



Review

Advances in Molecular Dynamics Simulations and Enhanced Sampling Methods for the Study of Protein Systems

Raudah Lazim [†], Donghyuk Suh [†] and Sun Choi ^{* ID}

College of Pharmacy and Graduate School of Pharmaceutical Sciences, Ewha Womans University, Seoul 03760, Korea; raudah@ewha.ac.kr (R.L.); dsuh@ewha.ac.kr (D.S.)

* Correspondence: sunchoi@ewha.ac.kr

[†] These authors contributed equally to this work.

Received: 17 July 2020; Accepted: 11 August 2020; Published: 1 September 2020



Abstract: Molecular dynamics (MD) simulation is a rigorous theoretical tool that when used efficiently could provide reliable answers to questions pertaining to the structure-function relationship of proteins. Data collated from protein dynamics can be translated into useful statistics that can be exploited to sieve thermodynamics and kinetics crucial for the elucidation of mechanisms responsible for the modulation of biological processes such as protein-ligand binding and protein-protein association. Continuous modernization of simulation tools enables accurate prediction and characterization of the aforementioned mechanisms and these qualities are highly beneficial for the expedition of drug development when effectively applied to structure-based drug design (SBDD). In this review, current all-atom MD simulation methods, with focus on enhanced sampling techniques, utilized to examine protein structure, dynamics, and functions are discussed. This review will pivot around computer calculations of protein-ligand and protein-protein systems with applications to SBDD. In addition, we will also be highlighting limitations faced by current simulation tools as well as the improvements that have been made to ameliorate their efficiency.

Keywords: molecular dynamics simulation; enhanced sampling; protein-protein interactions; binding free energy; protein-ligand binding affinity

1. Introduction

Proteins are vital constituents of living organisms, responsible for myriads of life-sustaining cellular processes such as molecular recognition, signal transduction, protein localization, and enzyme catalysis. These biological events are navigated by protein motions as well as physical interactions established between proteins and partnering molecules such as ligands, peptides, proteins and nucleic acids [1]. Progresses made in structural biology and biophysical characterization methods (X-ray crystallography, nuclear magnetic resonance (NMR) and cryo-electron microscopy) have resulted in the exponential growth in the number of three-dimensional structures available for proteins, protein-ligand and protein-protein complexes. This encourages the study of protein dynamics at the atomic level which is made facile with the advent of molecular dynamics (MD) simulation.

The introduction of MD simulation to molecular biology has enabled researchers to access a microscopic visualization of biological processes. While the history of MD simulation of macromolecules began with a simple structural investigation of a small protein, namely bovine pancreatic trypsin inhibitor, the field has matured to a stage that calls for the development of new technology to facilitate the study of larger proteins and protein complexes with more intricate dynamics [2–4]. MD simulations conducted on a variety of protein systems—single proteins, protein-ligand and protein complexes—have

provided valuable insights on protein stability, protein-ligand binding and protein-protein association, among others. These applications when assimilated into the drug discovery pipeline will empower researchers to identify crucial interactions necessary for the favorable binding of small molecules, peptides, and proteins to binding pockets and/or protein-protein interfaces. The synergistic use of MD simulations and free energy calculations will further aid in harnessing thermodynamic and kinetic parameters essential for the prediction of binding free energies. The abilities of these computational tools to determine binding interactions and binding affinities are beneficial for structure-based drug design (SBDD) as such predictions could be utilized to accelerate drug development by aiding hit-to-lead optimization to afford drugs with enhanced specificity and selectivity.

The advancement in computing architecture and algorithms has spurred the extensive growth of MD simulation tools over recent decades. Architectural developments of computers resulted in the invention of graphics processing units (GPUs) and parallel computing while the refinement of algorithms leads to a more accurate rendering of the potential energy surface (PES) of the system and better conformational sampling. The combination of both facets of computing has enabled a more reliable prediction of the physical properties of proteins as well as the observation of biologically relevant protein structures at the atomic level, a feat that requires timescales beyond milliseconds, expanded length scales, and enhanced sampling techniques that facilitate high free energy barrier crossing [5]. The burgeoning use of MD simulation as a main tool in the computational studies of proteins as well as a complementary machinery in experimental studies has propelled actions to improve various aspects of MD simulations, warranting better predictive power and more reliable *in silico* analysis of protein structures, dynamics and functions. In this review, we will provide an overview of simulation strategies, mainly all-atom MD simulation and enhanced sampling methods, used to examine protein structures and dynamics to facilitate the understanding of mechanisms regulating biological processes namely, protein-ligand binding and protein-protein association. The capitalization of the predictive power of MD simulation and enhanced sampling methods to determine binding free energies as well as potential binding sites are also discussed. Furthermore, limitations of current simulation techniques as well as improvements made to optimize the accuracy of MD simulation/enhanced sampling methods are highlighted. Figure 1 reflects a summary of the topics discussed in this review.

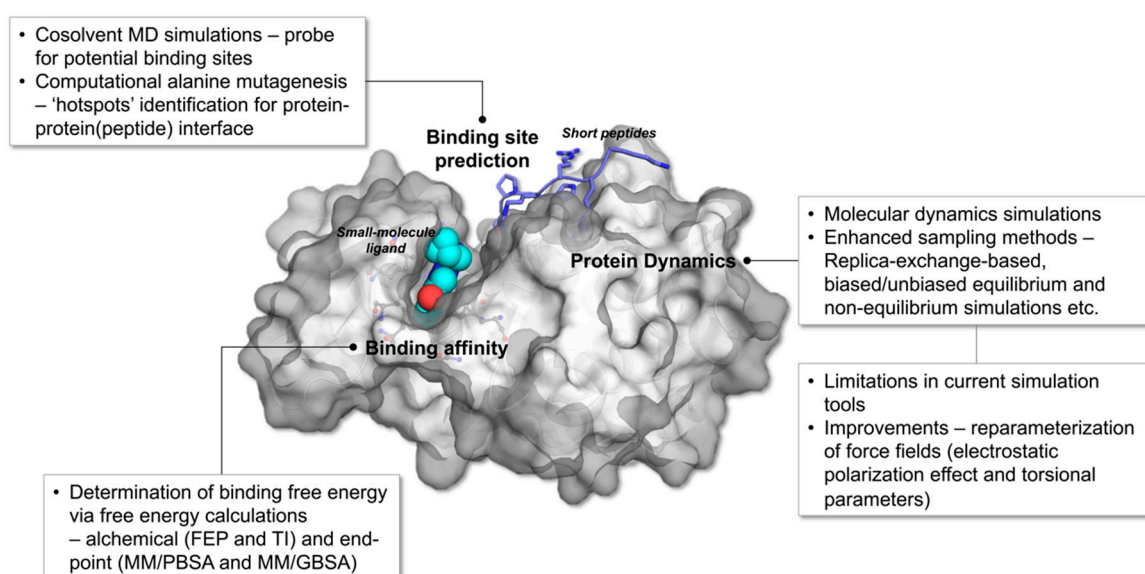


Figure 1. Overview of MD simulation and enhanced sampling methods utilized in the study of protein-protein (peptide) and protein-ligand complexes. The figure is constructed using the crystal structure of human PIM1 kinase in complexed with imidazopyridazin inhibitor and substrate peptide (PDB id: 2C3I).

