for the realization of hydrogen bonding. The region for donors or acceptors for hydrogen bonds in a protein corresponds to a typical distance of 2.5–3.5 Å, and the angle heteroatom donor-hydrogen-acceptor should be 120° or more;

– hydrogen bond acceptor regions. Recognize the elements of ligand and corresponding region with acceptor of hydrogen bond;

– hydrophobic interaction region. These regions are located near the hydrophobic fragments of the protein surface and correspond to the centers of the most hydrophobic groups of ligands – hydrocarbon rings (benzene, aliphatic), hydrocarbon radicals and others. The hydrophobic regions are located at a distance from 3.5 to 5.0 Å from the nearest protein atom;

– positively and negatively charged groups. Start with the new paragraph, there are several commercial programs that give a possibility to analyze a binding site and formulate pharmacophore hypothesis. Among them, the popular tools are MOE (Molecular Operating Environment) [81] and LigandScout [82]. As an example, we represent the pharmacophore structure obtained by means of LigandScout in Fig. 4, *a, b* for Mpro complex with a ligand. In Fig. 4, *a*, green arrows denote hydrogen bond donors, red arrows correspond to hydrogen bond acceptors, and yellow regions denote hydrophobic regions.

VS offers the possibility to significantly reduce the set of ligands; however, the most direct way to find active ligands is the docking procedure (structure-based approach) (step *c*). Docking calculations can be viewed as a model of the molecular recognition process that plays a crucial role in protein-ligand interactions [83].

The most crucial aspect of docking is the selection of the Scoring function. The scoring function gives an evaluation of a ligand-to-protein affinity. In modern docking studies, various scoring functions are usually calculated. Some of the most important functions are the following:

– force-field scoring. It describes the interaction energy as the sum of the Coulomb (electrostatic), van der Waals (Lennard-Jones potential), tensile, bending and angular torsional components of the coupling by means of a definitely chosen molecular mechanic force field (AMBER, TRIPOS, etc.) [84];

– knowledge-based scoring. This scoring function utilizes information from a database of ligand-protein complexes. It is designed to reproduce experimental structures rather than binding energies. Analysis of such information yields the paired atom-atom potential [85];

– empirical scoring. In this approach, scoring functions are fitted to reproduce experimental data, such as binding energies and conformations energies. It is based on counting and evaluating certain types of interactions corresponding to functional groups in contact. The total contribution can be divided into components: electrostatic, hydrophobic-hydrophilic, and hydrophobic-hydrophobic interactions. The enthalpy and entropic components are also taken into account in the calculation of Gibbs energy [86, 87];

– machine learning scoring. This approach, which is currently the most promising, does not assume an unambiguous functional form describing the ligand-protein interaction [19]. The results for the scoring function are estimated from previously obtained data for different supramolecular complexes. The machine learning approach to scoring functions demonstrated high accuracy. In particular, it was stated that the Random Forest (RF) approach was used to implicitly capture binding effects that are hard to model explicitly [88]. The SVM regression, as well as RF, were reported in [89];

– consensus docking (scoring). In this approach, several scoring functions are combined to balance errors from individual functions. The method is based on evaluating different poses to obtain the more correct one [90].
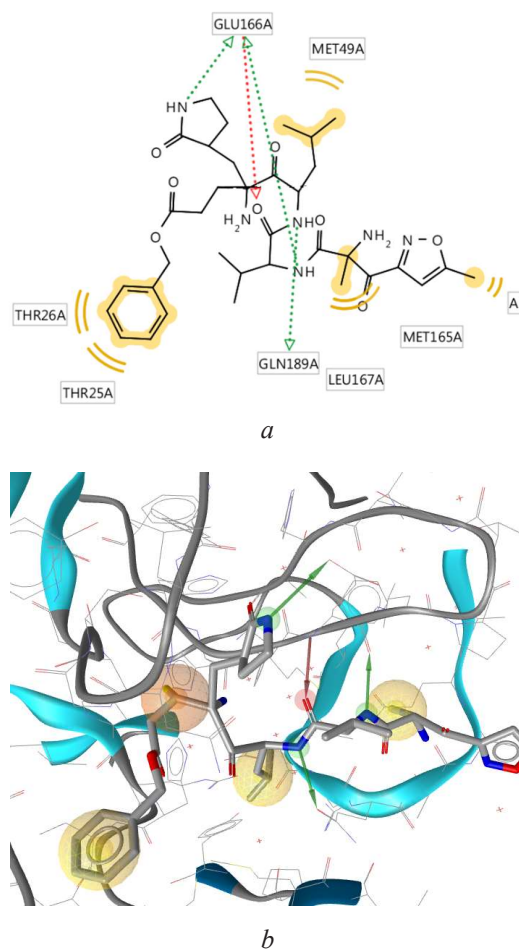

a

b

Fig. 4. Different pharmacophore models generated by LigandScout for 6LU7 Mpro complex with a ligand: *a* – 2D pharmacophore; *b* – 3D pharmacophore model

Comprehensive reviews of the description [91–94] and comparison of different scoring functions are given in [67–70] and references therein. A pharmacophore-guided deep learning approach for bioactive molecular generation has been discussed in [95].

An essential stage of rational drug design is lead generation (step *e*). At this stage, new structures are generated from hit structures obtained in the previous stages of the study (virtual screening, docking) [96]. Scaffold-constrained molecular generation is used to obtain a set of systems with higher activity. Usually, the term "scaffold" is defined as the basic (core) structure of a molecule or set of molecules. A scaffold can be thought of as a system with a specific shape and number of functional groups corresponding to a particular target. There are a number of approaches (usually implemented in corresponding computer programs) to efficiently work with chemical space. The most interesting approaches are the following:

– Scaffold Hunter is a computer program for hierarchical structuring, visualization, and analysis of complex structures and the bioactivity of data [97]. The program reads the data, identifies a biochemistry-related compound, and then iteratively removes one ring at a time from the "child" structure, creating a "parent" structure.

The importance of the scaffold concept is connected with a set of restrictions imposed on designed molecules. Without such restrictions, the probability of obtaining molecules with the desired activity corresponding to the target is low. There are several effective algorithms for scaffold generation. A new SOMOA (Scaffold Constrained Molecular Generation) algorithm has been proposed [98]. A SMILES-based recurrent neural network (RNN) [50] is used to generate new molecular structures optimized for different properties while exploring only the relevant chemical space;

– ScaffoldGraph is an open-source Python library for the generation and analysis of molecular scaffold networks that can be further optimized into drug molecules [99]. We use ScaffoldGraph to identify scaffolds that are statistically enriched for bioactivity utilizing a method called compound set enrichment;

– Scaffold Generator is an open-source Java library for generating, processing and displaying molecular scaffolds, scaffold trees, and networks [100]. The main function of the Scaffold Generator tools is to construct trees and networks of scaffolds from given collections of molecules;

– Scaffold hopping is a procedure of exchanging one scaffold for another while maintaining molecular features that are important for biological properties [101]. The scaffold hopping primarily focuses on existing structures [102, 103].

In practice, molecular docking calculations are significantly software-dependent (Table 2). The AutoDock suite (MGLTools) [104] requires that some settings, such as the selection of docking cell, grid space, etc., have to be performed manually. On the other hand, most of these procedures are implemented automatically in the LigandScout suite [82]. A critical assessment of docking programs and scoring functions in terms of reproducing crystallographically determined protein/ligand complex structures is available in [93].

Table 2

Available software and online tools for molecular docking calculations

Software	Description	References
AutoDock	AutoDock is open-source molecular modelling simulation software for protein-ligand docking	[105]
AutoDock Vina	AutoDock Vina uses an improved local search routine and a sophisticated gradient optimization method, achieving approximately two orders of magnitude improvement in speed and better accuracy in predicting binding modes compared to AutoDock	[106, 107]
GOLD	GOLD is the validated, configurable protein–ligand docking software for expert virtual screening, lead optimization and drug discovery	[108]
Schrödinger	The Schrodinger suite of software has a large number of applications for a variety of modeling, analysis and computational tasks. It enables using both ligand and structure-based methods. A license is required	Schrödinger Release 2023-3: Maestro, Schrödinger, LLC, New York, NY, 2023
LigandScout	LigandScout is computer software for an advanced molecular design that utilizes a concept of creating 3D pharmacophore models derived from structural data of macromolecule–ligand complexes or from sets of ligands. A license is required	[82]
SwissDock	SwissDock is a web service to predict the molecular interactions that may occur between a target protein and a small molecule. http://www.swissdock.ch	[109]
FlexX	FlexX is an automated docking tool using ultra-high-speed approach to predict the binding mode of ligands at a target binding site	[110]
MOE	Molecular Operating Environment (MOE) is a drug discovery software platform that integrates visualization, modeling and simulations, as well as methodology development, combined in one program package. A license is required	[81]

Numerous molecular docking studies have been reported in the context of repurposing and discovering potent non-covalent and covalent inhibitors for critical proteases of the virus SARS-CoV-2 [111–118]. Some excellent reviews of this topic are given elsewhere [1, 2, 7, 119].

In successful examples of drug repurposing studies, pharmacophore-based virtual screening has been utilized in combination with other tools, such as machine learning, MD simulations, and *in vitro* experimental validation assays (Fig. 5, A). Such a combined approach has been used for discovering small-molecule inhibitors of Mpro with micromolar activity by applying pharmacophore-based VS for identifying functional chemical groups responsible for the molecular recognition of ligands by the Mpro binding pocket (Fig. 5, B, C) [79].

In another instructive example, pharmacophore-based virtual screening of 2122 FDA-approved drugs was performed for repurposing against Mpro protease of SARS-CoV-2 [120]. The docking with flexible active site residues allowed to identify seven hit candidates, among which Boceprevir and Telaprevir (HCV drugs) and Nelfinavir (HIV drug) revealed high potency in an inhibition assay [120] (Fig. 6).

A series of computational tools were employed for redesigning perampanel, which is the known hit inhibitor of Mpro protease, for discovering multiple noncovalent, nonpeptidic inhibitors with ca. 20 nM IC_{50} values [121]. The study was initiated from 14 known drugs, as shown in Fig. 7. The suggested ligand designs were confirmed and augmented by resolving of high-resolution X-ray crystal structures.
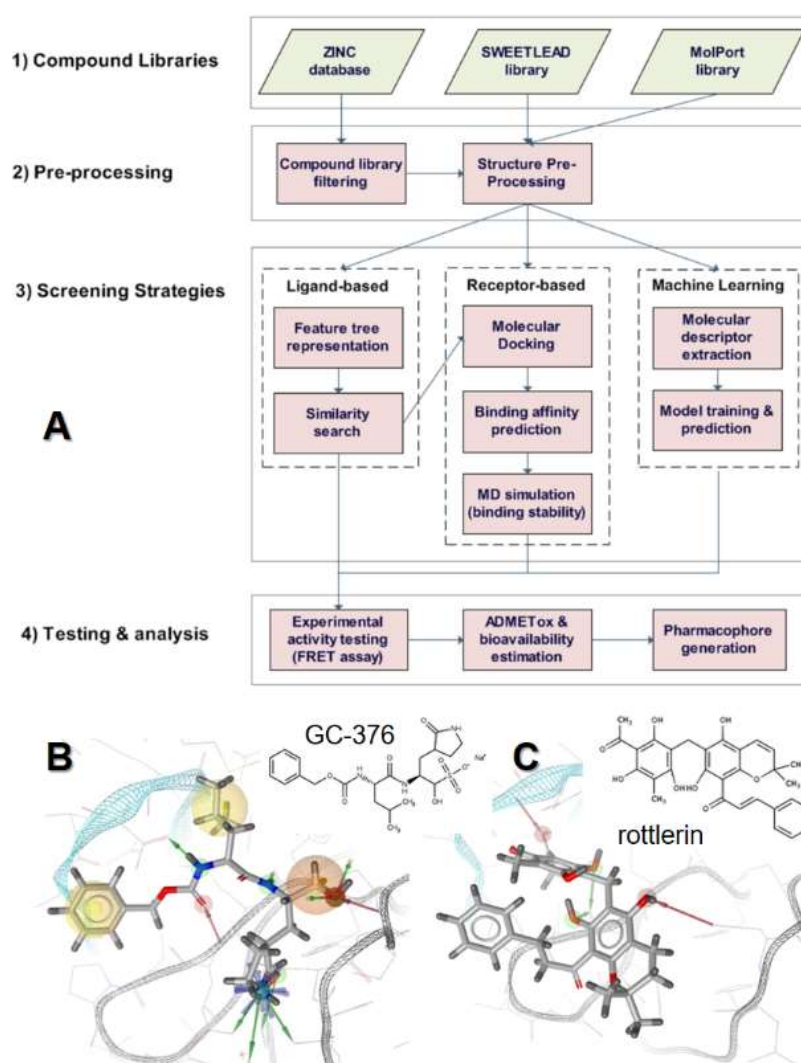

