A Pilot Study of All-Computational Drug Design Protocol–From Structure Prediction to Interaction Analysis

Yifei Wu, Lei Lou and Zhong-Ru Xie*

Computational Drug Discovery Laboratory, School of Electrical and Computer Engineering, College of Engineering, University of Georgia, Athens, GA, United States

Speeding up the drug discovery process is of great significance. To achieve that, high-efficiency methods should be exploited. The conventional wet-bench methods hardly meet the high-speed demand due to time-consuming experiments. Conversely, *in silico* approaches are much more efficient for drug discovery and design. However, *in silico* approaches usually serve as a supportive role in research processes. To fully exert the strength of computational methods, we propose a protocol which integrates various *in silico* approaches, from *de novo* protein structure prediction to ligand-protein interaction simulation. As a proof of concept, human SK2/calmodulin complex was used as a target for validation. First, we obtained a predicted structure of SK2/calmodulin and predicted binding sites which were consistent with the literature data. Then we investigated the ligand-protein interaction via virtual mutagenesis, flexible docking, and binding affinity calculation. As a result, the binding energies of mutants have similar trends compared with the EC_{50} values ($R = 0.6$ for NS309 in V481 mutants). The results indicate that our protocol can be applied to the drug design of structure unknown proteins. Our study also demonstrates that the integration of *in silico* approaches is feasible and it facilitates the acceleration of new drug discovery.

Keywords: ligand-protein docking, molecular dynamics simulation, computational drug discovery, SK2, structural prediction, binding site prediction, virtual mutation, pharmaceutical industry

1. INTRODUCTION

To discover a new drug is an urgent but time-consuming process (Zhou et al., 2016; Gómez-Bombarelli et al., 2018). In the process of new drug development, *in silico* approaches have been successfully exploited to perform multiple simulations, such as selecting drug candidates from a database via high throughput virtual screening (Aparna et al., 2014; de Ruyck et al., 2016; Vilar et al., 2016; Imam and Gilani, 2017). The application of *in silico* approaches not only speeds up new drug discovery, but also collects related information, reveals the mechanisms and proposes new hypotheses. Compared to bench research, computational experiments perform high-efficiency simulations which considerably reduce the research time and the cost of experiments (Abel et al., 2017). However, in most cases, *in silico* approaches played an assisting role in the process. In this study, we propose an all-computational protocol integrating multiple *in silico* approaches to simulate the entire drug discovery process from *de novo* protein structure prediction to drug-protein interaction disclosure.

Up to now, reliable *in silico* approaches, such as molecular dynamics simulation (MD simulation) and machine learning (Durrant and McCammon, 2011; Lavecchia, 2015; Mortier et al., 2015), have been increasingly developed and applied in finding new drugs and optimizing them for treatment of diseases. However, most research projects only use one single *in silico* approach. For instance, the homology models of CIB2, a calcium- and integrin- binding protein, were constructed based on CIB1 structures using SWISS-MODEL server (Waterhouse et al., 2018). Based on the models, the way how the point mutations affect the affinities of calcium- and integrin-binding were predicted and then validated by *in vitro* experiments (Riazuddin et al., 2012). Besides, the results of ligand-protein docking were used to test the substrate specificity of OCT-1 and OCT-2 (organic cation transporter) to guide the following *in vivo* experimental validation (Papaluca and Ramotar, 2016). These works simply employed the *in silico* approaches as supportive methods, which did not sufficiently leverage the high-speed advantages of computational methods. Integrating the established *in silico* researches into an all-computation pipeline and producing validated good results is a milestone in the "omics" era.

Human small conductance calcium-activated (SK2) ion channels, consisted of SK2 subunits and calmodulin molecule, have been proved to be therapeutic targets for treatment of neuronal diseases, such as Parkinson's and amyotrophic lateral sclerosis (ALS) (Bond et al., 2005; Lu et al., 2009). The crystalized structures of human SK2 bound with Riluzole, an approved drug for ALS (Romano et al., 2015), and an analog [pdb (Berman et al., 2003) ligand ID: 658] of the SK2 activator CyPPA, an anti-ataxic agent (Herrik et al., 2012), have been released recently. It was also reported that ligands of two different chemical classes, Riluzole and its analog NS309, and CyPPA and its analogs, all bind to the same binding site where the interface of SK2 and calmodulin is. In addition, two key residues in the binding site were mutated to investigate how different mutations affect the potency of two ligands (Cho et al., 2018). As a proof of concept, we chose SK2 as the target to examine the protocol we proposed. Hence, in this study, we basically repeated the entire procedure of the previous study of Cho et al. (2018) with consecutive *in silico* approaches and compared our results with those of the bench experiments. We first constructed the human SK2/calmodulin model (PSK2) using SWISS-MODEL server and docked the ligands Riluzole and CyPPA analog 1 (PDB ligand ID: 658, see section 2.1), onto the predicted binding site. The predicted site is consistent with the reported results (Cho et al., 2018). Then the residues V481 and A477 in the binding pocket were virtually mutated and the mutation effects were assessed via binding energies (MM-GBSA ΔG_{Bind}) calculation. The results show that the binding energies of mutants have similar trends compared with the EC_{50} (the concentration of a drug that gives half-maximal response) values ($R = 0.6$ for NS309 in V481 mutants). Overall, our results suggest that this protocol of *in silico* approach can provide a systematic prediction on the unknown structures of proteins and potential drugs, and they also demonstrate the ability of *in silico* approaches to speed up the new drug design process.

2. MATERIALS AND METHODS

2.1. Protein Structure Preparation

Having a determined or predicted structure of the drug target protein is the first prerequisite of structure-based drug design. To prove the all-computational protocol is valid, we started our process from structure prediction. To predict the structure of SK2 and calmodulin from *Homo sapiens*, its amino acid sequence was obtained from the UniProt website (uniprot ID: Q9H2S1). Then we uploaded the sequence onto the SWISS-MODEL server, the most widely-used and reliable structure prediction server, to build a structure model (Bienert et al., 2016; Waterhouse et al., 2018) and selected the 3-D complex structure of *Rattus norvegicus* SK2 ion channels with NS309 (PDB ID: 4J9Z) (Zhang et al., 2013) as the template of structure prediction. 4J9Z was downloaded from RCSB's Protein Data Bank (Berman et al., 2003), and the predicted structure of SK2 and calmodulin were combined in Maestro (11.5 version, Schrodinger). To test the accuracy of modeling, we uploaded the predicted protein structure and the crystalized complex structure of human SK2 and calmodulin with Riluzole (PDB ID: 5V02) onto Zhang's web server (Zhang and Skolnick, 2005) to calculate the TM-align score. 5V02 was downloaded from RCSB's Protein Data Bank. In addition, site-directed mutants were constructed using the Mutate Residue function of Maestro.

