Review

Molecular Docking: Shifting Paradigms in Drug Discovery

Luca Pinzi * and Giulio Rastelli

Department of Life Sciences, University of Modena and Reggio Emilia, Via Giuseppe Campi 103, 41125 Modena, Italy
* Correspondence: luca.pinzi@unimore.it

check for updates

Abstract: Molecular docking is an established in silico structure-based method widely used in drug discovery. Docking enables the identification of novel compounds of therapeutic interest, predicting ligand-target interactions at a molecular level, or delineating structure-activity relationships (SAR), without knowing *a priori* the chemical structure of other target modulators. Although it was originally developed to help understanding the mechanisms of molecular recognition between small and large molecules, uses and applications of docking in drug discovery have heavily changed over the last years. In this review, we describe how molecular docking was firstly applied to assist in drug discovery tasks. Then, we illustrate newer and emergent uses and applications of docking, including prediction of adverse effects, polypharmacology, drug repurposing, and target fishing and profiling, discussing also future applications and further potential of this technique when combined with emergent techniques, such as artificial intelligence.

Keywords: molecular docking; drug discovery; drug repurposing; reverse screening; target fishing; polypharmacology; adverse drug reactions

1. Introduction

The experimental screening of large libraries of compounds against panels of molecular targets, i.e., High-Throughput Screening (HTS), has represented the gold standard for discovering biologically active hits. However, the high costs required to establish and maintain these screening platforms often hamper their use for drug discovery [1]. Moreover, considering the recent developments in computer technology and the rapid increase of structural, chemical, and biological data available on an ever-growing number of therapeutic targets, it is easily understandable how the use of in silico approaches as chemoinformatics, molecular modeling, and artificial intelligence (AI) has significantly increased in the last decades [2–6]. Indeed, in silico approaches now enable the virtual screening of millions of compounds in an affordable time, thus reducing the initial costs of hit identification and improving chances of finding the desired drug candidates. At present, several molecular modeling techniques are available to facilitate drug discovery tasks, most of them being classified into structure-based and ligand-based approaches.

Structure-based methods rely on the information derived from the knowledge of the 3D structure of a target of interest, and they allow ranking databases of molecules according to the structural and electronic complementarity of ligands to a given target [7]. In this context, molecular docking is among one of the most popular and successful structure-based in silico methods, which help predict the interactions occurring between molecules and biological targets [7]. This process is generally accomplished by first predicting the molecular orientation of a ligand within a receptor, and then estimating their complementarity through the use of a scoring function [7].

Since its first appearance in the mid-1970s, docking has proved to be an important tool to help understanding how chemical compounds interact with their molecular targets, and for drug discovery and development. As a matter of fact, the number of studies reporting: (i) the use of molecular docking to identify structural determinants necessary for efficient ligand-receptor binding, and (ii) the development of more accurate docking methods, have heavily increased since its first appearance [7–21]. Among the first and more interesting studies on the use of docking in drug discovery and biology is the one from Kuntz et al. in the early 1980s [13]. In this study, the authors described a computational method enabling the exploration of geometrically feasible ligand-receptor alignments for the known heme-myoglobin/metmyoglobin and thyroxine/prealbumin structures [13]. This study was not the first to employ docking for predicting potential conformations of molecular complexes [9]. However, it reported for the first time the use of a simplified function containing solely the terms "hard sphere repulsions" and "hydrogen bonding" to describe protein-ligand interactions, which strongly differed from previous studies [9,11,12,22]. Moreover, the authors were also the first to consider the receptor as a solid rigid body, whose binding site is constituted by "pockets". Interestingly, the method adopted in this study was able to predict structures close to those of already reported X-ray complexes, and also to find protein conformations that could be used for energy refinement and eventually design novel ligands [13]. Since then, molecular docking underwent dramatic improvements, for example, by employing flexible algorithms in the calculations [21,23–26]. Moreover, it also started to be used for the design and optimization of compounds with therapeutic interest. An example of this comes from a study of Ring et al., in which several structure-based drug design methods, including docking, were performed to identify novel non-peptidic inhibitors of enzymes of the serine and cysteine protease families [27]. The results achieved in this work further consolidated the use of computer-aided structure-based drug-design methods for assisting the development of lead compounds [27].

Given the potential offered by this method, increasing efforts have been directed towards the improvement of docking algorithms and for overcoming its intrinsic limitations [28–30]. Indeed, major limitations characterizing docking include a restricted sampling of both ligand and receptor conformations in pose prediction, and the use of approximated scoring functions, which very often provide results that do not correlate with the experimental binding affinities [31,32]. Nevertheless, the application of docking in drug design is limited to biological targets for which crystal structures are known. Several approaches have been adopted to overcome this latter limitation. For example, the unavailability of 3D structures is often bypassed by building homology models derived from structural templates with highly-homologous sequences. Moreover, these methods could also be used in tandem with molecular dynamics (MD) to further validate and refine the in silico modeled complexes [33–35]. Nevertheless, the recent progress in structural biology and crystal structure determination, which are progressively increasing the accessibility to experimentally derived ligand-target complexes [36–39], will certainly mitigate this issue. In silico strategies, including molecular dynamics, have also been widely used to explore the conformational space of the investigated targets, ligands, and ligand-target complexes, and thus better describing the dynamic behavior of ligand-target complexes and refining the docking results [34,40,41]. More rigorous virtual screening methodologies have also been developed to improve the docking-based ligand-target complex predictions [42–46]. Indeed, these post-docking refinement and rescoring methods are of great interest in drug discovery because they usually provide higher hit rates in virtual screening campaigns and allow better correlation with experimental data [44,46].

A number of reviews discussing the role and applications of docking, and the possibilities it could offer in drug design and development, have been reported [7,18,47–50]. However, it should be noted that the uses and applications of docking have been changing since its first appearance. In fact, although it was first developed to investigate molecular recognition between large and small molecules, it is now also widely used to assist different tasks of drug discovery programs, such as hit identification and optimization, drug repositioning, *a posteriori* target identification (reverse screening), multi-target ligand design, and repositioning (Figure 1) [49,51–58]. Moreover, docking allows understanding the

relationships between different molecular targets involved in a given disease, which is also of high relevance for polypharmacology [59] and modern drug discovery in general.
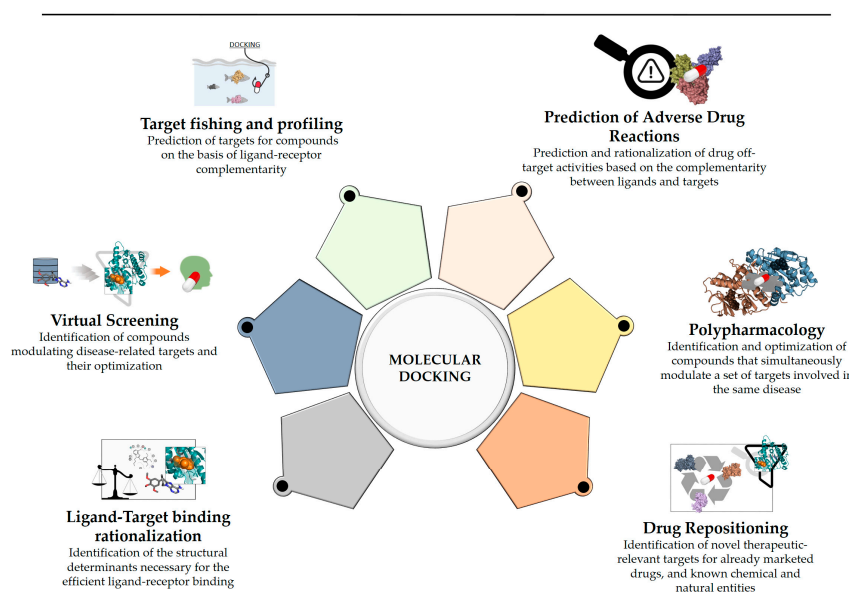

Figure 1. Main applications of molecular docking in current drug discovery. Molecular docking is currently employed to help rationalizing ligands activity towards a target of interest and to perform structure-based virtual screening campaigns, similarly to as when it was first developed. Besides these applications, it can also be used to identify series of targets for which the ligands present good complementarity (target fishing and profiling), some of them being potentially responsible for unexpected drug adverse reactions (off-targets prediction). Moreover, docking is also currently employed for the identification of ligands that simultaneously bind to a pool of selected targets of interest (polypharmacology) and for identifying novel uses for chemical compounds with already optimized safety profiles (drug repositioning).

In particular, the use of this technique has broadened towards novel drug discovery horizons, fueled by the improvement of docking algorithms and by the increase of the publicly accessible information on ligands and targets. For example, thanks to the improved speed and prediction power, docking has also been embedded into large-scale screening protocols to identify [60], e.g.: (i) protein binding sites in which ligands could bind [61]; (ii) novel molecular targets of known ligands [54]; (iii) potential adverse drug reactions (ADRs) [62], and; (iv) ligands with novel chemotypes active against a given target or a set of desired targets [55].

In this review, we will discuss how docking methods have been used to help assisting drug discovery tasks, giving particular emphasis on recent drug design strategies, including polypharmacology, drug repurposing, target identification, and prediction of adverse drug reactions.

2. Current Rational Design Approaches, Including Docking

The possibilities offered by molecular docking in drug discovery are well established [3,5,7,18,47,49,50,63]. However, docking presents intrinsic limitations that limit its prediction performances, the most relevant being reported in the previous section. Although docking has been mainly used as a standalone method for drug design, it is now often integrated into workflows that include other computational methods, such as ligand-based, structure-based, and AI approaches (Figure 2) [50,64]. This helps to account for some of the most relevant limitations characterizing this structure-based method.
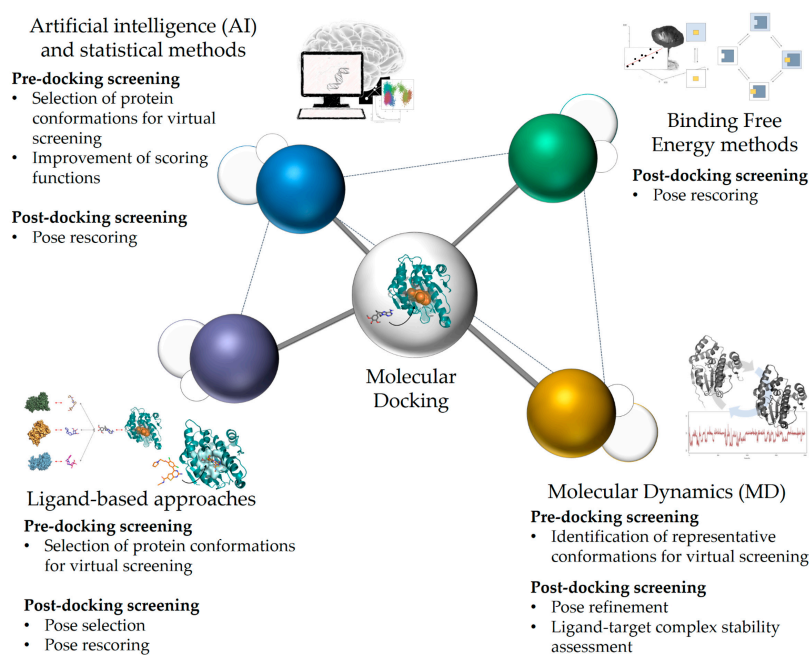
Figure 2. Integration of docking with ligand-based, molecular dynamics, binding free energy approaches, artificial intelligence (AI), and statistical methods. According to the available information, different in silico approaches can be combined with docking to generate integrated workflows with improved prediction performances. Different approaches can also be combined to integrate docking (e.g., molecular dynamics and binding free energy estimations can be combined with docking to improve virtual screening results). Likewise, different approaches can also be applied at different phases of the screening workflow to improve docking predictions. For example, molecular dynamics could be combined with AI-based methods to identify suitable receptor conformations for docking. Then, ligand-based approaches could be applied for rescoring the predicted docking poses [50,65,66].