Fig. 5. Example of pharmacophore-based *in silico* screening: A – a workflow for a screening procedure, operating with a compound library over 1 billion compounds from the databases ZINC, SWEETLEAD, and the MolPort. The selected compound repository was pre-filtered focusing on commercially available and drug-like chemicals. The reduced dataset was a subject of receptor-based screening followed by ligand-based screening. More computationally expensive approaches, such as receptor-based and machine learning-based screening were only applied to the smaller libraries; B – a pharmacophore model for Mpro derived from the reference inhibitor GC-376/GC-373 (PDB ID 7D1M) using the LigandScout suite; C – a pharmacophore model for Mpro derived from the inhibitor Rottlerin. Adopted with permission from [79]. Copyright © 2022 The Authors. Published by American Chemical Society
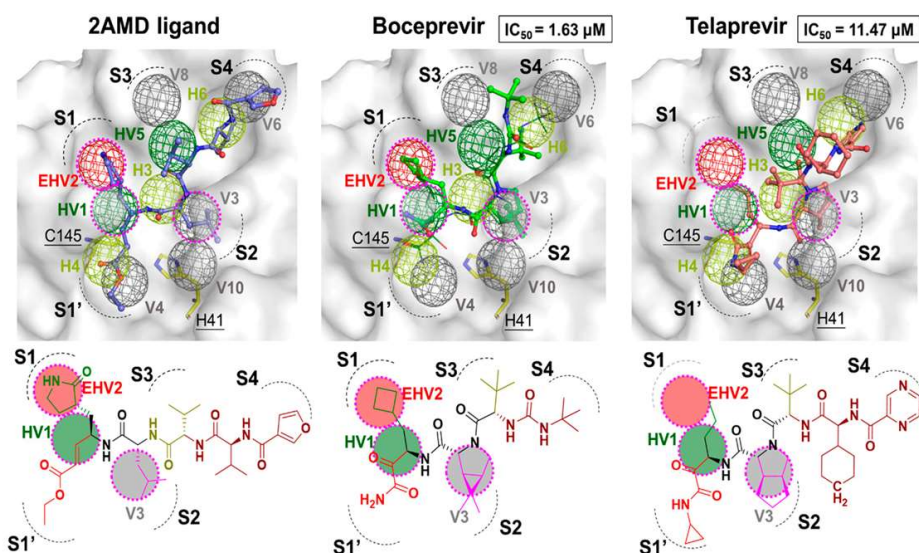
Fig. 6. Binding poses of Boceprevir (green ball and stick, 6WNP crystal pose, dark green line) and Telaprevir (salmon ball and stick) in the PPC1:2AMD cavity (with the reference 2AMD ligand crystal pose, violet ball and stick). 2D diagrams highlight the compound moieties. Adopted with permission from [120].
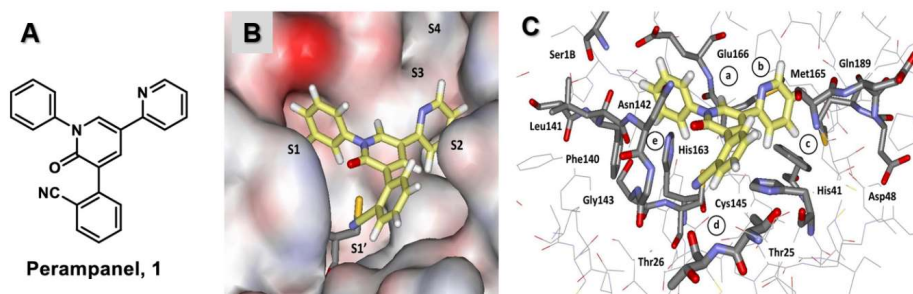Copyright © 2022 American Chemical Society

Fig. 7. Redesigning Perampanel using computational tools: A – molecular structure of Perampanel; B – docked structures for perampanel with Mpro. Surface rendering noting the binding pockets colored by atomic number; C – key interacting residues near the binding site including the catalytic residues His41 and Cys145. Adopted with permission from [121]. Copyright © 2022 The Authors. Published by American Chemical Society

3. 2. Molecular dynamics simulations of protein-ligand complexes

The molecular dynamics (MD) simulation method calculates and analyzes the physical movements of atoms and molecules for a fixed period of time. The MD trajectories of atoms and molecules are calculated by numerically solving Newton's equations of motion for a system of interacting particles, where forces between the particles and their potential energies are often calculated using interatomic potentials or molecular mechanical force fields [122]. The fundamentals of MD simulations are given elsewhere [123].

In the last few years, MD simulations have become one of the most widely used computational methods in biomolecular simulation. The main advantages of an MD technique are the following:

1) it uses empirical molecular mechanics (MM) force fields, which are very low-CPU/GPU consuming;

2) it can now be utilized to explore time-dependent phenomena in atomic detail over microseconds, which are within the time scale of many biochemical processes. Popular MD software and its capability are summarized in Table 3.

Atomistic MD simulations are a powerful tool to examine the stability and conformational dynamics of proteins and its complex with ligands. Root means square deviation (RMSD) measures the deviation of a set of coordinates of a protein to a reference set of coordinates. RMSD are calculated by least-square fitting the instant structure (τ_2) to its crystallographic or pre-equilibrated structure ($\tau_1 = 0$) by (9):

$$RMSD\left(\tau_1, \tau_2\right) = \left[\frac{1}{N}\sum_{i=1}^{N}\left\|r_i\left(\tau_1\right) - r_i\left(\tau_2\right)\right\|^2\right]^{\frac{1}{2}}, \qquad (9)$$

where N and $r_i(\tau)$ are the numbers of atoms, and the position atom i and its reference position at time τ [125, 131].

Numerous examples of using MD simulations for resolving SARS-CoV-2 protein structure, folding, binding hotspots and functions have been reported [117, 132–139]. In these MD studies, a typical sampling time scale was about hundreds of nanoseconds; however, in some reports, it has already reached up to 100 μs, as in the case of recent MD simulations of allosteric inhibition of the SARS-CoV-2 Mpro [140].

Table 3

Popular molecular dynamic (MD) simulation software

Software	Description	References
GROMOS	GROningen MOlecular Simulation (GROMOS) software package for classical MD simulations developed the University of Groningen and the Swiss Federal Institute of Technology (ETH Zurich). A license is required	[124]
GROMACS	GROMACS (Groningen MAchine for chemical simulation) is an efficient and versatile open-source program for classical MD simulations suited for the simulation of biological (macro) molecules in aqueous environments. It is compatible with various force fields such as GROMOS, OPLS, AMBER, and ENCAD. All-atom, atomistic and coarse-grained approaches are implemented	[125]
NAMD	NAMD is a high-performance biomolecular simulation program based on open-source codes optimized for executing at parallel supercomputers and workstation clusters. It is compatible with CHARMM force field	[126]
CHARMM	CHARMM (chemistry at HARvard molecular mechanics) is a MD simulation program that is primarily designed to study biological molecules, such as proteins, peptides, lipids, nucleic acids, carbohydrates, and small molecule ligands	[127]
AMBER	Amber is a biomolecular simulation suite divided into three major step-system preparation (antechamber, LEaP programs), simulation (sander), and trajectory analysis (ptraj analysis program). A license is required	[128]
LAMMPS	LAMMPS (large-scale atomic/molecular massively parallel simulator) is an open-source classical MD program for material science modeling. It has a library of intermolecular potentials for soft matter (biomolecules, polymers), solid-state materials (metals, semiconductors), and coarse-grained or mesoscopic systems	[129]
DL_POLY	DL_POLY is a general-purpose MD simulation package for studying of liquids of large complexity	[130]

Another example is given by homology modeling, combined with all-atom MD simulations and molecular mechanics Poisson-Boltzmann surface area (MMPBSA) scanning, which were performed to explore the trimeric form of the Spike protein and its interface with human ACE2 [141]. Twenty interacting residues in the Spike protein have been identified as responsible for tightly binding to ACE2 [141].

3. 2. 1. Stability of protein-ligand complexes

In the last two years, numerous studies of atomistic MD simulations of the structure and dynamics of Mpro and PLpro proteases, Spike-protein and their complexes with non-covalent inhibitors have been performed [117, 136, 139, 142–157]. Different physicochemical properties, such as site-mapping and pandemic mutations, were subject to investigation [142, 158]. The role of temperature in the thermal stability of the enzyme and its complexes was evaluated [148]. Depending on research strategies and used software, various biomolecular force fields were utilized, including AMBER [134, 141, 142, 147, 148, 152, 159], OPLS-AA [143, 144, 153], CHARMM27 [149, 158], CHARMM36 [145, 146, 151, 154], and Gromos G43a1 [156, 160] and G54a7 [150], respectively.

All-atom MD simulations have been performed to examine the conformational behavior of the monomeric and dimeric form of Mpro structure by microsecond time scale MD sampling [147]. The stability of the X-ray structure (PDB ID 6LU7) of Mpro was probed for monomeric and dimeric forms by utilizing the Amber99 force field (Fig. 8). It was found that the deviations of the equilibrated structures from their starting crystallographic coordinates measured as RMSD for the backbone C_α atoms were at 0.3–0.5 nm after

1.5 µs (Fig. 8, A), which suggest a relatively high structural stability. The residue-based fluctuations revealed that the most flexible protein regions were located at the loops within residues 40–65 and 180–200 (Fig. 8, B), respectively. The distance distribution for C_α atoms in the His-41 and Cys-145 demonstrated two distinct sub-states at 0.26 nm and 0.5 nm, which corresponded to some enlargement compared to the starting distance in the X-ray structure (0.19 nm in chain A and 0.26 nm in chain B) (Fig. 8, C–D).
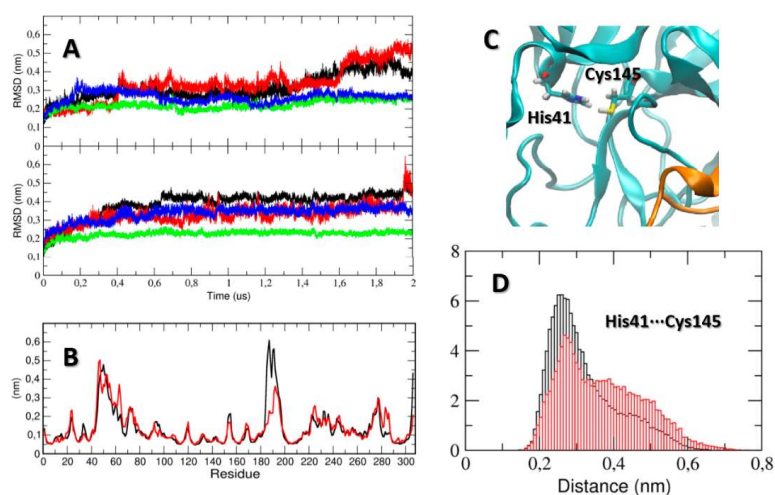

Fig. 8. All-atom MD simulations of the monomeric and dimeric form of Mpro: A – root means square deviation (RMSD) of C_α atoms during MD simulation of the monomer (upper graph) and the dimer (lower graph). The total RMSD is given in black, whereas those of the domain I, II, and III are color-coded; B – comparison of residue-based root mean square fluctuation (RMSF) of the monomeric and dimeric Mpro. RMSF plots for C_α carbons of chain A and B are presented in black and red, respectively; C – the catalytic dyad residue His41 and Cys145; D – the distance distribution between the residues His-41 and Cys-145 during MD simulations of the dimeric Mpro, in which chains A and B are plotted in black and red, respectively. Adapted with permission from [147]. © 2020 by the authors. Licensee MDPI, Basel, Switzerland

A typical protocol starts with the best affinity docking structure of a protein-ligand complex with the highest docking score and lowest root-mean-square deviation (RMSD), which is then employed for further evaluation by atomistic MD simulations in explicit aqueous solution [134, 135, 139, 156, 157, 161].