A PDB structure of the target protein cannot be directly used in molecular docking without preprocessing. In most of the cases, a PDB file does not include the information of hydrogens, the (potential) charges of atoms, or the bond orders between any two atoms. In addition, the protein structures may be determined with a missing fragment(s), a low resolution or alternate positions, or under an unnatural condition, for example, low or high pH values. To make sure molecular docking can simulate the binding between ligands and the target protein correctly and precisely, the protein and ligand preparation is necessary. The wild-type and mutant (predicted) protein structures to be used for docking were processed by protein preparation wizard in Maestro (Sastry et al., 2013). The workflow of protein preparation contains three steps as follows: (1) Preprocess: assigning bond orders, adding hydrogens, creating zero-order bonds to metals, creating disulfide, filling in missing side chains using Prime, deleting waters beyond 5.00 Å from het groups (to keep the water molecules which may form hydrogen-bond bridges between the protein and the ligand and remove those cannot form hydrogen-bond bridges) and generating het states using Epik (pH = 7.0 ± 2.0) (Shelley et al., 2007); (2) Optimization: setting pH = 7.0 and performing optimization; (3) Minimization: this step was performed using the OPLS3 force field (Harder et al., 2015). The converge heavy atoms to root-mean-square deviation (RMSD) is 0.30 Å.

2.2. Ligand Preparation

The 3-D molecular structures of Riluzole and NS309 were obtained from the PubChem database (Kim et al., 2018). The 3-D molecular structures of CyPPA analog 1 ((4-chloro-phenyl)-[2-(3,5-dimethyl-pyrazol-1-yl)-pyrimidin-4-yl]-amine) and

analog 2 (4-chloro-phenyl)-[2-(3,5-dimethyl-pyrazol-1-yl)-6-methyl-pyrimidin-4-yl]-amine) were built in Maestro (11.5 version, Schrodinger) based on a previous study (Cho et al., 2018). All the compounds were prepared using OPLS3 force field in Ligprep panel in Maestro (Sastry et al., 2013; Harder et al., 2015). The preparation process included converting 2D structures to 3D ones, adding hydrogens, computing correct partial charges, and optimizing the structures.

2.3. Binding Site Prediction

Knowing the potential ligand binding site(s) is also an important prerequisite prior to molecular docking. There are many well-developed binding site prediction methods and servers (Xie and Hwang, 2015); however, the predictions produced by different methods may disagree with each other. Therefore, researchers usually compare the prediction results produced by different methods to find the consensus among all predictions. In our study, binding site prediction process was completed using the LISE web server (http://lise.ibms.sinica.edu.tw/applet/) (Xie and Hwang, 2012), which is reported to have the highest accuracy (80–90% for a soluble protein), and the binding site prediction tool, Sitemap, in Maestro, which we used in the docking procedure after this step. The SK2 predicted structure was uploaded onto the LISE web server for binding site prediction. The top three predictions from LISE were then downloaded and imported into Maestro. Meanwhile, in Maestro, SiteMap (Halgren, 2007, 2009) predicted five ligand binding sites. After comparing the results obtained using two prediction methods, we used the consensus region to define the docking grid box.

2.4. Ligand-Protein Docking

In order to predict the details of the interaction between ligands and the target protein and to estimate their binding affinities (see section 2.5), ligand docking was conducted the extra-precision (XP) mode in Maestro. Maestro has three precision options for docking including high throughput virtual screening (HTVS), standard precision (SP), and extra precision (XP). Users can choose an option based on their need or the computational load. After the ligands and the target proteins were processed using Ligprep and protein preparation, respectively, a receptor grid box was generated according to the results of binding site prediction. The size of the receptor grid box was set as default. To investigate the interaction of the protein and ligands, Induced Fit Docking (IFD) (Farid et al., 2006; Sherman et al., 2006a,b; Clark et al., 2016) was performed in Maestro. Using the IFD, the Ligprep outputs were imported and docked to the target protein. The standard protocol was applied to generate up to 20 poses. The force field was OPLS3. Under the prime refinement tab, the conformations of binding site residues within 5 Å (default value) of the ligand were refined. In the Glide redocking process, the energy of each protein/ligand complex structure and the number of top structures were set as the default settings. The XP mode was used for all IFD process.

2.5. MM-GBSA Calculation

The binding energy (ΔG_{Bind}) between a protein and a ligand reflects how stable they bind to each other and how a point

mutation affects the ligand binding. Therefore, we examined if our model can correctly predict the trend of binding affinity changes of the mutations on the target protein. In this study, ΔG_{Bind} were estimated using the Prime MM-GBSA module in Maestro (Greenidge et al., 2012). In MM-GBSA panel, the pose viewer files of docked complex were uploaded. The solvation model was VSGB and the force field was OPLS3 (Li et al., 2011). Prime MM-GBSA ΔG_{Bind} was calculated using this equation:

$$\Delta G_{Bind} = E_{complex}(minimized) - [E_{ligand}(minimized) + E_{receptor}(minimized)] \qquad (1)$$

Where ΔG_{Bind} is binding free energy and $E_{complex}(minimized)$, $E_{ligand}(minimized)$, and $E_{receptor}(minimized)$ are minimized energies of receptor-ligand complex, ligand and receptor, respectively.

2.6. Virtual Mutation

Based on the literature, valine 481 (V481) of SK2 is a crucial residue which forms the hydrophobic core between SK2 and calmodulin (Cho et al., 2018). To investigate the impacts of V481 mutations in the binding pocket using *in silico* approaches, we first implemented the site-directed mutagenesis in Maestro. The V481 was virtually mutated to alanine, serine, threonine, aspartate, or phenylalanine. Then NS309 and CyPPA analog 2 were docked onto the mutated binding site of PSK2 using IFD and the binding free energies were calculated using MM-GBSA to reveal the effects of mutated residues.

Alanine 477 (A477) is another vital residue in the binding pocket (Cho et al., 2018). We exploited the same method mentioned above to virtually mutate A477 to isoleucine, leucine, valine, serine, threonine, arginine, and aspartate, docked NS309 and CyPPA analog 2 onto the mutated binding site of PSK2 and calculated the binding free energies using MM-GBSA.