In particular, ligand-based approaches have been used to select suitable protein conformations for docking screenings [67–70]. The ability of docking to discriminate active compounds from decoys can strongly depend on the used protein structures and the similarity degree of the screened ligands with those co-crystallized in the employed target conformations [69–72]. In this regard, Broccatelli et al. recently reported a study in which different ligand-based methods have been applied for the selection of protein conformations for docking, comparing the performance of different protocols in retrieving known CDK2 inhibitors within two distinct datasets [69]. Similar considerations arose also from the more recent studies of Xu et al. [73] and Kumar et al. [74], in which the authors successfully predicted the affinity and binding mode of a series of Hsp90 [73] and farnesoid X receptor ligands [74], by combining ligand-based approaches with docking.

Ligand-based approaches have also been used to improve the prediction performance of docking screenings, e.g., by measuring the 3D similarity between the binding conformation predicted by docking and the experimental conformation of the ligand co-crystallized in the employed protein conformation [75]. For example, Perryman et al. [76] demonstrated that pharmacophore-based rescoring can improve docking predictions in blind virtual screenings. Similar results were also obtained by Jiang et al. [77], who retrospectively evaluated the performance of DOCK [78,79] against three clinical drug targets (EGFR, IGF-1R, and HIVgp41). According to the obtained in silico results, in particular, the authors demonstrated that the combination of the standard DOCK energy function with an especially devised pharmacophore-based scoring function outperformed single approaches in discriminating active from inactive compounds [77]. Similarly, 3D- and shape-based similarity have also been used in tandem with docking for pose selection and virtual screening [80,81]. In particular, Kumar et al. demonstrated that 3D shape similarity matching is a promising approach for the selection of docking

poses [81]. Moreover, they also observed that docking-based scoring could be conveniently combined with 3D shape similarity to improve virtual screening results [81]. Anighoro et al. recently explored an alternative path to force field-based scoring functions to improve the ranking of candidate ligands of four therapeutic drug targets (DHFR, GR, VEGFR2, and HIV1PR) [80]. According to the obtained in silico results, in particular, the authors demonstrated that the use of 3D similarity approaches for rescoring docking poses could improve ranking lists and hit-rates in structure-based virtual screenings [80]. Nevertheless, methods based on the comparison of ligand-target binding features, e.g., through the use of the protein-ligand interaction fingerprints (PLIFs), have also successfully been explored for rescoring docking results [82–84].

Standing on current literature data, the combination of ligand-based and structure-based approaches allows to heavily improve the prediction power, and thus hit-rates, in virtual screening campaigns. However, it should also be noted that the possibility to apply ligand-centered methods in tandem with docking could be explored solely for those targets that have at least one reported co-crystallized ligand [70,81].

Structure-based approaches, such as molecular dynamics and binding free energy estimations, have also widely been used in combination with docking to improve virtual screening results. In particular, MD allows to evaluate residues flexibility in the target binding site, as well as to explore larger conformational changes potentially accessible to a given protein [33,34,41]. Therefore, it represents an efficient tool to identify receptor conformations for docking [70,85] and to evaluate the stability of the predicted complexes [33,34]. The possibilities offered by MD in prospective in silico screening are particularly appealing for flexible targets with a limited number of reported crystallographic conformations. An example of this comes from a study of Wang et al. [86], who have performed classical MD simulations in explicit solvent to evaluate the stability of the α-helical structure of amyloid β42 (Aβ42), thus identifying a representative protein conformation to perform virtual screening of commercially available compounds [86]. This approach allowed the selection of a set of compounds to be experimentally validated, five of them showing inhibition of Aβ42 aggregation in the micromolar range. Moreover, one of the identified hits also displayed inhibition of BACE1, which plays a key role in the pathogenesis of Alzheimer's disease [86]. Similar considerations could also be drawn from a more recent work of Spyrakis et al. [87], who demonstrated that standard molecular dynamics, clustering, and linear discriminant analysis (LDA) can be integrated to substantially improve structure-based virtual screening results. In this study, in particular, the authors firstly performed MD simulations on three flexible targets (the purine nucleoside phosphorylase PNP, the A_{2A} receptor, and the ABL1 tyrosine-protein kinase) in search of novel protein conformations. Then, they performed clustering via the K-medoids method on the calculated molecular dynamics trajectories [88] to identify MD-derived representative conformations of the investigated targets. Finally, the performance of the FLAP docking program [89] in discriminating active from inactive compounds extracted from DUD-e [90], which is a database useful to benchmark and validate docking protocols, was assessed. LDA was used to automatically select the best combination of protein templates among MD-derived representative and experimentally observed structures yielding the best screening results.

Altogether, the discussed examples demonstrated how the inclusion of classical MD in docking-based protocols could improve virtual screening performances, especially when dealing with highly flexible targets [86,87]. More advanced enhanced sampling techniques, such as umbrella sampling [91], metadynamics [92], and replica exchange MD [93], can also be applied to identify protein conformations for docking screening. Indeed, these techniques, which allow exploring a protein conformational landscape far larger with respect to that of standard MD simulations, have already been applied to study protein flexibility and function [94–101] and to identify additional binding pockets that could be exploited for the design of novel inhibitors [102,103]. However, it should be noted that the application of these advanced methods is computationally more demanding with respect to standard MD.

Combinations of docking with standard molecular dynamics and binding free energy estimations have also been recently explored to account for protein flexibility and to improve virtual screening predictions, respectively [44,46,104–106]. In fact, results of currently available docking algorithms might be affected by poor conformational sampling [7,31]. Moreover, they might provide inaccurate binding energy estimations [7,23] derived by approximate scoring functions [28,107,108]. Indeed, several scoring functions based on different algorithms and concepts, which can be classified into empirical, knowledge-based, and force field-based, have been developed for docking so far [28,30]. However, all of them employ a series of mathematical functions with approximations that do not accurately take into account some thermodynamic elements of the binding energy (e.g., entropy changes upon binding and solvation effects), to allow the fast prediction of ligand-target complex affinity [28,30]. Several approaches have been adopted to account for these issues so far. For example, Rastelli et al. [42] developed BEAR (Binding Estimation After Refinement), a post-docking tool that first performs MD-based structural refinement of ligand-protein complexes, and then predicts their binding free energy with the MM-PBSA and MM-GBSA methods [109]. Indeed, several studies, showing the improvements that the MM-PBSA and MM-GBSA methods, and BEAR provided, both in *a posteriori* and prospective virtual screenings, have been reported [44,46], demonstrating that the application of these approaches heavily improves docking results.

The more advanced and computationally expensive free energy prediction methods Free Energy Perturbation (FEP), Thermodynamic Integration (TI), and funnel metadynamics can also be used for the post-processing of docking results [110–112]. For example, Lee et al. [113] successfully applied a computational workflow that integrates molecular docking, MD simulation, and FEP calculations to predict the binding mode and ligand-protein affinity of a series of MDM2 and MDMX inhibitors, discussing also virtues and vices of the adopted protocol. More recently, Bhati et al. [114] proposed a method that combines MD and TI and was able to provide accurate binding affinities. The performance of their protocol was validated on five well-characterized proteins involved in several physiological processes. However, although the latter approaches are more accurate than docking scoring functions, or even other free energy methods as MM-PBSA and MM-GBSA, in predicting ligand-protein affinity, they are computationally expensive, therefore potentially less suitable for the screening of large libraries of compounds [115].

Very recently, statistical and Artificial Intelligence approaches have also gained a foothold in drug discovery [64]. In fact, these methods allow to easily exploit the ever-growing source of information contained in publicly available structural, chemical, and bioactivity databases, leading to more accurate binding affinity predictions. In particular, machine learning (ML) approaches, including Random Forest (RF) [116] and Support Vector Machines (SVM) [117], have been applied for improving the docking-based binding affinity predictions [29,118]. For example, Xie et al. [119] successfully developed and applied a computational workflow integrating SVM-based classification with docking calculations to identify novel inhibitors of the c-Met tyrosine kinase. In particular, the authors firstly developed and validated an SVM classification model able to discriminate between active and inactive c-Met ligands. Then, they compared the performances of the developed SVM-based model with respect to those of docking, and then the combination of the two methods. The combined SVM-model/docking approach, which provided the best hit-rates and enrichment factors, was finally used to screen a large library of ligands, leading to the identification of eight active compounds among those selected [119]. SVM techniques were also implemented in ensemble schemes to improve docking predictions on flexible targets. For example, Leong et al. [120] recently reported the development and retrospective validation of tailored SVM classification models able to improve docking posing and scoring predictions against the N-methyl-D-aspartate GluN1 receptor. In particular, the authors firstly developed SVM-based posing models by performing docking calculations on ligands reported in complex with seven crystal structures of the N-methyl-D-aspartate GluN1 receptor. Then, they derived a series of SVM-based scoring models from the poses predicted with the same procedure, for a dataset of compounds with reported bioactivity data. Finally, the authors performed a retrospective validation of the newly

developed SVM models by screening a further set of compounds, demonstrating that their approach outperformed standard docking in accurately predicting ligand poses and binding affinities [120].

Machine learning approaches have also been used to improve docking scoring functions. For example, Ballester et al. [121] developed among one of the first ML-based scoring function, called *"RF-Score"*, which uses Random Forest to improve protein-ligand binding affinity predictions. In particular, the authors firstly developed an RF-based scoring function by using different sets of ligand-protein complexes with known activity data reported within the PDBbind database [122]. Then, they compared the performance of *"RF-Score"* with that of other sixteen scoring functions implemented in currently available docking programs, demonstrating that it improved docking results both in virtual screening and lead optimization tasks and that its performances are independent from the employed training sets [121]. Similar considerations come also from a more recent study of Wang et al. [123], who developed a novel scoring function by re-parameterizing the one already implemented in AutoDock Vina [124], with random forest. According to the achieved results, the authors demonstrated that their method outperformed standard docking programs in predicting scores that correlate with experimentally derived binding affinities. Moreover, they also demonstrated that the prediction performances of their scoring function improved with the use of larger datasets for the RF-based scoring function correction [123].

Deep learning (DL) approaches have also been studied to improve docking results. For example, Pereira et al. [125] recently reported an approach based on Convolutional Neural Networks (CNN) called *"DeepVS"*, which learns the features relevant for the binding of a ligand to a target under study, given a set of docking results. In particular, they firstly developed a scoring function that generates DL-based docking scores based on structural data describing a ligand-protein complex (i.e., the atom types and their partial charges, distance between the atoms, and the amino acid types). Then, to demonstrate that the developed scoring function can provide improved results, the authors performed extensive validation screenings on forty different datasets reported in the DUD-e database [90]. According to the reported results, the newly developed *"DeepVS"* scoring function was able to provide higher screening performances with respect to those of standard scoring functions. Interestingly, this study was the first to report the application of deep learning for rescoring docking poses without requiring human-defined parameters, thus making it a particularly appealing approach for unsupervised, large scale, virtual screening campaigns [125].

Less sophisticated, but still robust, statistical techniques, such as score distribution data analysis, have also been examined. In particular, Wang et al. [126] recently reported the application of score distribution data analysis and docking for protein structure selection. Moreover, the adopted approach, which can be employed for a variety of virtual screening purposes, also allowed to successfully identify off-targets activities for 43 anticancer drugs approved by Food and Drug Administration (FDA) [126].

Although the recent introduction of statistical techniques allowed improving results of docking screenings, their applicability heavily depends on the availability of structural, chemical, or bioactivity data. Therefore, such approaches may not represent the optimal choice to improve docking prediction performances when dealing with recently identified, or not yet thoroughly studied, therapeutic targets.