2. Computational Methods for the Conformational Study of Protein-Protein & Protein-Ligand Systems

2.1. Enhanced Sampling Methods Used in Protein-Related Studies

MD simulation based on classical mechanics has been widely used for sampling conformational space of biomolecules, ranging from small molecules [6], peptides [7], to large proteins [8]. To observe important biological activities, such as enzymatic catalysis or protein-protein association/dissociation, one might need enhanced sampling methods rather than straight unbiased MD due to the rugged free energy surface of complex biological molecules and their extensive timescale of milliseconds to microseconds. Various methods have been proposed to overcome this shortcomings using multiple replicas of the same system [9–14], biasing the underlying potential energy function [15–21], using models that allow greater time-steps [22], and generating kinetic models followed by adaptive sampling [23,24].

Replica-exchange molecular dynamics (REMD) method has been suggested by Sugita and Okamoto [9]. Multiple copies of the system with different temperatures or Hamiltonian are run concurrently and exchanged at fixed interval in this generalized ensemble methodology (Figure 2a). The exchanges are permitted based on the Metropolis criterion, generating a generalized ensemble [25]. Since swapping with neighboring replicas of different temperature or Hamiltonian can enhance the convergence of MD simulation, REMD has been extensively adopted in various ensemble simulations. One could apply this method to many expanded ensembles when proper exchange acceptance criteria are used [26]. The applications of REMD are as follows: replica-exchange umbrella sampling (REUS) [10], temperature REMD [11,12], Hamiltonian REMD [13], free energy perturbation with REMD (FEP/REMD) [27], and replica-exchange constant-pH [14,28,29].

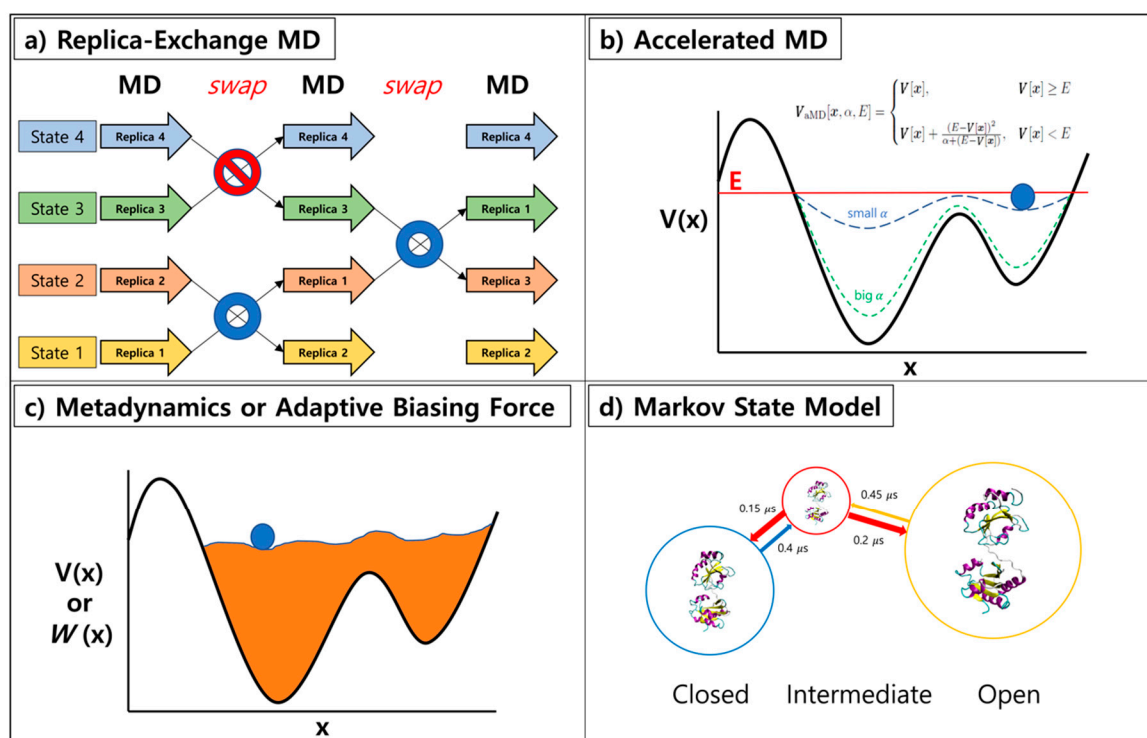


Figure 2. Overview of enhanced sampling methods: (a) replica-exchange MD simulation, (b) accelerated MD, (c) metadynamics or adaptive biasing force, and (d) Markov state model.

Hydrogen mass repartitioning (HMR) allows increment of intrinsic time-step from 2 fs to 4 fs for solvated biomolecular systems. The time-step of normal MD is determined by the fastest motions in the system, which is normally a vibrational motion of hydrogen bond, limited to 2 fs for a stable simulation.

The 4 fs time-step, a factor of 2 speed up, is granted by increasing the mass of hydrogen to about 3 amu and decreasing the equivalent mass from connected heavy atom. Hopkins et al. tested HMR on various benchmark systems, ranging from 3-residue peptide to 129-residue protein, hen egg white lysozyme (HEWL), and concluded there is no significant difference in kinetics and thermodynamics when HMR is used [30]. Balusek et al. assessed the validity of HMR for the membrane simulation, and figured out most structural properties (area-per-lipid, electron density profile, order parameters) being identical, but some kinetic properties such as diffusion constant showed some differences [31].