2.7. Molecular Dynamics Simulation

The molecular dynamics (MD) simulations were performed using GROMACS version 2018.1 and CHARMM36 all-atom force field (March 2019) (Vanommeslaeghe et al., 2010, 2012; Vanommeslaeghe and MacKerell, 2012; Yu et al., 2012; Gutiérrez et al., 2016). The starting coordinates of each protein-ligand complex were obtained from docking experiments. Then we defined a dodecahedral unit cell and filled it with water molecules. After adding ions, the complex was minimized for 50,000 steps of steepest descent minimization. Next, the complex was equilibrated using an NVT ensemble (constant Number of particles, Volume, and Temperature) and NPT ensemble (the Number of particles, Pressure, and Temperature). The target temperature for equilibration was 300 K. At last, the simulations were performed for 30 ns. After the MD simulations, we calculated the RMSD of the residues which were mentioned in the previous research in four trajectories (Cho et al., 2018). Then, we selected four time points of two residues: I100 on protein-NS309 complex (15,000, 18,000, 24,000, and 30,000 ps) and D64 on protein-CyPPA analog 1 (2,610, 6,000, 15,000, and 21,000 ps). Finally, four conformations of both residues were converted into PDB files and were superposed using PyMol.

3. RESULTS

3.1. Structure Prediction of Human SK2/Calmodulin Complex

The structure of human SK2/calmodulin was generated using template-based modeling on the SWISS-MODEL server (**Figure 1A**). Based on the structure of SK2/calmodulin from *Rattus norvegicus* (PDB ID: 4J9Z), the structure models were predicted using the amino acid sequences of human SK2 and calmodulin. Additionally, the loop (residue A403 to residue D413, the "intrinsically disordered fragment"—IDF), which had not been determined in structure of human SK2/calmodulin complex (PDB: 5V02), was obtained in the predicted model (**Figure 1A**). It suggests that the predicted structure can be used to supplement the crystallized structure.

To test the accuracy of the predicted model, we used TM-align server to calculate the TM-align score (Zhang and Skolnick, 2005). The predicted models were aligned with 5V02, and the TM-align scores of SK2 and calmodulin are 0.99124 and 0.89215. These TM-align scores show that the structure of human SK2/calmodulin complex has been accurately predicted.

3.2. Binding Site Prediction

To determine the binding pocket in SK2/calmodulin complex, the Sitemap and LISE were exploited to predict binding sites (Halgren, 2007, 2009; Xie and Hwang, 2012). The top three predicted results from LISE were overlapped with the results of

FIGURE 1 | Overlapped structures of the predicted model of human SK2 ion channels and 5V02. **(A)** The predicted structure of human SK2 ion channels (PSK2) in blue overlapped with its crystallized structure (5V02) in red. The structure of a missing loop (IDF) near the ligand binding pocket in 5V02/5V03 has also been predicted. The green compound was Riluzole, which was the ligand in 5V02. **(B)** The predicted binding sites (Blue dots represent the results from Sitemap; red dots represent the results from LISE) overlapped with the Riluzole binding site in 5V02.

Sitemap. After the comparison, we found that there was only one consensus. Then we selected this binding site to generate receptor grid for docking. To verify the accuracy of binding site prediction, we also overlapped the predicted binding site with that of 5V02 (**Figure 1B**). Notably, the predicted binding site is the same binding site of Riluzole in 5V02, which suggests that this binding site is the potential binding site for the ligands. Hence, the *in silico* approaches successfully predict the accurate binding site.

3.3. Molecular Docking

Based on the previous study (Cho et al., 2018), the interface of SK2 and calmodulin can be bound by Riluzole, NS309, CyPPA analog 1, and analog 2. To investigate whether we can obtain the same results using *in silico* approaches, we first docked Riluzole and CyPPA analog 1 onto the predicted model (PSK2) via IFD (Induced Fit Docking) in Maestro, and redocked these ligands onto the determined structures (5V02 and 5V03) and calculated the binding energies as the control. Then we calculated the binding energies using MM-GBSA to estimate the binding affinities (Greenidge et al., 2012). As a result, the MM-GBSA ΔG_{Bind} of PSK2 with Riluzole and CyPPA analog 1 are −40.19 and −56.11 kcal/mol, respectively. As indicated in **Table 1**, those results of PSK2 are compatible to the results of 5V02 and 5V03, which demonstrate that accurate results can be obtained using *in silico* approaches. Additionally, the docking pose of PSK2 with Riluzole is almost identical with that in 5V02 (**Figure 2A**, the RMSD between the ligand of crystal structure and docking poses on 5V02 or PSK2 is 0.43 or 0.72 Å), which indicates that the simulated result from IFD can obtain accurate data in comparison with the results of crystallization. In **Figure 2B**, the coordinates of docked and native ligands are almost the same even though the poses of two ligands are not completely overlapped (the RMSD between the ligand of crystal structure and docking poses on 5V03 or PSK2 is 1.59 or 0.87 Å). Analyzing the docking results, Riluzole and CyPPA analog 1 interact with residues in both SK2 and calmodulin (**Figure 3**). The interacting residues in the binding site are mostly hydrophobic residues. As CyPPA analogs are larger molecules, they interact with more residues. Compared the list of interacting residues (**Table S1**), our results are almost identical to those of the previous study (Cho et al., 2018). The discrepancy may be because Maestro analyzes the interactions between ligands and the protein based on the interaction energy and the previous
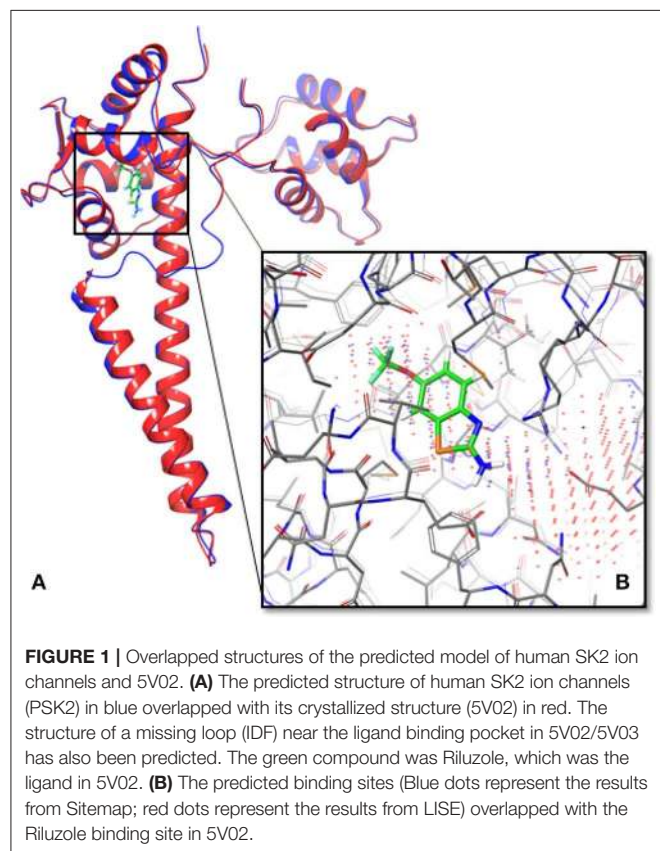
TABLE 1 | MM-GBSA ΔG_{Bind} of ligands bound to crystallized structures and predicted structures.