3. Reverse Screening for Target Fishing and Profiling

Docking has also been recently used for a variety of other purposes in drug discovery. In particular, Reverse Docking (RD), which allows predicting the biological targets of a molecule of interest [127], represents a valuable approach for computational target fishing and profiling [60].

Several docking approaches and algorithms are available to enable the reverse screening of a ligand towards a library of protein structures and to assess their binding affinity. However, the application of these approaches requires suitable libraries of targets [60,127]. Indeed, several databases are currently available to help performing RD screenings. Among one of the most known databases to facilitate computational target identification is PDTD [128], which provides information about protein structures, diseases, biological functions, and drugs. Moreover, tailored libraries of targets can also be manually

built upon publicly available databases of crystal structures and binding pockets, such as the Protein Data Bank (PDB) [129], sc-PDB [130], Pocketome [131], and Therapeutic Target Database (TTD) [132]. In particular, the PDB [129] and TTD [132] databases represent well-known reservoirs of information developed to help facilitating computational, molecular and structural biology, and to provide data about targets and diseases, respectively. The sc-PDB [130] and Pocketome [131] databases were instead developed for comparing protein cavities, better describing the ligand-protein pharmacophoric properties and for target identification via pocket-based virtual screening, and to benchmark docking screenings, respectively. Although these libraries of targets were not specifically developed for target fishing and profiling, they allow covering large structural spaces of the known proteome. However, it should be noted that the preparation of such libraries is a time-consuming task because each structure in the databases requires to be properly prepared for the docking calculations [60,127].

In RD screenings, potential targets of a ligand can be ranked according to scoring functions implemented in commonly used docking programs as Glide [133,134]. For example, Park and Cho [135] performed an extensive RD screening on a library of disease-related proteins, including entries from PDTD and around 500 kinases, to identify potential targets for 26 ginsenosides. Results of the screenings allowed the authors to identify potential ligand-protein interactions for some of the investigated natural products with anticancer targets, such as MEK1, EGFR, and Aurora A [135]. Interestingly, although the predicted ligand-protein interactions were not experimentally validated, anticancer activity has been reported for some of the investigated ginsenosides [136]. Moreover, the authors identified potential interactions with targets as acetylcholinesterase, human Carbonic Anhydrase II, and glutamate dehydrogenase, which might be responsible for cholinergic side effects and nephrotoxicity [135]. Although being extensively used in RD screenings, results of standard docking scoring functions can heavily depend on the investigated targets, or even on the employed structural conformations of a given protein. Therefore, normalization of the scores is generally recommended to avoid target-dependent biased results [60,127,137].

Integration of docking with more sophisticated ML-based methods has also very recently been explored for target predictions [138,139]. An example is a study of Nogueira and Koch [139], in which SVM and Neural Networks (NN) models were used to improve the results of reverse docking screenings, pre-processed with the PADIF (Protein Atom Score Contributions Derived Interaction Fingerprint) method [82]. In particular, the authors firstly built datasets of compounds for twenty biological targets with already reported bioactivity data on ChEMBL [140] and X-ray crystal structures. Then, for each target, they developed SVM and NN machine learning models able to discriminate active from non-active compounds based on the docking-derived PADIFs. Finally, they retrospectively validated their models, achieving notable prediction performances, both in terms of target ranking and on multi-target selectivity predictions [139]. Although these structure-based approaches might present several advantages over ligand-based methods and standard docking, they are time and computationally demanding. Moreover, they also need a large amount of bioactivity data to train the models, which sometimes could not be available for some of the targets under study. However, considering the recent advances in hardware and software [141,142] and the increased amount of publicly available bioactivity data, ML-based approaches will certainly play a central role in future target identification and profiling tasks.

Consensus approaches based on RD have also recently been reported for target fishing [143]. In particular, Lapillo and coworkers [143] performed an extensive benchmarking study on 13 different RD screening procedures to identify which method performed better in predicting targets for known ligands. Moreover, they also explored a docking-based consensus approach to improve RD target prediction performances [143]. According to the results reported in this study, target-fishing performances of the different docking procedures ranged around 25% and 35% (in terms of true predictions) for single approaches, and 36% for the consensus. However, this study also showed that the results of the applied RD approaches could be dependent on the features of the investigated proteins binding site [143].

As the set-up of reverse docking screening workflows requires more efforts and longer preparation with respect to standard virtual screening, various tools and web platforms have also been recently developed to facilitate RD. Most of them entrust in already compiled libraries of disease-relevant targets and implement standard programs (e.g., DOCK [79], AutoDock [144], and AutoDock Vina [124]) for performing reverse docking calculations. Therefore, they enable researches to easily identify the biological targets of their molecules of interest, even without massive computational efforts. Among the most relevant programs and web platforms currently in use for docking-based reverse screening are INVDOCK [145], TarFisDock [54], ACTP [146], and idTarget [147], which have been widely employed for target fishing and profiling in several studies, and for different purposes [148–152]. In particular, INVDOCK [145] is a software based on docking, which has been devised for the identification of potential targets for drugs, candidates under clinical trials and ligands, and to facilitate the study of their potential side effects. TarFisDock [54], ACTP [146], and idTarget [147] are web platforms that enable remote RD screening of a given ligand to a set of protein structures. In particular, TarFisDock [54] and idTarget [147] enable the screening of a given ligand to a set of proteins available within the PDTD and PDB by using the DOCK [79] and MEDock [153] docking programs, respectively. Similarly, the ACTP [123] web server enables the screening of a ligand by using the Libdock docking protocol [154]. However, reverse docking analyses of this latter web platform are restricted to a set of autophagy-related protein targets [146].

The utility of these platforms for drug discovery has already been reviewed elsewhere [60,127]. However, considering the relevance of the topic and the improvements in:

(i)   currently available computational techniques and software in general, which allow to more accurately screen larger databases;
(ii)   hardware facilities, which enable a faster screening of ligands to targets, to a larger public, and;
(iii) crystallography [36–39] and homology modeling [155,156] techniques, which allow expanding our knowledge on structural biology;

we envision that further advances in reverse docking will certainly play a central role for target fishing and profiling of ligands in future drug discovery.

4. Prediction of Adverse Drug Reactions

The early identification of drug side effects is of high interest in drug discovery. In fact, it is well known that most of the drug candidates fail clinical trials because of side effects deriving from unexpected interactions with off-targets. Moreover, post-marketing side effect analyses on approved drugs (i.e., pharmacovigilance) are also important because they allow revealing potential safety risks that often could not be detected within clinical trials [157]. Several computational approaches are currently available to assist this task [158–161]. However, most of them require a satisfactory amount of bioactivity data, or of already reported adverse effects as an input for the model training [158–161]. Interestingly, molecular docking needs solely the structural information of the targets to perform its predictions. Therefore, it represents a valuable approach to predict potential side effects of compounds at early phases of clinical and pre-clinical developments, or on marketed drugs with not yet reported exhaustive drug labels and bioactivity records. Indeed, applications of RD screening for identifying drug adverse effects have already been reported in the literature [127,149,162–164]. For example, Ji et al. [149] performed RD screenings to identify potential side effects for a series of anti-HIV drugs, highlighting several proteins whose modulation has already been associated with adverse reactions. Indeed, more than 85% of their target predictions found correspondence with clinical evidence reported in the literature [149]. However, it should also be noted that some of the targets known to be modulated by the investigated anti-HIV drugs could not be predicted within their study, due to the lack of structural information in the analyzed database [149] or because of intrinsic limitations of the adopted screening approach. Altogether, the results reported in this study demonstrated that docking-based reverse screening can be an efficient tool for identifying putative adverse drug reactions. However,

the obtainment of good performance predictions is tightly related to the information reported in the analyzed database of targets. Databases developed to facilitate adverse reactions recognition for drugs have already been reported, SIDER being one of the best known [165]. In particular, SIDER contains more than 140,000 drug-ADR pairs registered in clinical trials and post-marketing surveillance [165], which could facilitate the identification of target activities responsible for side effects. However, databases of targets can also be developed on purpose for performing more focused RD-based ADR analyses, as previously discussed for target fishing and profiling.

More sophisticated screening approaches based on consensus docking [166], or combination with machine learning techniques [138], have also been explored to improve the predictions of drug adverse events. For example, Jaundoo et al. [166] recently applied a consensus docking approach using three different programs, namely AutoDock, AutoDock Vina, and Glide, to predict putative adverse reactions of a set of 43 drugs approved by FDA for the treatment of Gulf War Illness (GWI)-related symptoms. Indeed, the authors performed their analyses to identify side effects potentially arising by the combined use of some of these drugs, suggesting caution for some of them. Moreover, strengths and limitations related to the use of consensus approaches with respect to standard docking were also discussed by the authors [166].

Combinations of docking with statistical approaches have been applied for predicting off-target activities of already reported drugs (see above) [126]. Moreover, a combination of docking with machine learning models has recently been explored. For example, Luo et al. [138] performed an extensive study on more than 1200 compounds extracted from DrugBank against 600 human proteins for predicting or rationalizing drug adverse effects. In particular, the authors firstly performed docking calculations with AutoDock Vina for assessing binding modes and affinity scores of the ligand-protein complexes, and they built ML-based models upon adverse reactions data reported within the SIDER database [165]. Then, they performed side effect predictions by comparing the evaluated ligand-protein scores with the trained ML models. Their analyses resulted in the prediction of 1533 putative adverse reactions and provided potential explanations on the biological mechanisms behind some of already reported drug side effects (e.g., irinotecan that induces decreased libido) [138], demonstrating that ADRs predictions on a large scale are feasible with computational workflows that integrate docking.

Based on the results of the aforementioned studies, it can be argued that consensus approaches with other molecular modeling methods, or integration with AI approaches, will play a pivotal role in future docking-based side effects prediction, allowing to overcome the main limitations currently affecting docking programs.

5. Polypharmacology

To avoid potentially harmful side effects, the pharmaceutical industry focused on the development of highly selective drugs. However, the high attrition rates in the late stages of clinical trials due to a lack of therapeutic efficacy have moved modern drug design towards polypharmacology, which refers to the identification of ligands that hit a set of selected, therapeutic-relevant targets [57,59,167,168]. In this context, molecular docking can provide valuable opportunities because it allows the identification of chemical scaffolds that efficiently and simultaneously bind to a pool of selected targets of interest. Indeed, several studies related to the use of docking for the design of novel multi-target ligands have already been reported [53,169,170]. Moreover, its utility for *de novo* polypharmacology design has also been reviewed [57,168,171,172]. The design of multi-target ligands on rational grounds is challenging [57]. Moreover, the selection of protein conformations to be used for docking can heavily affect the success of the design [70]. This is especially true when dealing with targets with structurally distant binding sites. Considering how difficult it can be to design multi-target ligands, docking is now generally applied in combination with other in silico approaches. In particular, several studies reporting the identification of multi-target ligands are based on the combination of docking screening with pharmacophore modeling [53,169,172]. For example, we recently reported the identification of the first Hsp90/B-Raf dual inhibitors, demonstrating that sub-structure pre-filtering

and pharmacophore-guided docking can be efficiently combined to search for polypharmacology ligands that bind to structurally unrelated targets [53]. However, workflows integrating docking with other in silico techniques have also been pursued for *de novo* multi-target drug design and polypharmacology in general [70,171,173]. For example, Selvam et al. [173] recently reported a study in which a combination of MD, probe mapping, and docking approaches was applied to investigate the selectivity of multi-target ligands towards a set of bioaminergic G-protein-coupled receptors [173]. The results showed that combined workflows, including docking, can be used to guide the design of selective, potentially safer, multi-target molecules able to circumvent side effects commonly associated with antipsychotic drugs. Moreover, the study also demonstrated how different in silico approaches could be efficiently combined to identify structural peculiarities of disease-related targets, for which selective ligands are not yet reported [173].