One can enhance the sampling by running simulations on modified potential energy functions, boosted and smoothened ones. Accelerated Molecular Dynamics (aMD) suggested by McCammon [16] directly modifies the potential energy function to expedite the transition of molecule from the local minimum. Simulating proteins and polypeptide are often accompanied with high potential energy barrier in between distinguished states of the system. When aMD is applied to the system, a robust bias energy is added to the potential energy function according to the set of variables (a threshold energy (E) and a boosting constant (α) to reduce the potential energy barrier (Figure 2b). Smoothened potential energy function of a system allows its exchange rate between various conformational states. Similar type of approaches with different rules for modification to the potential energy functions exist, namely local elevation and conformational flooding [32,33]. For methods above, reweighting of the perturbed free energy surface allows one to recover the original Boltzmann distribution of the system. However, proper selection of the boost potential is of importance as overestimation could invalidate events observed during the trajectory [34]. Nonetheless, usage of this algorithm has been successful on rare-event observations of many complex biological systems [35–37].

Another type of enhanced sampling method embraces continuous modifications of the energy profile during the simulation. Metadynamics (MTD) suggested by Parinello's group modifies potential energy function in a regular basis during the simulation. Addition of potential energy to visited states discourages repeated sampling of these states, hence promoting the inspection of unexplored conformational space (Figure 2c). Adaptive biasing force (ABF) method memorizes the applied force at each point during the simulation and accumulates these forces in bins along the transition coordinate to calculate the average force at each coordinate interval [38]. Flattening of the free energy landscape was subsequently performed by utilizing the potential derived from integrating the average force applied. Both methods are successful when few meaningful collective variables and proper transition coordinate are chosen. Recent work by Blanc et al. highlights the recovery stroke of myosin VI using ABF [39] and the folding mechanism of protein G from streptococcal bacteria has been studied by Granata et al. with metadynamics [40].

The Markov state model (MSM) is a different type of approach, generating a kinetic model from a long unbiased MD trajectory [23,41]. MSM is constructed by building microstates from the initial trajectory, then a transition matrix is calculated from the states with a carefully chosen lag time used to define transition (Figure 2d). The dimension reduction is often performed with principle component analysis (PCA) [42] or time-structure independent component analysis (tICA) [43]. PyEMMA by Noe's group [44] and MSMbuilder by Pande's group [45] are two mainstream open sources for MSM construction. Using MSM, the largely obscure multi-dimensional data is turned into a simplified model readily understood by human beings, and one can utilize this for adaptive sampling. The nature of protein/peptide folding has been readily studied by this approach [46]. MSM can be used to direct sampling to less explored regions, allowing users to run more simulations at these points resulting in an improved model of the system of interest [47].

2.2. Recent Applications of Enhanced Sampling Methods for Protein Complexes

Protein-protein association plays a pivotal role in a myriad of cellular signaling pathways. Misguided protein-protein interactions (PPIs) have been linked to neurodegenerative diseases such as Alzheimer's diseases and Parkinson's disease, and the inhibition of protein dimerization/oligomerization has been proposed as a potential therapeutic treatment for cancer [48–50]. The biological significance of

PPIs is further reflected by the increasing number of peptide-based drugs made available over the years, prompting the need to understand the mechanism governing the binding of peptide ligands to proteins. Numerous computational studies conducted to comprehend protein (peptide)-protein interactions have been conducted and oftentimes, these studies extended beyond traditional MD simulations. Wang et al., in order to precisely model protein-peptide binding, utilized the fast Fourier transform (FFT) based docking method, ClusPro PeptiDock, for the docking of peptide ligands, and refined the resulting protein-peptide complexes using Gaussian aMD (GaMD) simulation [51,52]. The GaMD simulation is analogous to the aMD simulation (vide supra) in that both enable unconstrained enhanced sampling without the need of a predefined collective variable. However, during the GaMD simulation, a boost potential in the form of a Gaussian function is implemented, to ensure more reliable reweighting during the construction of the potential of mean force (PMF) profile [53]. Using the aforementioned *PeptiDock+GaMD* method, Wang et al. successfully predicted the binding poses of three distinct peptides to their partnering proteins, namely, SH3 domain, XLP protein and PIM1 kinase [52]. The resulting protein-peptide complexes concurred well with their respective X-ray crystal structures, with lowest backbone root-mean-square-deviations (rmsds) ranging from 0.6 to 2.7 Å. When refinement was not performed, the protein-peptide complexes acquired using just PeptiDock recorded lowest backbone rmsds in the range of 3.3 to 4.8 Å. The better backbone rmsds obtained when refinement of the protein-peptide complexes was conducted after docking demonstrated the importance of considering protein flexibility and conformational heterogeneity when predicting binding poses of protein-peptide complexes. The same could be extended to protein-ligand and protein-protein complexes.

Besides modelling protein-peptide binding, mechanistic studies delineating protein-protein association and dissociation at the molecular level are also necessary to obtain a comprehensive picture of how native protein-protein contacts are established. Pan et al. introduced a protocol involving the synergistic use of long-timescale MD simulations with their newly developed enhanced sampling method, termed “tempered binding”, to investigate reversible protein-protein association [54]. The “tempered binding simulation” is comparable to simulated tempering, a global optimization method that dynamically changes the temperature of the system following the generalized Metropolis criterion [54,55]. However, instead of temperature, “tempered binding” updates the Hamiltonian function of the system by scaling the strength of interatomic interactions, such as near electrostatic interactions between proteins and between proteins and solvent molecules, through periodic Monte Carlo (MC) moves [54]. The on-the-fly scaling improves sampling efficiency by preventing conformational trapping, thus inducing the dissociation of long-lived states much faster than conventional MD simulation could. This consequently enabled the observation of reversible protein-protein association as demonstrated by the spontaneous association and dissociation of five protein-protein complexes to and from their respective native conformations. Conventional MD simulations performed on the same protein-protein complexes revealed similar spontaneous protein-protein association as well. However, protein-protein dissociation was not observed once the native conformation of the complex was achieved and kinetic trapping of the complex in non-native conformation was also discerned, corroborating the benefits of enhanced sampling methods in the study of reversible protein-protein association.