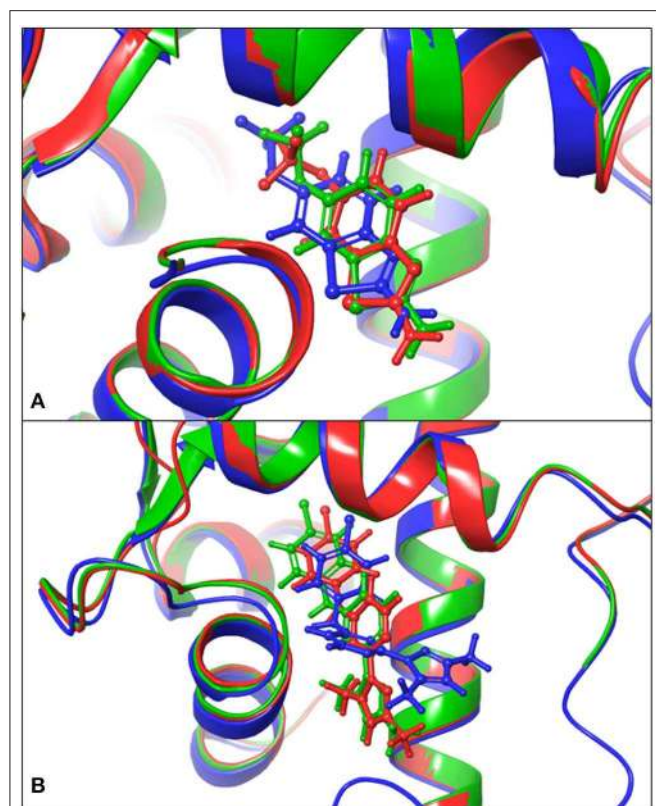| Protein	Ligand	MM-GBSA ΔG_{Bind} (kcal/mol)
5V02	Riluzole	−40.83
5V03	CyPPA analog 1	−62.72
PSK2	Riluzole	−40.19
PSK2	NS309	−49.97
PSK2	CyPPA analog 1	−56.11
PSK2	CyPPA analog 2	−64.84

FIGURE 2 | Comparison of the docked poses and the native structures. **(A)** Comparison of the docking results on the PSK2 in blue color, the docking results on 5V02 in green color, and determined structures 5V02 in red color; The blue-colored Riluzole is the docked pose on the PSK2, the green-colored Riluzole is the docked pose on 5V02, and the red-colored Riluzole is the determined structure in 5V02; **(B)** Comparison of the docking results on the PSK2 in blue color, the docking results on 5V03 in green color, and determined structure 5V03 in red color; The blue-colored CyPPA analog 1 is the docked pose on the PSK2, the green-colored CyPPA analog 1 is the docked pose on 5V03 and the red-colored CyPPA analog 1 is the determined structure in 5v03.

FIGURE 3 | 2D ligand-protein interaction of Riluzole **(A)** and CyPPA analog 1 **(B)** at PSK2 binding site. The pink arrow is referred to hydrogen bond. The red line is referred to Pi-cation interaction.

study (Cho et al., 2018) simply lists the residues within 5 Å of either ligand. According to the docking results of PSK2 with Riluzole and CyPPA analog 1, we docked NS309 and CyPPA analog 2 on the same binding site using the same methods mentioned above (**Figure 4** and **Figure S1**). The binding affinities MM-GBSA ΔG_{Bind} are showed in **Table 1**. The MM-GBSA ΔG_{Bind} values of CyPPA analog 1 and CyPPA analog 2 on the PSK2 are more negative than those of the other ligands, which suggests that CyPPA analog 1 and CyPPA analog 2 are promising drug candidates.

