

Review

Strategies for *in Silico* Drug Discovery to Modulate Macromolecular Interactions Altered by Mutations

Pitambar Poudel¹ , Maria A. Miteva^{2,*} , Emil Alexov^{1,*} ¹Department of Physics and Astronomy, Clemson University, Clemson, SC 29634, USA²Université Paris Cité, CNRS UMR 8038 CiTCoM, Inserm, U1268 MCTR Paris, France*Correspondence: maria.mitev@inserm.fr (Maria A. Miteva); ealexov@clemson.edu (Emil Alexov)

Academic Editor: Ilaria Baglivo

Submitted: 29 August 2024 Revised: 22 September 2024 Accepted: 9 October 2024 Published: 16 April 2025

Abstract

Most human diseases have genetic components, frequently single nucleotide variants (SNVs), which alter the wild type characteristics of macromolecules and their interactions. A straightforward approach for correcting such SNVs-related alterations is to seek small molecules, potential drugs, that can eliminate disease-causing effects. Certain disorders are caused by altered protein-protein interactions, for example, Snyder-Robinson syndrome, the therapy for which focuses on the development of small molecules that restore the wild type homodimerization of spermine synthase. Other disorders originate from altered protein-nucleic acid interactions, as in the case of cancer; in these cases, the elimination of disease-causing effects requires small molecules that eliminate the effect of mutation and restore wild type p53-DNA affinity. Overall, especially for complex diseases, pathogenic mutations frequently alter macromolecular interactions. This effect can be direct, i.e., the alteration of wild type affinity and specificity, or indirect via alterations in the concentration of the binding partners. Here, we outline progress made in methods and strategies to computationally identify small molecules capable of altering macromolecular interactions in a desired manner, reducing or increasing the binding affinity, and eliminating the disease-causing effect. When applicable, we provide examples of the outlined general strategy. Successful cases are presented at the end of the work.

Keywords: macromolecular interactions; protein-protein binding; protein-DNA binding; mutations; small molecules; drugs

1. Introduction

Macromolecular interactions, i.e., interactions involving proteins, DNA, or RNA, are crucial in various biological processes, serving as the molecular foundation for all essential cellular functions [1–3]. Among these interactions, receptor-ligand interactions are vital for cell signaling, enabling communication between cells and their environment. Indeed, ligands bind to their corresponding receptors with high specificity, triggering a cascade of cellular responses [4]. Protein-protein interactions (PPIs) are another example of macromolecular interactions involving two or more proteins forming a protein complex [5]. Protein complexes orchestrate various cellular activities by forming intricate networks that regulate pathways and processes [6]. It is estimated that a single protein can participate in dozens, possibly hundreds, of different interactions [7]. This has prompted the creation of many databases and tools for analyzing PPIs [8–10]. A subclass of PPIs involves membrane proteins that reside in cellular membranes and facilitate communications between the internal and external environments of the cell. These proteins mediate processes such as transport and signal transduction and are also frequently involved in interactions with other proteins or ligands [11,12]. Similarly, proteins interact with DNA and RNA to participate in the transcription and translation of genetic information, ultimately shaping the traits of an organism. Protein-nucleic acid interactions are also involved

in the formation of large macromolecular complexes such as ribosomes and histones/nucleosomes [13–15]. Examples of transient protein-nucleic acid interactions include DNA transcription factors and transport, translation, splicing and, in the case of RNA, silencing [16,17]. All the processes mentioned above are vital for the wild type function of the cell, and thus, understanding and manipulating these macromolecular interactions holds immense importance for drug discovery. For these reasons, many pharmaceutical developments target specific macromolecular interactions to modulate cellular functions. Here, we briefly mention a few examples of targeting such interactions which are altered by mutations.

Genetic disorders, i.e., diseases originating fully or partially from mutations of an individual's DNA, are also frequently associated with alterations of wild type macromolecular interactions [18–20]. In some cases, the mutation affects only one PPI (monogenic disorders), whereas in the case of complex disorders, mutations typically affect numerous PPIs. PPIs and, in particular, homodimerization, i.e., the interaction involving two identical monomers, is frequently seen as a key factor influencing the functionality of various proteins. However, mutations in the homodimer complex can lead to unregulated homodimerization, which in turn leads to pathogenicity. Indeed, in the case of Snyder-Robinson syndrome, it has been shown that most disease-inducing mutations lead to reduced homodimerization and



loss of function [21,22]. A particular example, the mutation F58L, located at the homodimer interface, is shown in Fig. 1A. This mutation was investigated computationally and experimentally, and it was demonstrated that it reduced homodimer affinity while having little effect on the stability of individual monomers [21]. In such cases, one can target mutations computationally to identify the small molecules required to restore the wild type homodimer affinity. In case of complex diseases, one should consider the effect of numerous mutations to understand PPIs at a cellular level and their association with disease [6]. Because of the complexity of interaction networks involving numerous macromolecules, interpreting the effects of genetic variants requires information from different levels, including structural information [23]. Perturbed PPIs may lead to aggregation, which is the main mechanism for many neurodegenerative diseases, including Alzheimer's disease, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis [24]. Of particular interest are mutations causing cancer since this is a complex disease influenced by many factors. Indeed, it has been demonstrated that cancer mutations frequently affect PPIs and are clustered within interaction hubs [25,26]. Due to the complexity of cancer causality, researchers have developed so-called "mutational signatures", i.e., sets of mutations that are associated with high cancer predisposition and provide insights into the causes of specific cancers [27]. One such example includes mutations in the p53 family members, i.e., p53, p63, and p73, which are common in many cancers, including head and neck tumors [28,29]. The study has shown that the mutations and the association between p53 members lead to cancers and diseases; consequently, p53 family cancers are a potential target for diagnostic and therapeutic approaches [30].

Genetic disorders are also associated with perturbed protein-nucleic acid interactions, and of particular interest are perturbed protein-DNA interactions resulting in an elevated risk of cancer [31]. A prominent example is p52 protein and its corresponding mutations, which lead to different cancers [32,33]. Another example of a disease caused by altered protein-DNA interactions is Rett syndrome, a neurological disorder caused by mutations in methyl-CpG-binding protein 2 (MeCp2) [34]. It has been shown computationally and experimentally that the most frequent mutations causing Rett syndrome reduce MeCp2-DNA affinity while having little effect on MeCp2 stability [35,36]. Fig. 1B shows a particular example of R133C mutation, which reduces MeCp2-DNA binding affinity with minimal effect on MeCp2 stability [35]. The list of examples of altered protein-DNA interactions leading to diseases is very long, mostly because protein-DNA interactions are involved in transcription, which is a key component of cellular function. Consequently, it is unsurprising that many diseases are associated with transcriptional disorders; however, developing drugs capable of modulating protein-DNA interactions remains challenging [37–39]. Altered protein-

RNA interactions are also among the leading causes of a range of diseases [40]. A particular example is metastasis associated lung adenocarcinoma transport 1 (MALAT1), a highly abundant nuclear long noncoding RNA that is overexpressed in several human cancers [41]. A single point mutation is assumed to disrupt the triple helix stability of MALAT1, suggesting a pivotal role of this structure in enabling MALAT1 accumulation [42,43]. Sometimes, altered macromolecular interactions in viruses can be targeted for antiviral drug discovery. For example, influenza A viruses are major human pathogens with the potential to cause periodic pandemics. The NS1 protein of the influenza A virus (NS1A) is responsible for protecting the virus from the host's immune defenses. The RNA-binding domain (RBD) of NS1A forms a homodimer to recognize double-stranded RNA (dsRNA); however, mutations in this domain may affect the dsRNA binding affinity. One such mutation is R38A, as shown in Fig. 1C, which eliminates the dsRNA binding ability by avoiding the unique Arg38-Arg38 pair that is crucial for dsRNA binding [44]. Hence, such a positive pocket in the RNA-binding domain of the influenza virus can be targeted computationally to identify small molecules capable of binding to this pocket to induce the disruption of the Arg38 pair.

Although our primary focus is on altered macromolecular interactions due to mutation, it is important to note that both overexpression and underexpression of the corresponding protein can sometimes lead to disease. An example of this phenomenon is calpain, a heterocomplex comprising two proteins. Both overexpression and underexpression of calpain have been associated with several neuropathological conditions, including Alzheimer's disease, cerebellar ataxia, and different cancers [45–47]. Another example involves the human epidermal growth factor receptor 2 (*HER2*) gene. Amplification of the *HER2* gene and overexpression of the *HER2* protein has been linked to several types of cancer, including breast, colon, gastric, esophageal, and endometrial [48]. On the other hand, underexpressed proteins, such as *PIK3R1*, are linked to breast cancer [49]. The underexpression of different tumor suppressors is also a leading cause of many cancers [50,51]. For example, Protein phosphatase 2A (*PP2A*) is a critical tumor suppressor, and its inhibition promotes various malignant characteristics of human cancer cells. One such cancerous inhibitor of *PP2A* is *CIP2A*, which supports the activity of several critical cancer drivers (*Akt*, *MYC*, *E2F1*) and promotes malignancy in most cancer types via *PP2A* inhibition [52]. All of the above-mentioned cases involving genetic disorders, cancers, and other diseases indicate that mutation-altered macromolecular interactions have pathogenic effects and indicate the necessity of the development of computational strategies to identify drugs capable of eliminating the effect of mutations on macromolecular interactions. Next, we will outline the computational methods and strategies involved in such a process.