Even though enhanced sampling methods advocate better sampling efficiency than conventional MD simulations in most cases, selecting the best enhanced sampling technique for your system of interest is also vital to ensure that a reliable free energy profile can be obtained. In a recently published study by Wingbermühle and Schäfer, a systematic comparison of the performance of four enhanced sampling techniques, namely, US, REUS, well-tempered MTD, replica exchange with solute tempering 2 (REST2), and biased exchange umbrella sampling (BEUS), was performed to examine the partial dissociation of the N-terminus of an antigenic peptide from Major Histocompatibility Complex class I (MHC I) protein [56]. Through this study, they concluded that BEUS afforded the most reliable PMF profile of the protein-peptide dissociation, given by the clear-cut segregation of the fully-bound and the partially dissociated states of the protein-peptide complex in the energy profile, as well as the

consistent sampling of conformations along the reaction coordinates (RC). This study also offered useful insights, which could be applied when simulating biological systems that are conflicted in terms of enthalpic and entropic contribution. A comprehensive understanding of the system enables proper selection of RC or collective variables that permits the sampling of entropy-driven events, such as the partially dissociated state. While this could elevate the element of bias during the simulation, BEUS was suggested to have mitigated this effect by using MC updates of RC values, thus providing room for other degrees of freedom to fluctuate.

3. Computational Methods for the Prediction of Protein-Protein (Peptide) and Protein-Ligand Binding

3.1. Free Energy Calculations for Prediction of Protein-Ligand and Protein-Protein Binding Affinity

The prediction of protein-ligand binding affinity has been a main goal in drug discovery field since the magnitudes of the affinity and the selectivity of designed ligands/drugs are crucial in SBDD. Free energy perturbation (FEP) and thermodynamic integration (TI) methods have been suitable for this aspect as both provide a relatively reliable prediction of protein-ligand binding affinity. Both require MD simulations at different values of lambdas ranging from 0 to 1 which represent their perturbation state, 0 being the original state and 1 being the fully perturbed state. Then, applying either Zwanzig's equation for FEP or thermodynamic integration formalism would disclose the associated free energy of perturbation [57,58]. Absolute binding free energy (ABFE) calculation, where total annihilation of the ligand in the binding pocket followed by its reappearance at bulk state where receptor is absent, is extensively used for small ligands. As free energy is a state function, the alchemical FEP route to getting binding free energy of ligand is as follows: (1) 'locking the ligand', restraining conformational, translational, and rotational degrees of freedom at bound state, (2) 'disappearing the locked ligand', turning off the interaction between ligand and its surroundings, (3) 'translocating ligand' of which corresponding free energy is zero, (4) 'reappearance of the locked ligand', turning on the interaction between ligand and its surroundings at bulk state, and (5) 'unlocking the ligand', releasing of the three restraints from (1). Aldeghi and co-workers performed ABFE simulations for a set of diverse inhibitors binding to bromodomain-containing protein 4 (BRD4) for actual pharmaceutical targets and drug-like molecules, and obtained a result giving excellent agreement with experimental data [59]. Lenselink et al. tested binding free energies of congeneric ligands to four different GPCRs using FEP+ from Schrödinger and obtained results in great agreement with experimental results [60–62].

This alchemical route can also be utilized when mutating a key residue in silico, obtaining the relative free energy difference upon mutation, $\Delta\Delta G$ [63]. Two types of topologies can be used for relative FEP where 'disappearing certain group and reappearing counterpart' takes place for "dual topology" [64], and 'perturbation of an existing atom to different type of atom' is done for "single topology" [65]. The difference between the two types of topologies is whether the starting and ending states share a perturbed atom. From methane to ethane perturbation, for example, dual topology will have another dummy methyl group attached to the carbon in methane. For single topology, one of the hydrogens in the methane becomes carbon while three hydrogens, non-interacting dummy atoms connected to the new carbon atom when system was methane, appear accordingly. Relative free energy calculations are cheaper than its absolute counterpart, of course, achieving convergence faster with fewer particles to be perturbed. Loeffler et al. studied the reproducibility of FEP among different MD packages to compare their reliability and to generate the benchmark protocol [66]. The results showed that the relative free energies calculated are equal to the difference in absolute free energies between the two states, hence proving the robustness of the methods and tools available in the field. Gapsys et al. successfully recovered the relative free energy from alchemically mutating residues [67]. The study was conducted towards 762 mutations and the changes in protein thermostability were predicted with remarkable accuracy.

To calculate protein-protein binding free energy, however, the alchemical FEP is not suitable because a simulation of 'annihilating a protein' from a solvated system is essentially impractical. If one

can observe multiple associations/dissociations of protein complex with unbiased MD, constructing MSM upon the trajectory would be advantageous [68]. On the other hand, a PMF method suggested by Woo and Roux is suitable where dissociation is a rare-event due to the large affinity between proteins or kinetic trap [69]. Using this method, one can physically pull one protein apart from another during the simulation with six relative translational and orientational degrees of freedom defined. The user chooses arbitrarily three points in each protein, (P1, P2, and P3) and (L1, L2, and L3), in order to specify the spherical coordinates $(r, \theta, \phi) = (\overline{P1L1}, \angle P1L1L2, \angle P1L1L2L3)$ and three Euler angles $(\Theta, \Phi, \Psi) = (\angle P2P1L1, \angle P2P1L1L2, \angle P3P2P1L1)$ [70]. String method application to this method was recently published from the same group [71]. The method quickly searches for the minimum free energy path (MFEP) along the separation path with implicit solvent model, generating a string of optimal separation path based on 6-dimensional relative position and orientation as defined above. The string is refined a few times with explicit solvent model, and MD simulation along the optimal path converges much faster than straight separation.

While many replicas are required in between the perturbation states for FEP methods and along the separation path for PMF-based and string methods, the computationally cheaper way to estimate binding free energy would be the molecular mechanics combined with Poisson-Boltzmann or generalized Born and surface area solvation (MM/PBSA or MM/GBSA). With the relatively short amount of sampling, one can utilize snapshots from the trajectory to obtain the binding free energy that is estimated by getting the difference between the free energy of the complex and that of each body. The free energies are calculated as $G = E_{MM} + G_{solv} - TS$, where E_{MM} stands for the bonded, electrostatic, and van der Waals energy terms from molecular mechanics, G_{solv} stands for solvation free energy, and TS stands for temperature and entropy respectively. The solvation free energy consists of polar and non-polar, where polar solvation free energy can be estimated using either Poisson-Boltzmann or generalized Born theorem and non-polar solvation free energy is predicted from a linear relation to the solvent accessible surface area (SASA). MM/PBSA and MM/GBSA have been widely used for estimating binding free energies and showed definitively better results when compared to empirical attempts such as docking and virtual screening [62]. Though affordable, unaccounted conformation entropy and free energy of water molecules in binding sites sometimes resulted in poor estimation [72].