Web tools and platforms based on docking are also available to explore polypharmacology and for identifying multi-target activities of ligands, such as the Computational Analysis of Novel Drug Opportunities (CANDO) platform [174,175], DRAR-CPI [176], and DPDR-CPI [177]. In particular, CANDO is a multicomponent platform, including molecular docking, that allows predicting potential multi-target interactions of compounds, whereas DRAR-CPI [176] and its upgraded version DPDR-CPI [177] are two web servers that allow identifying candidate targets for a given molecule, and also enable drug repositioning.

Considering the number of computational tools available for the in silico screenings and the challenges that must be faced in a multi-target drug discovery campaign, the best combination of methods should be selected, case by case, upon the available data on the targets [57,172], and hardware and software facilities.

6. Drug Repositioning

Drug repositioning, or repurposing, represents an established drug discovery approach that allows identifying novel therapeutic uses for already approved drugs, candidate compounds under clinical evaluation, natural products, or already synthesized ligands in general [178]. Given the wealth of information reported on ligands, targets and diseases into publicly available databases, increasing efforts have been made on the application of in silico repositioning-based discovery strategies over the last decades. Indeed, in silico repositioning approaches have already demonstrated to provide novel valuable opportunities for drug discovery and development [65,179–183].

In this context, molecular docking has become among one of the most popular computational approaches to repurpose compounds towards novel therapeutic targets. For example, docking can be applied in reverse screening approaches to identify novel molecular targets for known ligands, based on their structural complementarity [127,184]. Docking allows virtually screening databases of approved drugs, natural products, or already synthesized compounds into one or more biological targets of interest in an affordable time. An example of this comes from a study of Kinnings et al., who performed extensive structure-based studies on nine different *Mycobacterium tuberculosis* InhA structures to evaluate whether the entacapone and tolcapone drugs, approved for the treatment of Parkinson's disease, might be repurposed against tuberculosis [51]. Their results allowed the identification of entacapone as a promising lead compound against resistant strains of *Mycobacterium tuberculosis* [51]. Moreover, this study also demonstrated that the same drug could be potentially used for the treatment of unrelated disorders, e.g., Parkinson's disease and tuberculosis [51,185,186]. On the same line, Dakshanamurthy et al. [187] recently performed extensive docking-based virtual screenings on a subset of compounds taken from the DrugBank [188,189], BindingDB [190], and FDA (https://www.fda.gov/) databases against several X-ray crystal structures of human proteins reported in the Protein Data Bank [129]. According to the reported results, the authors discovered that the anti-parasitic drug mebendazole is also an anti-angiogenic VEGFR2 inhibitor. Moreover, they also successfully discovered that the COX-2 inhibitor celecoxib and dimethyl celecoxib bind to

Cadherin-11, which is a protein mediating calcium-dependent cell-cell adhesion that plays a crucial role in rheumatoid arthritis [187].

Despite their great potential for drug discovery [191–193], naturally occurring molecules have very seldom been explored for drug repositioning. Indeed, repositioning strategies based on the integration of chemocentric target identification with docking analyses have been recently explored for identifying novel therapeutic uses of natural compounds. For example, we have recently searched for novel therapeutic targets for the two non-psychoactive cannabinoids cannabigerol (CBG) and cannabichromene (CBC), by integrating shape-based similarity screening with rigid and flexible docking calculations [53]. In particular, in this study, a computational shape-based similarity screening was firstly performed within the DrugBank database. This analysis allowed the identification of InhA, which is an enzyme studied for the development of anti-tubercular drugs, as a potential target for repositioning of both CBG and CBC. Then, extensive docking analyses of the two cannabinoids were performed in the InhA binding site. Interestingly, docking calculations predicted that CBG, but not CBC, is a good candidate for the inhibition of InhA, a finding that was later confirmed by subsequent experimental *in vitro* assays. Altogether, results of this approach, which can be applied to repurpose both natural products and synthetic ligands, clearly demonstrate how molecular docking and ligand-based methods could complement each other in providing more accurate drug repositioning predictions [53]. The combined application of docking with machine learning approaches has also been explored for drug repurposing [194]. Moreover, molecular docking has also been efficiently integrated into web-based tools to enable remote structure-based virtual screenings of given libraries of compounds [195–197] and for drug repurposing predictions [176,177]. For example, Lagarde et al. [196] recently reported a retrospective repurposing screening on a curated library of already approved drugs by using the MTiOpenScreen docking web service [195]. In particular, the authors firstly built three different libraries of purchasable compounds containing approved drugs ("*Drugs-lib*"), food constituents ("*FOOD-lib*"), and natural products ("*NP-lib*"). Then, they performed docking screenings with the MTiOpenScreen web service into the developed "*Drugs-lib*" library to evaluate whether the adopted protocol was able to identify five approved ligands, for which drug reposition against cancer-related targets have already been reported. The performed analyses were able to identify the investigated drugs within the first 1500 top-scored ligands in all of the performed screening campaigns, demonstrating that the adopted protocol might be efficiently used to repurpose already marketed compounds [196].

Based on these promising results, docking represents a valuable approach also for predicting new therapeutic indications for already approved drugs, natural compounds, and already synthesized ligands, especially when used in tandem with other computational methods, such as ligand-based similarity approaches.

7. Concluding Remarks

Since its first appearance in the mid 1970's, molecular docking has represented a unique in silico tool to assist drug design and discovery. However, beyond the applications for which it was originally developed, docking is now also widely employed to assist a variety of other drug discovery tasks, such as the identification of novel chemical scaffolds within large libraries of compounds, to perform in silico target fishing and profiling for drug repositioning, polypharmacology, prediction of adverse effects and beyond, as described in this review article. Being a versatile tool, docking will certainly find application also in other fields of drug discovery. Moreover, docking has been successfully embedded within automated workflows for the screening of large libraries of compounds and targets [54,145–147,195,198]. Of course, the recent advancements in the field of high-performance computing played a key role in this respect. For example, they enabled the in silico screening of millions of compounds in an affordable time [141,142]. Moreover, the recent advancements on Graphics Processing Units (GPUs) have also provided remarkable improvements, both in data-driven drug discovery and in molecular dynamics simulations [34,199]. Indeed, GPU calculations enabled a large

exploration of the conformational landscape potentially accessible to proteins, in shorter times with respect to CPUs [200]. Finally, GPU computing made big data-driven computation tasks accessible to a larger public [201], and it is expected to play a prominent role, not only in docking but in future in silico drug design in general [202].

The modalities by which docking is used to assist the different tasks of drug discovery have also changed along the years. In particular, although it was initially developed and used as a standalone method, docking is now mostly employed in combination with other computational approaches within integrated workflows. This allows to overcome some of the most relevant intrinsic limitations characterizing molecular docking, such as the non-exhaustive conformational sampling and the use of approximate scoring functions [7,31,32,46]. The application of combined approaches usually results in improved prediction performances and allows to better exploit the information coming from different sources. Indeed, applications of combined workflows, including docking, have been explored to assist different tasks of drug discovery. For example, docking has been used in tandem with ligand-based, molecular dynamics, binding free energy calculations, and AI approaches to improve the prediction performances in *de novo* virtual screening, as well as to assist target fishing, ADRs prediction, polypharmacology, and drug repurposing, as discussed. However, one should be aware of the fact that each computational method has its limitations, which might hamper its integration with docking in combined workflows, or even reduce the prediction power of the adopted protocol. For example, methods that provide predictions upon already available bioactivity and chemical data (e.g., AI techniques) may not be the best approach to integrate with docking for less-characterized molecular targets. Likewise, ligand-based approaches might not represent the optimal choice for improving docking predictions when a sufficient number of ligands for the target(s) under study is not available. However, when the combination is feasible, these methods demonstrated to heavily improve docking predictions, both in terms of hit-rates and enrichment factors. On the contrary, molecular dynamics and binding free energy estimations could help improving docking predictions even for less-characterized targets, for example, through the identification of conformational ensembles to be used for structure-based analyses [85], and to more accurately evaluate the ligand-protein binding affinity [46]. Considering the number of in silico tools and techniques currently available, there are still countless opportunities for docking to be explored in integrated workflows. Moreover, their integration will also be facilitated by the continuous improvements in hardware and software engineering, as discussed. Besides, novel valuable opportunities for data and methods integration will certainly come along the increase of the publicly available structural, chemical, and biological information, and its implementation within databases, web platforms, and automated workflows. Further efforts should be directed toward a better integration of the different approaches with the publicly available information reported in these databases. This is expected to provide novel valuable opportunities in future drug discovery and development and, in particular, in the design of challenging and innovative drugs (i.e., multi-target ligands), as well as in assisting ligand profiling and repositioning. Considering the high attrition rates characterizing drug discovery [203,204], the possibilities offered by docking in combination with the approaches outlined here will be important to reduce time and costs in both the development of clinical candidates with better safety profiles (target profiling and ADRs findings) and for the identification of novel applications of already known drugs (target profiling and drug repositioning).

Abbreviations

ADR   Adverse Drug Reaction
AI   Artificial Intelligence
BEAR   Binding Estimation After Refinement
CANDO   Computational Analysis of Novel Drug Opportunities
CBC   Cannabichromene
CBG   Cannabigerol
CNN   Convolutional Neural Networks
CPU   Central Processing Unit
DL   Deep Learning
FDA   Food and Drug Administration
FEP   Free Energy Perturbation
GPU   Graphics Processing Units
GWI   Gulf War Illness
HTS   High-Throughput Screening
LDA   Linear Discriminant Analysis
MD   Molecular Dynamics
ML   Machine Learning
NN   Neural Networks
PADIF   Protein Atom Score Contributions Derived Interaction Fingerprint
PLIFs   Protein-Ligand Interaction Fingerprints
PNP   Purine Nucleoside Phosphorylase
RD   Reverse Docking
RF   Random Forest
SAR   Structure-Activity Relationships
SVM   Support Vector Machines
TI   Thermodynamic Integration

References

1. Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E.W. Computational methods in drug discovery. *Pharmacol. Rev.* **2014**, *66*, 334–395. [CrossRef] [PubMed]
2. Song, C.M.; Lim, S.J.; Tong, J.C. Recent advances in computer-aided drug design. *Brief. Bioinform.* **2009**, *10*, 579–591. [CrossRef] [PubMed]
3. Macalino, S.J.Y.; Gosu, V.; Hong, S.; Choi, S. Role of computer-aided drug design in modern drug discovery. *Arch. Pharm. Res.* **2015**, *38*, 1686–1701. [CrossRef] [PubMed]
4. D'Agostino, D.; Clematis, A.; Quarati, A.; Cesini, D.; Chiappori, F.; Milanesi, L.; Merelli, I. Cloud Infrastructures for In Silico Drug Discovery: Economic and Practical Aspects. *Biomed Res. Int.* **2013**, *2013*, 138012. [PubMed]
5. Jorgensen, W.L. The Many Roles of Computation in Drug Discovery. *Science* **2004**, *303*, 1813–1818. [CrossRef] [PubMed]
6. Kapetanovic, I.M. Computer-aided drug discovery and development (CADDD): In silico-chemico-biological approach. *Chem. Biol. Interact.* **2008**, *171*, 165–176. [CrossRef]
7. Kitchen, D.B.; Decornez, H.; Furr, J.R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **2004**, *3*, 935–949. [CrossRef]
8. DesJarlais, R.L.; Sheridan, R.P.; Dixon, J.S.; Kuntz, I.D.; Venkataraghavan, R. Docking flexible ligands to macromolecular receptors by molecular shape. *J. Med. Chem.* **1986**, *29*, 2149–2153. [CrossRef] [PubMed]
9. Levinthal, C.; Wodak, S.J.; Kahn, P.; Dadivanian, A.K. Hemoglobin interaction in sickle cell fibers. I: Theoretical approaches to the molecular contacts. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1330–1334. [CrossRef]
10. Goodsell, D.S.; Olson, A.J. Automated docking of substrates to proteins by simulated annealing. *Proteins* **1990**, *8*, 195–202. [CrossRef]
11. Salemme, F.R. An hypothetical structure for an intermolecular electron transfer complex of cytochromes c and b5. *J. Mol. Biol.* **1976**, *102*, 563–568. [CrossRef]