To verify the ligand-induced perturbations, NMR spectrum was used to identify residues with significant chemical shifts in previous experiments (Cho et al., 2018). With computational approaches, we ran MD simulations for each protein-ligand complex to simulate the conformational changes after the binding of ligands. A previous research reports that the residues on EF hands of calmodulin had conformational changes due to the ligand binding (Cho et al., 2018). Therefore, we calculated
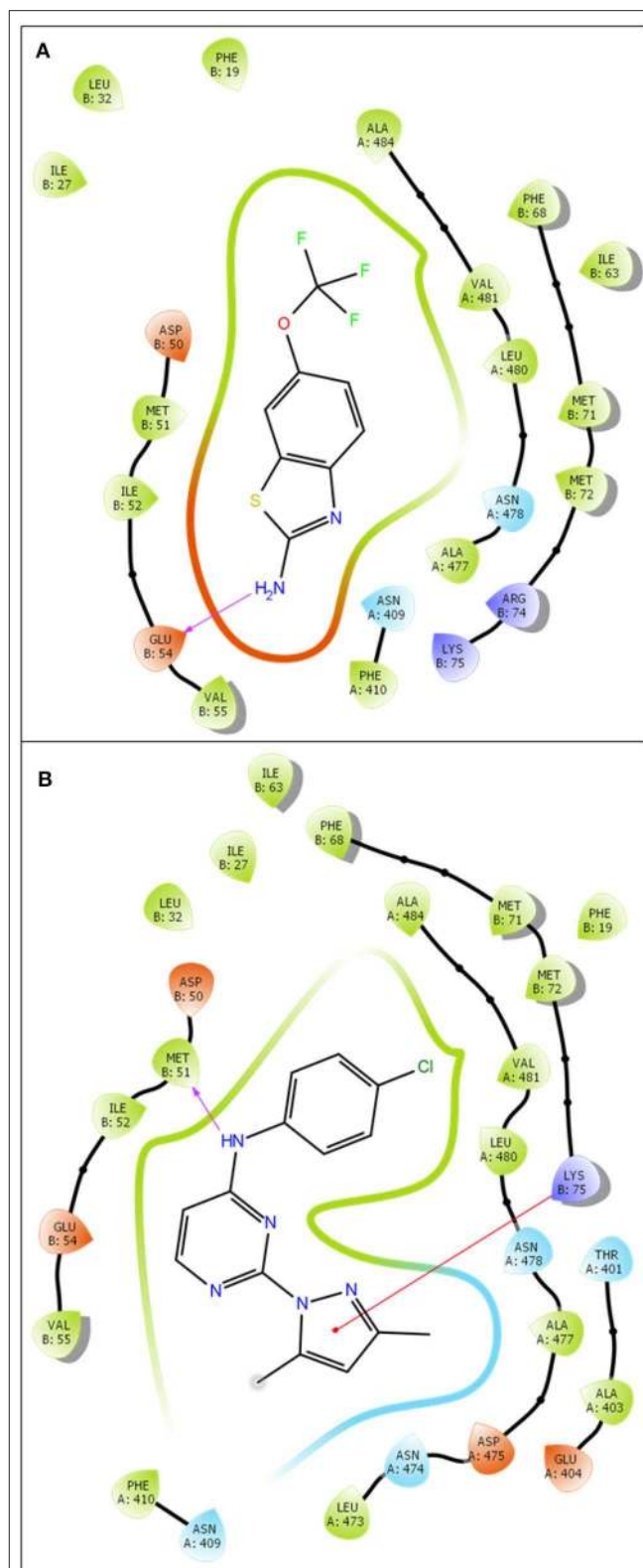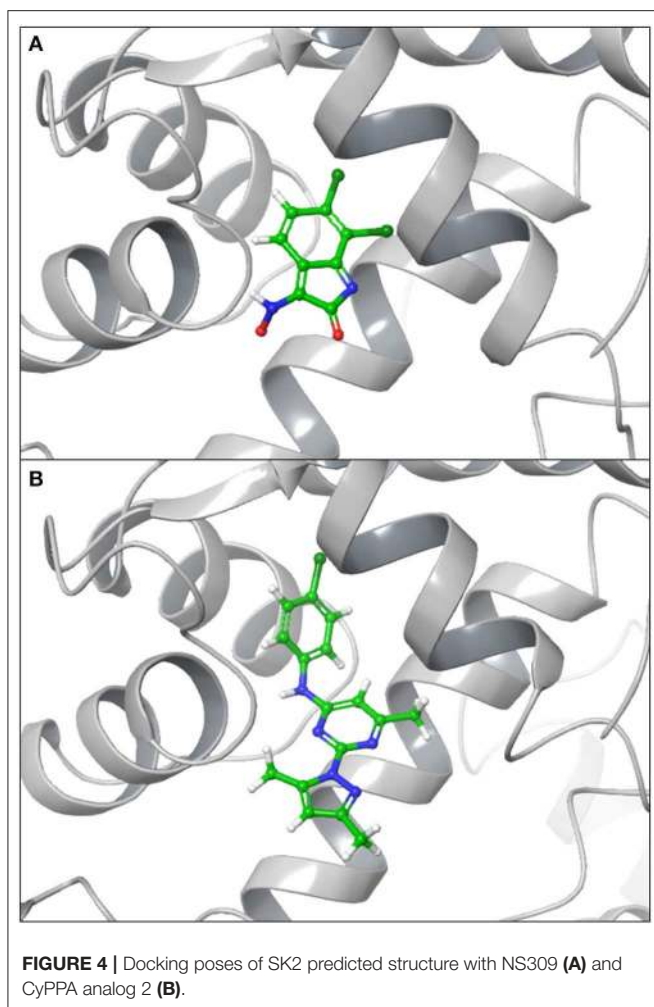
FIGURE 4 | Docking poses of SK2 predicted structure with NS309 **(A)** and CyPPA analog 2 **(B)**.

RMSD for those residues. According to the results of RMSD, we compared four trajectories of each residue and superposed their different poses at different time points. As a result, we found I100, on the complex of NS309, had obvious perturbations on four time points (**Figure 5** and **Figure S2**). In addition, D64 which was a critical residue for calcium sensing, also showed dramatically conformational changing on the complex structure of PSK2 and CyPPA analog 1 (**Figure 6** and **Figure S3**). Hence, the different poses of those residues demonstrated that the ligand could induce the perturbations of the calmodulin, which was consistent with the conclusion in the previous research (Cho et al., 2018).

3.4. Virtual Mutations at V481 and A477 in the Binding Pocket

The results of a previous study (Cho et al., 2018) show that the site-directed mutations on the key residues V481 and A477 in the binding site result in changes in the binding affinities. To validate that the *in silico* approach can simulate and predict the impacts of these mutations, we performed the virtual mutation experiments, docked the ligands on all mutants, and calculated the corresponding binding energies of each mutant. The MM-GBSA ΔG_{Bind} of V481 mutants are shown in **Table 2**. To verify the accuracy of *in silico* approaches, the Pearson's

correlation coefficients between MM-GBSA ΔG_{Bind} and EC_{50} were calculated ($R = 0.6$ for NS309 in V481 mutants). **Table 2** indicates that substitutions with small side chains (V481S) or a charged amino acid (V481D) significantly decrease the binding affinities of NS309. Conversely, the binding affinity of NS309 in PSK2 V481F is close to that in wild-type. The results above are consistent with the conclusion in literature, that is, this position requires a non-charged residue with a bigger side chain (Cho et al., 2018).

Similarly, we find that CyPPA analog 2 in PSK2 V481F mutant with large aromatic side chains also shows a close binding affinity compared to that in PSK2 WT (**Table 2**), which demonstrates a good correlation between calculated binding affinities and EC_{50}. The mutants with small side chains (V481A and V481S) or a charged amino acid (V481D) also shows the relatively lower binding affinities of CyPPA analog 2. Those results demonstrate that the simulation results are compatible with the data from biological experiments. The consistent conclusion has successfully proved that the all-computational protocol can be widely applied in future biomedical research.

In **Table 3**, all PSK2 A477 mutants have slightly lower NS309 potency than that of PSK2 WT and the predicted binding affinities MM-GBSA ΔG_{Bind} have similar results. According to the results in literature (Cho et al., 2018), CyPPA analog 2 should exhibit shifted potencies at PSK2 A477L, PSK2 A477V, PSK2 A477S, PSK2 A477T, PSK2 A477R, and PSK2 A477D, but not at PSK2 A477I. However, in **Table 3**, the MM-GBSA ΔG_{Bind} of PSK2 A477I is not different from those of other mutants. As an isoleucine has a long side chain, different rotamers may largely change the estimated binding affinities. Performing an MD simulation may be a good solution to optimize the structures of mutants and improve the docked poses and estimated binding energies.