Atomistic MD simulations of the effects of the ensemble-based mutations on binding for the SARS-CoV-2 Spike-protein/RBD complexes to various nanobodies were performed to identify dynamic, energetics, and binding affinity fingerprints [132].

3. 2. 2. Binding free energy calculations of protein-ligand complexes

One of the remaining bottlenecks in computer-aided drug design is the initial selection and further optimization of the identified hits. One of the computational approaches that can guide these screening efforts is binding free energy calculations of ligand-protein complexes for newly selected hit candidates [162]. In addition, alchemical free energy (AFE) calculations can offer high accuracy at a low computational cost, keeping a growing interest in their application during hit-to-lead optimization campaigns [163].

Free energy calculations based on MD simulations are computationally expensive, and, therefore, they are often applied to explore smaller sub-sets of compounds preselected by *in silico* screening of the larger chemical spaces. This technique is up-and-coming in hit-to-lead and lead optimization studies.

In thermodynamics, the binding free energy ΔG_b (in kcal/mol) is the measure of the interaction affinity between a ligand and target protein (10), which is defined by the dissociation constant K_D (in mol/L) of a complex:

$$\Delta G_b = k_B T^* \ln K_D, \tag{10}$$

where k_B is the Boltzmann constant in kcal/K and T is the temperature in K. Therefore, K_D is a measure of the affinity of a ligand towards its protein target. In equilibrium, K_D equals the ligand concentration at which half of the protein molecules are ligand-bound. ΔG_b is related through the K_D to the binding affinity of a drug candidate.

Several computational methods have been suggested for calculating ΔG using MD simulations:

– the linear interaction energy (LIE) [164];

– fast pulling of ligand (FPL) [165];

– molecular mechanism-Poisson−Boltzmann (generalized Born) surface area (MM-PB(GB)SA) [166];

– non-equilibrium molecular dynamics (NEMD) [167];

– thermodynamic integration [168];

– free energy perturbation (FEP) approaches [169].

Accurate estimation of absolute binding free energy relies on several subprocesses that require human intervention for introducing geometric restraints, defining the collective variables, and performing post-treatments.

Therefore, a novel approach for automating and streamlining free-energy calculations, referred to as the binding free energy estimator (BFEE2) [170], has been suggested. It provides both standardized alchemical and geometrical workflows (Fig. 9) for protein–ligand and host–guest complexes. In addition, BFEE2 supports such popular force fields as CHARMM, Amber, OPLS, and GROMOS, respectively [170].
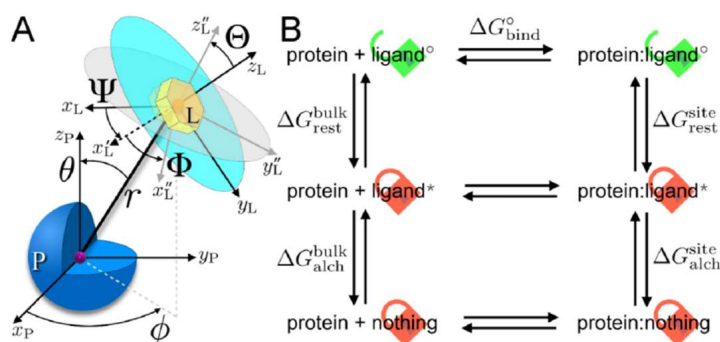

Fig. 9. Alchemical and geometrical workflows for binding free energy estimation: A – definition of collective variables describing the degrees of freedoms of the protein−ligand complex; B – the alchemical absolute binding free-energy calculation workflow. Reproduced with permission from [170]. Copyright © 2021, American Chemical Society
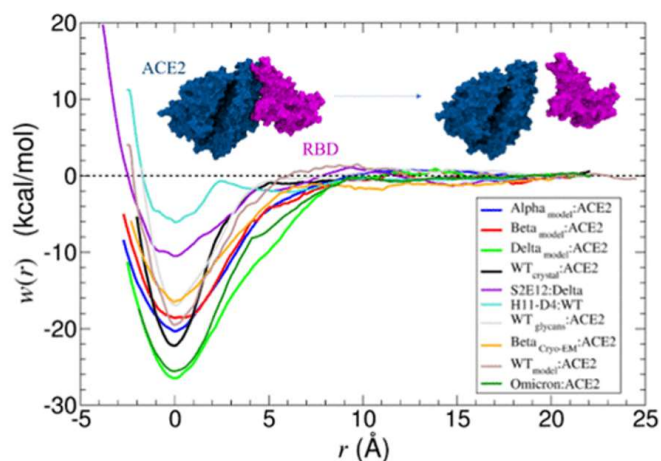

Fig. 10. Free-energy profiles for separation of ACE2 and RBD SAR-CoV-2 proteins estimated by potential-of-mean-force MD simulations. Adopted with permission from [172]. Copyright © 2022, American Chemical Society

One of a strategy for free-binding energy estimation is the so-called "geometrical route", in which two molecular components (protein-ligand, protein-protein) are progressively separated from one another in the presence of conformational restraints [171]. Recently, this approach has been used for studying key SARS-CoV-2 proteins. The standard binding free energy of the receptor-binding domain of the most widespread variants, namely, Alpha, Beta, Delta, and Omicron BA.2, as well as the wild type (WT) in complex with ACE2 have been determined using a rigorous theoretical framework that combines MD simulations and potential-of-mean-force calculations, as shown in Fig. 10 [172].

3. 2. 3. Hybrid MM/QM Approaches

Popular protocols based on molecular docking of best protein-ligand structures in combination with MD simulations of their stability in explicit water solutions can reveal interaction fingerprints that potentially hold promises for designing novel potent drugs. Moreover, the modern biomolecular simulations based on classical molecular mechanics (MM) force fields, such as OPLS-AA, CHARMM, and AMBER, can reproduce the majority of experimental properties of biomolecular systems with high accuracy [139]. However, some crucial aspects of covalent bond formation and breakage within catalytic pockets are still beyond the realm of classical force fields, so these calculations require quantum-chemical formalism (i.e., electronic and nuclear interactions) to be taken into account. Therefore, some hybrid MM/QM approaches have also been suggested.

A recent study used a hybrid MM/QM approach to select a subset of FDA-approved drugs against Mpro protease [173]. The workflow is summarized in Fig. 11, A and initiated with molecular docking of 1615 FDA-approved drugs, ending up with pre-selected 62 candidates. In the next step, classical MD simulations based on the CHARMM36 force field sampled these docked protein-drug complexes (Fig. 11, B), further reducing the candidate pool up to 26 molecules. In the final stage, the selected 26 drug molecules were treated by a pseudo-quantum mechanical (ANI) force field, which was trained by neural network models on DFT (wB97X/6-31G(d)) calculation data points so that free energy MD analysis (MM/PBSA) reduced the final selection up to 9 drugs only [173].

Another example is given by long-scale 5 μs steered-MD simulations aiming to examine the binding of PF-07321332 inhibitor to SARS-CoV-2 Mpro. These steered potential guided MD simulations revealed the important role of the ligand pyrrolidinyl group and the protease residues Glu166 and Gln189 in the ligand-binding process (Fig. 12). The MD-refined structures were further studied by QM/MM calculations to unravel the reaction mechanism for the formation of the thioimidate product between Mpro and the PF-07321332 inhibitor [174].

The thermodynamics of the catalytic reaction in SARS-CoV-2 Mpro, activated by a proton transfer (PT) from Cys145 to His41, has been studied by a hybrid quantum/classical MD approach and the perturbed matrix method (PMM) [175]. The proposed approach offered a cost-effective procedure for identifying a few key sites and some specific water molecules, which play a significant role in enhancing or reducing the thermodynamic feasibility of the PT reaction by selective desolvation of the active site of Mpro [175].
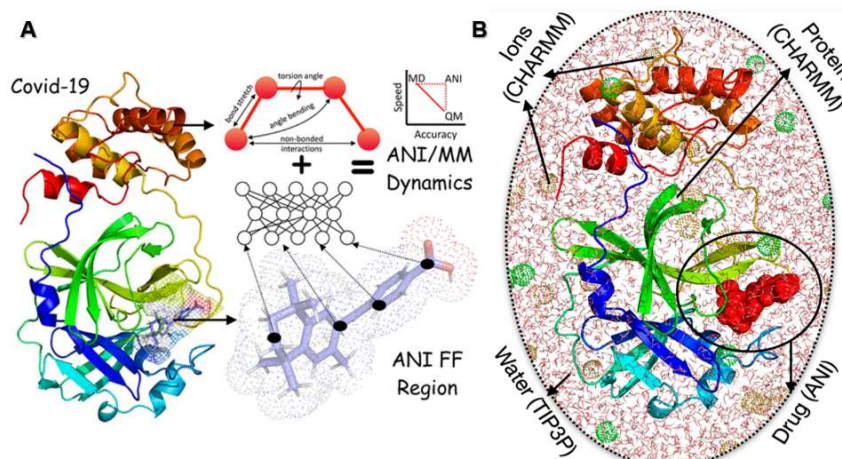

Fig. 11. Workflow of a hybrid MM/QM approach: A – scheme of hybrid ANI/MM MD simulations of Mpro; B – an example of the simulated system consisting of a drug (red surface) bound to the enzyme and solvated with TIP3P water molecules and Na$^+$ and Cl$^-$ ions. Reproduced with permission from [173]. Copyright © 2020 American Chemical Society
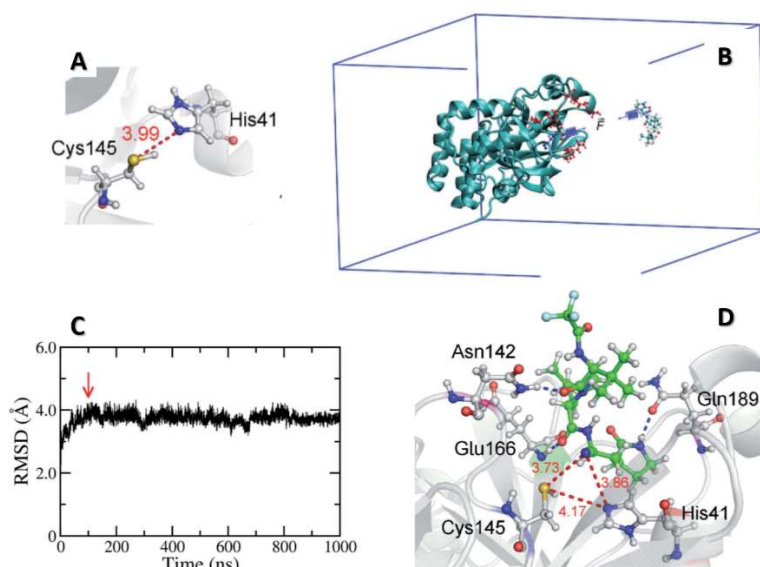

Fig. 12. Steered-MD simulations of inhibitor PF-07321332 binding to SARS-CoV-2 Mpro: A – scheme of the protonation states of the catalytic dyad in Mpro. The distance between the Cys145-Sg and His41-N3 atoms are given from the X-ray data (PDB ID 7JYC); B – 3D-structure of the SARS-CoV-2 Mpro+PF-07321332. Scheme of setting up a constant force for steered MD simulations; C – a RMSD plot of the complex SARS-CoV-2 Mpro+PF-07321332 over MD simulations; D – the representative structure of the equilibrated complex. Dotted lines show H-bonds between the ligand and residue Asn142, Glu166, and Q189, respectively. The distances between atoms are given in Å. Adapted and reproduced from [174] with permission from RSC Advances and the Royal Society of Chemistry

4. Conclusions

Current computational modelling methodologies for the discovery of novel agents against COVID-19 utilize complex workflows combining several tools, such as ultra-virtual pharmacophore-based screening of large chemical spaces, molecular docking of pre-screened candidates and atomistic MD simulations of protein-ligand complexes for hit-ligands, followed by binding free energy calculations [176].