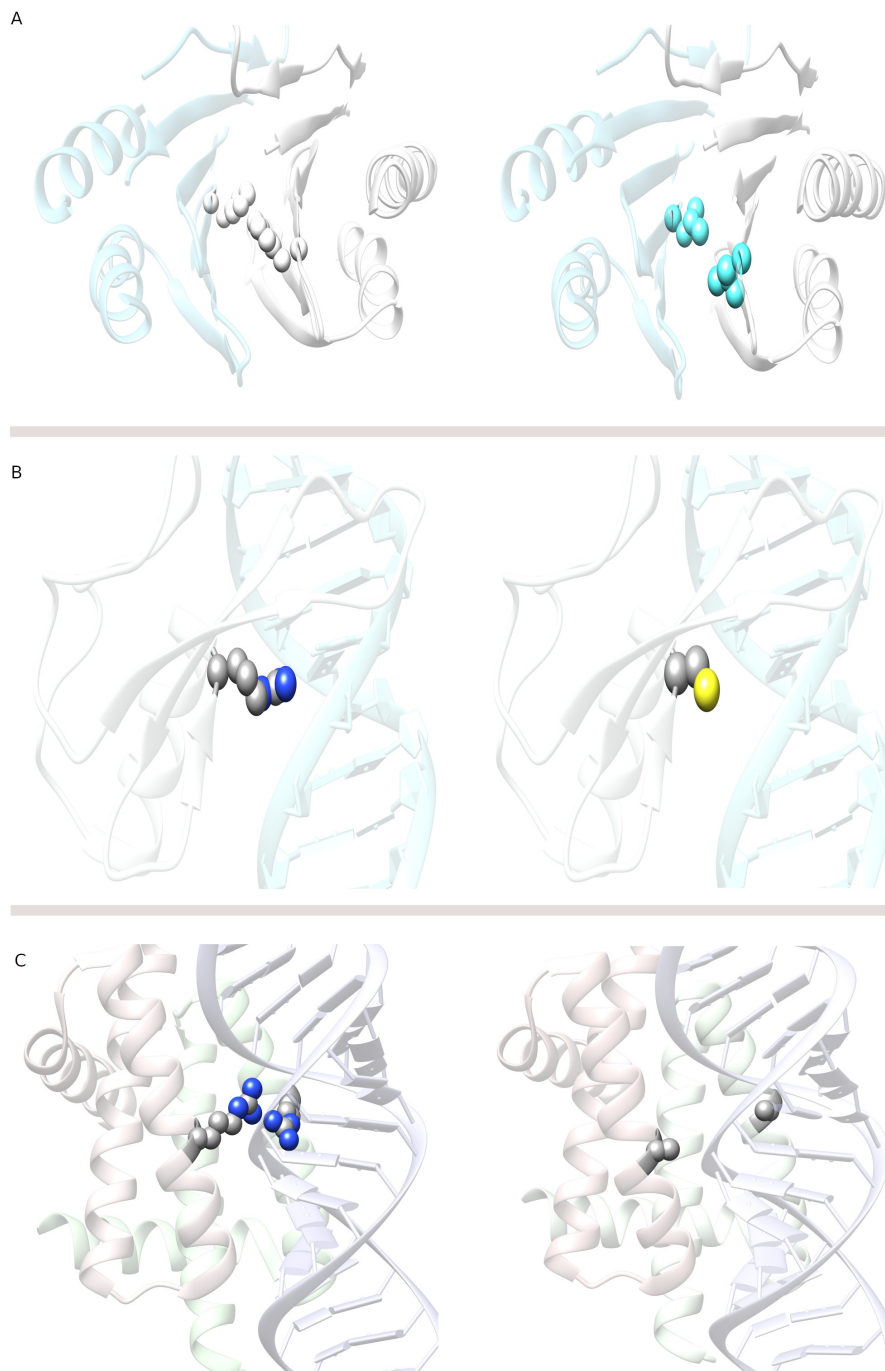


Fig. 1. Mutations at the interface alter wild type macromolecular interactions. (A) Altered protein-protein (spermine synthase homodimer) interaction. The left side shows a wild type spermine synthase, and the right side shows a mutant spermine synthase. The mutation at the homodimer interface is F58L (shown in spherical representation), which has been shown to cause Snyder-Robinson syndrome by altering homodimerization. (B) Altered protein-DNA (MeCp2-DNA) interaction. The left side shows a wild type MeCp2-DNA complex, and the right side shows a mutant MeCp2-DNA. The mutation at the MeCp2-DNA interface is R133C (shown in spherical representation), which has been shown to cause Rett syndrome by altering MeCp2-DNA binding. (C) Altered protein-dsRNA (NS1-RNA) interaction. On the left side, a wild type NS1-dsRNA complex is shown, and on the right, mutant NS1-dsRNA. The mutation at the NS1-dsRNA interface is R38A (shown in spherical representation), which has been shown to alter NS1-dsRNA binding in the virus. Such mutation reveals how the virus can be targeted by searching for small molecules that can mimic the effect of such mutations and inhibit viral immune evasion.

2. Overview of Computational Methods

2.1 Reducing/Eliminating Effects of Mutations With Small Molecules

The above-mentioned alterations of macromolecular interactions caused by missense mutations can be targeted by small molecules or peptides to restore wild type interactions. However, it should be clarified that the first step to achieve this requires the identification of a consensus outcome for all mutations affecting particular macromolecular interactions. Returning to the Rett syndrome case, it has been shown that all pathogenic mutations occurring at the protein-DNA interface reduce binding affinity when compared to the wild type protein [35]. Thus, it is necessary to develop a small molecule or peptide that enhances the binding affinity of the mutants to the affinity of the wild type protein. However, identifying such a universal small molecule requires targeting an interfacial patch that does not include any of the mutation sites, thereby making it mutant-independent. Such an approach would result in the development of an enhancer that restores the wild type affinity for all known mutants. On the other side of the spectrum are mutations that either increase the affinity of the mutants or indirectly facilitate the macromolecular interactions of overexpressed partners. To eliminate such effects, one should develop small molecules that reduce either interactions or the overexpression of the partners. Similarly, as above, this should be done in the context of the effects of mutations and the locations of the mutation sites. Below, we briefly outline the details of computational approaches for designing small molecules for use as inhibitors and enhancers.

2.2 *In Silico* Considerations for the Design of Small Molecules as Inhibitors

Recent work has summarized the current state-of-the-art in designing inhibitors without specifically considering that the altered macromolecular interactions are caused by mutations [53]. Small interfaces have been shown to be more effectively inhibited than interactions involving large interfacial areas. It should be mentioned that the inhibition of macromolecular interactions differs from the inhibition of the active site of an enzyme. In the former case, one knows the location of the active site and needs to develop a small molecule capable of binding there with the desired affinity. In the case of inhibiting interactions, especially those altered by mutation, the investigator should first determine the best binding site. Since the binding interfaces of the partners are of similar size, one can target both interfaces and identify small molecules that have a high affinity to binding to both partners, and for which the binding sites are aligned across the interface. Such a scenario will represent an inhibitor that sterically hinders the binding since the presence of the small molecule, especially at the core of the interface, will introduce van der Waals (vdW) overlaps with binding partners. However, such interfaces are

typically much larger than the size of small molecules, and macromolecules can reduce the vdW overlaps by adopting conformational changes (this will be discussed later). We illustrate this concept for a particular heterocomplex, calpains, which consist of large and small domains (L-S domains) (Fig. 2).

An alternative scenario for developing inhibitors involves considering small molecules with a high affinity to one of the partners but repulsive interactions toward the other partner. In such a case, the small molecule is expected to bind tightly at the interface of one of the partners while pushing away the other partner, thereby reducing the affinity. The main challenge for this scenario centers on how to determine that a given small molecule has a repulsive interaction with one of the partners. One can explore general biophysical considerations such as charge-charge interactions or hydrophobic *versus* hydrophilic properties; however, in many cases, this will not provide satisfactory results. Perhaps the most applicable approach will involve protein-DNA/RNA interactions. Since DNA/RNA molecules are highly negatively charged, one would expect that any small molecule with a negative net charge should be repelled by DNA/RNA.

2.3 *In Silico* Considerations for the Design of Small Molecules as Enhancers

To computationally design macromolecular interaction enhancers, the protocols outlined above can be applied even when the objective is different. To identify a small molecule serving as molecular glue, researchers should seek small molecule candidates that have a strong affinity to both partners and are relatively small. Thus, the binding of the small molecule should have little to no vdW overlaps with any of the partners. The optimal solution may involve a small molecule capable of binding at the periphery of the macromolecular interface. This is an important distinction compared with inhibitors and opens the possibility for small soluble molecules since binding at the periphery does not involve a large desolvation penalty. Here, we will illustrate this concept with a particular case mentioned above, namely Rett syndrome involving altered MeCp2-DNA interactions [35]. Since the mutation sites predominantly affecting MeCp2-DNA binding are located at the protein-DNA interface, the best strategy is to target the interface periphery with small molecules and thus potentially eliminate the effects of all mutations. *In silico* screening against Food and Drug Administration (FDA) approved drugs has resulted in several promising hits, one of which is shown in Fig. 3 (left panel). One can appreciate that the small molecule makes several hydrogen bonds with both MeCp2 and DNA, presumably stabilizing the mutant MeCp2-DNA complex and restoring its wild type binding affinity (Fig. 3, right panel).

It should be noted that the enhancement or stabilization of PPIs, including dimers or oligomers, can lead to

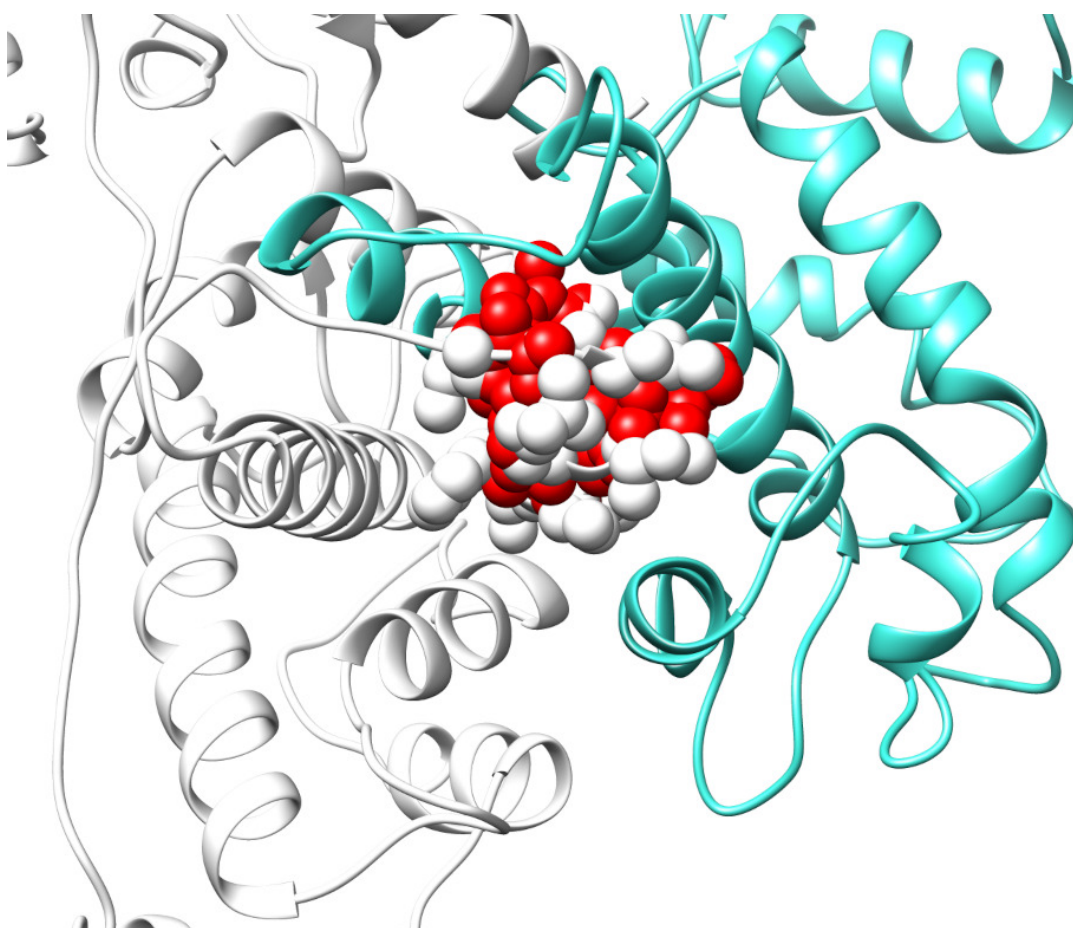


Fig. 2. van der Waals clashes at the interface of calpain large and small (L-S) domains due to the presence of the small molecule. L and S domains in calpain are represented in white and cyan, respectively. The red-colored spheres represent atoms of the small molecule, which exhibit van der Waals (vdW) overlaps with the atoms (white) in the L domain. The vdW clashes are expected to alter the associations of the S- and L-domains of the calpain.

either activation or inhibition of biological mechanisms [54,55]. Targeting areas at or near the interfaces of two or multiple proteins by small molecules to stabilize their interactions can benefit from a significantly higher specificity [56–58]. Enhancing PPIs can be more promising than inhibiting because, as opposed to the large number of PPIs targeted by inhibitors, the binding pockets formed by the association of two or several partners shows a more favorable drug profile [56,59]. Currently, most of PPI stabilizers are so-called orthosteric modulators of PPIs, as in the case of Rett syndrome illustrated above, since they bind close to the PPI interface. Other examples are ebselen-based molecules that stabilize the superoxide dismutase 1 (SOD1) dimer interface, stabilizers of 14-3-3 PPIs, among others [54,60,61]. There are, however, other possibilities to stabilize multimers via allosteric mechanisms, including the case of taxol allosterically modulating the interdimer interface of the complex of the $\alpha\beta$ tubulin dimer and the case of allosteric stabilization of the SOD1 dimer interface by small molecules [62,63]. Another interesting case is the complex inhibition of nuclear hormone receptors (NRs)

and co-regulators (Co-R) of gene expression, which partially stabilize the retinoic acid receptor(RAR)-CoR complex [64]. Inverse agonist binding also exerts allosteric effects on the NR-NR dimer interface [64]. Inducers of dimerization, also termed dimerizers, are molecules capable of binding to two proteins, utilizing two distinct binding pockets, thereby bringing the proteins in close spatial proximity. Examples of the latter include the immunosuppressant drugs cyclosporin A and FK-506, which are potent inhibitors of T cell activation [65].