3.2. Recent Methods for the Prediction of Binding 'Hotspots' for Protein-Ligand and Protein-Protein Association

The prediction of residues crucial for the formation of favorable protein-ligand or protein-protein interactions is an asset in the field of rational drug design. These decisive sets of amino acids are called 'hotspots' and numerous studies with the objective of quantifying the contribution of residues towards binding affinity have been conducted to facilitate precise identification of 'hotspot' residues [73–77]. Experimentally, the determination of 'hotspots' proceeds through traditional methods such as alanine scanning mutagenesis, structure-activity relationship by NMR and multiple solvent crystal structures (MSCS) [78–80]. However, these experimental approaches are often time-consuming and labor-intensive. Hence, computational tools that alleviate the workload involved in mapping binding pockets and protein-protein interfaces are highly sought to expedite the process of developing ligand/drugs that preferentially bind at these 'hotspots'.

The cosolvent or mixed-solvent MD simulation is one of the most widely used computational tool in SBDD for its accuracy in identifying ligand binding sites [73,81]. This method is the theoretical equivalent of the MSCS method in that it uses small organic probes to sweep protein surfaces to locate binding 'hotspots', which include orthosteric (active site), allosteric and cryptic pockets. Allosteric binding site define a binding pocket that is located away from the active site and can modulate protein activity through allostery. The process of allostery involves the propagation of signals from the allosteric pocket—initiated through ligand binding or mutation of residues—to the active site. This allosteric communication proceeds through a network of residue-residue interactions and/or structural fluctuations [82–86]. Numerous studies on the allosteric regulation of proteins have been

conducted in recent years for its potential application towards the design of novel therapeutic ligands with minimal side effects, as allosteric sites, contrary to orthosteric binding sites, are less conserved and are topographically more varied [87,88]. These studies reflected the importance of considering structural dynamics to acquire a better understanding of allosteric communications [86,89,90]. However, before jumping into the allosteric modulation, it is imperative for researchers to identify and characterize accurately the potential allosteric sites that could be targeted for drug design. Cryptic sites, a subset of allosteric binding sites, are dynamic binding pockets which appearance are typically induced through the binding of orthosteric ligand that effectuates conformational changes via induced-fit mechanism, ligand-induced conformational state transitions or the simultaneous occurrence of both events [91–93]. In recent years, the SBDD of orthosteric ligands has been slightly upstaged by the design of ligands targeting allosteric and cryptic binding sites—potential druggable targets that are gaining attention for its plausible contribution to the development of ligands with better specificity and selectivity with attenuated side effects. This makes cosolvent MD simulation an ideal tool for SBDD as it enables the exploration of the overall druggable space of proteins and could be a solution for therapeutically attractive proteins that previously were deemed ‘undruggable’ targets [94].

As the name suggests, cosolvent MD simulation uses solvent mixtures that usually consist of organic molecules or fragments that possessed key pharmacophore features crucial for ligand binding such as hydrophobic centers, aromatic rings, hydrogen bond donors and hydrogen bond acceptors [81]. There are various successful protocols utilizing cosolvent MD simulation for binding ‘hotspot’ prediction, namely, SWISH, MixMD, MDMix, SILCS, mCSM and many others [93,95–101]. A recent development of SWISH—a Hamiltonian replica-exchange-based method developed to survey hydrophobic patches on proteins for druggability—demonstrated proficiency of using cosolvent probe in precisely identifying cryptic pockets [93,95]. The SWISH with probe approach successfully uncovered the cryptic pocket of the Niemann-Pick Type C2 protein, which was alluded to be challenging as the cryptic pocket is flanked by β -sheets and is highly specific towards sterol-like molecules. The use of SWISH successfully prompted the opening of the cryptic site, allowing the binding of benzene probes (analogous to a hydrophobic/aromatic ring pharmacophore feature) which further exposed the cryptic pocket while simultaneously stabilizing its opened state [95].

CrypticScout, accessible through PlayMolecule platform (www.playmolecule.org), formulated a similar protocol, which engages benzene probes to facilitate the identification of cryptic pockets [91,101,102]. The cryptic sites, after the simulation, were characterized through the solvent accessible surface area (SASA) calculation, ‘hotspot’ overlaps (with reference to apo structure) and scoring functions defined by free energy grids and occupancy. In addition to benzene probes, there are also other organic probes available through CrypticScout, namely, acetone, imidazole, isopropanol, and phenol, a good representation of pharmacophore features (vide supra) important for ligand binding [101]. The development of analytical tools that aid in the analysis of cosolvent MD simulation, such as the cosolvent analysis toolkit [73], further engendered the accessibility of cosolvent MD simulation to both expert and non-experts alike, expediting the exploration of protein’s chemical space for drug design. These studies highlighted the importance of computer simulations to continuously evolve to meet current needs in protein research. MD simulations and enhanced sampling methods are highly productive for studies related to protein folding, conformational change, protein-ligand binding, PPIs and enzymatic reactions; whereby efficient conformational sampling is of the essence. However, in situations such as the discovery of cryptic pockets (vide supra), enhanced sampling may not be sufficient for the unravelling of these obscured binding sites. Instead, augmentation of MD simulations/enhanced sampling methods with organic fragment probes may be necessary to maneuver the exposure of these cryptic binding sites [93,95,101].

Alanine scanning mutagenesis is another experimental technique that has been computerized to accelerate ‘hotspot’ prediction, mostly at protein-protein interface. This method involved the site-directed mutagenesis of specific residues to alanine to assess the contribution of the amino acid towards protein stability and function [103–105]. Computational alanine scanning (CAS), a robust

alternative to its experimental counterpart, permits the rapid scanning of protein surface for amino acid clusters vital for protein-protein association [74]. Over the years, several CAS methods have been developed, centering around free energy calculations performed on single structures or structure ensembles from NMR or MD trajectories to predict 'hotspots' at protein-protein interface [106,107]. PPIs mostly occur at large, shallow regions on the protein surface (Figure 1) and the interactions established are usually transient, making it difficult to distinguish protein-protein binding interface. Therefore, it is crucial to consider protein dynamics to determine residues that contribute significantly towards binding free energies. For instance, the flex ddG method in Rosetta utilized a synergy of global optimization, Monte Carlo sampling with 'backrub' implementation, and advanced energy functions to predict changes in the binding free energy upon mutation ($\Delta\Delta G$) [108]. Rosetta 'backrub' simulation accounts for the conformational flexibility of proteins through the sampling of backbone and side chain rotamers. This method has been demonstrated to reproduce localized structural heterogeneity observed experimentally via high-resolution crystal structures and NMR structures. The utilization of the 'backrub' method in flex ddG to sample degrees of freedom at the mutation site and the neighboring residues permitted accurate depiction of protein flexibility in nature, which translates to better estimation of $\Delta\Delta G$ [109].