From a short-term perspective, *in silico* screening approaches are going to use a synergetic effect of both physics-based molecular docking and data-based scoring functions, which was reflected in the latest 3DR Grand Challenge 4 results for ligand IC_{50} predictions [177]. To accelerate *in silico* screening of ultra-large chemical libraries, some hybrid iterative approaches have been suggested, such as DeepDocking [178], in which structure-based docking of a sparse library subset is further utilized for training machine-leaning models, which, in turn, are used for filtering the whole library to reduce additionally its chemical space size [8].

From a long-term perspective, further growth of readily accessible commercial and proprietary chemical space libraries is expected [51, 179]. It has been reported that the xREAL extension of Enamine REAL Space currently consists of 173 billion compounds [8]. This number is expected to be further expanded up to 10^{15} compounds by utilizing larger building block sets and four/five-component scaffolds [8]. However, focusing on specialized ultra-large libraries designed for specific scaffolds over general-purpose on-demand spaces seems more promising. In the last decade, the synergy between experiments and atomic-scale MD simulations, accelerated by dramatic increases in computational power with the advent of GPUs and cloud computing, has opened opportunities for mesoscale all-atom MD simulations of the whole viruses (Fig. 13) [180].

Taking into account all the click-like chemistry challenges, the enormous abundance and diversity of drug-like on-demand chemical spaces, some transformations are now occurring from computer-aided to computer-driven drug design [181]. Instead of a traditional view at computational chemistry and structure-based design, as a subset of tools capable of aiding acceleration of the drug discovery process, they become now regarded as a driving force in small molecule drug discovery for urgent therapy against coronavirus SARS-CoV-2.
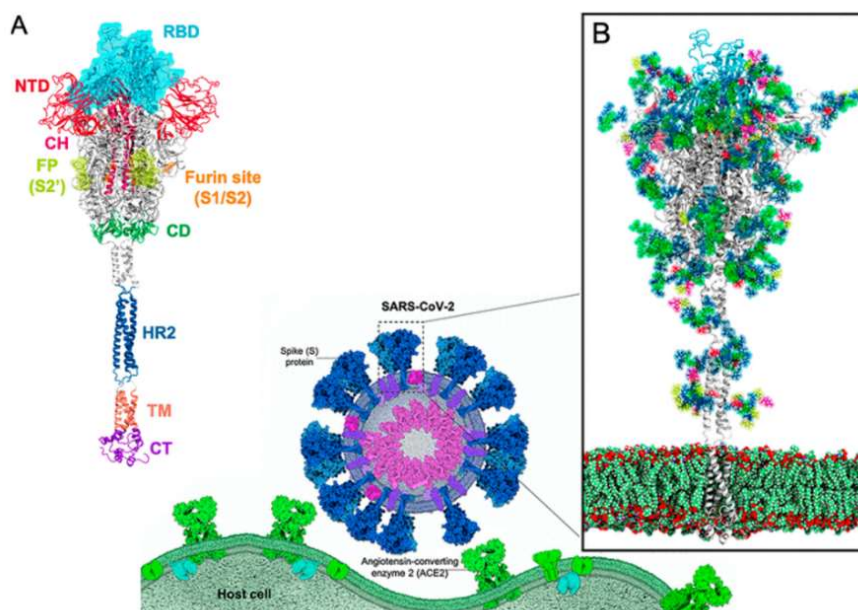

Fig. 13. Mesoscale all-atom MD simulations of viruses: A – 3D-structure of the full-length Spike protein in the open conformation (PDB ID: 6VSB); B – MD snapshot of glycosylated Spike protein embedded in a lipid membrane. (Front) Scheme of recognizing coronavirus SARS-CoV-2 of the host cell by interacting with the glycosylated Spike protein homotrimer. Adapted with permission from [180]. Copyright © 2022 American Chemical Society

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this article.

Data availability

The manuscript has no associated data.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

References

1. Vincenzi, M., Mercurio, F. A., Leone, M. (2023). Looking for SARS-CoV-2 Therapeutics Through Computational Approaches. Current Medicinal Chemistry, 30 (28), 3158–3214. doi: https://doi.org/10.2174/0929867329666221004104430

2. Muratov, E. N., Amaro, R., Andrade, C. H., Brown, N., Ekins, S., Fourches, D. et al. (2021). A critical overview of computational approaches employed for COVID-19 drug discovery. Chemical Society Reviews, 50 (16), 9121–9151. doi: https://doi.org/10.1039/d0cs01065k

3. Middha, S. K., Usha, T., Sukhralia, S., Pareek, C., Yadav, R., Agnihotri, R. et al.; Parihar, A., Khan, R., Kumar, A., Kaushik, A. K., Gohel, H. (Eds.) (2022). Prediction of drug-target interaction – a helping hand in drug repurposing. Computational Approaches for Novel Therapeutic and Diagnostic Designing to Mitigate SARS-CoV-2 Infection. Academic Press, 519–536. doi: https://doi.org/10.1016/b978-0-323-91172-6.00006-6

4. Nandi, S., Nayak, B. S., Khede, M. K., Saxena, A. K. (2022). Repurposing of Chemotherapeutics to Combat COVID-19. Current Topics in Medicinal Chemistry, 22 (32), 2660–2694. doi: https://doi.org/10.2174/1568026623666221130142517

5. Yadav, M., Dhagat, S., Eswari, J. S. (2020). Emerging strategies on in silico drug development against COVID-19: challenges and opportunities. European Journal of Pharmaceutical Sciences, 155, 105522. doi: https://doi.org/10.1016/j.ejps.2020.105522

6. Gao, K., Wang, R., Chen, J., Cheng, L., Frishcosy, J., Huzumi, Y. et al. (2022). Methodology-Centered Review of Molecular Modeling, Simulation, and Prediction of SARS-CoV-2. Chemical Reviews, 122 (13), 11287–11368. doi: https://doi.org/10.1021/acs.chemrev.1c00965

7. Kumar, S., Kovalenko, S., Bhardwaj, S., Sethi, A., Gorobets, N. Yu., Desenko, S. M. et al. (2022). Drug repurposing against SARS-CoV-2 using computational approaches. Drug Discovery Today, 27 (7), 2015–2027. doi: https://doi.org/10.1016/j.drudis.2022.02.004

8. Sadybekov, A. V., Katritch, V. (2023). Computational approaches streamlining drug discovery. Nature, 616 (7958), 673–685. doi: https://doi.org/10.1038/s41586-023-05905-z

9. Shah, A., Jain, M.; Rudrapal, M., Egbuna, C. (Eds.) (2022). Limitations and future challenges of computer-aided drug design methods. Computer Aided Drug Design (CADD): From Ligand-Based Methods to Structure-Based Approaches. Elsevier, 283–297. doi: https://doi.org/10.1016/b978-0-323-90608-1.00006-x

10. Chen, L., Fan, Z., Chang, J., Yang, R., Hou, H., Guo, H. et al. (2023). Sequence-based drug design as a concept in computational drug design. Nature Communications, 14 (1). doi: https://doi.org/10.1038/s41467-023-39856-w

11. Bon, M., Bilsland, A., Bower, J., McAulay, K. (2022). Fragment-based drug discovery –the importance of high-quality molecule libraries. Molecular Oncology, 16 (21), 3761–3777. doi: https://doi.org/10.1002/1878-0261.13277

12. Green, H., Koes, D. R., Durrant, J. D. (2021). DeepFrag: a deep convolutional neural network for fragment-based lead optimization. Chemical Science, 12 (23), 8036–8047. doi: https://doi.org/10.1039/d1sc00163a

13. Shi, X.-X., Wang, Z.-Z., Wang, F., Hao, G.-F., Yang, G.-F. (2023). ACFIS 2.0: an improved web-server for fragment-based drug discovery via a dynamic screening strategy. Nucleic Acids Research, 51 (W1), W25–W32. doi: https://doi.org/10.1093/nar/gkad348

14. Gurung, A. B., Ali, M. A., Lee, J., Farah, M. A., Al-Anazi, K. M. (2021). An Updated Review of Computer-Aided Drug Design and Its Application to COVID-19. BioMed Research International, 2021, 1–18. doi: https://doi.org/10.1155/2021/8853056

15. Bajad, N. G., Rayala, S., Gutti, G., Sharma, A., Singh, M., Kumar, A., Singh, S. K. (2021). Systematic review on role of structure based drug design (SBDD) in the identification of anti-viral leads against SARS-Cov-2. Current Research in Pharmacology and Drug Discovery, 2, 100026. doi: https://doi.org/10.1016/j.crphar.2021.100026

16. Zhang, S., Krumberger, M., Morris, M. A., Parrocha, C. M. T., Kreutzer, A. G., Nowick, J. S. (2021). Structure-based drug design of an inhibitor of the SARS-CoV-2 (COVID-19) main protease using free software: A tutorial for students and scientists. European Journal of Medicinal Chemistry, 218, 113390. doi: https://doi.org/10.1016/j.ejmech.2021.113390

17. Varela-Rial, A., Majewski, M., De Fabritiis, G. (2021). Structure based virtual screening: Fast and slow. WIREs Computational Molecular Science, 12 (2). doi: https://doi.org/10.1002/wcms.1544

18. Chen, W., Liu, X., Zhang, S., Chen, S. (2023). Artificial intelligence for drug discovery: Resources, methods, and applications. Molecular Therapy – Nucleic Acids, 31, 691–702. doi: https://doi.org/10.1016/j.omtn.2023.02.019

19. Nasser, M., Yusof, U. K., Salim, N. (2023). Deep Learning Based Methods for Molecular Similarity Searching: A Systematic Review. Processes, 11 (5). doi: https://doi.org/10.3390/pr11051340

20. Amaro, R. E., Mulholland, A. J. (2018). Multiscale methods in drug design bridge chemical and biological complexity in the search for cures. Nature Reviews Chemistry, 2 (4). doi: https://doi.org/10.1038/s41570-018-0148

21. Amin, Sk. A., Banerjee, S., Singh, S., Qureshi, I. A., Gayen, S., Jha, T. (2021). First structure–activity relationship analysis of SARS-CoV-2 virus main protease (Mpro) inhibitors: an endeavor on COVID-19 drug discovery. Molecular Diversity, 25 (3), 1827–1838. doi: https://doi.org/10.1007/s11030-020-10166-3

22. Luo, D., Tong, J.-B., Zhang, X., Xiao, X.-C., Bian, S. (2022). Computational strategies towards developing novel SARS-CoV-2 Mpro inhibitors against COVID-19. Journal of Molecular Structure, 1247, 131378. doi: https://doi.org/10.1016/j.molstruc.2021.131378