2.4 *In Silico* Screening via Docking and Machine Learning

Molecular docking is a common method in computational drug design, especially when the 3D structure of the protein target is known. Its main purpose is to understand and predict how small molecules will recognize and bind to a target, both in terms of structure (finding possible binding modes) and energy (predicting binding strength). This process involves two main steps: predicting how a small molecule (ligand) fits and orients itself within the protein's binding site and evaluating the quality of this fit using a

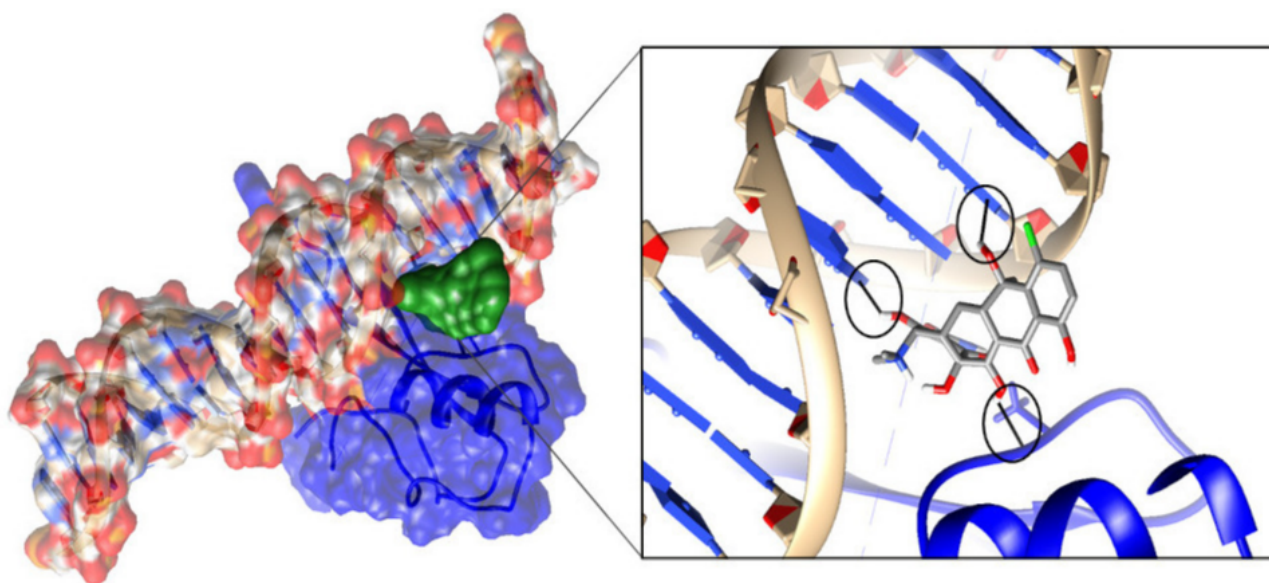


Fig. 3. MeCp2 (blue) bound to cognate DNA (ball and stick presentation), showing molecular surfaces. The small molecule bound at the periphery is shown in green (left panel). Zoom view of the interactions between the small molecule and MeCp2 and DNA (right panel). Hydrogen bonds are circled in black. The hydrogen bonds formed between the small molecule and MeCp2 and DNA are expected to restore the wild type binding affinity of the MeCp2-DNA complex.

scoring function. Ideally, the method should replicate the experimental binding mode and rank it highest among all possible fits. Additionally, it aims to score active molecules higher than inactive ones.

While docking is often used as a stand-alone method in drug design, it is increasingly combined with other computational methods, including artificial intelligence (AI), to obtain better results [66,67]. Since the 2000s, machine learning (ML) approaches based on nonlinear modeling algorithms, such as k-nearest neighbors, support vector machines (SVM), random forest (RF), and artificial neural network (ANN), have been frequently used for drug discovery [68]. Several studies have combined docking and ML to identify new hit molecules [69–72]. For example, a recent study exploited ML for virtual screening of an ensemble of protein conformations (EnOpt method) to effectively rank and prioritize the best candidates [73]. During the last decade, advanced AI approaches, such as deep learning (DL), have enabled the implementation of more innovative and efficient modeling methods using large and heterogeneous datasets of small molecules and drugs [74,75]. DL utilizes an ANN with more than two hidden layers that can characterize structures in higher-dimensional data representations. In deep neural network (DNN), each neuron receives input signals from its connected neurons, which are then multiplied by their respective weights and summed up. A recent example of this approach is a new protocol combining generative molecular AI and physics-based absolute binding free energy molecular dynamics simulations, which have been proposed as a means for discovering new ligands for different target proteins [76]. Several studies have

reported the use of docking for inhibitor discovery [77–83]. For example, two proteasome inhibitors, carfilzomib, and bortezomib, a boronic acid-based small molecule, were successfully identified using structure-based virtual screening and used to treat refractory multiple myeloma [77]. Similarly, a highly promising novel drug-like inhibitor of lymphocyte-specific protein tyrosine kinase (LCK) was recently identified using integrated DL and docking methodologies [78]. Another example involves the use of molecular docking and ML-based virtual screening to find the inhibitors for Tankyrase enzymes (TNKS) [79]. Similarly, molecular docking and molecular dynamics (MD) simulations were used to study the theoretical interaction mechanism between Janus kinase 3 (JAK3) kinase and its inhibitors [80]. Another investigation reported a series of reversible non-covalent Monoacylglycerol lipase (MAGL) inhibitors identified via virtual screening [81]. Similarly, a flexible-receptor docking protocol was used to identify the possible inhibitors for Pim-1 Kinase [82]. The progress made to identify new small molecules targeting acetylcholinesterase in Alzheimer’s disease is reviewed here [83]. Most of the above-mentioned studies were carried out using tools for *in silico* docking; the most frequently used are listed in Table 1 (Ref. [84–98]).

The interaction energy of docking software is largely evaluated by scoring functions, which present major challenges in the docking-scoring methodology [99–102]. Scoring functions are trained and benchmarked in most cases involving proteins and small molecules, where the binding site is *a priori* known, and the binding site is typically quite druggable. Furthermore, their applicability for

Table 1. Docking Tools.

Docking tools	Scoring function	Search algorithm	Type	URL ^a	Reference
AutoDock Vina	Force-Field Based	MC/BGFS (Monte Carlo/Broyden–Fletcher–Goldfarb–Shanno algorithm)	Academic	https://github.com/ccsb-scripps/AutoDock-Vina	[84,85]
AutoDock4	Force-Field Based	Lamarckian Genetic Algorithm	Academic	https://github.com/ccsb-scripps/AutoDock4	[86]
Dock	Force-Field Based	Matching Algorithm	Academic	https://dock.compbio.ucsf.edu/	[87]
GLIDE	Empirical	Hybrid (Exhaustive Sampling Method)	Commercial	https://www.schrodinger.com/platform/products/glide/	[88]
FlexX	Empirical	Fragmentation Sampling Method	Commercial	https://biosolveit.de/download/?product=flexx	[89]
GOLD	Empirical	Genetic Algorithm	Commercial	https://www.ccdc.cam.ac.uk/solutions/software/gold/	[90]
SwissDock	Force-Field Based	Evolutionary Optimization	Academic	https://www.swissdock.ch/	[91]
rDock	Empirical	Hybrid (Genetic Algorithm optimization followed by Monte Carlo and Simplex Minimization)	Academic	https://rdock.github.io/	[92]
PLANTS	Empirical	ACO (Ant Colony Optimization Method)	Academic	Currently Not Accessible	[93]
MedusaDock	Force-Field Based	GPU Based Parallel Genetic Search Algorithm	Academic	https://bitbucket.org/dokhlab/medusadock-web/src/master/	[94]
HADDOCK	Empirical	Hybrid (Chemical Shift Perturbation and NMR results)	Academic	https://www.bonvinlab.org/software/#haddock	[95]
Surflex	Empirical	Shape Matching	Commercial	https://www.biopharmics.com	[96]
ROSETTALIGAND	Force-Field Based	Monte Carlo Minimization	Academic	https://meilerlab.org/rosetta-ligand/	[97]
Gnina	Convolutional Neural Networks	MC/BGFS (Monte Carlo/Broyden–Fletcher–Goldfarb–Shanno algorithm)	Academic	https://github.com/gnina/gnina	[98]

^a Source code URL if available or download link/webserver from its developers.

in silico screening involving DNA or RNA molecules is questionable despite recent developments in tools to study RNA-protein complexes and multiple ligand docking methods [103,104]. Leclerc and Cedergren were among the first to model RNA ligands using a 3D-SAR-based approach coupled with molecular dynamics (MD) calculations to optimize small molecule poses [104,105]. The site identification by ligand competitive saturation (SILCS) computational approach was recently extended to target RNA, termed SILCS-RNA, and was applied to seven RNA systems for which small molecule-RNA interaction data were available [104]. The issues with the scoring function and other points deserving attention will be discussed in the “challenges” section below.

3. Challenges in Modulating Macromolecular Interactions

Modulating macromolecular interactions with small molecules is difficult. Challenges facing such approaches include limitations inherent to docking tools and scoring functions, available experimental data, the typical flatness of the interface, and large interfacial areas.

3.1 Challenges in Docking

3.1.1 Scoring Function

Molecular docking is a valuable tool for structure-based drug discovery. One of the most critical components of a docking application is the scoring function. Structure-based virtual screening (SBVS) relies on traditional scoring functions, which often fail to distinguish between binders and non-binders. Therefore, an effective scoring function must be both rapid and accurate when scoring and ranking the numerous docking poses generated during simulations. It should efficiently screen small molecule libraries, select the correct binding poses, and identify active molecules. A good scoring function should also ensure strong binding between a small molecule and its native receptor while exhibiting weak binding to off-target proteins. However, the scoring functions used in most docking tools are based on force fields, statistics, or empirical methods and are mainly trained on datasets obtained on high-quality protein-ligand (small molecule) complex structures. These complex structures have been determined experimentally, and the datasets do not include docking poses. Consequently, when ranking docking poses, traditional scoring functions often select false positives (poses that are far from the native pose) because they are unable to distinguish them from true native poses. Traditional scoring functions also typically rely on overly simplified linear energy combinations; as such, they cannot effectively learn from unusual cases or large-scale decoy datasets. In some cases, ML classifications can reduce false positives [101]. Therefore, to overcome the limitations of current docking-scoring approaches, the integration of AI, such as ML and DL, can be highly valuable in improving the success rate of scoring when targeting macro-

molecular interactions, in particular PPIs [67,74,102,106–111]. Due to the particular physicochemical properties of small-molecule inhibitors of PPIs discussed above, focused scoring functions for targeting PPIs have been developed using ML techniques [112]. The integration of ligand entropy and solvation considerations in empirical ML-based scoring functions has been shown to improve the performance of the scoring functions for the screening of PPIs [113]. Similarly, classical scoring functions for docking are also unable to exploit large volumes of structural and interaction data [99]. These traditional docking tools face challenges when designing covalent small molecules because the formation, breaking, and rearrangements of covalent bonds involve quantum mechanical (QM) phenomena. These processes typically cannot be accurately represented by the force fields or empirical approaches used for non-covalent protein-ligand interactions [114]. Although QM contributions are increasingly integrated into docking applications, fully implementing QM methods remains impractical for routine use due to the computational demands imposed by the size and complexity of molecular systems and the vast number of configurations of molecules involved [114]. However, faster and simpler modeling approaches are available that can often address the need for QM calculations in covalent docking. In many cases, a full QM treatment of the docking process may not be necessary, offering practical alternatives for designing covalent small molecules.

3.1.2 Conformational Sampling

Another challenge in docking is sampling the conformational space. Docking aims to identify the best binding pose between a protein and a small molecule and is more accurate if both are flexible. Typically, identifying a small molecule involves screening across millions of molecules and carrying out flexible protein and flexible small molecule docking. This requires an extensive conformational search, making the process computationally intensive. In such cases, one aims to carry out rigid protein and flexible small molecule docking, which only requires a conformational search for the small molecule, not the receptor, and making it computationally more efficient than the previous approach. However, this approach also has limitations; for example, it does not account for the flexibility of the protein, which is crucial for accurate protein-small molecule interactions. The optimal binding conformation of a small molecule depends on both its internal structure and its interactions with the binding site of the protein.