4. DISCUSSION

In the field of new drug discovery, research efficiency is particularly essential. On one hand, the speed of new drug development is of great importance to patients, especially the ones with fatal diseases such as cancers or acute infectious diseases (Shi et al., 2015). Statistics show that there will be around 1.7 million new cancer cases and 600 thousands cancer deaths in the United States in 2019 (Siegel et al., 2019). This race against time has always been a huge challenge for the researchers in this field. On the other hand, to bring a new drug to the market from compound identification to final FDA approval usually costs up to billions of dollars (Cleary et al., 2018). Therefore, the time and cost-efficient virtual process of drug development will benefit many pharmaceutical companies and our society. *In silico* approaches which can save considerable amount of research time and cost should be applied to drug design. With the rapid development of computer science and engineering, the availability and accuracy of *in silico* approaches have been constantly improving. Although many progresses have been made in utilizing *in silico* approaches to simulate certain biological experiments, the whole experimental processes completed in the virtual way from protein structure simulation
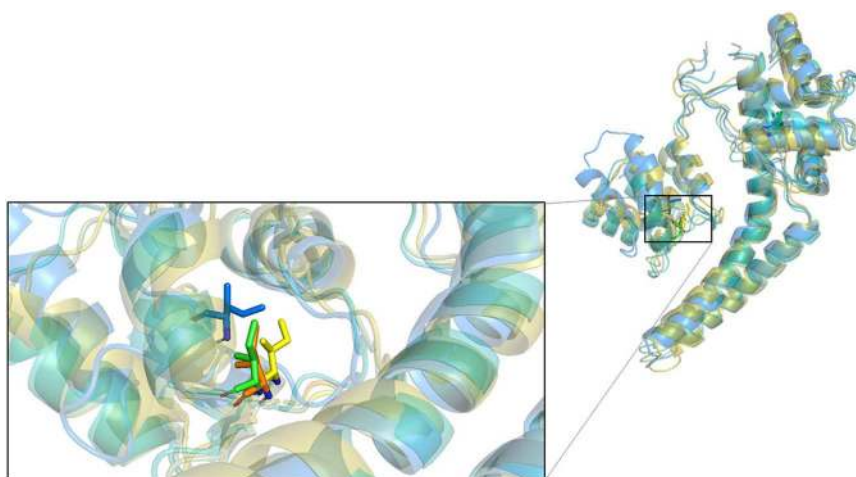
FIGURE 5 | Superposition of four conformations of I100 at four time points on the simulation trajectory of the complex of PSK2 and NS309.
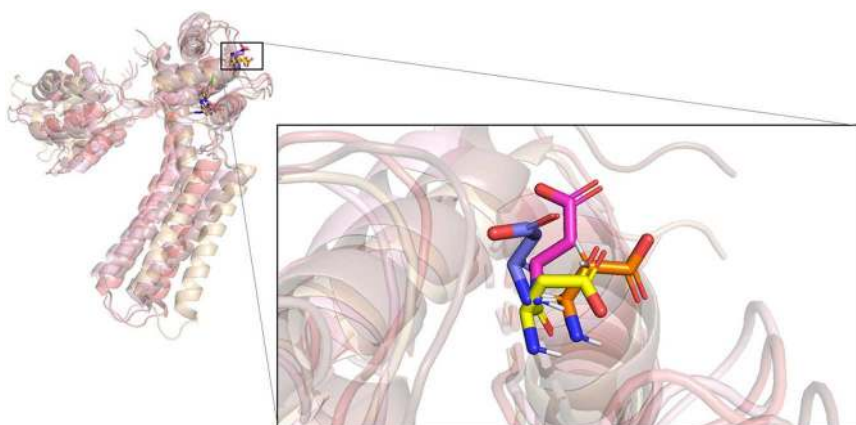

FIGURE 6 | Superposition of four conformations of D64 at four time points on the simulation trajectory of the complex of PSK2 and CyPPA analog 1.

TABLE 2 | EC_{50} and MM-GBSA ΔG_{Bind} of NS309 and CyPPA analog 2 bound to the V481 mutants.

Protein	NS309		CyPPA analog 2	
	EC_{50} (μM)	MM-GBSA ΔG_{Bind} (kcal/mol)	EC_{50} (μM)	MM-GBSA ΔG_{Bind} (kcal/mol)
WT	1.1	−49.97	2.5	−64.84
V481A	4.3	−54.28	>250	−29.05
V481S	5.1	−32.65	>250	−49.94
V481T	1.2	−32.77	46.6	−49.97
V481D	7.4	−25.97	>250	−46.75
V481F	0.8	−57.77	3.3	−64.09

TABLE 3 | EC_{50} and MM-GBSA ΔG_{Bind} of NS309 and CyPPA analog 2 bound to the A477 mutants.

Protein	NS309		CyPPA analog 2	
	EC_{50} (μM)	MM-GBSA ΔG_{Bind} (kcal/mol)	EC_{50} (μM)	MM-GBSA ΔG_{Bind} (kcal/mol)
WT	1.1	−49.97	2.5	−67.14
A477I	2.0	−37.26	2.0	−47.89
A477L	1.7	−24.71	99.0	−37.60
A477V	2.0	−39.67	>250	−45.63
A477S	3.0	−36.65	>250	−46.05
A477T	2.0	−23.36	>250	−51.04
A477R	2.1	−31.24	>250	−51.50
A477D	3.9	−30.94	>250	−49.71

to protein-drug interaction characterization has never been achieved before.

In this study, we proposed a protocol of integration of *in silico* approaches to simulate the process from protein structure determination to key residues mutagenesis and characterization. To validate this strategy, we selected the human SK2 ion channels as our target protein. With successful prediction, we obtained

accurate protein structures (TM-align score >0.5) and the same binding site as the crystallized structure. Furthermore, the docking poses of Riluzole and CyPPA analog 1 are consistent with the ligand bound conformations in the crystallized structures. We also successfully reproduced similar effects of site-directed mutagenesis on the ligand binding, which demonstrated great potential of the integration of *in silico* approaches. However, the purpose of integrating *in silico* approaches is not to completely replace biological experiments but to speed up the drug discovery process with the continuous and automatic computational process.