23. Golbraikh, A., Shen, M., Xiao, Z., Xiao, Y.-D., Lee, K.-H., Tropsha, A. (2003). Rational selection of training and test sets for the development of validated QSAR models. Journal of Computer-Aided Molecular Design, 17 (2/4), 241–253. doi: https://doi.org/10.1023/a:1025386326946

24. Golbraikh, A., Tropsha, A. (2000). Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. Molecular Diversity, 5 (4), 231–243. doi: https://doi.org/10.1023/a:102137210868

25. Tetko, I. V., Gasteiger, J., Todeschini, R., Mauri, A., Livingstone, D., Ertl, P. et al. (2005). Virtual Computational Chemistry Laboratory – Design and Description. Journal of Computer-Aided Molecular Design, 19 (6), 453–463. doi: https://doi.org/10.1007/s10822-005-8694-y

26. Yap, C. W. (2011). PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. Journal of Computational Chemistry, 32 (7), 1466–1474. doi: https://doi.org/10.1002/jcc.21707

27. Todeschini, R., Consonni, V. (2000). Handbook of Molecular Descriptors. Weinheim: Wiley-VCH. doi: https://doi.org/10.1002/9783527613106

28. Devillers, J., Balaban, A. T. (1999). Topological Indices and Related Descriptors in QSAR and QSPR. London: CRC Press, 811. doi: https://doi.org/10.1201/9781482296945

29. Malinowski, E. R. (2002). Factor Analysis in Chemistry. Wiley-VCH, 432.

30. Devillers, J. (1996). Genetic Algorithms in Molecular Modeling: Academic Press. doi: https://doi.org/10.1016/b978-0-12-213810-2.x5000-2

31. Solnon, C., Jussien, N. (2013). Ant Colony Optimization and Constraint Programming. John Wiley & Sons. doi: https://doi.org/10.1002/9781118557563

32. Breiman, L. (2001). Random Forests. Machine Learning, 45, 5–32. doi: https://doi.org/10.1023/a:1010933404324

33. Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58 (1), 267–288. doi: https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

34. Hesterberg, T., Choi, N. H., Meier, L., Fraley, C. (2008). Least angle and ℓ1 penalized regression: A review. Statistics Surveys, 2, 61–93. doi: https://doi.org/10.1214/08-ss035

35. Zhao, P., Yu, B. (2007). Stagewise Lasso. Journal of Machine Learning Research, 8, 2701–2726.

36. Hastie, T., Tibshirani, R. (2004). Efficient quadratic regularization for expression arrays. Biostatistics, 5 (3), 329–340. doi: https://doi.org/10.1093/biostatistics/kxh010

37. Roy, K., Kar, S., Das, R. N. (2015). A Primer on QSAR/QSPR Modeling: Fundamental Concepts. Cham: Springer International Publishing. doi: https://doi.org/10.1007/978-3-319-17281-1

38. Khaldan, A., Bouamrane, S., En-Nahli, F., El-mernissi, R., El khatabi, K., Hmamouchi, R. et al. (2021). Prediction of potential inhibitors of SARS-CoV-2 using 3D-QSAR, molecular docking modeling and ADMET properties. Heliyon, 7 (3), e06603. doi: https://doi.org/10.1016/j.heliyon.2021.e06603

39. Jayakumar, J., Ebanesar, A., Gautam, S. (2023). Predictive analysis, diagnosis of COVID-19 through computational screening and validation with spectro photometrical approach. Toxicology and Environmental Health Sciences, 15 (2), 109–117. doi: https://doi.org/10.1007/s13530-023-00172-x

40. Beger, R. D., Young, J. F., Fang, H. (2004). Discriminant Function Analyses of Liver-Specific Carcinogens. Journal of Chemical Information and Computer Sciences, 44 (3), 1107–1110. doi: https://doi.org/10.1021/ci0342829

41. Hosmer, D. W. J., Lemeshow, S. (2000). Applied Logistic Regression. John Wiley & Sons. doi: https://doi.org/10.1002/0471722146

42. Xu, Y., Zomer, S., Brereton, R. G. (2006). Support Vector Machines: A Recent Method for Classification in Chemometrics. Critical Reviews in Analytical Chemistry, 36 (3-4), 177–188. doi: https://doi.org/10.1080/10408340600969486

43. Herbrich, R. (2001). Learning Kernel Classifiers Theory and Algorithms. Cambridge: The MIT Press. doi: https://doi.org/10.7551/mitpress/4170.001.0001

44. Müller, K.-R., Rätsch, G., Sonnenburg, S., Mika, S., Grimm, M., Heinrich, N. (2005). Classifying 'Drug-likeness' with Kernel-Based Learning Methods. Journal of Chemical Information and Modeling, 45 (2), 249–253. doi: https://doi.org/10.1021/ci049737o

45. Rokach, L., Maimon, O. (2014). Data Mining with Decision Trees. Vol. 81. World Scientific, 328. doi: https://doi.org/10.1142/9097

46. Santana, M. V. S., Silva-Jr, F. P.; Parihar, A., Khan, R., Kumar, A., Kaushik, A. K., Gohel, H. (Eds.) (2022). Artificial intelligence methods to repurpose and discover new drugs to fight the Coronavirus disease-2019 pandemic. Computational Approaches for Novel Therapeutic and Diagnostic Designing to Mitigate SARS-CoV-2 Infection. Academic Press, 537–557. doi: https://doi.org/10.1016/b978-0-323-91172-6.00016-9

47. Krishnamurthy, A., Goel, P.; Parihar, A., Khan, R., Kumar, A., Kaushik, A. K., Gohel, H. (Eds.) (2022). Artificial intelligence-based drug screening and drug repositioning tools and their application in the present scenario. Computational Approaches for Novel Therapeutic and Diagnostic Designing to Mitigate SARS-CoV-2 Infection. Academic Press, 379–398. doi: https://doi.org/10.1016/b978-0-323-91172-6.00010-8

48. Murcia-Soler, M., Pérez-Giménez, F., García-March, F. J., Salabert-Salvador, M. T., Díaz-Villanueva, W., Castro-Bleda, M. J., Villanueva-Pareja, A. (2004). Artificial Neural Networks and Linear Discriminant Analysis: A Valuable Combination in the Selection of New Antibacterial Compounds. Journal of Chemical Information and Computer Sciences, 44 (3), 1031–1041. doi: https://doi.org/10.1021/ci030340e

49. Gabrielsson, R. B., Carlsson, G. (2019). Exposition and Interpretation of the Topology of Neural Networks. Paper presented at the 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA). doi: https://doi.org/10.1109/icmla.2019.00180

50. Skinnider, M. A., Stacey, R. G., Wishart, D. S., Foster, L. J. (2021). Chemical language models enable navigation in sparsely populated chemical space. Nature Machine Intelligence, 3 (9), 759–770. doi: https://doi.org/10.1038/s42256-021-00368-1

51. Tay, D. W. P., Yeo, N. Z. X., Adaikkappan, K., Lim, Y. H., Ang, S. J. (2023). 67 million natural product-like compound database generated via molecular language processing. Scientific Data, 10 (1). doi: https://doi.org/10.1038/s41597-023-02207-x

52. Chen, H., Bajorath, J. (2023). Designing highly potent compounds using a chemical language model. Scientific Reports, 13 (1). doi: https://doi.org/10.1038/s41598-023-34683-x

53. Bilodeau, C., Jin, W., Jaakkola, T., Barzilay, R., Jensen, K. F. (2022). Generative models for molecular discovery: Recent advances and challenges. WIREs Computational Molecular Science, 12 (5). doi: https://doi.org/10.1002/wcms.1608

54. Pyzer-Knapp, E., Laino, T., Grisafi, A., Wilkins, D., Willatt, M., Ceriotti, M. et al. (2019). Machine Learning in Chemistry: Data-Driven Algorithms, Learning Systems, and Predictions. American Chemical Society. doi: https://doi.org/10.1021/bk-2019-1326

55. Varmuza, K., Filzmoser, P. (2009). Introduction to Multivariate Statistical Analysis in Chemometrics. Boca Raton: CRC Press, 336. doi: https://doi.org/10.1201/9781420059496

56. Pen, L. Z., Xian Xian, K., Yew, C. F., Hau, O. S., Sumari, P., Abualigah, L. et al. (2023). Classification Applications with Deep Learning and Machine Learning Technologies. Cham: Springer International Publishing. doi: https://doi.org/10.1007/978-3-031-17576-3

57. Veerasamy, R., Rajak, H., Jain, A. K., Sivadasan, S., Varghese, C., K., A. R. (2011). Validation of QSAR Models-Strategies and Importance. International Journal of Drug Design and Discovery, 2 (3), 511–518.

58. Alexander, D. L. J., Tropsha, A., Winkler, D. A. (2015). Beware of R2: Simple, Unambiguous Assessment of the Prediction Accuracy of QSAR and QSPR Models. Journal of Chemical Information and Modeling, 55 (7), 1316–1322. doi: https://doi.org/10.1021/acs.jcim.5b00206

59. Consonni, V., Ballabio, D., Todeschini, R. (2009). Comments on the Definition of the Q2 Parameter for QSAR Validation. Journal of Chemical Information and Modeling, 49 (7), 1669–1678. doi: https://doi.org/10.1021/ci900115y

60. Liu, H., Cocea, M. (2017). Semi-random partitioning of data into training and test sets in granular computing context. Granular Computing, 2 (4), 357–386. doi: https://doi.org/10.1007/s41066-017-0049-2

61. Joseph, V. R. (2022). Optimal ratio for data splitting. Statistical Analysis and Data Mining: The ASA Data Science Journal, 15 (4), 531–538. doi: https://doi.org/10.1002/sam.11583

62. Arlot, S., Celisse, A. (2010). A survey of cross-validation procedures for model selection. Statistics Surveys, 4, 40–79. doi: https://doi.org/10.1214/09-ss054

63. Espinosa, G., Arenas, A., Giralt, F. (2001). Fundamentals of Molecular Similarity. Boston: Springer US.

64. Mestres, J., Rohrer, D. C., Maggiora, G. M. (1997). MIMIC: A molecular-field matching program. Exploiting applicability of molecular similarity approaches. Journal of Computational Chemistry, 18 (7), 934–954. doi: https://doi.org/10.1002/(sici)1096-987x(199705)18:7<934::aid-jcc6>3.0.co;2-s

65. Klebe, G. (Ed.) (2007). Virtual Screening: An Alternative or Complement to High Throughput Screening? Kluwer Academic Publishers.