3.1.3 Identification of the Docking Site

In traditional protein targeting, the objective is to bind potential small molecules to locations in macromolecules capable of modifying their activity. A protein typically has a key binding site for a specific substrate but may also have other allosteric sites. Identifying these sites is essen-

tial and is usually performed experimentally to obtain the structure of the protein-small molecule complex. If known binding small molecules exist, they can serve as a reference for defining the docking cavity. However, defining the cavity poses a challenge: if it only includes residues close to a known binder, the software's performance may be overestimated when redocking the known binder. Conversely, a very large cavity increases the search space for the algorithm. On the other hand, when the bound small molecule is absent from the crystal structure, the identification of the active site in proteins becomes critical. One approach to mitigate this challenge is blind docking, which involves defining a single site or a docking box around the receptor protein. The docking program then explores all sites within the docking box to identify the best site for the small molecule. However, blind docking predicts the bound conformations with no *a priori* knowledge; as such, it requires extensive conformational searches and is computationally challenging, as outlined above. Additionally, defining a large docking box can lead to errors in the scoring function calculation. This occurs because the docking box is represented as a grid, with the scoring function calculated at each grid point. For larger proteins, a larger docking box is needed, which in turn may require larger spacing between the grid points. This increased spacing results in less accurate docking scores. A large docking box may introduce more potential binding sites, many of which may not be relevant. This increases the chance of identifying false positives where the predicted binding poses are not biologically meaningful. To overcome these limitations, one can define several patches around the protein structure instead of a single large docking box. This involves setting up multiple smaller docking boxes around the target protein and merging the results obtained from these patches to identify the optimal binding pose.

However, applying the same approach to modulate macromolecular interactions in which wild type characteristics are altered by mutations is even more complicated. The additional challenges derive from the lack of prior knowledge of the location of the binding site; indeed, very few experimentally determined structures exist for macromolecular complexes with a small molecule at or near the binding interface. Furthermore, identifying a putative binding site located away from all known missense mutations is challenging because it implies additional restriction of plausible good binding pockets. In addition, if one seeks to design binding enhancers, binding site(s) should be identified for both partners and should be geometrically complementary.

3.1.4 Availability of 3D Structure

Docking relies on the availability of the 3D structure of the target protein and macromolecular complex, but not all proteins have had their structures experimentally determined, including the 3D structures of their mutants. In

such cases, the first step in drug development should be modeling the 3D structures of unbound and bound partners. Progress made in 3D structure prediction, for example using AlphaFold 1, 2, and 3, makes this step straightforward [115,116]. However, it is quite unlikely that modeling tools will be able to correctly capture the effects of a single mutation on the 3D structure of both unbound and bound partners.

3.1.5 Structural Water

Structural water molecules bound to proteins are commonly observed in crystallographic structures, posing a challenge for drug discovery. In approximately 65% of protein-small molecule complexes resolved using crystallography, at least one water molecule plays a role in protein-small molecule recognition [117]. These labile water molecules can stabilize the complex through hydrogen bonding between the small molecule and the protein, or may be displaced due to small molecule binding. They contribute significantly to the entropic and enthalpic changes of the binding, where the loss of entropy from transferring water molecules from the bulk solvent to the binding site is balanced by the enthalpic contribution of additional hydrogen bonds. Available docking software may or may not be able to account for these active water molecules. Several studies have shown that the inclusion of active water molecules in docking simulations can improve the accuracy of determination of small molecule docking poses while enhancing small molecule enrichment. However, during virtual screening, they are typically neglected due to the computational costs associated with modeling these water molecules [118,119].

3.2 Challenges Regarding Available Experimental Data

Available experimental data for small molecules bound to macromolecular complexes serve two roles: (a) as a benchmarking set to assess the performance of empirical and first principle docking methods; and (b) as a training set for ML and adjustable empirical methods. In this regard, significant challenges remain to further improve the prediction of drug candidates, in particular, data availability for macromolecular interactions [68]. For many targets, the available experimental data lack either quantity or quality. One may expect a rapid increase in high-quality data because of initiatives such as the Library of Integrated Network-Based Cellular Signatures (LINCS) developed by the US National Institutes of Health (NIH). Innovative high-throughput (HTP) experimental technologies, such as the CRISPR knockout method, which tracks genetic modifications due to protein-small molecule interactions, and the cost-effective HTP reduced-representation expression profiling method, L1000, can efficiently provide large quantities of new data [120]. A unique chemical library dedicated to the inhibition of PPIs has also been created and is accessible for *in silico* and experimental screening [121]. Several

databases have been developed containing information for small-molecule modulators of PPIs. The Timbal database automatically extracts relevant data from ChEMBL following the identification of PPI targets [122]. 2P2Idb is a hand-curated structural database dedicated to PPIs with known small-molecule orthosteric modulators containing more than 270 protein-inhibitor complexes corresponding to 242 unique small-molecule inhibitors derived from the Protein Data Bank [123]. Pharmacological data available in the iPPI-DB database have been retrieved from more than 2000 peer-reviewed publications and patents [124].

3.3 Experimental Challenges in Validating Computational Findings

In the best-case scenario, computational findings should be validated experimentally. This includes validation that the small molecule binds to the target, binds at the predicted position, and adopts the predicted pose. While this can be done experimentally for a limited number of small molecule candidates, it is not feasible to perform such operations for hundreds or thousands of candidates, especially regarding binding position and pose. Carrying out thousands of X-ray or nuclear magnetic resonance (NMR) experiments for a particular target is not expected to be of high priority for any experimental lab. This is a crucial limitation for optimizing drug candidates because one can further optimize small molecules based on their experimentally determined binding mode. Other experimental validations, such as the effect of small molecules on macromolecular binding, can be done on a large scale; however, this process does not reveal the cause of the observed effect.

3.4 Challenges Arising From the Large Interfacial Area of Macromolecular Complexes

Traditional strategies for small-molecule drug discovery focus mostly on targeting enzymes, ion channels, and/or receptors (G Protein-Coupled Receptors (GPCRs) or nuclear receptors) due to the fact that their deep binding sites easily accommodate small molecules [125]. Similarly, one can target defective PPIs with small molecules, particularly using orally bioavailable molecules (the most convenient and safest for patients) [126]. Targeting modulation of PPIs using small molecules was challenging for a very long time due to their relatively flat and large surfaces [127]. Although PPIs are diverse in terms of binding affinity and duration (i.e., permanent or transient interaction), and their interface areas (1500–3000 Å²) are much larger than those of the protein-small molecule contact area (300–1000 Å²), they are flexible and contain aromatic residues [128–130]. The large surface area of macromolecules is usually buried on each side of the interface, and small molecules do not cover the entire protein-binding surface; thus, a subset of the interface that contributes to high-affinity binding may still be active (contributing to the binding and not influenced by the presence of the small molecule) [131–134].

4. Success Stories

Most recently reported success stories have involved the development of small molecules capable of modulating the macromolecular interactions of wild type proteins via HTP experimental screening [55,58,135,136]. Much less progress was made in modulating mutant macromolecular interactions [137]. One of the reasons for this is that at the current state of drug development, researchers are unwilling to invest the required time and resources to develop a drug for a particular mutation causing genetic disorders unless the mutation has a very high frequency (i.e., it is seen in many patients).

As discussed above, computational approaches are very useful for the prediction of small-molecule drug candidates targeting macromolecular interactions. Although the development of small molecules capable of disrupting PPIs is generally considered outside the remit of mainstream drug discovery, recent developments have achieved this. Below, we outline several successes in identifying small molecules via computational approaches that modulate macromolecular interactions. In the examples presented, the disease can be due to mutations present in pathways related to the target of interest and not necessarily on the targeted protein. Virtual screening of large databases occasionally identifies small molecules that inhibit and even recruit protein-protein complexes. For example, small molecules that affect the equilibrium between tubulin and microtubules are effective anticancer drugs [138]. By performing molecular docking-based virtual screening combined with an *in vitro* tubulin polymerization inhibition assay, two novel and potent tubulin inhibitors targeting the colchicine binding site in tubulin were identified [139]. These molecules displayed potent antiproliferative activity with IC₅₀ values in the nM range. In another study, molecular dynamic simulations combined with binding free energy calculations were performed on seven tubulin models to understand the mechanism of action of three natural drugs: plinabulin, docetaxel, and vinblastine [140]. The results obtained indicated that the drugs disrupted microtubule polymerization by altering its conformation and the binding free energy of the neighboring tubulin subunits. Another important anticancer target is neuropilin-1 (NRP-1), among the most important co-receptors of vascular endothelial growth factor-A (VEGF-A); the binding of VEGF-A to NRP-1 plays a critical role in angiogenesis and tumor progression. Promising small drug-like molecules disrupting the binding of VEGF-A to NRP-1 were found by performing structure-based virtual screening on the VEGF-A binding pocket of NRP-1 [141]. Four molecules with a common chlorobenzyloxy alkyloxy halogenobenzyl amine scaffold inhibited the binding of biotinylated VEGF-A to recombinant NRP-1 with a Ki of approximately 10 μM, which can serve as a base for the further development of new NRP-1 inhibitors [141]. Similarly, to target the programmed cell death-1/programmed cell death-ligand 1 (PD-1/PD-L1)

pathway in anticancer immunotherapy, small-molecule inhibitors capable of blocking the PD-1/PD-L1 axis have been identified using molecular docking methods and screening the ChEMBL. The small molecule used was a bifunctional inhibitor of PD-1/PD-L1, which inhibited the PD-1/PD-L1 interactions and promoted the dimerization, internalization, and degradation of PD-L1 [142].

In addition to cancer, successful examples of small molecules inhibiting macromolecular interactions have been reported for other diseases identified by *in silico* approaches. Structure-based virtual screening was employed to discover a potent inhibitor of the Grb14-Insulin receptor interactions that can play a role in the treatment of Diabetes mellitus type 2 [143]. Successful *in silico* screening strategies were also employed for the stabilization of native PPIs via small molecules with a molecular glue mode of action. A list of some successful glues has been reviewed [144]. Recently, small-molecule stabilizers of the complex formed by carbohydrate-response element-binding protein (ChREBP regulating transcription of glucose-responsive genes) and 14-3-3 protein were identified via virtual screening that engaged a composite interface pocket formed by both protein partners. Structure-based optimization and X-ray crystallography revealed a distinct difference in the binding modes of weak, general, and specific potent stabilizers of the ChREBP/14-3-3 complex. A structure-guided stabilizer optimization was performed on selected stabilizers, which resulted in up to 26-fold enhancement of the 14-3-3/ChREBP interaction [145]. This study demonstrated the potential of rational design for the development of selective PPI stabilizers starting from weak PPI inhibitors.

To the best of our knowledge, relatively few computational investigations have focused directly on the mutant protein. For example, a study was carried out to restore the enzymatic activity of the malfunctioning spermine synthase (SMS) mutant G56S. The purpose of this study was to stabilize the dimer by binding a small molecule at the mutant homo-dimer interface. Computational and experimental analyses were carried out, leading to the identification of five stabilizers with distinctive chemical scaffolds. These chemical scaffolds are drug-like and can serve as starting points for the development of lead molecules to further address the disease-causing effects linked to Snyder-Robinson syndrome [146].

Toxic gain-of-function (not loss-of-function) of SOD1 mutants causes motor neuron degeneration. PPI inhibitors of mutant SOD1 can be used as novel therapeutics for amyotrophic lateral sclerosis (ALS). Docking-based virtual screening of approximately 32,000 small molecules was performed on the N-terminal binding site of SOD1 interacting with tubulin combined with an assay system for interaction inhibition between mutant SOD1 and tubulin to identify mutant SOD1 PPI inhibitors [147]. Consequently, five of the six assay-executable small molecules among the top-ranked small molecules during docking simulations

were found to inhibit the mutant SOD1-tubulin interaction *in vitro*. Binding mode analysis predicted that certain of these inhibitors might bind the tubulin binding site of G85R SOD1.