Another recently established CAS tool is the BUDE Alanine Scanning (BudeAlaScan) method, which was built upon the empirical free energy approach of a molecular docking algorithm, BUDE (Bristol University Docking Engine) [74,110]. Similar to flex ddG, this method considers conformational plasticity by considering a set of side chain rotamers for charged residues (Asp, Glu, Arg, Lys and His) to approximate the entropy loss resulting from the formation of inter-protein salt bridges between the side chains of the aforementioned charged residues and a fixed protein backbone. Ibarra et al. in introducing the BudeAlaScan method has also simultaneously compared the performance of several CAS methods, including flex ddG, in accurately predicting 'hotspot' residues. The comparable results achieved for flex ddG and BudeAlaScan emphasized the importance of considering conformational heterogeneity to obtain a more precise prediction of $\Delta\Delta G$ [74]. This comparative study also observed quality $\Delta\Delta G$ prediction from mutation Cutoff Scanning Matrix (mCSM), a machine learning approach that was trained based on several curated databases documenting changes in protein stability and protein-protein affinity upon mutation, such as SKEMPI and ProTherm [111–114]. The mCSM method uses structural graph-based signatures to represent the changes in the protein environment upon mutation. The graph-based signature denotes atoms as nodes and interatomic interactions as edges and encodes changes in the physicochemical properties of residues through the difference in pharmacophore count between wild type (WT) and mutated proteins. In the recently refurbished mCSM method, called mCSM-PPI2, an additional six features were utilized to train and test the prediction model namely: (i) the environment of the WT residues such as solvent accessibility, torsional fluctuations, depth of residue, and amino acid content in the neighborhood of the mutated residue, (ii) detection of mutation from or to glycine (flexible side chain) or proline (rigid side chain), (iii) evolutionary conservation, (iv) non-covalent interaction network analysis, (v) energy terms (interaction energies between protein chains and changes in the predicted folding free energy on mutation) and, (vi) atomic fluctuations [112]. All of these features aid in improving the predictive power of mCSM by considering protein flexibility and environment effect, both of which undergo considerable changes when mutations are introduced.

Besides the utilization of CAS in determining 'hotspots' at protein-protein interface, CAS has also been adapted for protein-ligand systems [76,77,115]. Liu et al. presents an efficient protocol that combines MM/GBSA and the interaction entropy (IE) approach developed by Duan et al. to calculate the contribution of specific residues to the overall protein-ligand binding free energy [77,115]. Typically, the inclusion of entropic contribution in end-point free energy calculation such as MM/GBSA and MM/PBSA involved the use of expensive normal mode analysis (NMA). However, the use of NMA in CAS will incur a huge computational cost and therefore is not practical for the assessment of $\Delta\Delta G$ especially when a sizable mutation scan is required. The introduction of the IE approach offers a more rigorous and less expensive computation of binding entropy as it explicitly calculates the

entropic contribution by directly observing the fluctuation in interaction energy between individual residues and ligand from a single MD simulation, hence eliminating the additional computational cost required for entropy calculation [115]. The implementation of this CAS approach in the investigation of the binding of three drugs to anaplastic lymphoma kinase (ALK) and its mutants revealed 'hotspot' residues that were critical for protein-ligand binding. Considerable agreement between the calculated and the experimentally reported total binding free energies of the ALK-ligand systems were also observed [77].

4. Limitations and Improvements in Current Computational Approach

4.1. Limitations and Challenges of Current Computational Methods

Current computational approaches utilizing all-atom MD simulations, despite its fruitful aspects, are still limited due to high computational cost. The technological advances and usage of GPU significantly aid in speeding up the sampling [116]. The benchmark in 2011 showed about 10 ns/day using 24 processors of CPU [117], and the benchmark in 2018 shows a significant boost-up of 330 ns/day with a single GPU [116], both applying a 2 fs time step and considering approximately 25,000 atom system. The nature of taking $O(N \log N)$, where N is the number of particles in the system, for particle mesh Ewald (PME) long range interaction computation, however, is unavoidable. Using the state-of-the-art machines and algorithms of unbiased MD, about 2 milliseconds of sampling was needed to generate a reasonable kinetic model to observe associations and dissociations of two proteins showing femtomolar binding affinities [68]. Considering that ordinary experiments trace macroscopic number of proteins with timescales of tens of microseconds to minutes, there is a lot more room for growth for MD simulation that computes single macromolecule for hundreds of nanoseconds to a few microseconds only.

Another concern for MD simulation would be the validity of the result. An ultimate goal of MD would be to yield the same thermodynamic and kinetic observables, which one would find in experiments, but with more atomistic details. To draw this outcome, one would need sufficient amount of sampling from the correct force field. While more precise force fields are being developed, MD users might want to focus on if the right answer, a converged result, has been drawn out from the simulation or not. Evaluating ergodicity from the trajectory and determining whether convergence has been achieved are difficult problems. Sampling complete 3N-dimensional configuration of complex biomolecular system with N particles is simply impossible [118]. When the transitions between states are slow or rare event exists, insufficient sampling will make ensemble averages very sensitive to the number of slow transition or rare events. Even after ensemble averages seem to have converged, one would never know if there is any other unobserved rare event remaining. Many attempts to justify and gauge the quality of MD result have been made [119–122].

One can argue the convergence of the result using the statistical tools, but it can never be a definitive statement.

4.2. Recent Improvements in MD Simulations and Enhanced Sampling Methods

The predictive capacity of MD simulations lies in the reliability of classical potential energy functions, i.e., force fields, to describe the inter- and intramolecular interactions between particles in a system of interest. Force fields encompassed empirically derived energy terms that approximate torsional potentials and interatomic interactions, namely bonded and non-bonded interactions, in a condensed phase. The establishment of a force field that realistically represents the PES of a system is one of the main concerns in MD simulation that developers try to ameliorate through several alterations such as the inclusion of the polarization effect of proteins and solvent, refinement of torsional parameters, and improvement of water models [123,124]. Of particular interest in force field development efforts is the reparameterization of electrostatic interactions (non-bonded) that are governed by the Coulomb's law equation to characterize interactions between charged atoms [125].