66. Bajusz, D., Rácz, A., Héberger, K.; Chackalamannil, S., Rotella, D., Ward, S. E. (Eds.) (2017). Chemical Data Formats, Fingerprints, and Other Molecular Descriptions for Database Analysis and Searching. Comprehensive Medicinal Chemistry III. Oxford: Elsevier, 329–378. doi: https://doi.org/10.1016/b978-0-12-409547-2.12345-5

67. Capecchi, A., Probst, D., Reymond, J.-L. (2020). One molecular fingerprint to rule them all: drugs, biomolecules, and the metabolome. Journal of Cheminformatics, 12 (1). doi: https://doi.org/10.1186/s13321-020-00445-4

68. Willett, P. (2006). Similarity-based virtual screening using 2D fingerprints. Drug Discovery Today, 11 (23-24), 1046–1053. doi: https://doi.org/10.1016/j.drudis.2006.10.005

69. Scior, T., Bender, A., Tresadern, G., Medina-Franco, J. L., Martínez-Mayorga, K., Langer, T. et al. (2012). Recognizing Pitfalls in Virtual Screening: A Critical Review. Journal of Chemical Information and Modeling, 52 (4), 867–881. doi: https://doi.org/10.1021/ci200528d

70. Landrum, G. (2020). RDKit: Open-source cheminformatics. Available at: https://www.rdkit.org/

71. Alves, V. M., Bobrowski, T., Melo-Filho, C. C., Korn, D., Auerbach, S., Schmitt, C. et al. (2020). QSAR Modeling of SARS-CoV Mpro Inhibitors Identifies Sufugolix, Cenicriviroc, Proglumetacin, and other Drugs as Candidates for Repurposing against SARS-CoV-2. Molecular Informatics, 40 (1). doi: https://doi.org/10.1002/minf.202000113

72. Morgan, H. L. (1965). The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. Journal of Chemical Documentation, 5 (2), 107–113. doi: https://doi.org/10.1021/c160017a018

73. Willett, P. (2003). Similarity-based approaches to virtual screening. Biochemical Society Transactions, 31 (3), 603–606. doi: https://doi.org/10.1042/bst0310603

74. Raymond, J. W., Willett, P. (2002). Effectiveness of graph-based and fingerprint-based similarity measures for virtual screening of 2D chemical structure databases. Journal of Computer-Aided Molecular Design, 16 (1), 59–71. doi: https://doi.org/10.1023/a:1016387816342

75. Gossen, J., Albani, S., Hanke, A., Joseph, B. P., Bergh, C., Kuzikov, M. et al. (2021). A Blueprint for High Affinity SARS-CoV-2 Mpro Inhibitors from Activity-Based Compound Library Screening Guided by Analysis of Protein Dynamics. ACS Pharmacology & Translational Science, 4 (3), 1079–1095. doi: https://doi.org/10.1021/acsptsci.0c00215

76. Nikolova, N., Jaworska, J. (2003). Approaches to Measure Chemical Similarity – a Review. QSAR & Combinatorial Science, 22 (9-10), 1006–1026. Portico. doi: https://doi.org/10.1002/qsar.200330831

77. Stumpfe, D., Ripphausen, P., Bajorath, J. (2012). Virtual compound screening in drug discovery. Future Medicinal Chemistry, 4 (5), 593–602. doi: https://doi.org/10.4155/fmc.12.19

78. Giordano, D., Biancaniello, C., Argenio, M. A., Facchiano, A. (2022). Drug Design by Pharmacophore and Virtual Screening Approach. Pharmaceuticals, 15 (5), 646. doi: https://doi.org/10.3390/ph15050646

79. Glaab, E., Manoharan, G. B., Abankwa, D. (2021). Pharmacophore Model for SARS-CoV-2 3CLpro Small-Molecule Inhibitors and in Vitro Experimental Validation of Computationally Screened Inhibitors. Journal of Chemical Information and Modeling, 61 (8), 4082–4096. doi: https://doi.org/10.1021/acs.jcim.1c00258

80. Wermuth, C. G., Ganellin, C. R., Lindberg, P., Mitscher, L. A. (1998). Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). Pure and Applied Chemistry, 70 (5), 1129–1143. doi: https://doi.org/10.1351/pac199870051129

81. Molecular Operating Environment (MOE). (2023). 2022.02 Chemical Computing Group ULC. Montreal.

82. Wolber, G., Langer, T. (2004). LigandScout: 3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters. Journal of Chemical Information and Modeling, 45 (1), 160–169. doi: https://doi.org/10.1021/ci049885e

83. Brooijmans, N., Kuntz, I. D. (2003). Molecular Recognition and Docking Algorithms. Annual Review of Biophysics and Biomolecular Structure, 32 (1), 335–373. doi: https://doi.org/10.1146/annurev.biophys.32.110601.142532

84. Weiner, S. J., Kollman, P. A., Nguyen, D. T., Case, D. A. (1986). An all atom force field for simulations of proteins and nucleic acids. Journal of Computational Chemistry, 7 (2), 230–252. doi: https://doi.org/10.1002/jcc.540070216

85. Gohlke, H., Hendlich, M., Klebe, G. (2000). Knowledge-based scoring function to predict protein-ligand interactions. Journal of Molecular Biology, 295 (2), 337–356. doi: https://doi.org/10.1006/jmbi.1999.3371

86. Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V., Mee, R. P. (1997). Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. Journal of Computer-Aided Molecular Design, 11 (5), 425–445. doi: https://doi.org/10.1023/a:1007996124545

87. Alex, A., Finn, P. (1997). Fast and accurate predictions of relative binding energies. Journal of Molecular Structure: THEO-CHEM, 398-399, 551–554. doi: https://doi.org/10.1016/s0166-1280(97)00031-6

88. Ballester, P. J., Mitchell, J. B. O. (2010). A machine learning approach to predicting protein–ligand binding affinity with applications to molecular docking. Bioinformatics, 26 (9), 1169–1175. doi: https://doi.org/10.1093/bioinformatics/btq112

89. Guedes, I. A., Barreto, A. M. S., Marinho, D., Krempser, E., Kuenemann, M. A., Sperandio, O. et al. (2021). New machine learning and physics-based scoring functions for drug discovery. Scientific Reports, 11 (1). doi: https://doi.org/10.1038/s41598-021-82410-1

90. Charifson, P. S., Corkery, J. J., Murcko, M. A., Walters, W. P. (1999). Consensus Scoring: A Method for Obtaining Improved Hit Rates from Docking Databases of Three-Dimensional Structures into Proteins. Journal of Medicinal Chemistry, 42 (25), 5100–5109. doi: https://doi.org/10.1021/jm990352k

91. Kumar, T. D. A. (2022). Drug Design. A Conceptual Overview. Abingdon: CRC Press. doi: https://doi.org/10.1201/9781003298755

92. Kitchen, D. B., Decornez, H., Furr, J. R., Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. Nature Reviews Drug Discovery, 3 (11), 935–949. doi: https://doi.org/10.1038/nrd1549

93. Warren, G. L., Andrews, C. W., Capelli, A.-M., Clarke, B., LaLonde, J., Lambert, M. H. et al. (2005). A Critical Assessment of Docking Programs and Scoring Functions. Journal of Medicinal Chemistry, 49 (20), 5912–5931. doi: https://doi.org/10.1021/jm050362n

94. Stahl, M., Rarey, M. (2001). Detailed Analysis of Scoring Functions for Virtual Screening. Journal of Medicinal Chemistry, 44 (7), 1035–1042. doi: https://doi.org/10.1021/jm0003992

95. Zhu, H., Zhou, R., Cao, D., Tang, J., Li, M. (2023). A pharmacophore-guided deep learning approach for bioactive molecular generation. Nature Communications, 14 (1). doi: https://doi.org/10.1038/s41467-023-41454-9

96. Mazanetz, M., Law, R., Whittaker, M. (2013). Hit and Lead Identification from Fragments. De novo Molecular Design, Weinheim: Wiley-VCH Verlag GmbH & Co.KGaA, 143–200. doi: https://doi.org/10.1002/9783527677016.ch6

97. Wetzel, S., Klein, K., Renner, S., Rauh, D., Oprea, T. I., Mutzel, P., Waldmann, H. (2009). Interactive exploration of chemical space with Scaffold Hunter. Nature Chemical Biology, 5 (8), 581–583. doi: https://doi.org/10.1038/nchembio.187

98. Langevin, M., Minoux, H., Levesque, M., Bianciotto, M. (2020). Scaffold-Constrained Molecular Generation. Journal of Chemical Information and Modeling, 60 (12), 5637–5646. doi: https://doi.org/10.1021/acs.jcim.0c01015

99. Scott, O. B., Edith Chan, A. W. (2020). ScaffoldGraph: an open-source library for the generation and analysis of molecular scaffold networks and scaffold trees. Bioinformatics, 36 (12), 3930–3931. doi: https://doi.org/10.1093/bioinformatics/btaa219

100. Schaub, J., Zander, J., Zielesny, A., Steinbeck, C. (2022). Scaffold Generator: a Java library implementing molecular scaffold functionalities in the Chemistry Development Kit (CDK). Journal of Cheminformatics, 14 (1). doi: https://doi.org/10.1186/s13321-022-00656-x

101. Buckle, D. R., Erhardt, P. W., Ganellin, C. R., Kobayashi, T., Perun, T. J., Proudfoot, J., Senn-Bilfinger, J.; Desai, M. C. (Ed.) (2013). Glossary of Terms Used in Medicinal Chemistry Part II (IUPAC Recommendations 2013). Annual Reports in Medicinal Chemistry. Academic Press, 48, 387–418. doi: https://doi.org/10.1016/b978-0-12-417150-3.00024-7

102. Hessler, G., Baringhaus, K.-H. (2010). The scaffold hopping potential of pharmacophores. Drug Discovery Today: Technologies, 7 (4), e263–e269. doi: https://doi.org/10.1016/j.ddtec.2010.09.001

103. Brown, N. (Ed.) (2013). Scaffold Hopping in Medicinal Chemistry. Wiley-VCH. doi: https://doi.org/10.1002/9783527665143

104. Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., Olson, A. J. (2016). Computational protein–ligand docking and virtual drug screening with the AutoDock suite. Nature Protocols, 11 (5), 905–919. doi: https://doi.org/10.1038/nprot.2016.051

105. Goodsell, D. S., Morris, G. M., Olson, A. J. (1996). Automated docking of flexible ligands: Applications of AutoDock. Journal of Molecular Recognition, 9 (1), 1–5. doi: https://doi.org/10.1002/(sici)1099-1352(199601)9:1<1::aid-jmr241>3.0.co;2-6

106. Trott, O., Olson, A. J. (2009). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry, 31 (2), 455–461. doi: https://doi.org/10.1002/jcc.21334

107. Eberhardt, J., Santos-Martins, D., Tillack, A. F., Forli, S. (2021). AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. Journal of Chemical Information and Modeling, 61 (8), 3891–3898. doi: https://doi.org/10.1021/acs.jcim.1c00203

108. Jones, G., Willett, P., Glen, R. C., Leach, A. R., Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. Journal of Molecular Biology, 267 (3), 727–748. doi: https://doi.org/10.1006/jmbi.1996.0897

109. Grosdidier, A., Zoete, V., Michielin, O. (2011). SwissDock, a protein-small molecule docking web service based on EA-Dock DSS. Nucleic Acids Research, 39, W270–W277. doi: https://doi.org/10.1093/nar/gkr366

110. Kramer, B., Rarey, M., Lengauer, T. (1999). Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. Proteins: Structure, Function, and Genetics, 37 (2), 228–241. doi: https://doi.org/10.1002/(sici)1097-0134(19991101)37:2<228::aid-prot8>3.0.co;2-8

111. Saeedi, S., Rahmati, A., Chavoshpour-Natanzi, Z. (2022). Synthesis of pyrazolo[5′,1′:2,3]imidazo[1,5-c]quinazolin-6(5H)-ones and molecular docking study of their affinity against the COVID-19 main protease. RSC Advances, 12 (30), 19579–19589. doi: https://doi.org/10.1039/d2ra03179e

112. Amendola, G., Ettari, R., Previti, S., Di Chio, C., Messere, A., Di Maro, S. et al. (2021). Lead Discovery of SARS-CoV-2 Main Protease Inhibitors through Covalent Docking-Based Virtual Screening. Journal of Chemical Information and Modeling, 61 (4), 2062–2073. doi: https://doi.org/10.1021/acs.jcim.1c00184

113. Porto, V. A., Porto, R. S. (2022). In silico Studies of Novel Synthetic Compounds as Potential Drugs to Inhibit Coronavirus (SARS-CoV-2): A Systematic Review. Biointerface Research in Applied Chemistry, 12 (4), 4293–4306. doi: https://doi.org/10.33263/briac124.42934306

114. Hou, N., Shuai, L., Zhang, L., Xie, X., Tang, K., Zhu, Y. et al. (2023). Development of Highly Potent Noncovalent Inhibitors of SARS-CoV-2 3CLpro. ACS Central Science, 9 (2), 217–227. doi: https://doi.org/10.1021/acscentsci.2c01359