For anticancer candidates, a combined ML and docking-based screening was employed to identify small molecules inhibiting the PPI site of the chaperone DNAJA1 and mutant p53. The hit small molecules (GY1-22) were validated *in vitro* and *in vivo*, which may lead to an effective treatment for preventing carcinogenesis [148]. Likewise, increased activity of lysine methyltransferase NSD2 carrying activating mutations is associated with multiple myeloma and acute lymphoblastic leukemia. The first antagonists that block the PPI between the N-terminal proline-tryptophan-tryptophan-proline domain (PWWP) domain of nuclear receptor-binding SET domain-containing 2 (NSD2) and the H3K36me2 histone and abrogate H3K36me2 binding to the PWWP1 domain in cells were identified by virtual screening and experimental validation (showing a K_d of 3.4 μ M) [149].

5. Discussion

It is important to note that altering the binding of mutant macromolecules with small molecules will have a therapeutic effect only if this alteration results in the restoration of the wild type function. Many macromolecular complexes are formed to activate the corresponding reaction by inducing conformational change. For example, calpain protease requires catalytic and regulatory subunits to bind first then, on calcium binding, conformational changes are induced, which reorient the catalytic triad residues into an active state [150,151]. Similarly, protein kinase-like CaMKII on phosphorylation induces conformational changes in the phosphorylated protein that alter its activity or ability to interact with other molecules [152]. Likewise, the epidermal growth factor receptor (EGFR) is a dimeric membrane protein that is activated by changes in the conformation of the transmembrane (TM) domain caused by changes in the interactions of the TM helices of the membrane lipid bilayer [153]. Similarly, the p53-DNA complex, which undergoes conformation changes on mutations, is associated with different forms of human cancer [154]. Thus, if one intends to enhance mutant activity, the small molecule should induce conformational change similarly to that induced by the wild type and should not necessarily influence the binding. Similarly, if one wants to reduce the functionality of the mutant, the small molecule should not allow for wild type conformation to occur. Thus, the overall therapeutic effect does not necessarily involve complete restoration of the wild type binding affinity. This is one of the main ideas developed in this review, i.e., drug efficacy is evaluated by the ability of the drug to restore functionality associated with the corresponding macromolecular interaction(s), not necessarily by restoring the wild type binding.

We have not discussed mutations that influence the pH and salt dependence of binding despite that this factor is very important for many diseases, especially cancer, since cancer tissue has a different pH from normal tissue. The available data to address this issue is very scarce, and little work has been done in this direction. We simply outline several prominent contributions in this field, mostly from Barber's lab. Typically, cancer cells have a higher pH than normal cells; this was identified by Barber and coworkers to enable metastatic progression [155,156]. They analyzed 29 cancers and identified six amino acid mutation signatures, including four signatures that were dominated by either arginine to histidine (Arg > His) or glutamate to lysine (Glu > Lys) mutations [157]. Of particular interest is the mutation R273H in p53, which alters the pH-dependence of p53-DNA binding and causes cancer progression [158].

Overall, altered macromolecular interactions are frequently found to be the major component causing pathogenicity. Therefore, it is crucial to further enhance our ability to modulate defective interactions with small molecules and potential drugs. This holds more promise than using small molecules to stabilize individual mutation-destabilized macromolecules because once the macromolecules are unstable, they are quickly degraded by proteases, and small molecules may not be able to prevent the degradation process. Furthermore, as mentioned above, in many cases, the therapeutic effect may be achievable without complete restoration of the wild type affinity, which relaxes the objectives of *in silico* drug discovery.

It should also be mentioned that the review does exclude an important component of drug discovery, namely the validation of the suggested small molecules that modulate macromolecular interactions. Depending on the computational resources available, researchers typically select the top N small molecules (N in the range of hundreds or more) and subject these small molecules and their corresponding macromolecular complexes to further analysis. Typical analyses include energy calculations with either molecular mechanics Poisson-Boltzmann (Generalized Born) surface area (MMPB(GB)SA) or quantum mechanics (QM) methods to estimate the effect of small molecules on the binding free energy. Furthermore, one typically applies molecular dynamics simulations to evaluate whether the small molecule remains bound at the predicted position, either at the binding interface or at its periphery, and thus to verify the effects of the small molecule on the binding.

6. Conclusion

This review outlined the importance of modulating macromolecular interactions with small molecules as promising therapeutic solutions for genetic disorders caused by altered macromolecular interactions. This is particularly promising for complex diseases where mutations affect the macromolecular interactions but not completely abolishing them, thus the small molecules effect could be

modest while still capable of restoring wild type binding. At the same time, it is pointed out that the modulation of macromolecular interactions with small molecules is a challenge because of the plasticity of macromolecular interfaces and the large interfacial area, however, in some cases this could be beneficial allowing small molecules to be accommodated deeply into the interface and to serve as enhancers. Having in mind that disease-causing mutations in some cases reduce binding affinity or result in under-expression of the partners, while in other cases may increase the binding affinity or cause overexpression of the partners, it is equally important to design small molecules that serve as inhibitors or enhancers. The success stories presented in the manuscript indicate that despite the challenges, this is an accomplishable goal.

Author Contributions

PP carried out the investigation, collected data, and wrote the paper; MM and EA carried out the investigation, wrote the paper, and supervised the research. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

Clemson Palmetto Cluster is acknowledged.

Funding

The work of P.P. and E.A. was supported by a grant from NIH, grant number R35GM151964.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] Anashkina AA. Protein-DNA recognition mechanisms and specificity. *Biophysical Reviews*. 2023; 15: 1007–1014. <https://doi.org/10.1007/s12551-023-01137-7>
- [2] Corley M, Burns MC, Yeo GW. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Molecular Cell*. 2020; 78: 9–29. <https://doi.org/10.1016/j.molcel.2020.03.011>
- [3] Vazquez A. *Protein Interaction Networks*. CRC Press/Taylor & Francis: Boca Raton (FL), USA. 2010.
- [4] Weis WI, Kobilka BK. The Molecular Basis of G Protein-Coupled Receptor Activation. *Annual Review of Biochemistry*. 2018; 87: 897–919. <https://doi.org/10.1146/annurev-biochem-060614-033910>
- [5] De Las Rivas J, Fontanillo C. Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS Computational Biology*. 2010; 6: e1000807. <https://doi.org/10.1371/journal.pcbi.1000807>
- [6] Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, *et al.* Architecture of the human interactome defines protein

- communities and disease networks. *Nature*. 2017; 545: 505–509. <https://doi.org/10.1038/nature22366>
- [7] Stumpf MPH, Thorne T, de Silva E, Stewart R, An HJ, Lappe M, *et al.* Estimating the size of the human interactome. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105: 6959–6964. <https://doi.org/10.1073/pnas.0708078105>
- [8] Nakajima N, Akutsu T, Nakato R. Databases for Protein-Protein Interactions. *Methods in Molecular Biology* (Clifton, N.J.). 2021; 2361: 229–248. https://doi.org/10.1007/978-1-0716-1641-3_14
- [9] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*. 2003; 13: 2498–2504. <https://doi.org/10.1101/gr.1239303>
- [10] Majeed A, Mukhtar S. Protein-Protein Interaction Network Exploration Using Cytoscape. *Methods in Molecular Biology* (Clifton, N.J.). 2023; 2690: 419–427. https://doi.org/10.1007/978-1-0716-3327-4_32
- [11] Cho W, Stahelin RV. Membrane-protein interactions in cell signaling and membrane trafficking. *Annual Review of Biophysics and Biomolecular Structure*. 2005; 34: 119–151 <https://doi.org/10.1146/annurev.biophys.33.110502.133337>
- [12] Moore DT, Berger BW, DeGrado WF. Protein-protein interactions in the membrane: sequence, structural, and biological motifs. *Structure*. 2008; 16: 991–1001. <https://doi.org/10.1016/j.str.2008.05.007>
- [13] Ramakrishnan V. Ribosome structure and the mechanism of translation. *Cell*. 2002; 108: 557–572. [https://doi.org/10.1016/s0092-8674\(02\)00619-0](https://doi.org/10.1016/s0092-8674(02)00619-0)
- [14] Onufriev AV, Schiessel H. The nucleosome: from structure to function through physics. *Current Opinion in Structural Biology*. 2019; 56: 119–130. <https://doi.org/10.1016/j.sbi.2018.11.003>
- [15] Tsunaka Y, Furukawa A, Nishimura Y. Histone tail network and modulation in a nucleosome. *Current Opinion in Structural Biology*. 2022; 75: 102436. <https://doi.org/10.1016/j.sbi.2022.102436>
- [16] Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, *et al.* The Human Transcription Factors. *Cell*. 2018; 172: 650–665. <https://doi.org/10.1016/j.cell.2018.01.029>
- [17] Catalanotto C, Cogoni C, Zardo G. MicroRNA in Control of Gene Expression: An Overview of Nuclear Functions. *International Journal of Molecular Sciences*. 2016; 17: 1712. <https://doi.org/10.3390/ijms17101712>
- [18] Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. *Cell*. 2011; 144: 986–998. <https://doi.org/10.1016/j.cell.2011.02.016>
- [19] Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature*. 2009; 461: 218–223. <https://doi.org/10.1038/nature08454>
- [20] Sahni N, Yi S, Taipale M, Fuxman Bass JI, Coulombe-Huntington J, Yang F, *et al.* Widespread macromolecular interaction perturbations in human genetic disorders. *Cell*. 2015; 161: 647–660. <https://doi.org/10.1016/j.cell.2015.04.013>
- [21] Peng Y, Norris J, Schwartz C, Alexov E. Revealing the Effects of Missense Mutations Causing Snyder-Robinson Syndrome on the Stability and Dimerization of Spermine Synthase. *International Journal of Molecular Sciences*. 2016; 17: 77. <https://doi.org/10.3390/ijms17010077>
- [22] Zhang Z, Teng S, Wang L, Schwartz CE, Alexov E. Computational analysis of missense mutations causing Snyder-Robinson syndrome. *Human Mutation*. 2010; 31: 1043–1049. <https://doi.org/10.1002/humu.21310>
- [23] Laddach A, Ng JCF, Chung SS, Fraternali F. Genetic variants and protein-protein interactions: a multidimensional network-centric view. *Current Opinion in Structural Biology*. 2018; 50: 82–90. <https://doi.org/10.1016/j.sbi.2017.12.006>
- [24] Calabrese G, Molzahn C, Mayor T. Protein interaction networks in neurodegenerative diseases: From physiological function to aggregation. *The Journal of Biological Chemistry*. 2022; 298: 102062. <https://doi.org/10.1016/j.jbc.2022.102062>
- [25] Nishi H, Tyagi M, Teng S, Shoemaker BA, Hashimoto K, Alexov E, *et al.* Cancer missense mutations alter binding properties of proteins and their interaction networks. *PLoS One*. 2013; 8: e66273. <https://doi.org/10.1371/journal.pone.0066273>
- [26] Iengar P. Identifying pathways affected by cancer mutations. *Genomics*. 2018; 110: 318–328. <https://doi.org/10.1016/j.ygeno.2017.12.004>
- [27] Koh G, Degasperis A, Zou X, Momen S, Nik-Zainal S. Mutational signatures: emerging concepts, caveats and clinical applications. *Nature Reviews. Cancer*. 2021; 21: 619–637. <https://doi.org/10.1038/s41568-021-00377-7>
- [28] Balz V, Scheckenbach K, Götte K, Bockmühl U, Petersen I, Bier H. Is the p53 inactivation frequency in squamous cell carcinomas of the head and neck underestimated? Analysis of p53 exons 2–11 and human papillomavirus 16/18 E6 transcripts in 123 unselected tumor specimens. *Cancer Research*. 2003; 63: 1188–1191.
- [29] Kropveld A, Rozemuller EH, Leppers FG, Scheidel KC, de Weger RA, Koole R, *et al.* Sequencing analysis of RNA and DNA of exons 1 through 11 shows p53 gene alterations to be present in almost 100
- [30] Partridge M, Costea DE, Huang X. The changing face of p53 in head and neck cancer. *International Journal of Oral and Maxillofacial Surgery*. 2007; 36: 1123–1138. <https://doi.org/10.1016/j.ijom.2007.06.006>
- [31] Yan L, Geng Q, Cao Z, Liu B, Li L, Lu P, *et al.* Insights into DNMT1 and programmed cell death in diseases. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*. 2023; 168: 115753. <https://doi.org/10.1016/j.biopha.2023.115753>
- [32] Pan W, Meshcheryakov VA, Li T, Wang Y, Ghosh G, Wang VYF. Structures of NF- κ B p52 homodimer-DNA complexes rationalize binding mechanisms and transcription activation. *eLife*. 2023; 12: e86258. <https://doi.org/10.7554/eLife.86258>
- [33] Ghosh G, Wang VYF. Origin of the Functional Distinctiveness of NF- κ B/p52. *Frontiers in Cell and Developmental Biology*. 2021; 9: 764164. <https://doi.org/10.3389/fcell.2021.764164>
- [34] Kyle SM, Vashi N, Justice MJ. Rett syndrome: a neurological disorder with metabolic components. *Open Biology*. 2018; 8: 170216. <https://doi.org/10.1098/rsob.170216>
- [35] Yang Y, Kucukkal TG, Li J, Alexov E, Cao W. Binding Analysis of Methyl-CpG Binding Domain of MeCP2 and Rett Syndrome Mutations. *ACS Chemical Biology*. 2016; 11: 2706–2715. <https://doi.org/10.1021/acscchembio.6b00450>
- [36] Kucukkal TG, Alexov E. Structural, Dynamical, and Energetical Consequences of Rett Syndrome Mutation R133C in MeCP2. *Computational and Mathematical Methods in Medicine*. 2015; 2015: 746157. <https://doi.org/10.1155/2015/746157>
- [37] Camero S, Benítez MJ, Jiménez JS. Anomalous protein-DNA interactions behind neurological disorders. *Advances in Protein Chemistry and Structural Biology*. 2013; 91: 37–63. <https://doi.org/10.1016/B978-0-12-411637-5.00002-0>
- [38] Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. *Cell*. 2013; 152: 1237–1251. <https://doi.org/10.1016/j.cell.2013.02.014>
- [39] Hong EJ, West AE, Greenberg ME. Transcriptional control of cognitive development. *Current Opinion in Neurobiology*. 2005; 15: 21–28. <https://doi.org/10.1016/j.conb.2005.01.002>
- [40] Liu Y, Shi SL. The roles of hnRNP A2/B1 in RNA biology and disease. *Wiley Interdisciplinary Reviews. RNA*. 2021; 12: e1612. <https://doi.org/10.1002/wrna.1612>
- [41] Gutschner T, Hämmerle M, Diederichs S. MALAT1 – a