12. Wodak, S.J.; Janin, J. Computer analysis of protein-protein interaction. *J. Mol. Biol.* **1978**, *124*, 323–342. [CrossRef]

13. Kuntz, I.D.; Blaney, J.M.; Oatley, S.J.; Langridge, R.; Ferrin, T.E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982**, *161*, 269–288. [CrossRef]

14. Kuhl, F.S.; Crippen, G.M.; Friesen, D.K. A combinatorial algorithm for calculating ligand binding. *J. Comput. Chem.* **1984**, *5*, 24–34. [CrossRef]

15. DesJarlais, R.L.; Sheridan, R.P.; Seibel, G.L.; Dixon, J.S.; Kuntz, I.D.; Venkataraghavan, R. Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. *J. Med. Chem.* **1988**, *31*, 722–729. [CrossRef] [PubMed]

16. Warwicker, J. Investigating protein-protein interaction surfaces using a reduced stereochemical and electrostatic model. *J. Mol. Biol.* **1989**, *206*, 381–395. [CrossRef]

17. Jiang, F.; Kim, S.H. "Soft docking": Matching of molecular surface cubes. *J. Mol. Biol.* **1991**, *219*, 79–102. [CrossRef]

18. Meng, X.Y.; Zhang, H.X.; Mezei, M.; Cui, M. Molecular docking: A powerful approach for structure-based drug discovery. *Curr. Comput. Aided. Drug Des.* **2011**, *7*, 146–157. [CrossRef] [PubMed]

19. Amaro, R.E.; Baudry, J.; Chodera, J.; Demir, Ö.; McCammon, J.A.; Miao, Y.; Smith, J.C. Ensemble Docking in Drug Discovery. *Biophys. J.* **2018**, *114*, 2271–2278. [CrossRef]

20. Abagyan, R.; Totrov, M. High-throughput docking for lead generation. *Curr. Opin. Chem. Biol.* **2001**, *5*, 375–382. [CrossRef]

21. Carlson, H.A. Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol.* **2002**, *6*, 447–452. [CrossRef]

22. Greer, J.; Bush, B.L. Macromolecular shape and surface maps by solvent exclusion. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 303–307. [CrossRef] [PubMed]

23. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489. [CrossRef] [PubMed]

24. Rosenfeld, R.; Vajda, S.; DeLisi, C. Flexible docking and design. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 677–700. [CrossRef] [PubMed]

25. Sherman, W.; Beard, H.S.; Farid, R. Use of an induced fit receptor structure in virtual screening. *Chem. Biol. Drug Des.* **2006**, *67*, 83–84. [CrossRef] [PubMed]

26. Leach, A.R. Ligand docking to proteins with discrete side-chain flexibility. *J. Mol. Biol.* **1994**, *235*, 345–356. [CrossRef]

27. Ring, C.S.; Sun, E.; McKerrow, J.H.; Lee, G.K.; Rosenthal, P.J.; Kuntz, I.D.; Cohen, F.E. Structure-based inhibitor design by using protein models for the development of antiparasitic agents. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3583–3587. [CrossRef]

28. Coupez, B.; Lewis, R.A. Docking and scoring–theoretically easy, practically impossible? *Curr. Med. Chem.* **2006**, *13*, 2995–3003.

29. Ain, Q.U.; Aleksandrova, A.; Roessler, F.D.; Ballester, P.J. Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2015**, *5*, 405–424. [CrossRef]

30. Guedes, I.A.; Pereira, F.S.S.; Dardenne, L.E. Empirical scoring functions for structure-based virtual screening: Applications, critical aspects, and challenges. *Front. Pharmacol.* **2018**, *9*, 1089. [CrossRef]

31. Elokely, K.M.; Doerksen, R.J. Docking challenge: Protein sampling and molecular docking performance. *J. Chem. Inf. Model.* **2013**, *53*, 1934–1945. [CrossRef] [PubMed]

32. Pantsar, T.; Poso, A. Binding affinity via docking: fact and fiction. *Molecules* **2018**, *23*, 1899. [CrossRef] [PubMed]

33. Salmaso, V.; Moro, S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. *Front. Pharmacol.* **2018**, *9*, 923. [CrossRef] [PubMed]

34. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of molecular dynamics and related methods in drug discovery. *J. Med. Chem.* **2016**, *59*, 4035–4061. [CrossRef] [PubMed]

35. Alonso, H.; Bliznyuk, A.A.; Gready, J.E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* **2006**, *26*, 531–568. [CrossRef] [PubMed]

36. Bard, J.; Ercolani, K.; Svenson, K.; Olland, A.; Somers, W. Automated systems for protein crystallization. *Methods* **2004**, *34*, 329–347. [CrossRef] [PubMed]

37. Gavira, J.A. Current trends in protein crystallization. *Arch. Biochem. Biophys.* **2016**, *602*, 3–11. [CrossRef] [PubMed]

38. Dauter, Z.; Wlodawer, A. Progress in protein crystallography. *Protein Pept. Lett.* **2016**, *23*, 201–210. [CrossRef]

39. Grimes, J.M.; Hall, D.R.; Ashton, A.W.; Evans, G.; Owen, R.L.; Wagner, A.; McAuley, K.E.; von Delft, F.; Orville, A.M.; Sorensen, T.; et al. Where is crystallography going? *Acta Crystallogr. Sect. D* **2018**, *74*, 152–166. [CrossRef]

40. Durrant, J.D.; McCammon, J.A. Molecular dynamics simulations and drug discovery. *BMC Biol.* **2011**, *9*, 71. [CrossRef]

41. Karplus, M.; McCammon, J.A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652. [CrossRef] [PubMed]

42. Rastelli, G.; Degliesposti, G.; Del Rio, A.; Sgobba, M. Binding estimation after refinement, a new automated procedure for the refinement and rescoring of docked ligands in virtual screening. *Chem. Biol. Drug Des.* **2009**, *73*, 283–286. [CrossRef] [PubMed]

43. Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. *J. Comput. Chem.* **2011**, *32*, 866–877. [CrossRef] [PubMed]

44. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, *10*, 449–461. [CrossRef] [PubMed]

45. Pu, C.; Yan, G.; Shi, J.; Li, R. Assessing the performance of docking scoring function, FEP, MM-GBSA, and QM/MM-GBSA approaches on a series of PLK1 inhibitors. *Medchemcomm* **2017**, *8*, 1452–1458. [CrossRef] [PubMed]

46. Rastelli, G.; Pinzi, L. Refinement and rescoring of virtual screening results. *Front. Chem.* **2019**, *7*, 498. [CrossRef] [PubMed]

47. Gschwend, D.A.; Good, A.C.; Kuntz, I.D. Molecular docking towards drug discovery. *J. Mol. Recognit.* **1996**, *9*, 175–186. [CrossRef]

48. De Vivo, M.; Cavalli, A. Recent advances in dynamic docking for drug discovery. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1320. [CrossRef]

49. Shoichet, B.K.; McGovern, S.L.; Wei, B.; Irwin, J.J. Lead discovery using molecular docking. *Curr. Opin. Chem. Biol.* **2002**, *6*, 439–446. [CrossRef]

50. Ferreira, L.G.; Dos Santos, R.N.; Oliva, G.; Andricopulo, A.D. Molecular docking and structure-based drug design strategies. *Molecules* **2015**, *20*, 13384–13421. [CrossRef]

51. Kinnings, S.L.; Liu, N.; Buchmeier, N.; Tonge, P.J.; Xie, L.; Bourne, P.E. Drug discovery using chemical systems biology: Repositioning the safe medicine Comtan to treat multi-drug and extensively drug resistant tuberculosis. *PLoS Comput. Biol.* **2009**, *5*, e1000423. [CrossRef] [PubMed]

52. Pinzi, L.; Lherbet, C.; Baltas, M.; Pellati, F.; Rastelli, G. In silico repositioning of cannabigerol as a novel inhibitor of the enoyl acyl carrier protein (ACP) reductase. *Molecules* **2019**, *24*, 2567. [CrossRef] [PubMed]

53. Anighoro, A.; Pinzi, L.; Marverti, G.; Bajorath, J.; Rastelli, G. Heat shock protein 90 and serine/threonine kinase B-Raf inhibitors have overlapping chemical space. *RSC Adv.* **2017**, *7*, 31069–31074. [CrossRef]

54. Li, H.; Gao, Z.; Kang, L.; Zhang, H.; Yang, K.; Yu, K.; Luo, X.; Zhu, W.; Chen, K.; Shen, J.; et al. TarFisDock: a web server for identifying drug targets with docking approach. *Nucleic Acids Res.* **2006**, *34*, W219–W224. [CrossRef]

55. Lyu, J.; Wang, S.; Balius, T.E.; Singh, I.; Levit, A.; Moroz, Y.S.; O'Meara, M.J.; Che, T.; Algaa, E.; Tolmachova, K.; et al. Ultra-large library docking for discovering new chemotypes. *Nature* **2019**, *566*, 224–229. [CrossRef]

56. Graziani, D.; Caligari, S.; Callegari, E.; De Toma, C.; Longhi, M.; Frigerio, F.; Dilernia, R.; Menegon, S.; Pinzi, L.; Pirona, L.; et al. Evaluation of amides, carbamates, sulfonamides, and ureas of 4-Prop-2-ynylidenecycloalkylamine as potent, selective, and bioavailable negative allosteric modulators of metabotropic glutamate receptor 5. *J. Med. Chem.* **2019**, *62*, 1246–1273. [CrossRef]

57. Ramsay, R.R.; Popovic-Nikolic, M.R.; Nikolic, K.; Uliassi, E.; Bolognesi, M.L. A perspective on multi-target drug discovery and design for complex diseases. *Clin. Transl. Med.* **2018**, *7*, 3. [CrossRef]

58. Lee, A.; Lee, K.; Kim, D. Using reverse docking for target identification and its applications for drug discovery. *Expert Opin. Drug Discov.* **2016**, *11*, 707–715. [CrossRef]

59. Anighoro, A.; Bajorath, J.; Rastelli, G. Polypharmacology: Challenges and opportunities in drug discovery. *J. Med. Chem.* **2014**, *57*, 7874–7887. [CrossRef]

60. Gloriam, D.E. Bigger is better in virtual drug screens. *Nature* **2019**, *566*, 193–194. [CrossRef]

61. Hazarika, R.R.; Sostaric, N.; Sun, Y.; van Noort, V. Large-scale docking predicts that sORF-encoded peptides may function through protein-peptide interactions in Arabidopsis thaliana. *PLoS ONE* **2018**, *13*, e0205179. [CrossRef] [PubMed]

62. LaBute, M.X.; Zhang, X.; Lenderman, J.; Bennion, B.J.; Wong, S.E.; Lightstone, F.C. Adverse drug reaction prediction using scores produced by large-scale drug-protein target docking on high-performance computing machines. *PLoS ONE* **2014**, *9*, e106298. [CrossRef] [PubMed]

63. de Ruyck, J.; Brysbaert, G.; Blossey, R.; Lensink, M.F. Molecular docking as a popular tool in drug design, an in silico travel. *Adv. Appl. Bioinform. Chem.* **2016**, *9*, 1–11. [CrossRef] [PubMed]

64. Lavecchia, A. Machine-learning approaches in drug discovery: methods and applications. *Drug Discov. Today* **2015**, *20*, 318–331. [CrossRef] [PubMed]