115. Evseeva, L. V., Ivanov, V. V., Karpina, V. R., Kovalenko, S. S., Kuznetsov, I. E., Langer, T. et al. (2020). The virtual screening application for searching potential antiviral agents to treat COVID-19 disease. Journal of Organic and Pharmaceutical Chemistry, 18 (2 (70)), 3–15. doi: https://doi.org/10.24959/ophcj.20.200019

116. Kulkarni, S. A., Nagarajan, S. K., Ramesh, V., Palaniyandi, V., Selvam, S. P., Madhavan, T. (2020). Computational evaluation of major components from plant essential oils as potent inhibitors of SARS-CoV-2 spike protein. Journal of Molecular Structure, 1221, 128823. doi: https://doi.org/10.1016/j.molstruc.2020.128823

117. Ahmad, M., Dwivedy, A., Mariadasse, R., Tiwari, S., Kar, D., Jeyakanthan, J., Biswal, B. K. (2020). Prediction of Small Molecule Inhibitors Targeting the Severe Acute Respiratory Syndrome Coronavirus-2 RNA-dependent RNA Polymerase. ACS Omega, 5 (29), 18356–18366. doi: https://doi.org/10.1021/acsomega.0c02096

118. Lohachova, K. O., Sviatenko, A. S., Kyrychenko, A., Ivanov, V. V., Langer, T., Kovalenko, S. M., Kalugin, O. N. (2023). Computer-aided drug design of novel nirmatrelvir analogs inhibiting main protease of Coronavirus SARS-CoV-2. Journal of Applied Pharmaceutical Science, 13. In press

119. Tejera, E., Munteanu, C. R., López-Cortés, A., Cabrera-Andrade, A., Pérez-Castillo, Y. (2020). Drugs Repurposing Using QSAR, Docking and Molecular Dynamics for Possible Inhibitors of the SARS-CoV-2 Mpro Protease. Molecules, 25 (21), 5172. doi: https://doi.org/10.3390/molecules25215172

120. Pathak, N., Chen, Y.-T., Hsu, Y.-C., Hsu, N.-Y., Kuo, C.-J., Tsai, H. P. (2020). Uncovering Flexible Active Site Conformations of SARS-CoV-2 3CL Proteases through Protease Pharmacophore Clusters and COVID-19 Drug Repurposing. ACS Nano, 15 (1), 857–872. doi: https://doi.org/10.1021/acsnano.0c07383

121. Zhang, C.-H., Stone, E. A., Deshmukh, M., Ippolito, J. A., Ghahremanpour, M. M., Tirado-Rives, J. et al. (2021). Potent Noncovalent Inhibitors of the Main Protease of SARS-CoV-2 from Molecular Sculpting of the Drug Perampanel Guided by Free Energy Perturbation Calculations. ACS Central Science, 7 (3), 467–475. doi: https://doi.org/10.1021/acscentsci.1c00039

122. Salo-Ahen, O. M. H., Alanko, I., Bhadane, R., Bonvin, A. M. J. J., Honorato, R. V., Hossain, S. et al. (2020). Molecular Dynamics Simulations in Drug Discovery and Pharmaceutical Development. Processes, 9 (1), 71. doi: https://doi.org/10.3390/pr9010071

123. Allen, M. P., Tildesley, D. J. (2017). Computer Simulation of Liquids: Second Edition. Oxford: Oxford University Press. doi: https://doi.org/10.1093/oso/9780198803195.001.0001

124. Christen, M., Hünenberger, P. H., Bakowies, D., Baron, R., Bürgi, R., Geerke, D. P. et al. (2005). The GROMOS software for biomolecular simulation: GROMOS05. Journal of Computational Chemistry, 26 (16), 1719–1751. doi: https://doi.org/10.1002/jcc.20303

125. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. Journal of Computational Chemistry, 26 (16), 1701–1718. doi: https://doi.org/10.1002/jcc.20291

126. Kalé, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N. et al. (1999). NAMD2: Greater Scalability for Parallel Molecular Dynamics. Journal of Computational Physics, 151 (1), 283–312. doi: https://doi.org/10.1006/jcph.1999.6201

127. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. Journal of Computational Chemistry, 4 (2), 187–217. doi: https://doi.org/10.1002/jcc.540040211

128. Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M. et al. (2005). The Amber biomolecular simulation programs. Journal of Computational Chemistry, 26 (16), 1668–1688. doi: https://doi.org/10.1002/jcc.20290

129. Thompson, A. P., Aktulga, H. M., Berger, R., Bolintineanu, D. S., Brown, W. M., Crozier, P. S. et al. (2022). LAMMPS – a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. Computer Physics Communications, 271, 108171. doi: https://doi.org/10.1016/j.cpc.2021.108171

130. Smith, W., Yong, C. W., Rodger, P. M. (2002). DL_POLY: Application to molecular simulation. Molecular Simulation, 28 (5), 385–471. doi: https://doi.org/10.1080/08927020290018769

131. Van Der Spoel, D., Lindahl, E., Hess, B., van Buuren, A. R., Apol, E., Meulenhoff, P. J. et al. (2010). Gromacs User Manual version 4.5.4. Available at: www.gromacs.org

132. Verkhivker, G. M., Agajanian, S., Oztas, D. Y., Gupta, G. (2021). Atomistic Simulations and In Silico Mutational Profiling of Protein Stability and Binding in the SARS-CoV-2 Spike Protein Complexes with Nanobodies: Molecular Determinants of Mutational Escape Mechanisms. ACS Omega, 6 (40), 26354–26371. doi: https://doi.org/10.1021/acsomega.1c03558

133. Domingo, M., Faraudo, J. (2023). Effect of surfactants on SARS-CoV-2: Molecular dynamics simulations. The Journal of Chemical Physics, 158 (11). doi: https://doi.org/10.1063/5.0135251

134. Samanta, P. N., Majumdar, D., Leszczynski, J. (2023). Elucidating Atomistic Insight into the Dynamical Responses of the SARS-CoV-2 Main Protease for the Binding of Remdesivir Analogues: Leveraging Molecular Mechanics To Decode the Inhibition Mechanism. Journal of Chemical Information and Modeling, 63 (11), 3404–3422. doi: https://doi.org/10.1021/acs.jcim.3c00105

135. Rashid, H. ur, Ahmad, N., Abdalla, M., Khan, K., Martines, M. A. U., Shabana, S. (2022). Molecular docking and dynamic simulations of Cefixime, Etoposide and Nebrodenside A against the pathogenic proteins of SARS-CoV-2. Journal of Molecular Structure, 1247, 131296. doi: https://doi.org/10.1016/j.molstruc.2021.131296

136. Muhseen, Z. T., Hameed, A. R., Al-Hasani, H. M. H., Tahir ul Qamar, M., Li, G. (2020). Promising terpenes as SARS-CoV-2 spike receptor-binding domain (RBD) attachment inhibitors to the human ACE2 receptor: Integrated computational approach. Journal of Molecular Liquids, 320, 114493. doi: https://doi.org/10.1016/j.molliq.2020.114493

137. Wakchaure, P. D., Ghosh, S., Ganguly, B. (2020). Revealing the Inhibition Mechanism of RNA-Dependent RNA Polymerase (RdRp) of SARS-CoV-2 by Remdesivir and Nucleotide Analogues: A Molecular Dynamics Simulation Study. The Journal of Physical Chemistry B, 124 (47), 10641–10652. doi: https://doi.org/10.1021/acs.jpcb.0c06747

138. Byléhn, F., Menéndez, C. A., Perez-Lemus, G. R., Alvarado, W., de Pablo, J. J. (2021). Modeling the Binding Mechanism of Remdesivir, Favilavir, and Ribavirin to SARS-CoV-2 RNA-Dependent RNA Polymerase. ACS Central Science, 7 (1), 164–174. doi: https://doi.org/10.1021/acscentsci.0c01242

139. Ghorbani, M., Brooks, B. R., Klauda, J. B. (2020). Critical Sequence Hotspots for Binding of Novel Coronavirus to Angiotensin Converter Enzyme as Evaluated by Molecular Simulations. The Journal of Physical Chemistry B, 124 (45), 10034–10047. doi: https://doi.org/10.1021/acs.jpcb.0c05994

140. Carli, M., Sormani, G., Rodriguez, A., Laio, A. (2020). Candidate Binding Sites for Allosteric Inhibition of the SARS-CoV-2 Main Protease from the Analysis of Large-Scale Molecular Dynamics Simulations. The Journal of Physical Chemistry Letters, 12 (1), 65–72. doi: https://doi.org/10.1021/acs.jpclett.0c03182

141. Sakkiah, S., Guo, W., Pan, B., Ji, Z., Yavas, G., Azevedo, M. et al. (2021). Elucidating Interactions Between SARS-CoV-2 Trimeric Spike Protein and ACE2 Using Homology Modeling and Molecular Dynamics Simulations. Frontiers in Chemistry, 8, 622632.. doi: https://doi.org/10.3389/fchem.2020.622632

142. Sheik Amamuddy, O., Verkhivker, G. M., Tastan Bishop, Ö. (2020). Impact of Early Pandemic Stage Mutations on Molecular Dynamics of SARS-CoV-2 Mpro. Journal of Chemical Information and Modeling, 60 (10), 5080–5102. doi: https://doi.org/10.1021/acs.jcim.0c00634

143. Rajagopal, K., Varakumar, P., Aparna, B., Byran, G., Jupudi, S. (2020). Identification of some novel oxazine substituted 9-anilinoacridines as SARS-CoV-2 inhibitors for COVID-19 by molecular docking, free energy calculation and molecular dynamics studies. Journal of Biomolecular Structure and Dynamics, 39 (15), 5551–5562. doi: https://doi.org/10.1080/07391102.2020.1798285

144. Yoshino, R., Yasuo, N., Sekijima, M. (2020). Identification of key interactions between SARS-CoV-2 main protease and inhibitor drug candidates. Scientific Reports, 10 (1), 12493. doi: https://doi.org/10.1038/s41598-020-69337-9

145. Sang, P., Tian, S.-H., Meng, Z.-H., Yang, L.-Q. (2020). Anti-HIV drug repurposing against SARS-CoV-2. RSC Advances, 10 (27), 15775–15783. doi: https://doi.org/10.1039/d0ra01899f

146. Mukherjee, S., Dasgupta, S., Adhikary, T., Adhikari, U., Panja, S. S. (2020). Structural insight to hydroxychloroquine-3C-like proteinase complexation from SARS-CoV-2: inhibitor modelling study through molecular docking and MD-simulation study. Journal of Biomolecular Structure and Dynamics, 39 (18), 7322–7334. doi: https://doi.org/10.1080/07391102.2020.1804458

147. Grottesi, A., Bešker, N., Emerson, A., Manelfi, C., Beccari, A. R., Frigerio, F. et al. (2020). Computational Studies of SARS-CoV-2 3CLpro: Insights from MD Simulations. International Journal of Molecular Sciences, 21 (15), 5346. doi: https://doi.org/10.3390/ijms21155346

148. Kumar, D., Kumari, K., Jayaraj, A., Kumar, V., Kumar, R. V., Dass, S. K. et al. (2020). Understanding the binding affinity of noscapines with protease of SARS-CoV-2 for COVID-19 using MD simulations at different temperatures. Journal of Biomolecular Structure and Dynamics, 39 (7), 2659–2672. doi: https://doi.org/10.1080/07391102.2020.1752310

149. Al-Khafaji, K., AL-Duhaidahawi, D., Taskin Tok, T. (2021). Using integrated computational approaches to identify safe and rapid treatment for SARS-CoV-2. Journal of Biomolecular Structure and Dynamics, 39 (9), 3387-3395.. doi: https://doi.org/10.1080/07391102.2020.1764392