- paradigm for long noncoding RNA function in cancer. *Journal of Molecular Medicine* (Berlin, Germany). 2013; 91: 791–801. <https://doi.org/10.1007/s00109-013-1028-y>
- [42] Brown JA, Bulkley D, Wang J, Valenstein ML, Yario TA, Steitz TA, *et al.* Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nature Structural & Molecular Biology*. 2014; 21: 633–640. <https://doi.org/10.1038/nsmb.2844>
- [43] Brown JA, Kinzig CG, DeGregorio SJ, Steitz JA. Hoogsteen-position pyrimidines promote the stability and function of the MALAT1 RNA triple helix. *RNA* (New York, N.Y.). 2016; 22: 743–749. <https://doi.org/10.1261/rna.055707.115>
- [44] Cheng A, Wong SM, Yuan YA. Structural basis for dsRNA recognition by NS1 protein of influenza A virus. *Cell Research*. 2009; 19: 187–195. <https://doi.org/10.1038/cr.2008.288>
- [45] Mahaman YAR, Huang F, Kessete Afewerky H, Maibouge TMS, Ghose B, Wang X. Involvement of calpain in the neuropathogenesis of Alzheimer's disease. *Medicinal Research Reviews*. 2019; 39: 608–630. <https://doi.org/10.1002/med.21534>
- [46] Baudry M. Calpain-1 and Calpain-2 in the Brain: Dr. Jekyll and Mr Hyde? *Current Neuropharmacology*. 2019; 17: 823–829. <https://doi.org/10.2174/1570159X17666190228112451>
- [47] Wang Y, Hersheson J, Lopez D, Hammer M, Liu Y, Lee KH, *et al.* Defects in the CAPN1 Gene Result in Alterations in Cerebellar Development and Cerebellar Ataxia in Mice and Humans. *Cell Reports*. 2016; 16: 79–91. <https://doi.org/10.1016/j.celrep.2016.05.044>
- [48] Growdon WB, Groeneweg J, Byron V, DiGloria C, Borger DR, Tambouret R, *et al.* HER2 over-expressing high grade endometrial cancer expresses high levels of p95HER2 variant. *Gynecologic Oncology*. 2015; 137: 160–166. <https://doi.org/10.1016/j.ygyno.2015.01.533>
- [49] Cizkova M, Vacher S, Meseure D, Trassard M, Susini A, Mluchova D, *et al.* PIK3R1 underexpression is an independent prognostic marker in breast cancer. *BMC Cancer*. 2013; 13: 545. <https://doi.org/10.1186/1471-2407-13-545>
- [50] Wang J, Zhang K, Wang J, Wu X, Liu X, Li B, *et al.* Underexpression of LKB1 tumor suppressor is associated with enhanced Wnt signaling and malignant characteristics of human intrahepatic cholangiocarcinoma. *Oncotarget*. 2015; 6: 18905–18920. <https://doi.org/10.18632/oncotarget.4305>
- [51] Mlakar V, Berginc G, Volavsek M, Stor Z, Rems M, Glavac D. Presence of activating KRAS mutations correlates significantly with expression of tumour suppressor genes DCN and TPM1 in colorectal cancer. *BMC Cancer*. 2009; 9: 282. <https://doi.org/10.1186/1471-2407-9-282>
- [52] Wang J, Okkeri J, Pavic K, Wang Z, Kauko O, Halonen T, *et al.* Oncoprotein CIP2A is stabilized via interaction with tumor suppressor PP2A/B56. *EMBO Reports*. 2017; 18: 437–450. <https://doi.org/10.15252/embr.201642788>
- [53] Ran X, Gestwicki JE. Inhibitors of protein-protein interactions (PPIs): an analysis of scaffold choices and buried surface area. *Current Opinion in Chemical Biology*. 2018; 44: 75–86. <https://doi.org/10.1016/j.cbpa.2018.06.004>
- [54] Zarzycka B, Kuenemann MA, Miteva MA, Nicoles GAF, Vriend G, Sperandio O. Stabilization of protein-protein interaction complexes through small molecules. *Drug Discovery Today*. 2016; 21: 48–57. <https://doi.org/10.1016/j.drudis.2015.09.011>
- [55] Konstantinidou M, Arkin MR. Molecular glues for protein-protein interactions: Progressing toward a new dream. *Cell Chemical Biology*. 2024; 31: 1064–1088. <https://doi.org/10.1016/j.chembiol.2024.04.002>
- [56] Thiel P, Kaiser M, Ottmann C. Small-molecule stabilization of protein-protein interactions: an underestimated concept in drug discovery? *Angewandte Chemie (International Ed. in English)*. 2012; 51: 2012–2018. <https://doi.org/10.1002/anie.201107616>
- [57] Arkin MR, Randal M, DeLano WL, Hyde J, Luong TN, Oslob JD, *et al.* Binding of small molecules to an adaptive protein-protein interface. *Proceedings of the National Academy of Sciences*. 2003; 100: 1603–1608. <https://doi.org/10.1073/pnas.252756299>
- [58] Kenanova DN, Visser EJ, Virta JM, Sijbesma E, Centorrino F, Vickery HR, *et al.* A Systematic Approach to the Discovery of Protein-Protein Interaction Stabilizers. *ACS Central Science*. 2023; 9: 937–946. <https://doi.org/10.1021/acscentsci.2c01449>
- [59] Block P, Weskamp N, Wolf A, Klebe G. Strategies to search and design stabilizers of protein-protein interactions: a feasibility study. *Proteins*. 2007; 68: 170–186. <https://doi.org/10.1002/prot.21296>
- [60] Ampornpanai K, Rogers M, Watanabe S, Yamanaka K, O'Neill PM, Hasnain SS. Novel Selenium-based compounds with therapeutic potential for SOD1-linked amyotrophic lateral sclerosis. *EBioMedicine*. 2020; 59: 102980. <https://doi.org/10.1016/j.ebiom.2020.102980>
- [61] Somsen BA, Cossar PJ, Arkin MR, Brunsveld L, Ottmann C. 14-3-3 Protein-Protein Interactions: From Mechanistic Understanding to Their Small-Molecule Stabilization. *Chembiochem: a European Journal of Chemical Biology*. 2024; 25: e202400214. <https://doi.org/10.1002/cbic.202400214>
- [62] Alushin GM, Lander GC, Kellogg EH, Zhang R, Baker D, Nogales E. High-resolution microtubule structures reveal the structural transitions in $\alpha\beta$ -tubulin upon GTP hydrolysis. *Cell*. 2014; 157: 1117–1129. <https://doi.org/10.1016/j.cell.2014.03.053>
- [63] Wright GSA, Antonyuk SV, Kershaw NM, Strange RW, Samar Hasnain S. Ligand binding and aggregation of pathogenic SOD1. *Nature Communications*. 2013; 4: 1758. <https://doi.org/10.1038/ncomms2750>
- [64] Gronemeyer H, Gustafsson JA, Laudet V. Principles for modulation of the nuclear receptor superfamily. *Nature Reviews. Drug Discovery*. 2004; 3: 950–964. <https://doi.org/10.1038/nrd1551>
- [65] Siekierka JJ, Hung SH, Poe M, Lin CS, Sigal NH. A cytosolic binding protein for the immunosuppressant FK506 has peptidyl-prolyl isomerase activity but is distinct from cyclophilin. *Nature*. 1989; 341: 755–757. <https://doi.org/10.1038/341755a0>
- [66] Bande AY, Baday S. Accelerating Molecular Docking using Machine Learning Methods. *Molecular Informatics*. 2024; 43: e202300167. <https://doi.org/10.1002/minf.202300167>
- [67] Schneider P, Walters WP, Plowright AT, Sieroka N, Listgarten J, Goodnow RA, Jr, *et al.* Rethinking drug design in the artificial intelligence era. *Nature Reviews. Drug Discovery*. 2020; 19: 353–364. <https://doi.org/10.1038/s41573-019-0050-3>
- [68] Bagherian M, Sabeti E, Wang K, Sartor MA, Nikolovska-Coleska Z, Najarian K. Machine learning approaches and databases for prediction of drug-target interaction: a survey paper. *Briefings in Bioinformatics*. 2021; 22: 247–269. <https://doi.org/10.1093/bib/bbz157>
- [69] Dudas B, Bagdad Y, Picard M, Perahia D, Miteva MA. Machine learning and structure-based modeling for the prediction of UDP-glucuronosyltransferase inhibition. *iScience*. 2022; 25: 105290. <https://doi.org/10.1016/j.isci.2022.105290>
- [70] Chongjun Y, Nasr AMS, Latif MAM, Rahman MBA, Marlisah E, Tejo BA. Predicting repurposed drugs targeting the NS3 protease of dengue virus using machine learning-based QSAR, molecular docking, and molecular dynamics simulations. *SAR and QSAR in Environmental Research*. 2024; 35: 707–728. <https://doi.org/10.1080/1062936X.2024.2392677>
- [71] Zulfat M, Hakami MA, Hazazi A, Mahmood A, Khalid A, Alqurashi RS, *et al.* Identification of novel NLRP3 inhibitors as therapeutic options for epilepsy by machine learning-based virtual screening, molecular docking and biomolecular simulation studies. *Heliyon*. 2024; 10: e34410.