65. March-Vila, E.; Pinzi, L.; Sturm, N.; Tinivella, A.; Engkvist, O.; Chen, H.; Rastelli, G. On the integration of in silico drug design methods for drug repurposing. *Front. Pharmacol.* **2017**, *8*, 298. [CrossRef] [PubMed]

66. Wilson, G.L.; Lill, M.A. Integrating structure-based and ligand-based approaches for computational drug design. *Future Med. Chem.* **2011**, *3*, 735–750. [CrossRef]

67. Maggiora, G.; Vogt, M.; Stumpfe, D.; Bajorath, J. Molecular similarity in medicinal chemistry. *J. Med. Chem.* **2014**, *57*, 3186–3204. [CrossRef]

68. Sutherland, J.J.; Nandigam, R.K.; Erickson, J.A.; Vieth, M. Lessons in molecular recognition. 2. Assessing and improving cross-docking accuracy. *J. Chem. Inf. Model.* **2007**, *47*, 2293–2302. [CrossRef]

69. Broccatelli, F.; Brown, N. Best of both worlds: on the complementarity of ligand-based and structure-based virtual screening. *J. Chem. Inf. Model.* **2014**, *54*, 1634–1641. [CrossRef]

70. Pinzi, L.; Caporuscio, F.; Rastelli, G. Selection of protein conformations for structure-based polypharmacology studies. *Drug Discov. Today* **2018**, *23*, 1889–1896. [CrossRef]

71. Jain, A.N. Effects of protein conformation in docking: improved pose prediction through protein pocket adaptation. *J. Comput. Aided. Mol. Des.* **2009**, *23*, 355–374. [CrossRef] [PubMed]

72. Verdonk, M.L.; Mortenson, P.N.; Hall, R.J.; Hartshorn, M.J.; Murray, C.W. Protein−ligand docking against non-native protein conformers. *J. Chem. Inf. Model.* **2008**, *48*, 2214–2225. [CrossRef] [PubMed]

73. Xu, X.; Yan, C.; Zou, X. Improving binding mode and binding affinity predictions of docking by ligand-based search of protein conformations: evaluation in D3R grand challenge 2015. *J. Comput. Aided. Mol. Des.* **2017**, *31*, 689–699. [CrossRef] [PubMed]

74. Kumar, A.; Zhang, K.Y. A cross docking pipeline for improving pose prediction and virtual screening performance. *J. Comput. Aided. Mol. Des.* **2018**, *32*, 163–173. [CrossRef] [PubMed]

75. Kumar, A.; Zhang, K.Y.J. Advances in the development of shape similarity methods and their application in drug discovery. *Front. Chem.* **2018**, *6*, 315. [CrossRef] [PubMed]

76. Perryman, A.L.; Santiago, D.N.; Forli, S.; Martins, D.S.; Olson, A.J. Virtual screening with AutoDock Vina and the common pharmacophore engine of a low diversity library of fragments and hits against the three allosteric sites of HIV integrase: participation in the SAMPL4 protein-ligand binding challenge. *J. Comput. Aided. Mol. Des.* **2014**, *28*, 429–441. [CrossRef] [PubMed]

77. Jiang, L.; Rizzo, R.C. Pharmacophore-based similarity scoring for DOCK. *J. Phys. Chem. B* **2015**, *119*, 1083–1102. [CrossRef]

78. Moustakas, D.T.; Lang, P.T.; Pegg, S.; Pettersen, E.; Kuntz, I.D.; Brooijmans, N.; Rizzo, R.C. Development and validation of a modular, extensible docking program: DOCK 5. *J. Comput. Aided. Mol. Des.* **2006**, *20*, 601–619. [CrossRef]

79. Lang, P.T.; Brozell, S.R.; Mukherjee, S.; Pettersen, E.F.; Meng, E.C.; Thomas, V.; Rizzo, R.C.; Case, D.A.; James, T.L.; Kuntz, I.D. DOCK 6: Combining techniques to model RNA-small molecule complexes. *RNA* **2009**, *15*, 1219–1230. [CrossRef]

80. Anighoro, A.; Bajorath, J. Three-dimensional similarity in molecular docking: prioritizing ligand poses on the basis of experimental binding modes. *J. Chem. Inf. Model.* **2016**, *56*, 580–587. [CrossRef]

81. Kumar, A.; Zhang, K.Y.J. Application of shape similarity in pose selection and virtual screening in CSARdock2014 exercise. *J. Chem. Inf. Model.* **2016**, *56*, 965–973. [CrossRef] [PubMed]

82. Jasper, J.B.; Humbeck, L.; Brinkjost, T.; Koch, O. A novel interaction fingerprint derived from per atom score contributions: exhaustive evaluation of interaction fingerprint performance in docking based virtual screening. *J. Cheminform.* **2018**, *10*, 15. [CrossRef] [PubMed]

83. Da, C.; Kireev, D. Structural protein–ligand interaction fingerprints (SPLIF) for structure-based virtual screening: method and benchmark study. *J. Chem. Inf. Model.* **2014**, *54*, 2555–2561. [CrossRef]

84. Liu, J.; Su, M.; Liu, Z.; Li, J.; Li, Y.; Wang, R. Enhance the performance of current scoring functions with the aid of 3D protein-ligand interaction fingerprints. *BMC Bioinform.* **2017**, *18*, 343. [CrossRef] [PubMed]

85. Caporuscio, F.; Rastelli, G. Exploiting computationally derived out-of-the-box protein conformations for drug design. *Future Med. Chem.* **2016**, *8*, 1887–1897. [CrossRef] [PubMed]

86. Wang, Y.Y.; Li, L.; Chen, T.; Chen, W.; Xu, Y. Microsecond molecular dynamics simulation of Aβ42 and identification of a novel dual inhibitor of Aβ42 aggregation and BACE1 activity. *Acta Pharmacol. Sin.* **2013**, *34*, 1243–1250. [CrossRef] [PubMed]

87. Spyrakis, F.; Benedetti, P.; Decherchi, S.; Rocchia, W.; Cavalli, A.; Alcaro, S.; Ortuso, F.; Baroni, M.; Cruciani, G. A pipeline to enhance ligand virtual screening: Integrating molecular dynamics and fingerprints for ligand and proteins. *J. Chem. Inf. Model.* **2015**, *55*, 2256–2274. [CrossRef]

88. Park, H.S.; Jun, C.H. A simple and fast algorithm for K-medoids clustering. *Expert Syst. Appl.* **2009**, *36*, 3336–3341. [CrossRef]

89. Baroni, M.; Cruciani, G.; Sciabola, S.; Perruccio, F.; Mason, J.S. A common reference framework for analyzing/comparing proteins and ligands. fingerprints for ligands and proteins (FLAP): Theory and application. *J. Chem. Inf. Model.* **2007**, *47*, 279–294. [CrossRef]

90. Mysinger, M.M.; Carchia, M.; Irwin, J.J.; Shoichet, B.K. Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *J. Med. Chem.* **2012**, *55*, 6582–6594. [CrossRef]

91. Torrie, G.M.; Valleau, J.P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187–199. [CrossRef]

92. Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562–12566. [CrossRef] [PubMed]

93. Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151. [CrossRef]

94. Oleinikovas, V.; Saladino, G.; Cossins, B.P.; Gervasio, F.L. Understanding cryptic pocket formation in protein targets by enhanced sampling simulations. *J. Am. Chem. Soc.* **2016**, *138*, 14257–14263. [CrossRef] [PubMed]

95. Leontiadou, H.; Galdadas, I.; Athanasiou, C.; Cournia, Z. Insights into the mechanism of the PIK3CA E545K activating mutation using MD simulations. *Sci. Rep.* **2018**, *8*, 15544. [CrossRef]

96. Yang, L.Q.; Sang, P.; Tao, Y.; Fu, Y.X.; Zhang, K.Q.; Xie, Y.H.; Liu, S.Q. Protein dynamics and motions in relation to their functions: several case studies and the underlying mechanisms. *J. Biomol. Struct. Dyn.* **2014**, *32*, 372–393. [CrossRef]

97. Herbert, C.; Schieborr, U.; Saxena, K.; Juraszek, J.; De Smet, F.; Alcouffe, C.; Bianciotto, M.; Saladino, G.; Sibrac, D.; Kudlinzki, D.; et al. Molecular mechanism of SSR128129E, an extracellularly acting, small-molecule, allosteric inhibitor of FGF receptor signaling. *Cancer Cell* **2013**, *23*, 489–501. [CrossRef]

98. D'Abramo, M.; Besker, N.; Chillemi, G.; Grottesi, A. Modeling conformational transitions in kinases by molecular dynamics simulations: achievements, difficulties, and open challenges. *Front. Genet.* **2014**, *5*, 128. [CrossRef]

99. Meng, Y.; Lin, Y.; Roux, B. Computational study of the "DFG-flip" conformational transition in c-Abl and c-Src tyrosine kinases. *J. Phys. Chem. B.* **2015**, *119*, 1443–1456. [CrossRef]

100. Berteotti, A.; Cavalli, A.; Branduardi, D.; Gervasio, F.L.; Recanatini, M.; Parrinello, M. Protein conformational transitions: The closure mechanism of a kinase explored by atomistic simulations. *J. Am. Chem. Soc.* **2009**, *131*, 244–250. [CrossRef]

101. Morando, M.A.; Saladino, G.; D'Amelio, N.; Pucheta-Martinez, E.; Lovera, S.; Lelli, M.; López-Méndez, B.; Marenchino, M.; Campos-Olivas, R.; Gervasio, F.L. Conformational selection and induced fit mechanisms in the binding of an anticancer drug to the c-src kinase. *Sci. Rep.* **2016**, *6*, 24439. [CrossRef] [PubMed]

102. Comitani, F.; Gervasio, F.L. Exploring cryptic pockets formation in targets of pharmaceutical interest with SWISH. *J. Chem. Theory Comput.* **2018**, *14*, 3321–3331. [CrossRef] [PubMed]

103. Pisani, P.; Caporuscio, F.; Carlino, L.; Rastelli, G. Molecular dynamics simulations and classical multidimensional scaling unveil new metastable states in the conformational landscape of CDK2. *PLoS ONE* **2016**, *11*, e0154066. [CrossRef] [PubMed]

104. Gioia, D.; Bertazzo, M.; Recanatini, M.; Masetti, M.; Cavalli, A. Dynamic docking: A paradigm shift in computational drug discovery. *Molecules* **2017**, *22*, 2029. [CrossRef] [PubMed]

105. Degliesposti, G.; Portioli, C.; Parenti, M.D.; Rastelli, G. BEAR, a novel virtual screening methodology for drug discovery. *J. Biomol. Screen.* **2011**, *16*, 129–133. [CrossRef] [PubMed]

106. Parenti, M.D.; Rastelli, G. Advances and applications of binding affinity prediction methods in drug discovery. *Biotechnol. Adv.* **2012**, *30*, 244–250. [CrossRef]

107. Rastelli, G. Emerging topics in structure-based virtual screening. *Pharm. Res.* **2013**, *30*, 1458–1463. [CrossRef]

108. Gilson, M.K.; Zhou, H.X. Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42. [CrossRef]

109. Kollman, P.A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897. [CrossRef]

110. Jorgensen, W.L.; Thomas, L.L. Perspective on free-energy perturbation calculations for chemical equilibria. *J. Chem. Theory Comput.* **2008**, *4*, 869–876. [CrossRef]

111. Limongelli, V.; Bonomi, M.; Parrinello, M. Funnel metadynamics as accurate binding free-energy method. *Proc. Natl. Acad. Sci. USA.* **2013**, *110*, 6358–6363. [CrossRef] [PubMed]