150. Muhammad, I. A., Muangchoo, K., Muhammad, A., Ajingi, Y. S., Muhammad, I. Y., Umar, I. D., Muhammad, A. B. (2020). A Computational Study to Identify Potential Inhibitors of SARS-CoV-2 Main Protease (Mpro) from Eucalyptus Active Compounds. Computation, 8 (3), 79. doi: https://doi.org/10.3390/computation8030079

151. Maurya, A. K., Mishra, N. (2020). In silico validation of coumarin derivatives as potential inhibitors against Main Protease, NSP10/NSP16-Methyltransferase, Phosphatase and Endoribonuclease of SARS CoV-2. Journal of Biomolecular Structure and Dynamics, 39 (18), 7306–7321. doi: https://doi.org/10.1080/07391102.2020.1808075

152. Alamri, M. A., Tahir ul Qamar, M., Mirza, M. U., Bhadane, R., Alqahtani, S. M., Muneer, I. et al. (2020). Pharmacoinformatics and molecular dynamics simulation studies reveal potential covalent and FDA-approved inhibitors of SARS-CoV-2 main protease 3CLpro. Journal of Biomolecular Structure and Dynamics, 39 (13), 4936–4948. doi: https://doi.org/10.1080/07391102.2020.1782768

153. Stoddard, S. V., Stoddard, S. D., Oelkers, B. K., Fitts, K., Whalum, K., Whalum, K. et al. (2020). Optimization Rules for SARS-CoV-2 Mpro Antivirals: Ensemble Docking and Exploration of the Coronavirus Protease Active Site. Viruses, 12 (9), 942. doi: https://doi.org/10.3390/v12090942

154. Rout, J., Swain, B. C., Tripathy, U. (2020). In silicoinvestigation of spice molecules as potent inhibitor of SARS-CoV-2. Journal of Biomolecular Structure and Dynamics, 40 (2), 860–874. doi: https://doi.org/10.1080/07391102.2020.1819879

155. Umesh, Kundu, D., Selvaraj, C., Singh, S. K., Dubey, V. K. (2021). Identification of new anti-nCoV drug chemical compounds from Indian spices exploiting SARS-CoV-2 main protease as target. Journal of Biomolecular Structure and Dynamics, 39 (9), 3428–3434. doi: https://doi.org/10.1080/07391102.2020.1763202

156. El Hassab, M. A., Shoun, A. A., Al-Rashood, S. T., Al-Warhi, T., Eldehna, W. M. (2020). Identification of a New Potential SARS-COV-2 RNA-Dependent RNA Polymerase Inhibitor via Combining Fragment-Based Drug Design, Docking, Molecular Dynamics, and MM-PBSA Calculations. Frontiers in Chemistry, 8, 584894. doi: https://doi.org/10.3389/fchem.2020.584894

157. Gangadevi, S., Badavath, V. N., Thakur, A., Yin, N., De Jonghe, S., Acevedo, O. et al. (2021). Kobophenol A Inhibits Binding of Host ACE2 Receptor with Spike RBD Domain of SARS-CoV-2, a Lead Compound for Blocking COVID-19. The Journal of Physical Chemistry Letters, 12 (7), 1793–1802. doi: https://doi.org/10.1021/acs.jpclett.0c03119

158. Liang, J., Karagiannis, C., Pitsillou, E., Darmawan, K. K., Ng, K., Hung, A., Karagiannis, T. C. (2020). Site mapping and small molecule blind docking reveal a possible target site on the SARS-CoV-2 main protease dimer interface. Computational Biology and Chemistry, 89, 107372. doi: https://doi.org/10.1016/j.compbiolchem.2020.107372

159. Khan, M. A., Meena, S., Alam, Md. A., Ghosh, S. (2020). A solvent sensitive coumarin derivative coupled with gold nanoparticles as selective fluorescent sensor for Pb2+ ions in real samples. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 243, 118810. doi: https://doi.org/10.1016/j.saa.2020.118810

160. Kumar, K. A., Sharma, M., Dalal, V., Singh, V., Tomar, S., Kumar, P. (2021). Multifunctional inhibitors of SARS-CoV-2 by MM/PBSA, essential dynamics, and molecular dynamic investigations. Journal of Molecular Graphics and Modelling, 107, 107969. doi: https://doi.org/10.1016/j.jmgm.2021.107969

161. Eltayb, W. A., Abdalla, M., Rabie, A. M. (2023). Novel Investigational Anti-SARS-CoV-2 Agent Ensitrelvir "S-217622": A Very Promising Potential Universal Broad-Spectrum Antiviral at the Therapeutic Frontline of Coronavirus Species. ACS Omega, 8 (6), 5234–5246. doi: https://doi.org/10.1021/acsomega.2c03881

162. King, E., Aitchison, E., Li, H., Luo, R. (2021). Recent Developments in Free Energy Calculations for Drug Discovery. Frontiers in Molecular Biosciences, 8, 712085. doi: https://doi.org/10.3389/fmolb.2021.712085

163. Zara, L., Efrém, N.-L., van Muijlwijk-Koezen, J. E., de Esch, I. J. P., Zarzycka, B. (2021). Progress in Free Energy Perturbation: Options for Evolving Fragments. Drug Discovery Today: Technologies, 40, 36–42. doi: https://doi.org/10.1016/j.ddtec.2021.10.001

164. Hansson, T., Marelius, J., Åqvist, J. (1998). Ligand binding affinity prediction by linear interaction energy methods. Journal of Computer-Aided Molecular Design, 12 (1), 27–35. doi: https://doi.org/10.1023/a:1007930623000

165. Ngo, S. T., Hung, H. M., Nguyen, M. T. (2016). Fast and accurate determination of the relative binding affinities of small compounds to HIV-1 protease using non-equilibrium work. Journal of Computational Chemistry, 37 (31), 2734–2742. doi: https://doi.org/10.1002/jcc.24502

166. Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L. et al. (2000). Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. Accounts of Chemical Research, 33 (12), 889–897. doi: https://doi.org/10.1021/ar000033j

167. Jarzynski, C. (1997). Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. Physical Review E, 56 (5), 5018–5035. doi: https://doi.org/10.1103/physreve.56.5018

168. Kollman, P. (1993). Free energy calculations: Applications to chemical and biochemical phenomena. Chemical Reviews, 93 (7), 2395–2417. doi: https://doi.org/10.1021/cr00023a004

169. Zwanzig, R. W. (2004). High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. The Journal of Chemical Physics, 22 (8), 1420–1426. doi: https://doi.org/10.1063/1.1740409

170. Fu, H., Chen, H., Cai, W., Shao, X., Chipot, C. (2021). BFEE2: Automated, Streamlined, and Accurate Absolute Binding Free-Energy Calculations. Journal of Chemical Information and Modeling, 61 (5), 2116–2123. doi: https://doi.org/10.1021/acs.jcim.1c00269

171. Blazhynska, M., Goulard Coderc de Lacam, E., Chen, H., Roux, B., Chipot, C. (2022). Hazardous Shortcuts in Standard Binding Free Energy Calculations. The Journal of Physical Chemistry Letters, 13 (27), 6250–6258. doi: https://doi.org/10.1021/acs.jpclett.2c01490

172. Coderc de Lacam, E. G., Blazhynska, M., Chen, H., Gumbart, J. C., Chipot, C. (2022). When the Dust Has Settled: Calculation of Binding Affinities from First Principles for SARS-CoV-2 Variants with Quantitative Accuracy. Journal of Chemical Theory and Computation, 18 (10), 5890–5900. doi: https://doi.org/10.1021/acs.jctc.2c00604

173. Gupta, A., Zhou, H.-X. (2020). Profiling SARS-CoV-2 Main Protease (MPRO) Binding to Repurposed Drugs Using Molecular Dynamics Simulations in Classical and Neural Network-Trained Force Fields. ACS Combinatorial Science, 22 (12), 826–832. doi: https://doi.org/10.1021/acscombsci.0c00140

174. Ngo, S. T., Nguyen, T. H., Tung, N. T., Mai, B. K. (2022). Insights into the binding and covalent inhibition mechanism of PF-07321332 to SARS-CoV-2 Mpro. RSC Advances, 12 (6), 3729–3737. doi: https://doi.org/10.1039/d1ra08752e

175. Zanetti-Polzi, L., Smith, M. D., Chipot, C., Gumbart, J. C., Lynch, D. L., Pavlova, A. et al. (2021). Tuning Proton Transfer Thermodynamics in SARS-CoV-2 Main Protease: Implications for Catalysis and Inhibitor Design. The Journal of Physical Chemistry Letters, 12 (17), 4195–4202. doi: https://doi.org/10.1021/acs.jpclett.1c00425

176. Qayed, W. S., Ferreira, R. S., Silva, J. R. A. (2022). In Silico Study towards Repositioning of FDA-Approved Drug Candidates for Anticoronaviral Therapy: Molecular Docking, Molecular Dynamics and Binding Free Energy Calculations. Molecules, 27 (18), 5988. doi: https://doi.org/10.3390/molecules27185988

177. Parks, C. D., Gaieb, Z., Chiu, M., Yang, H., Shao, C., Walters, W. P. et al. (2020). D3R grand challenge 4: blind prediction of protein–ligand poses, affinity rankings, and relative binding free energies. Journal of Computer-Aided Molecular Design, 34 (2), 99–119. doi: https://doi.org/10.1007/s10822-020-00289-y

178. Gentile, F., Yaacoub, J. C., Gleave, J., Fernandez, M., Ton, A.-T., Ban, F. et al. (2022). Artificial intelligence-enabled virtual screening of ultra-large chemical libraries with deep docking. Nature Protocols, 17 (3), 672–697. doi: https://doi.org/10.1038/s41596-021-00659-2

179. Gorgulla, C., Boeszoermenyi, A., Wang, Z.-F., Fischer, P. D., Coote, P. W., Padmanabha Das, K. M. et al. (2020). An open-source drug discovery platform enables ultra-large virtual screens. Nature, 580 (7805), 663–668. doi: https://doi.org/10.1038/s41586-020-2117-z

180. Arantes, P. R., Saha, A., Palermo, G. (2020). Fighting COVID-19 Using Molecular Dynamics Simulations. ACS Central Science, 6 (10), 1654–1656. doi: https://doi.org/10.1021/acscentsci.0c01236

181. Frye, L., Bhat, S., Akinsanya, K., Abel, R. (2021). From computer-aided drug discovery to computer-driven drug discovery. Drug Discovery Today: Technologies, 39, 111–117. doi: https://doi.org/10.1016/j.ddtec.2021.08.001

Volodymyr Ivanov, Doctor of Chemical Sciences, Professor, Department of Materials Chemistry, School of Chemistry, V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, Ukraine, 61022

Kateryna Lohachova, Postgraduate Student, Department of Inorganic Chemistry
School of Chemistry, V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, Ukraine, 61022

Yaroslav Kolesnyk, PhD, Department of Inorganic Chemistry, School of Chemistry, V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, Ukraine, 61022

Anton Zakharov, PhD, Department of Materials Chemistry, School of Chemistry, V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, Ukraine, 61022

Larysa Yevsieieva, Senior Lecturer, Department of Organic Chemistry
School of Chemistry, V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, Ukraine, 61022

Alexander Kyrychenko*, Doctor of Chemical Sciences, Senior Researcher, Department of Inorganic Chemistry, School of Chemistry, V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, Ukraine, 61022

Thierry Langer, PhD, Professor, Department of Pharmaceutical Chemistry , University of Vienna, Universitätsring 1, Vienna, Austria, A-1010

Sergiy M. Kovalenko, Doctor of Science, Professor, Department of Organic Chemistry, School of Chemistry, V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, Ukraine, 61022

Oleg N. Kalugin*, PhD, Professor, Department of Inorganic Chemistry, School of Chemistry, V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkiv, Ukraine, 61022

Corresponding authors: Oleg Kalugin, e-mail: onkalugin@gmail.com; Alexander Kyrychenko, e-mail: a.v.kyrychenko@karazin.ua