- <https://doi.org/10.1016/j.heliyon.2024.e34410>
- [72] Bazuhair MA, Alghamdi AA, Baothman O, Afzal M, Alzarea SI, Imam F, *et al.* Chemical analogue based drug design for cancer treatment targeting PI3K: integrating machine learning and molecular modeling. *Molecular Diversity*. 2024; 28: 2345–2364. <https://doi.org/10.1007/s11030-024-10966-x>
- [73] Bhatt R, Wang A, Durrant JD. Teaching old docks new tricks with machine learning enhanced ensemble docking. *Scientific Reports*. 2024; 14: 20722. <https://doi.org/10.1038/s41598-024-71699-3>
- [74] Tropsha A, Isayev O, Varnek A, Schneider G, Cherkasov A. Integrating QSAR modelling and deep learning in drug discovery: the emergence of deep QSAR. *Nature Reviews. Drug Discovery*. 2024; 23: 141–155. <https://doi.org/10.1038/s41573-023-00832-0>
- [75] Du BX, Qin Y, Jiang YF, Xu Y, Yiu SM, Yu H, *et al.* Compound-protein interaction prediction by deep learning: Databases, descriptors and models. *Drug Discovery Today*. 2022; 27: 1350–1366. <https://doi.org/10.1016/j.drudis.2022.02.023>
- [76] Loeffler HH, Wan S, Klähn M, Bhati AP, Coveney PV. Optimal Molecular Design: Generative Active Learning Combining REINVENT with Precise Binding Free Energy Ranking Simulations. *Journal of Chemical Theory and Computation*. 2024; 20: 10.1021/acs.jctc.4c00576. <https://doi.org/10.1021/acs.jctc.4c00576>
- [77] Miller Z, Kim KS, Lee DM, Kasam V, Baek SE, Lee KH, *et al.* Proteasome inhibitors with pyrazole scaffolds from structure-based virtual screening. *Journal of Medicinal Chemistry*. 2015; 58: 2036–2041. <https://doi.org/10.1021/jm501344n>
- [78] Guo H, Shen ZY, Yuan YY, Chen RF, Yang JY, Liu XC, *et al.* Discovery of a Novel and Potent LCK Inhibitor for Leukemia Treatment via Deep Learning and Molecular Docking. *Journal of Chemical Information and Modeling*. 2024; 64: 4835–4849. <https://doi.org/10.1021/acs.jcim.4c00151>
- [79] Berishvili VP, Kuimov AN, Voronkov AE, Radchenko EV, Kumar P, Choonara YE, *et al.* Discovery of Novel Tankyrase Inhibitors through Molecular Docking-Based Virtual Screening and Molecular Dynamics Simulation Studies. *Molecules (Basel, Switzerland)*. 2020; 25: 3171. <https://doi.org/10.3390/molecules25143171>
- [80] Zhong Q, Qin J, Zhao K, Guo L, Li D. Molecular Docking and Molecular Dynamics Simulation of New Potential JAK3 Inhibitors. *Current Computer-aided Drug Design*. 2024; 20: 764–772. <https://doi.org/10.2174/1573409919666230525154120>
- [81] Xiong F, Ding X, Zhang H, Luo X, Chen K, Jiang H, *et al.* Discovery of novel reversible monoacylglycerol lipase inhibitors via docking-based virtual screening. *Bioorganic & Medicinal Chemistry Letters*. 2021; 41: 127986. <https://doi.org/10.1016/j.bmcl.2021.127986>
- [82] Li G, Li W, Xie Y, Wan X, Zheng G, Huang N, *et al.* Discovery of Novel Pim-1 Kinase Inhibitors with a Flexible-Receptor Docking Protocol. *Journal of Chemical Information and Modeling*. 2019; 59: 4116–4119. <https://doi.org/10.1021/acs.jcim.9b00494>
- [83] Peitzika SC, Pontiki E. A Review on Recent Approaches on Molecular Docking Studies of Novel Compounds Targeting Acetylcholinesterase in Alzheimer Disease. *Molecules (Basel, Switzerland)*. 2023; 28: 1084. <https://doi.org/10.3390/molecules28031084>
- [84] Eberhardt J, Santos-Martins D, Tillack AF, Forli S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling*. 2021; 61: 3891–3898. <https://doi.org/10.1021/acs.jcim.1c00203>
- [85] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*. 2010; 31: 455–461. <https://doi.org/10.1002/jcc.21334>
- [86] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*. 2009; 30: 2785–2791. <https://doi.org/10.1002/jcc.21256>
- [87] Allen WJ, Balias TE, Mukherjee S, Brozell SR, Moustakas DT, Lang PT, *et al.* DOCK 6: Impact of new features and current docking performance. *Journal of Computational Chemistry*. 2015; 36: 1132–1156. <https://doi.org/10.1002/jcc.23905>
- [88] Stanzione F, Giangreco I, Cole JC. Use of molecular docking computational tools in drug discovery. *Progress in Medicinal Chemistry*. 2021; 60: 273–343. <https://doi.org/10.1016/bs.pmch.2021.01.004>
- [89] Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology*. 1996; 261: 470–489. <https://doi.org/10.1006/jmbi.1996.0477>
- [90] Jones G, Willett P, Glen RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *Journal of Molecular Biology*. 1995; 245: 43–53. [https://doi.org/10.1016/s0022-2836\(95\)80037-9](https://doi.org/10.1016/s0022-2836(95)80037-9)
- [91] Grosdidier A, Zoete V, Michielin O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Research*. 2011; 39: W270–W277. <https://doi.org/10.1093/nar/gkr366>
- [92] Ruiz-Carmona S, Alvarez-Garcia D, Foloppe N, Garmendia-Doval AB, Juhos S, Schmidtke P, *et al.* rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Computational Biology*. 2014; 10: e1003571. <https://doi.org/10.1371/journal.pcbi.1003571>
- [93] Korb O, Stützle T, Exner TE. PLANTS: Application of ant colony optimization to structure-based drug design. *Ant Colony Optimization and Swarm Intelligence*. 5th International Workshop. ANTS 2006. Springer: Berlin, Heidelberg. 2006.
- [94] Fan M, Wang J, Jiang H, Feng Y, Mahdavi M, Madduri K, *et al.* GPU-Accelerated Flexible Molecular Docking. *The Journal of Physical Chemistry. B*. 2021; 125: 1049–1060. <https://doi.org/10.1021/acs.jpcc.0c09051>
- [95] Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*. 2003; 125: 1731–1737. <https://doi.org/10.1021/ja026939x>
- [96] Spitzer R, Jain AN. Surflex-Dock: Docking benchmarks and real-world application. *Journal of Computer-aided Molecular Design*. 2012; 26: 687–699. <https://doi.org/10.1007/s10822-011-9533-y>
- [97] Meiler J, Baker D. ROSETTALIGAND: protein-small molecule docking with full side-chain flexibility. *Proteins*. 2006; 65: 538–548. <https://doi.org/10.1002/prot.21086>
- [98] McNutt AT, Francoeur P, Aggarwal R, Masuda T, Meli R, Ragoza M, *et al.* GNINA 1.0: molecular docking with deep learning. *Journal of Cheminformatics*. 2021; 13: 43. <https://doi.org/10.1186/s13321-021-00522-2>
- [99] Li H, Peng J, Sidorov P, Leung Y, Leung KS, Wong MH, *et al.* Classical scoring functions for docking are unable to exploit large volumes of structural and interaction data. *Bioinformatics (Oxford, England)*. 2019; 35: 3989–3995. <https://doi.org/10.1093/bioinformatics/btz183>
- [100] Tran-Nguyen VK, Bret G, Rognan D. True Accuracy of Fast Scoring Functions to Predict High-Throughput Screening Data from Docking Poses: The Simpler the Better. *Journal of Chemical Information and Modeling*. 2021; 61: 2788–2797. <https://doi.org/10.1021/acs.jcim.1c00292>
- [101] Adeshina YO, Deeds EJ, Karanicolas J. Machine learning clas-