112. Clark, A.J.; Tiwary, P.; Borrelli, K.; Feng, S.; Miller, E.B.; Abel, R.; Friesner, R.A.; Berne, B.J. Prediction of protein–ligand binding poses via a combination of induced fit docking and metadynamics simulations. *J. Chem. Theory Comput.* **2016**, *12*, 2990–2998. [CrossRef] [PubMed]

113. Lee, H.S.; Jo, S.; Lim, H.S.; Im, W. Application of binding free energy calculations to prediction of binding modes and affinities of MDM2 and MDMX inhibitors. *J. Chem. Inf. Model.* **2012**, *52*, 1821–1832. [CrossRef] [PubMed]

114. Bhati, A.P.; Wan, S.; Wright, D.W.; Coveney, P.V. Rapid, accurate, precise, and reliable relative free energy prediction using ensemble based thermodynamic integration. *J. Chem. Theory Comput.* **2017**, *13*, 210–222. [CrossRef] [PubMed]

115. Wang, C.; Greene, D.; Xiao, L.; Qi, R.; Luo, R. Recent developments and applications of the MMPBSA method. *Front. Mol. Biosci.* **2018**, *4*, 87. [CrossRef] [PubMed]

116. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

117. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

118. Wojcikowski, M.; Ballester, P.J.; Siedlecki, P.; Wójcikowski, M.; Ballester, P.J.; Siedlecki, P. Performance of machine-learning scoring functions in structure-based virtual screening. *Sci. Rep.* **2017**, *7*, 46710. [CrossRef]

119. Xie, Q.Q.; Zhong, L.; Pan, Y.L.; Wang, X.Y.; Zhou, J.P.; Di-wu, L.; Huang, Q.; Wang, Y.L.; Yang, L.L.; Xie, H.Z.; et al. Combined SVM-based and docking-based virtual screening for retrieving novel inhibitors of c-Met. *Eur. J. Med. Chem.* **2011**, *46*, 3675–3680. [CrossRef]

120. Leong, M.K.; Syu, R.G.; Ding, Y.L.; Weng, C.F. Prediction of N-methyl-D-aspartate receptor GluN1-ligand binding affinity by a novel SVM-Pose/SVM-score combinatorial ensemble docking scheme. *Sci. Rep.* **2017**, *7*, 40053. [CrossRef]

121. Ballester, P.J.; Mitchell, J.B. A machine learning approach to predicting protein–ligand binding affinity with applications to molecular docking. *Bioinformatics* **2010**, *26*, 1169–1175. [CrossRef] [PubMed]

122. Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind database: collection of binding affinities for protein−ligand complexes with known three-dimensional structures. *J. Med. Chem.* **2004**, *47*, 2977–2980. [CrossRef] [PubMed]

123. Wang, C.; Zhang, Y. Improving scoring-docking-screening powers of protein-ligand scoring functions using random forest. *J. Comput. Chem.* **2017**, *38*, 169–177. [CrossRef] [PubMed]

124. Trott, O.; Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [CrossRef] [PubMed]

125. Pereira, J.C.; Caffarena, E.R.; dos Santos, C.N. Boosting docking-based virtual screening with deep learning. *J. Chem. Inf. Model.* **2016**, *56*, 2495–2506. [CrossRef] [PubMed]

126. Wang, N.; Wang, L.; Xie, X.Q. ProSelection: A novel algorithm to select proper protein structure subsets for in silico target identification and drug discovery research. *J. Chem. Inf. Model.* **2017**, *57*, 2686–2698. [CrossRef] [PubMed]

127. Xu, X.; Huang, M.; Zou, X. Docking-based inverse virtual screening: methods, applications, and challenges. *Biophys. Rep.* **2018**, *4*, 1–16. [CrossRef] [PubMed]

128. Gao, Z.; Li, H.; Zhang, H.; Liu, X.; Kang, L.; Luo, X.; Zhu, W.; Chen, K.; Wang, X.; Jiang, H. PDTD: a web-accessible protein database for drug target identification. *BMC Bioinform.* **2008**, *9*, 104. [CrossRef] [PubMed]

129. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [CrossRef]

130. Kellenberger, E.; Muller, P.; Schalon, C.; Bret, G.; Foata, N.; Rognan, D. sc-PDB: An annotated database of druggable binding sites from the Protein Data Bank. *J. Chem. Inf. Model.* **2006**, *46*, 717–727. [CrossRef]

131. Kufareva, I.; Ilatovskiy, A.V.; Abagyan, R. Pocketome: An encyclopedia of small-molecule binding sites in 4D. *Nucleic Acids Res.* **2012**, *40*, D535–D540. [CrossRef] [PubMed]

132. Chen, X.; Ji, Z.L.; Chen, Y.Z. TTD: Therapeutic Target Database. *Nucleic Acids Res.* **2002**, *30*, 412–415. [CrossRef] [PubMed]

133. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K.; et al. Glide: A new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749. [CrossRef] [PubMed]

134. Halgren, T.A.; Murphy, R.B.; Friesner, R.A.; Beard, H.S.; Frye, L.L.; Pollard, W.T.; Banks, J.L. Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759. [CrossRef] [PubMed]

135. Park, K.; Cho, A.E. Using reverse docking to identify potential targets for ginsenosides. *J. Ginseng Res.* **2017**, *41*, 534–539. [CrossRef] [PubMed]

136. Nag, S.A.; Qin, J.J.; Wang, W.; Wang, M.H.; Wang, H.; Zhang, R. Ginsenosides as anticancer agents: in vitro and in vivo activities, structure-activity relationships, and molecular mechanisms of action. *Front. Pharmacol.* **2012**, *3*, 25. [CrossRef] [PubMed]

137. Luo, Q.; Zhao, L.; Hu, J.; Jin, H.; Liu, Z.; Zhang, L. The scoring bias in reverse docking and the score normalization strategy to improve success rate of target fishing. *PLoS ONE* **2017**, *12*, e0171433. [CrossRef]

138. Luo, H.; Fokoue-Nkoutche, A.; Singh, N.; Yang, L.; Hu, J.; Zhang, P. Molecular docking for prediction and interpretation of adverse drug reactions. *Comb. Chem. High. Throughput Screen.* **2018**, *21*, 314–322. [CrossRef]

139. Nogueira, M.S.; Koch, O. The development of target-specific machine learning models as scoring functions for docking-based target prediction. *J. Chem. Inf. Model.* **2019**, *59*, 1238–1252. [CrossRef]

140. Gaulton, A.; Bellis, L.J.; Bento, A.P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; et al. ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **2011**, *40*, D1100–D1107. [CrossRef]

141. Perez-Sanchez, H.; Wenzel, W. Optimization methods for virtual screening on novel computational architectures. *Curr. Comput. Aided. Drug Des.* **2011**, *7*, 44–52. [CrossRef] [PubMed]

142. Dong, D.; Xu, Z.; Zhong, W.; Peng, S. Parallelization of molecular docking: A review. *Curr. Top. Med. Chem.* **2018**, *18*, 1015–1028. [CrossRef] [PubMed]

143. Lapillo, M.; Tuccinardi, T.; Martinelli, A.; Macchia, M.; Giordano, A.; Poli, G. Extensive reliability evaluation of docking-based target-fishing strategies. *Int. J. Mol. Sci.* **2019**, *20*, 1023. [CrossRef] [PubMed]

144. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791. [CrossRef] [PubMed]

145. Chen, Y.Z.; Zhi, D.G. Ligand–protein inverse docking and its potential use in the computer search of protein targets of a small molecule. *Proteins* **2001**, *43*, 217–226. [CrossRef]

146. Xie, T.; Zhang, L.; Zhang, S.; Ouyang, L.; Cai, H.; Liu, B. ACTP: A webserver for predicting potential targets and relevant pathways of autophagy-modulating compounds. *Oncotarget* **2016**, *7*, 10015–10022. [CrossRef]

147. Wang, J.C.; Chu, P.Y.; Chen, C.M.; Lin, J.H. idTarget: A web server for identifying protein targets of small chemical molecules with robust scoring functions and a divide-and-conquer docking approach. *Nucleic Acids Res.* **2012**, *40*, W393–W399. [CrossRef]

148. Wang, Z.Y.; Kang, H.; Ji, L.L.; Yang, Y.Q.; Liu, T.; Cao, Z.W.; Morahan, G.; Wang, Z.T. Proteomic characterization of the possible molecular targets of pyrrolizidine alkaloid isoline-induced hepatotoxicity. *Environ. Toxicol. Pharmacol.* **2012**, *34*, 608–617. [CrossRef]

149. Ji, Z.L.; Wang, Y.; Yu, L.; Han, L.Y.; Zheng, C.J.; Chen, Y.Z. In silico search of putative adverse drug reaction related proteins as a potential tool for facilitating drug adverse effect prediction. *Toxicol. Lett.* **2006**, *164*, 104–112. [CrossRef]

150. Eric, S.; Ke, S.; Barata, T.; Solmajer, T.; Antic Stankovic, J.; Juranic, Z.; Savic, V.; Zloh, M. Target fishing and docking studies of the novel derivatives of aryl-aminopyridines with potential anticancer activity. *Bioorg. Med. Chem.* **2012**, *20*, 5220–5228. [CrossRef]

151. Zhang, L.; Fu, L.; Zhang, S.; Zhang, J.; Zhao, Y.; Zheng, Y.; He, G.; Yang, S.; Ouyang, L.; Liu, B. Discovery of a small molecule targeting ULK1-modulated cell death of triple negative breast cancer in vitro and in vivo. *Chem. Sci.* **2017**, *8*, 2687–2701. [CrossRef] [PubMed]

152. Hassan, A.H.E.; Yoo, S.Y.; Lee, K.W.; Yoon, Y.M.; Ryu, H.W.; Jeong, Y.; Shin, J.S.; Kang, S.Y.; Kim, S.Y.; Lee, H.H.; et al. Repurposing mosloflavone/5,6,7-trimethoxyflavone-resveratrol hybrids: Discovery of novel p38-α MAPK inhibitors as potent interceptors of macrophage-dependent production of proinflammatory mediators. *Eur. J. Med. Chem.* **2019**, *180*, 253–267. [CrossRef] [PubMed]

153. Chang, D.T.H.; Oyang, Y.J.; Lin, J.H. MEDock: A web server for efficient prediction of ligand binding sites based on a novel optimization algorithm. *Nucleic Acids Res.* **2005**, *33*, W233–W238. [CrossRef] [PubMed]

154. Diller, D.J.; Merz, K.M.J. High throughput docking for library design and library prioritization. *Proteins* **2001**, *43*, 113–124. [CrossRef]

155. Muhammed, M.T.; Aki-Yalcin, E. Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chem. Biol. Drug Des.* **2019**, *93*, 12–20. [CrossRef] [PubMed]

156. Daga, P.R.; Patel, R.Y.; Doerksen, R.J. Template-based protein modeling: recent methodological advances. *Curr. Top. Med. Chem.* **2010**, *10*, 84–94. [CrossRef] [PubMed]

157. Klein, E.; Bourdette, D. Postmarketing adverse drug reactions: A duty to report? *Neurol. Clin. Pract.* **2013**, *3*, 288–294. [CrossRef]

158. Yoo, S.; Noh, K.; Shin, M.; Park, J.; Lee, K.H.; Nam, H.; Lee, D. In silico profiling of systemic effects of drugs to predict unexpected interactions. *Sci. Rep.* **2018**, *8*, 1612. [CrossRef]

159. Dimitri, G.M.; Lió, P. DrugClust: A machine learning approach for drugs side effects prediction. *Comput. Biol. Chem.* **2017**, *68*, 204–210. [CrossRef]

160. Li, Z.; Han, P.; You, Z.H.; Li, X.; Zhang, Y.; Yu, H.; Nie, R.; Chen, X. In silico prediction of drug-target interaction networks based on drug chemical structure and protein sequences. *Sci. Rep.* **2017**, *7*, 11174. [CrossRef]