The long-range nature of electrostatic interactions engendered high computational cost. Therefore, to circumvent expensive calculations, traditional force fields have adopted the mean-field treatment of amino acids by assigning fixed partial charges at each atom center. However, the use of fixed point charges overlooked the intrinsic traits of charged atoms to assume lower energy states through the distortion of their electron density cloud and charge redistribution when approached by oppositely charged atoms or in the presence of an external dielectric field. This limitation has been actively rectified over the years through explicit consideration of polarizability, which consequently permitted a more precise electrostatic description of protein atoms in a non-homogeneous natural environment [123,126].

Developers of well-established force fields such as AMBER, AMOEBA, CHARMM, and OPLS customarily utilized three popular methods, *viz.* induced dipole model, the Drude oscillator model, and the fluctuating charge model for straightforward inclusion of electronic polarizability in MD simulations [127–132]. The induced dipole model, used in AMBER and AMOEBA polarizable force fields, generally propagates polarizability by assigning inducible dipoles by means of self-consistent field (SCF) calculations based on fixed charges at atom centers, which simultaneously were allocated monopoles. Mutual dipole-monopole interactions as well as dipole-dipole interactions fluctuate with changes in their immediate surrounding electric field, contributing to the additional polarization potential energy term of the force field [130,133–137]. The AMOEBA force field went a step further by considering multipoles through quadrupole moments to reproduce more precisely the anisotropic molecular polarizabilities of non-bonded interactions. On the other hand, the Drude oscillator model, which was established by CHARMM developers, involves the inclusion of a supplementary charged atom attached to atom centers via a harmonic spring. Its displacement from original coordinates when exposed to external electric field mirrors atomic polarizability [128]. The fluctuating charge model, another polarizability model implemented in CHARMM, depicts atomic polarizability by following the principles of electronegativity equalization, whereby charge distribution proceeds until equal chemical potential is achieved between two polarizable cores [129]. Despite differences in the rendition of atomic polarizability among the three methods highlighted (*vide supra*), a common purpose was to initiate a molecular response towards fluctuations in the surrounding environment during computer simulations, an undertaking that aimed to emulate the intrinsic behavior of particles in nature.

Numerous tinkering has been done to further justify the superiority of polarizable force fields over classical force fields in precisely modelling the physical processes of proteins. Recent improvements in the aforementioned polarizable force fields include the incorporation of screening effects and charge penetration corrections to precisely model the electrostatic energy of short-range interactions, and the consideration of anisotropic polarization of water/solvent molecules to further justify the thesis of polarizability being a many-body effect [127,132,138–140]. Consideration of the electrostatic polarization effect has positive implications on numerous protein-related studies [141–144]. The use of AMOEBA polarizable force field to reproduce the free energy profiles of the permeation of K⁺ and Na⁺ ions through the Gramicidin A channel showed good agreement with experimental observations, in comparison to classical MD simulations [145]. The utilization of Drude polarizable force field has enabled the accurate observation of the mechanism governing the partial activation of a voltage-gated sodium channel. The conformational change of the voltage-sensing domain from resting to the pre-active state and the more precise construction of the free energy profile afforded by the polarizable force field as compared to its non-polarizable counterpart indicated the importance of electrostatic polarization effect in providing a reliable thermodynamic picture of the voltage channel activation [146]. Another recent study exploring the stacking of the aromatic rings of Phe16 and Tyr37 in HIV-1 nucleocapsid protein (NCp7) showed that the explicit treatment of electrostatic polarizability was vital for the accurate reconstruction of the energy profile [147]. Using AMOEBA polarizable force field, the unstacked form of NCp7, characterized by the aromatic rings of Phe16 and Tyr37, was favored and this concurred well with experiment. This is contrary to the simulation performed using a non-polarizable force field, which preferred the stacked conformation—a transient structure according to experiment.

Besides refinement related to polarizability, there are also other improvements, which specifically target certain issues related to current interest in protein research. For instance, the growing focus on PPIs in computational SBDD have unraveled a shortcoming of existing force fields in accurately representing intrinsically disordered proteins (IDPs) or regions (IDRs). IDRs is one of the major contributors towards protein-protein association and current torsion angle potential in existing force fields are not sufficient to precisely depict the flexibility of IDRs. Yu et al. recently developed a non-polarizable, residue-specific force field, called AMBER ff99SBnmr2, to balance the torsional potential to cater to both folded and disordered protein regions. This force field assimilated per-residue dihedral angle modifications, based on experimental information curated from PDB structures of protein fragments (coil library), into the extant ff99SBnmr1—force field that was formulated based on the backbone dihedral potential of ff99SB and optimized using NMR chemical shifts of fully solved proteins [148]. Another recent force field that was developed for protein-protein complexes, and concomitantly achieving optimal description for both folded and disordered regions of a protein, is the *DES-Amber* force field [149]. This force field was developed by reparametrizing dihedral angle parameters, similar to ff99SBnmr2, and optimizing non-bonded interactions through ab initio QM calculations and experimental data. Despite the significant improvement observed for the modelling of protein-protein complexes and individual ordered or disordered proteins, a caveat was added suggesting the necessity of considering polarization effect explicitly for better predictions of protein-protein binding free energies, which in this study showed significant variance from experimental values.

5. Prospective Outlook

Recent improvements in MD simulation accompanying the state-of-the-art sampling methods and their applications reveal the robust potential of MD in the field, and the remaining limitations points out where MD needs to be further developed. The unique insights upon the dynamics given by all-atomistic MD simulations are crucial for drug discovery and elucidation of mechanism for many biological processes. MD is an interdisciplinary output from numerous theoretician and experimentalists standing on the fields of physics, chemistry, mathematics, computer science, and many more. As a recent illustration, the burgeoning field of artificial intelligence has now started aiding in boosting the efficiency and accuracy of MD simulation [150]. Machine learning and deep learning techniques have assisted the development of more accurate force fields [151], enhanced sampling through learning algorithms to find optimal reaction coordinates [152], and analyzed data to figure out the underlying free energy surface of the system [153]. The collaborative work among miscellaneous fields has made MD simulation more powerful, and MD will continue benefiting from the growth of affiliated fields. One can run longer simulation with the same amount of real time with the advancements of technology, get a more accurate result with the development of force fields and sampling methods, and tackle various biological problems with recently unveiled crystal structures. However, a careful setup of MD simulations is required to resolve existing biological problems effectively and accurately, and thorough understanding upon components of MD simulations would be beneficial in the long run.