- sification can reduce false positives in structure-based virtual screening. *Proceedings of the National Academy of Sciences of the United States of America*. 2020; 117: 18477–18488. <https://doi.org/10.1073/pnas.2000585117>
- [102] Singh N, Villoutreix BO. A Hybrid Docking and Machine Learning Approach to Enhance the Performance of Virtual Screening Carried out on Protein-Protein Interfaces. *International Journal of Molecular Sciences*. 2022; 23: 14364. <https://doi.org/10.3390/ijms232214364>
- [103] Bheemireddy S, Sandhya S, Srinivasan N, Sowdhamini R. Computational tools to study RNA-protein complexes. *Frontiers in Molecular Biosciences*. 2022; 9: 954926. <https://doi.org/10.3389/fmolb.2022.954926>
- [104] Kognole AA, Hazel A, MacKerell AD, Jr. SILCS-RNA: Toward a Structure-Based Drug Design Approach for Targeting RNAs with Small Molecules. *Journal of Chemical Theory and Computation*. 2022; 18: 5672–5691. <https://doi.org/10.1021/acs.jctc.2c00381>
- [105] Leclerc F, Cedergren R. Modeling RNA-ligand interactions: the Rev-binding element RNA-aminoglycoside complex. *Journal of Medicinal Chemistry*. 1998; 41: 175–182. <https://doi.org/10.1021/jm970372r>
- [106] Reker D, Schneider P, Schneider G. Multi-objective active machine learning rapidly improves structure-activity models and reveals new protein-protein interaction inhibitors. *Chemical Science*. 2016; 7: 3919–3927. <https://doi.org/10.1039/c5sc04272k>
- [107] Gao M, Zhao L, Zhang Z, Wang J, Wang C. Using a stacked ensemble learning framework to predict modulators of protein-protein interactions. *Computers in Biology and Medicine*. 2023; 161: 107032. <https://doi.org/10.1016/j.combiomed.2023.107032>
- [108] Gehlhaar DK, Mermelstein DJ. FitScore: a fast machine learning-based score for 3D virtual screening enrichment. *Journal of Computer-aided Molecular Design*. 2024; 38: 29. <https://doi.org/10.1007/s10822-024-00570-4>
- [109] Crivelli-Decker JE, Beckwith Z, Tom G, Le L, Khuttan S, Salomon-Ferrer R, *et al.* Machine Learning Guided AQFEP: A Fast and Efficient Absolute Free Energy Perturbation Solution for Virtual Screening. *Journal of Chemical Theory and Computation*. 2024; 20: 10.1021/acs.jctc.4c00399. <https://doi.org/10.1021/acs.jctc.4c00399>
- [110] Liu H, Hu B, Chen P, Wang X, Wang H, Wang S, *et al.* Docking Score ML: Target-Specific Machine Learning Models Improving Docking-Based Virtual Screening in 155 Targets. *Journal of Chemical Information and Modeling*. 2024; 64: 5413–5426. <https://doi.org/10.1021/acs.jcim.4c00072>
- [111] de Azevedo WF, Jr, Quiroga R, Villarreal MA, da Silveira NJF, Bitencourt-Ferreira G, da Silva AD, *et al.* SANdReS 2.0: Development of machine-learning models to explore the scoring function space. *Journal of Computational Chemistry*. 2024; 45: 2333–2346. <https://doi.org/10.1002/jcc.27449>
- [112] Kastritis PL, Rodrigues JPGLM, Bonvin AMJJ. HADDOCK(2P2I): a biophysical model for predicting the binding affinity of protein-protein interaction inhibitors. *Journal of Chemical Information and Modeling*. 2014; 54: 826–836. <https://doi.org/10.1021/ci4005332>
- [113] Guedes IA, Barreto AMS, Marinho D, Krempser E, Kuenemann MA, Sperandio O, *et al.* New machine learning and physics-based scoring functions for drug discovery. *Scientific Reports*. 2021; 11: 3198. <https://doi.org/10.1038/s41598-021-82410-1>
- [114] Sottriffer C. Docking of Covalent Ligands: Challenges and Approaches. *Molecular Informatics*. 2018; 37: e1800062. <https://doi.org/10.1002/minf.201800062>
- [115] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021; 596: 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- [116] Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024; 630: 493–500. <https://doi.org/10.1038/s41586-024-07487-w>
- [117] Klebe G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today*. 2006; 11: 580–594. <https://doi.org/10.1016/j.drudis.2006.05.012>
- [118] Zhong H, Wang Z, Wang X, Liu H, Li D, Liu H, *et al.* Importance of a crystalline water network in docking-based virtual screening: a case study of BRD4. *Physical Chemistry Chemical Physics: PCCP*. 2019; 21: 25276–25289. <https://doi.org/10.1039/c9cp04290c>
- [119] Huang N, Shoichet BK. Exploiting ordered waters in molecular docking. *Journal of Medicinal Chemistry*. 2008; 51: 4862–4865. <https://doi.org/10.1021/jm8006239>
- [120] Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, *et al.* A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell*. 2017; 171: 1437–1452.e17. <https://doi.org/10.1016/j.cell.2017.10.049>
- [121] Bosc N, Muller C, Hoffer L, Lagorce D, Bourg S, Derviaux C, *et al.* Fr-PPIChem: An Academic Compound Library Dedicated to Protein-Protein Interactions. *ACS Chemical Biology*. 2020; 15: 1566–1574. <https://doi.org/10.1021/acschembio.0c00179>
- [122] Higuero AP, Jubb H, Blundell TL. TIMBAL v2: update of a database holding small molecules modulating protein-protein interactions. *Database: the Journal of Biological Databases and Curation*. 2013; 2013: bat039. <https://doi.org/10.1093/database/bat039>
- [123] Basse MJ, Betzi S, Morelli X, Roche P. 2P2Idb v2: update of a structural database dedicated to orthosteric modulation of protein-protein interactions. *Database: the Journal of Biological Databases and Curation*. 2016; 2016: baw007. <https://doi.org/10.1093/database/baw007>
- [124] Torchet R, Druart K, Ruano LC, Moine-Franel A, Borges H, Doppelt-Azeroual O, *et al.* The iPPI-DB initiative: a community-centered database of protein-protein interaction modulators. *Bioinformatics (Oxford, England)*. 2021; 37: 89–96. <https://doi.org/10.1093/bioinformatics/btaa1091>
- [125] Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, *et al.* A comprehensive map of molecular drug targets. *Nature Reviews. Drug Discovery*. 2017; 16: 19–34. <https://doi.org/10.1038/nrd.2016.230>
- [126] Ivanov AA, Khuri FR, Fu H. Targeting protein-protein interactions as an anticancer strategy. *Trends in Pharmacological Sciences*. 2013; 34: 393–400. <https://doi.org/10.1016/j.tips.2013.04.007>
- [127] Rehman AU, Khurshid B, Ali Y, Rasheed S, Wadood A, Ng HL, *et al.* Computational approaches for the design of modulators targeting protein-protein interactions. *Expert Opinion on Drug Discovery*. 2023; 18: 315–333. <https://doi.org/10.1080/17460441.2023.2171396>
- [128] Nooren IMA, Thornton JM. Diversity of protein-protein interactions. *The EMBO Journal*. 2003; 22: 3486–3492. <https://doi.org/10.1093/emboj/cdg359>
- [129] Arkin MR, Randal M, DeLano WL, Hyde J, Luong TN, Oslob JD, *et al.* Binding of small molecules to an adaptive protein-protein interface. *Proceedings of the National Academy of Sciences of the United States of America*. 2003; 100: 1603–1608. <https://doi.org/10.1073/pnas.252756299>
- [130] Smith MC, Gestwicki JE. Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert Reviews in Molecular Medicine*. 2012; 14: e16. <https://doi.org/10.1017/erm.2012.10>
- [131] Bogan AA, Thorn KS. Anatomy of hot spots in protein

- interfaces. *Journal of Molecular Biology*. 1998; 280: 1–9. <https://doi.org/10.1006/jmbi.1998.1843>
- [132] Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor interface. *Science (New York, N.Y.)*. 1995; 267: 383–386. <https://doi.org/10.1126/science.7529940>
- [133] DeLano WL. Unraveling hot spots in binding interfaces: progress and challenges. *Current Opinion in Structural Biology*. 2002; 12: 14–20. [https://doi.org/10.1016/s0959-440x\(02\)00283-x](https://doi.org/10.1016/s0959-440x(02)00283-x)
- [134] Ma B, Elkayam T, Wolfson H, Nussinov R. Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces. *Proceedings of the National Academy of Sciences of the United States of America*. 2003; 100: 5772–5777. <https://doi.org/10.1073/pnas.1030237100>
- [135] Zhong M, Lee GM, Sijbesma E, Ottmann C, Arkin MR. Modulating protein-protein interaction networks in protein homeostasis. *Current Opinion in Chemical Biology*. 2019; 50: 55–65. <https://doi.org/10.1016/j.cbpa.2019.02.012>
- [136] Arkin MR, Tang Y, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chemistry & Biology*. 2014; 21: 1102–1114. <https://doi.org/10.1016/j.chembiol.2014.09.001>
- [137] Markossian S, Ang KK, Wilson CG, Arkin MR. Small-Molecule Screening for Genetic Diseases. *Annual Review of Genomics and Human Genetics*. 2018; 19: 263–288. <https://doi.org/10.1146/annurev-genom-083117-021452>
- [138] Hadfield JA, Ducki S, Hirst N, McGown AT. Tubulin and microtubules as targets for anticancer drugs. *Progress in Cell Cycle Research*. 2003; 5: 309–325.
- [139] Liu G, Jiao Y, Huang C, Chang P. Identification of novel and potent small-molecule inhibitors of tubulin with antitumor activities by virtual screening and biological evaluations. *Journal of Computer-aided Molecular Design*. 2019; 33: 659–664. <https://doi.org/10.1007/s10822-019-00206-y>
- [140] Chu Y, Tian Z, Yang M, Li W. Conformation and energy investigation of microtubule longitudinal dynamic instability induced by natural products. *Chemical Biology & Drug Design*. 2023; 102: 444–456. <https://doi.org/10.1111/cbdd.14189>
- [141] Starzec A, Miteva MA, Ladam P, Villoutreix BO, Perret GY. Discovery of novel inhibitors of vascular endothelial growth factor-A-Neuropilin-1 interaction by structure-based virtual screening. *Bioorganic & Medicinal Chemistry*. 2014; 22: 4042–4048. <https://doi.org/10.1016/j.bmc.2014.05.068>
- [142] Wang T, Cai S, Cheng Y, Zhang W, Wang M, Sun H, *et al.* Discovery of Small-Molecule Inhibitors of the PD-1/PD-L1 Axis That Promote PD-L1 Internalization and Degradation. *Journal of Medicinal Chemistry*. 2022; 65: 3879–3893. <https://doi.org/10.1021/acs.jmedchem.1c01682>
- [143] Gondoin A, Hampe C, Eudes R, Fayolle C, Pierre-Eugène C, Miteva M, *et al.* Identification of insulin-sensitizing molecules acting by disrupting the interaction between the Insulin Receptor and Grb14. *Scientific Reports*. 2017; 7: 16901. <https://doi.org/10.1038/s41598-017-17122-6>
- [144] Dewey JA, Delalande C, Azizi SA, Lu V, Antonopoulos D, Babnigg G. Molecular Glue Discovery: Current and Future Approaches. *Journal of Medicinal Chemistry*. 2023; 66: 9278–9296. <https://doi.org/10.1021/acs.jmedchem.3c00449>
- [145] Sijbesma E, Visser E, Plitzko K, Thiel P, Milroy LG, Kaiser M, *et al.* Structure-based evolution of a promiscuous inhibitor to a selective stabilizer of protein-protein interactions. *Nature Communications*. 2020; 11: 3954. <https://doi.org/10.1038/s41467-020-17741-0>
- [146] Zhang Z, Martiny V, Lagorce D, Ikeguchi Y, Alexov E, Miteva MA. Rational design of small-molecule stabilizers of spermine synthase dimer by virtual screening and free energy-based approach. *PloS One*. 2014; 9: e110884. <https://doi.org/10.1371/journal.pone.0110884>
- [147] Hirayama K, Fujiwara Y, Terada T, Shimizu K, Wada K, Kabuta T. Virtual screening identification of novel chemical inhibitors for aberrant interactions between pathogenic mutant SOD1 and tubulin. *Neurochemistry International*. 2019; 126: 19–26. <https://doi.org/10.1016/j.neuint.2019.02.020>
- [148] Jacobsen D, Bushara O, Mishra RK, Sun L, Liao J, Yang GY. Druggable sites/pockets of the p53-DNAJA1 protein-protein interaction: In silico modeling and in vitro/in vivo validation. *Methods in Enzymology*. 2022; 675: 83–107. <https://doi.org/10.1016/bs.mie.2022.07.005>
- [149] Ferreira de Freitas R, Liu Y, Szweczyk MM, Mehta N, Li F, McLeod D, *et al.* Discovery of Small-Molecule Antagonists of the PWWP Domain of NSD2. *Journal of Medicinal Chemistry*. 2021; 64: 1584–1592. <https://doi.org/10.1021/acs.jmedchem.0c01768>
- [150] Strobl S, Fernandez-Catalan C, Braun M, Huber R, Masumoto H, Nakagawa K, *et al.* The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. *Proceedings of the National Academy of Sciences of the United States of America*. 2000; 97: 588–592. <https://doi.org/10.1073/pnas.97.2.588>
- [151] Ono Y, Sorimachi H, Suzuki K. Structure and physiology of calpain, an enigmatic protease. *Biochemical and Biophysical Research Communications*. 1998; 245: 289–294. <https://doi.org/10.1006/bbrc.1998.8085>
- [152] Rostas JAP, Skelding KA. Calcium/Calmodulin-Stimulated Protein Kinase II (CaMKII): Different Functional Outcomes from Activation, Depending on the Cellular Microenvironment. *Cells*. 2023; 12: 401. <https://doi.org/10.3390/cells12030401>
- [153] Lelimosin M, Limongelli V, Sansom MSP. Conformational Changes in the Epidermal Growth Factor Receptor: Role of the Transmembrane Domain Investigated by Coarse-Grained MetaDynamics Free Energy Calculations. *Journal of the American Chemical Society*. 2016; 138: 10611–10622. <https://doi.org/10.1021/jacs.6b05602>
- [154] Chen JM, Rosal R, Smith S, Pincus MR, Brandt-Rauf PW. Common conformational effects of p53 mutations. *Journal of Protein Chemistry*. 2001; 20: 101–105. <https://doi.org/10.1023/a:1011065022283>
- [155] Webb BA, Chimenti M, Jacobson MP, Barber DL. Dysregulated pH: a perfect storm for cancer progression. *Nature Reviews. Cancer*. 2011; 11: 671–677. <https://doi.org/10.1038/nrc3110>
- [156] White KA, Kisor K, Barber DL. Intracellular pH dynamics and charge-changing somatic mutations in cancer. *Cancer Metastasis Reviews*. 2019; 38: 17–24. <https://doi.org/10.1007/s10555-019-09791-8>
- [157] Szpiech ZA, Strauli NB, White KA, Ruiz DG, Jacobson MP, Barber DL, *et al.* Prominent features of the amino acid mutation landscape in cancer. *PloS One*. 2017; 12: e0183273. <https://doi.org/10.1371/journal.pone.0183273>
- [158] White KA, Ruiz DG, Szpiech ZA, Strauli NB, Hernandez RD, Jacobson MP, *et al.* Cancer-associated arginine-to-histidine mutations confer a gain in pH sensing to mutant proteins. *Science Signaling*. 2017; 10: eaam9931. <https://doi.org/10.1126/scisignal.aam9931>