161. Ekins, S. Predicting undesirable drug interactions with promiscuous proteins in silico. *Drug Discov. Today* **2004**, *9*, 276–285. [CrossRef]

162. Chen, Y.Z.; Ung, C.Y. Prediction of potential toxicity and side effect protein targets of a small molecule by a ligand-protein inverse docking approach. *J. Mol. Graph. Model.* **2001**, *20*, 199–218. [CrossRef]

163. Yang, L.; Chen, J.; He, L. Harvesting candidate genes responsible for serious adverse drug reactions from a chemical-protein interactome. *PLoS Comput. Biol.* **2009**, *5*, e1000441. [CrossRef] [PubMed]

164. Li, C.Y.; Yu, Q.; Ye, Z.Q.; Sun, Y.; He, Q.; Li, X.M.; Zhang, W.; Luo, J.; Gu, X.; Zheng, X.; et al. A nonsynonymous SNP in human cytosolic sialidase in a small Asian population results in reduced enzyme activity: Potential link with severe adverse reactions to oseltamivir. *Cell Res.* **2007**, *17*, 357–362. [CrossRef] [PubMed]

165. Kuhn, M.; Letunic, I.; Jensen, L.J.; Bork, P. The SIDER database of drugs and side effects. *Nucleic Acids Res.* **2016**, *44*, D1075–D1079. [CrossRef] [PubMed]

166. Jaundoo, R.; Bohmann, J.; Gutierrez, G.E.; Klimas, N.; Broderick, G.; Craddock, T.J.A. Using a consensus docking approach to predict adverse drug reactions in combination drug therapies for gulf war illness. *Int. J. Mol. Sci.* **2018**, *19*, 3355. [CrossRef]

167. Kola, I.; Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* **2004**, *3*, 711–716. [CrossRef]

168. Zhang, W.; Bai, Y.; Wang, Y.; Xiao, W. Polypharmacology in drug discovery: A review from systems pharmacology perspective. *Curr. Pharm. Des.* **2016**, *22*, 3171–3181. [CrossRef]

169. Lepailleur, A.; Freret, T.; Lemaître, S.; Boulouard, M.; Dauphin, F.; Hinschberger, A.; Dulin, F.; Lesnard, A.; Bureau, R.; Rault, S. Dual histamine H3R/serotonin 5-HT4R ligands with antiamnesic properties: pharmacophore-based virtual screening and polypharmacology. *J. Chem. Inf. Model.* **2014**, *54*, 1773–1784. [CrossRef] [PubMed]

170. Wei, D.; Jiang, X.; Zhou, L.; Chen, J.; Chen, Z.; He, C.; Yang, K.; Liu, Y.; Pei, J.; Lai, L. Discovery of multitarget inhibitors by combining molecular docking with common pharmacophore matching. *J. Med. Chem.* **2008**, *51*, 7882–7888. [CrossRef] [PubMed]

171. Rastelli, G.; Pinzi, L. Computational polypharmacology comes of age. *Front. Pharmacol.* **2015**, *6*, 157. [CrossRef] [PubMed]

172. Zhang, W.; Pei, J.; Lai, L. Computational multitarget drug design. *J. Chem. Inf. Model.* **2017**, *57*, 403–412. [CrossRef] [PubMed]

173. Selvam, B.; Porter, S.L.; Tikhonova, I.G. Addressing selective polypharmacology of antipsychotic drugs targeting the bioaminergic receptors through receptor dynamic conformational ensembles. *J. Chem. Inf. Model.* **2013**, *53*, 1761–1774. [CrossRef] [PubMed]

174. Minie, M.; Chopra, G.; Sethi, G.; Horst, J.; White, G.; Roy, A.; Hatti, K.; Samudrala, R. CANDO and the infinite drug discovery frontier. *Drug Discov. Today* **2014**, *19*, 1353–1363. [CrossRef] [PubMed]

175. Chopra, G.; Samudrala, R. Exploring polypharmacology in drug discovery and repurposing using the CANDO platform. *Curr. Pharm. Des.* **2016**, *22*, 3109–3123. [CrossRef] [PubMed]

176. Luo, H.; Chen, J.; Shi, L.; Mikailov, M.; Zhu, H.; Wang, K.; He, L.; Yang, L. DRAR-CPI: A server for identifying drug repositioning potential and adverse drug reactions via the chemical-protein interactome. *Nucleic Acids Res.* **2011**, *39*, W492–W498. [CrossRef]

177. Luo, H.; Zhang, P.; Cao, X.H.; Du, D.; Ye, H.; Huang, H.; Li, C.; Qin, S.; Wan, C.; Shi, L.; et al. DPDR-CPI, a server that predicts drug positioning and drug repositioning via chemical-protein interactome. *Sci. Rep.* **2016**, *6*, 35996. [CrossRef]

178. Hurle, M.R.; Yang, L.; Xie, Q.; Rajpal, D.K.; Sanseau, P.; Agarwal, P. Computational drug repositioning: From data to therapeutics. *Clin. Pharmacol. Ther.* **2013**, *93*, 335–341. [CrossRef] [PubMed]

179. Keiser, M.J.; Setola, V.; Irwin, J.J.; Laggner, C.; Abbas, A.I.; Hufeisen, S.J.; Jensen, N.H.; Kuijer, M.B.; Matos, R.C.; Tran, T.B.; et al. Predicting new molecular targets for known drugs. *Nature* **2009**, *462*, 175–181. [CrossRef]

180. Liu, X.; Zhu, F.; Ma, X.H.; Shi, Z.; Yang, S.Y.; Wei, Y.Q.; Chen, Y.Z. Predicting targeted polypharmacology for drug repositioning and multi- target drug discovery. *Curr. Med. Chem.* **2013**, *20*, 1646–1661. [CrossRef]

181. Nugent, T.; Plachouras, V.; Leidner, J.L. Computational drug repositioning based on side-effects mined from social media. *PeerJ Comput. Sci.* **2016**, *2*, e46. [CrossRef]

182. Defranchi, E.; Schalon, C.; Messa, M.; Onofri, F.; Benfenati, F.; Rognan, D. Binding of protein kinase inhibitors to synapsin I inferred from pair-wise binding site similarity measurements. *PLoS ONE* **2010**, *5*, e12214. [CrossRef]

183. Barneh, F.; Jafari, M.; Mirzaie, M. Updates on drug-target network; facilitating polypharmacology and data integration by growth of DrugBank database. *Brief. Bioinform.* **2016**, *17*, 1070–1080. [CrossRef] [PubMed]

184. Kharkar, P.S.; Warrier, S.; Gaud, R.S. Reverse docking: A powerful tool for drug repositioning and drug rescue. *Future Med. Chem.* **2014**, *6*, 333–342. [CrossRef] [PubMed]

185. Wenning, G.K.; O'Connell, M.T.; Patsalos, P.N.; Quinn, N.P. A clinical and pharmacokinetic case study of an interaction of levodopa and antituberculous therapy in Parkinson's disease. *Mov. Disord.* **1995**, *10*, 664–667. [CrossRef] [PubMed]

186. Shen, C.H.; Chou, C.H.; Liu, F.C.; Lin, T.Y.; Huang, W.Y.; Wang, Y.C.; Kao, C.H. Association Between Tuberculosis and Parkinson Disease: A Nationwide, Population-Based Cohort Study. *Medicine* **2016**, *95*, e2883. [CrossRef] [PubMed]

187. Dakshanamurthy, S.; Issa, N.T.; Assefnia, S.; Seshasayee, A.; Peters, O.J.; Madhavan, S.; Uren, A.; Brown, M.L.; Byers, S.W. Predicting new indications for approved drugs using a proteochemometric method. *J. Med. Chem.* **2012**, *55*, 6832–6848. [CrossRef] [PubMed]

188. Wishart, D.S.; Knox, C.; Guo, A.C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: A comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* **2006**, *34*, D668–D672. [CrossRef] [PubMed]

189. Wishart, D.S.; Feunang, Y.D.; Guo, A.C.; Lo, E.J.; Marcu, A.; Grant, J.R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; et al. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074–D1082. [CrossRef] [PubMed]

190. Liu, T.; Lin, Y.; Wen, X.; Jorissen, R.N.; Gilson, M.K. BindingDB: A web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201. [CrossRef] [PubMed]

191. Koehn, F.E.; Carter, G.T. The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discov.* **2005**, *4*, 206–220. [CrossRef]

192. Harvey, A.L. Natural products in drug discovery. *Drug Discov. Today* **2008**, *13*, 894–901. [CrossRef] [PubMed]

193. Thomford, N.E.; Senthebane, D.A.; Rowe, A.; Munro, D.; Seele, P.; Maroyi, A.; Dzobo, K. Natural products for drug discovery in the 21st century: innovations for novel drug discovery. *Int. J. Mol. Sci.* **2018**, *19*, 1578. [CrossRef] [PubMed]

194. Kinnings, S.L.; Liu, N.; Tonge, P.J.; Jackson, R.M.; Xie, L.; Bourne, P.E. A machine learning-based method to improve docking scoring functions and its application to drug repurposing. *J. Chem. Inf. Model.* **2011**, *51*, 408–419. [CrossRef] [PubMed]

195. Labbé, C.M.; Rey, J.; Lagorce, D.; Vavruša, M.; Becot, J.; Sperandio, O.; Villoutreix, B.O.; Tufféry, P.; Miteva, M.A. MTiOpenScreen: A web server for structure-based virtual screening. *Nucleic Acids Res.* **2015**, *43*, W448–W454. [CrossRef] [PubMed]

196. Lagarde, N.; Rey, J.; Gyulkhandanyan, A.; Tufféry, P.; Miteva, M.A.; Villoutreix, B.O. Online structure-based screening of purchasable approved drugs and natural compounds: Retrospective examples of drug repositioning on cancer targets. *Oncotarget* **2018**, *9*, 32346–32361. [CrossRef] [PubMed]

197. Grosdidier, A.; Zoete, V.; Michielin, O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* **2011**, *39*, W270–W277. [CrossRef]

198. Irwin, J.J.; Shoichet, B.K.; Mysinger, M.M.; Huang, N.; Colizzi, F.; Wassam, P.; Cao, Y. Automated docking screens: A feasibility study. *J. Med. Chem.* **2009**, *52*, 5712–5720. [CrossRef]

199. Gawehn, E.; Hiss, J.A.; Brown, J.B.; Schneider, G. Advancing drug discovery via GPU-based deep learning. *Expert Opin. Drug Discov.* **2018**, *13*, 579–582. [CrossRef]

200. Loukatou, S.; Papageorgiou, L.; Fakourelis, P.; Filntisi, A.; Polychronidou, E.; Bassis, I.; Megalooikonomou, V.; Makałowski, W.; Vlachakis, D.; Kossida, S. Molecular dynamics simulations through GPU video games technologies. *J. Mol. Biochem.* **2014**, *3*, 64–71.

201. Chen, X.W.; Lin, X. Big data deep learning: challenges and perspectives. *IEEE Access* **2014**, *2*, 514–525. [CrossRef]

202. Stone, J.E.; Hardy, D.J.; Ufimtsev, I.S.; Schulten, K. GPU-accelerated molecular modeling coming of age. *J. Mol. Graph. Model.* **2010**, *29*, 116–125. [CrossRef] [PubMed]

203. Dowden, H.; Munro, J. Trends in clinical success rates and therapeutic focus. *Nat. Rev. Drug Discov.* **2019**, *18*, 495–496. [CrossRef] [PubMed]

204. Kinch, M.S.; Haynesworth, A.; Kinch, S.L.; Hoyer, D. An overview of FDA-approved new molecular entities: 1827–2013. *Drug Discov. Today* **2014**, *19*, 1033–1039. [CrossRef] [PubMed]