A possible reason why an all-computation protocol of drug design has not been proposed and implemented is the inaccuracy or uncertainty of prediction results might accumulate in the sequential computational pipeline. However, the state-of-the-art bioinformatics algorithms, software or servers have been highly accurate in many cases so that it is time to integrate them into a fully computational process or even a fully automatic process. This is the first study to demonstrate the feasibility of an all-computational protocol in drug design. To achieve the ultimate goal of "automatic" drug discovery, more online servers or computational algorithms like PROCHECK (Laskowski et al., 1993), which can be used to assess or estimate the reliability of prediction results generated by each *in silico* approach, need to be developed. Despite its innovative approach, there are a few limitations of this study. First, the accuracy of protein structure prediction relies on the methods and/or templates. In our research, we selected the SK2/calmodulin from *Rattus norvegicus* (PDB ID: 4J9Z) as the template to build protein structure on SWISS MODEL, whose results are more accurate than the results from other webservers (data not shown). Hence, a reliable tool or method is critical to the accuracy of the final simulated results. Second, the binding affinity changes are hard to be precisely reproduced, especially those of the mutants, because considering possible conformational changes on target proteins is still the biggest challenge of docking. This suggests that MD simulation which can simulate the conformational changes should be used to further improve the precision of the predictions.

In conclusion, this work established and demonstrated an integrated protocol of *in silico* approaches for the first time. Its applicability can be potentially extended beyond the characterizing of SK2 ion channels to investigating other proteins with unknown structures, such as the Alzheimer's disease related proteins (Fitzpatrick et al., 2017; Hatami et al., 2017), which are also treatment targets of neural degenerative diseases. Although there are challenges to the *in silico* approaches, our work still paves a new way toward an automatic procedure of drug design.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

YW carried out the research, analyzed the data, and drafted the manuscript. LL analyzed the data. Z-RX conceived the project, guided the research, and revised the manuscript. All authors read and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2020.00081/full#supplementary-material

REFERENCES

Abel, R., Mondal, S., Masse, C., Greenwood, J., Harriman, G., Ashwell, M. A., et al. (2017). Accelerating drug discovery through tight integration of expert molecular design and predictive scoring. *Curr. Opin. Struct. Biol.* 43, 38–44. doi: 10.1016/j.sbi.2016.10.007

Aparna, V., Dineshkumar, K., Mohanalakshmi, N., Velmurugan, D., and Hopper, W. (2014). Identification of natural compound inhibitors for multidrug efflux pumps of *Escherichia coli* and pseudomonas aeruginosa using *in silico* high-throughput virtual screening and *in vitro* validation. *PLoS ONE* 9:e101840. doi: 10.1371/journal.pone.0101840

Berman, H. M., Bourne, P. E., Westbrook, J., and Zardecki, C. (2003). "The protein data bank," in *Protein Structure*, ed D. Chasman (Boca Raton, FL: CRC Press), 394–410. doi: 10.1201/9780203911327.ch14

Bienert, S., Waterhouse, A., de Beer, T. A., Tauriello, G., Studer, G., Bordoli, L., et al. (2016). The swiss-model repository—new features and functionality. *Nucleic Acids Res.* 45, D313–D319. doi: 10.1093/nar/gkw1132

Bond, C. T., Maylie, J., and Adelman, J. P. (2005). Sk channels in excitability, pacemaking and synaptic integration. *Curr. Opin. Neurol.* 15, 305–311. doi: 10.1016/j.conb.2005.05.001

Cho, L. T.-Y., Alexandrou, A. J., Torella, R., Knafels, J., Hobbs, J., Taylor, T., et al. (2018). An intracellular allosteric modulator binding pocket in SK2 ion channels is shared by multiple chemotypes. *Structure* 26, 533–544. doi: 10.1016/j.str.2018.02.017

Clark, A. J., Tiwary, P., Borrelli, K., Feng, S., Miller, E. B., Abel, R., et al. (2016). Prediction of protein–ligand binding poses via a combination of induced fit docking and metadynamics simulations. *J. Chem. Theory Comput.* 12, 2990–2998. doi: 10.1021/acs.jctc.6b00201

Cleary, E. G., Beierlein, J. M., Khanuja, N. S., McNamee, L. M., and Ledley, F. D. (2018). Contribution of NIH funding to new drug approvals 2010–2016. *Proc. Natl. Acad. Sci. U.S.A.* 115, 2329–2334. doi: 10.1073/pnas.1715368115

de Ruyck, J., Brysbaert, G., Blossey, R., and Lensink, M. F. (2016). Molecular docking as a popular tool in drug design, an *in silico* travel. *Adv. Appl. Bioinform. Chem.* 9:1. doi: 10.2147/AABC.S105289

Durrant, J. D., and McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biol.* 9:71. doi: 10.1186/1741-7007-9-71

Farid, R., Day, T., Friesner, R. A., and Pearlstein, R. A. (2006). New insights about herg blockade obtained from protein modeling, potential energy mapping, and docking studies. *Bioorg. Med. Chem.* 14, 3160–3173. doi: 10.1016/j.bmc.2005.12.032

Fitzpatrick, A. W., Falcon, B., He, S., Murzin, A. G., Murshudov, G., Garringer, H. J., Crowther, R. A., et al. (2017). Cryo-em structures of tau filaments from Alzheimer's disease. *Nature* 547:185. doi: 10.1038/nature23002

Gómez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hernández-Lobato, J. M., Sánchez-Lengeling, B., Sheberla, D., et al. (2018). Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent. Sci.* 4, 268–276. doi: 10.1021/acscentsci.7b00572

Greenidge, P. A., Kramer, C., Mozziconacci, J.-C., and Wolf, R. M. (2012). MM/GBSA binding energy prediction on the PDBbind data set: successes, failures, and directions for further improvement. *J. Chem. Inf. Model.* 53, 201–209. doi: 10.1021/ci300425v

Gutiérrez, I. S., Lin, F.-Y., Vanommeslaeghe, K., Lemkul, J. A., Armacost, K. A., Brooks, C. L. III., et al. (2016). Parametrization of halogen bonds in the charmm general force field: improved treatment of ligand–protein interactions. *Bioorg. Med. Chem.* 24, 4812–4825. doi: 10.1016/j.bmc.2016.06.034

Halgren, T. (2007). New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug Des.* 69, 146–148. doi: 10.1111/j.1747-0285.2007.00483.x

Halgren, T. A. (2009). Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* 49, 377–389. doi: 10.1021/ci800324m

Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J. Y., et al. (2015). OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* 12, 281–296. doi: 10.1021/acs.jctc.5b00864

Hatami, A., Monjazeb, S., Milton, S., and Glabe, C. G. (2017). Familial Alzheimer's disease mutations within the amyloid precursor protein alter the aggregation and conformation of the amyloid-β peptide. *J. Biol. Chem.* 292, 3172–3185. doi: 10.1074/jbc.M116.755264

Herrik, K. F., Redrobe, J. P., Holst, D., Hougaard, C., Nielsen, K. S., Nielsen, A. N., et al. (2012). CyPPA, a positive SK3/SK2 modulator, reduces activity of dopaminergic neurons, inhibits dopamine release, and counteracts hyperdopaminergic behaviors induced by methylphenidate1. *Front Pharmacol.* 3:11. doi: 10.3389/fphar.2012.00011

Imam, S. S., and Gilani, S. J. (2017). Computer aided drug design: a novel loom to drug discovery. *Org. Med. Chem.* 1, 1–6. doi: 10.19080/OMCIJ.2017.01.555567

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2018). Pubchem 2019 update: improved access to chemical data. *Nucleic Acids Res.* 47, D1102–D1109. doi: 10.1093/nar/gky1033

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). Procheck: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291. doi: 10.1107/S0021889892009944

Lavecchia, A. (2015). Machine-learning approaches in drug discovery: methods and applications. *Drug Discov. Today* 20, 318–331. doi: 10.1016/j.drudis.2014.10.012

Li, J., Abel, R., Zhu, K., Cao, Y., Zhao, S., and Friesner, R. A. (2011). The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins* 79, 2794–2812. doi: 10.1002/prot.23106

Lu, L., Timofeyev, V., Li, N., Rafizadeh, S., Singapuri, A., Harris, T. R., et al. (2009). α-actinin2 cytoskeletal protein is required for the functional membrane localization of a Ca2+-activated K+ channel (SK2 channel). *Proc. Natl. Acad. Sci. U.S.A.* 106, 18402–18407. doi: 10.1073/pnas.0908207106

Mortier, J., Rakers, C., Bermudez, M., Murgueitio, M. S., Riniker, S., and Wolber, G. (2015). The impact of molecular dynamics on drug design: applications for the characterization of ligand–macromolecule complexes. *Drug Discov. Today* 20, 686–702. doi: 10.1016/j.drudis.2015.01.003

Papaluca, A., and Ramotar, D. (2016). A novel approach using *C. elegans* DNA damage-induced apoptosis to characterize the dynamics of uptake transporters for therapeutic drug discoveries. *Sci. Rep.* 6:36026. doi: 10.1038/srep36026

Riazuddin, S., Belyantseva, I. A., Giese, A. P., Lee, K., Indzhykulian, A. A., Nandamuri, S. P., et al. (2012). Alterations of the CIB2 calcium-and integrin-binding protein cause usher syndrome type 1J and nonsyndromic deafness DFNB48. *Nat. Genet.* 44, 1265–1271. doi: 10.1038/ng.2426

Romano, S., Coarelli, G., Marcotulli, C., Leonardi, L., Piccolo, F., Spadaro, M., et al. (2015). Riluzole in patients with hereditary cerebellar ataxia: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol.* 14, 985–991. doi: 10.1016/S1474-4422(15)00201-X

Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* 27, 221–234. doi: 10.1007/s10822-013-9644-8

Shelley, J. C., Cholleti, A., Frye, L. L., Greenwood, J. R., Timlin, M. R., and Uchimaya, M. (2007). Epik: a software program for pK a prediction and protonation state generation for drug-like molecules. *J. Comput. Aided Mol. Des.* 21, 681–691. doi: 10.1007/s10822-007-9133-z

Sherman, W., Beard, H. S., and Farid, R. (2006a). Use of an induced fit receptor structure in virtual screening. *Chem. Biol. Drug Des.* 67, 83–84. doi: 10.1111/j.1747-0285.2005.00327.x

Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A., and Farid, R. (2006b). Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* 49, 534–553. doi: 10.1021/jm050540c

Shi, J., Wang, E., Milazzo, J. P., Wang, Z., Kinney, J. B., and Vakoc, C. R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* 33, 661–667. doi: 10.1038/nbt.3235

Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer statistics, 2019. *CA Cancer J. Clin.* 69, 7–34. doi: 10.3322/caac.21551

Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., et al. (2010). Charmm general force field: a force field for drug-like molecules compatible with the charmm all-atom additive biological force fields. *J. Comput. Chem.* 31, 671–690. doi: 10.1002/jcc.21367

Vanommeslaeghe, K., and MacKerell, A. D. Jr. (2012). Automation of the CHARMM general force field (CGenFF) I: bond perception and atom typing. *J. Chem. Inf. Model.* 52, 3144–3154. doi: 10.1021/ci300363c

Vanommeslaeghe, K., Raman, E. P., and MacKerell, A. D. Jr. (2012). Automation of the charmm general force field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* 52, 3155–3168. doi: 10.1021/ci3003649

Vilar, S., Quezada, E., Uriarte, E., Costanzi, S., Borges, F., Viña, D., et al. (2016). Computational drug target screening through protein interaction profiles. *Sci. Rep.* 6:36969. doi: 10.1038/srep36969

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). Swiss-model: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi: 10.1093/nar/gky427

Xie, Z.-R., and Hwang, M.-J. (2012). Ligand-binding site prediction using ligand-interacting and binding site-enriched protein triangles. *Bioinformatics* 28, 1579–1585. doi: 10.1093/bioinformatics/bts182

Xie, Z.-R., and Hwang, M.-J. (2015). "Methods for predicting protein–ligand binding sites," in *Methods Mol. Biol.*, ed A. Kukol (New York, NY: Springer), 383–398.

Yu, W., He, X., Vanommeslaeghe, K., and MacKerell, A. D. Jr. (2012). Extension of the charmm general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem.* 33, 2451–2468. doi: 10.1002/jcc.23067

Zhang, M., Pascal, J. M., and Zhang, J.-F. (2013). Unstructured to structured transition of an intrinsically disordered protein peptide in coupling Ca2+-sensing and SK channel activation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4828–4833. doi: 10.1073/pnas.1220253110

Zhang, Y., and Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* 33, 2302–2309. doi: 10.1093/nar/gki524

Zhou, W., Wang, Y., Lu, A., and Zhang, G. (2016). Systems pharmacology in small molecular drug discovery. *Int. J. Mol. Sci.* 17:246. doi: 10.3390/ijms17020246