Author Contributions: Conceptualization, visualization, and writing—original draft preparation, review, and editing, R.L. and D.S.; Supervision, review, and editing, S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Bio and Medical Technology Development Program (NRF-2019M3E5D4065251) funded by the Ministry of Science and ICT (MSIT) and the Ministry of Health and Welfare (MOHW) through the National Research Foundation of Korea (NRF), Medical Research Center (MRC) grant (No. 2018R1A5A2025286), and the Mid-career Researcher Program (NRF-2017R1A2B4010084). It was also supported by RP-Grant 2020 of Ewha Womans University.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Du, X.; Li, Y.; Xia, Y.-L.; Ai, S.-M.; Liang, J.; Sang, P.; Ji, X.-L.; Liu, S.-Q. Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int. J. Mol. Sci.* **2016**, *17*, 144. [[CrossRef](#)] [[PubMed](#)]
2. McCammon, J.A.; Gelin, B.R.; Karplus, M. Dynamics of folded proteins. *Nature* **1977**, *267*, 585–590. [[CrossRef](#)] [[PubMed](#)]
3. Levitt, M.; Huber, R. Molecular dynamics of native protein: II. Analysis and nature of motion. *J. Mol. Biol.* **1983**, *168*, 621–657. [[CrossRef](#)]
4. Karplus, M.; McCammon, J.A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652. [[CrossRef](#)]
5. Jing, Z.; Liu, C.; Cheng, S.Y.; Qi, R.; Walker, B.D.; Piquemal, J.-P.; Ren, P. Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. *Annu. Rev. Biophys.* **2019**, *48*, 371–394. [[CrossRef](#)]
6. Sun, Y.; Kollman, P.A. Conformational sampling and ensemble generation by molecular dynamics simulations: 18-Crown-6 as a test case. *J. Comput. Chem.* **1992**, *13*, 33–40. [[CrossRef](#)]
7. Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W.F.; Mark, A.E. Reversible peptide folding in solution by molecular dynamics simulation¹¹ Edited by R. Huber. *J. Mol. Biol.* **1998**, *280*, 925–932. [[CrossRef](#)]
8. Karplus, M.; Kuriyan, J. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6679–6685. [[CrossRef](#)]
9. Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151. [[CrossRef](#)]
10. Sugita, Y.; Kitao, A.; Okamoto, Y. Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* **2000**, *113*, 6042–6051. [[CrossRef](#)]
11. Earl, D.J.; Deem, M.W. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3910–3916. [[CrossRef](#)] [[PubMed](#)]
12. Wang, L.; Friesner, R.A.; Berne, B.J. Replica exchange with solute scaling: A more efficient version of replica exchange with solute tempering (REST2). *J. Phys. Chem. B* **2011**, *115*, 9431–9438. [[CrossRef](#)] [[PubMed](#)]
13. Fukunishi, H.; Watanabe, O.; Takada, S. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J. Chem. Phys.* **2002**, *116*, 9058–9067. [[CrossRef](#)]
14. Swails, J.M.; York, D.M.; Roitberg, A.E. Constant pH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation. *J. Chem. Theory Comput.* **2014**, *10*, 1341–1352. [[CrossRef](#)] [[PubMed](#)]
15. Kumar, S.; Rosenberg, J.M.; Bouzida, D.; Swendsen, R.H.; Kollman, P.A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011–1021. [[CrossRef](#)]
16. Hamelberg, D.; Mongan, J.; McCammon, J.A. Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919–11929. [[CrossRef](#)] [[PubMed](#)]
17. Pang, Y.T.; Miao, Y.; Wang, Y.; McCammon, J.A. Gaussian Accelerated Molecular Dynamics in NAMD. *J. Chem. Theory Comput.* **2017**, *13*, 9–19. [[CrossRef](#)]
18. Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *Wires Comput. Mol. Sci.* **2011**, *1*, 826–843. [[CrossRef](#)]
19. Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. *Annu. Rev. Phys. Chem.* **2016**, *67*, 159–184. [[CrossRef](#)]
20. Comer, J.; Gumbart, J.C.; Henin, J.; Lelievre, T.; Pohorille, A.; Chipot, C. The adaptive biasing force method: Everything you always wanted to know but were afraid to ask. *J. Phys. Chem. B* **2015**, *119*, 1129–1151. [[CrossRef](#)]
21. Fu, H.; Shao, X.; Chipot, C.; Cai, W. Extended Adaptive Biasing Force Algorithm. An On-the-Fly Implementation for Accurate Free-Energy Calculations. *J. Chem. Theory Comput.* **2016**, *12*, 3506–3513. [[CrossRef](#)] [[PubMed](#)]
22. Pomès, R.; McCammon, J.A. Mass and step length optimization for the calculation of equilibrium properties by molecular dynamics simulation. *Chem. Phys. Lett.* **1990**, *166*, 425–428. [[CrossRef](#)]
23. Pande, V.S.; Beauchamp, K.; Bowman, G.R. Everything you wanted to know about Markov State Models but were afraid to ask. *Methods* **2010**, *52*, 99–105. [[CrossRef](#)] [[PubMed](#)]

24. Chodera, J.D.; Noe, F. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* **2014**, *25*, 135–144. [[CrossRef](#)] [[PubMed](#)]
25. Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092. [[CrossRef](#)]
26. Radak, B.K.; Suh, D.; Roux, B. A generalized linear response framework for expanded ensemble and replica exchange simulations. *J. Chem. Phys.* **2018**, *149*, 072315. [[CrossRef](#)]
27. Jiang, W.; Roux, B. Free Energy Perturbation Hamiltonian Replica-Exchange Molecular Dynamics (FEP/H-REMD) for Absolute Ligand Binding Free Energy Calculations. *J. Chem. Theory Comput.* **2010**, *6*, 2559–2565. [[CrossRef](#)]
28. Stern, H.A. Molecular simulation with variable protonation states at constant pH. *J. Chem. Phys.* **2007**, *126*, 164112. [[CrossRef](#)]
29. Radak, B.K.; Chipot, C.; Suh, D.; Jo, S.; Jiang, W.; Phillips, J.C.; Schulten, K.; Roux, B. Constant-pH Molecular Dynamics Simulations for Large Biomolecular Systems. *J. Chem. Theory Comput.* **2017**, *13*, 5933–5944. [[CrossRef](#)]
30. Hopkins, C.W.; Le Grand, S.; Walker, R.C.; Roitberg, A.E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11*, 1864–1874. [[CrossRef](#)]
31. Balusek, C.; Hwang, H.; Lau, C.H.; Lundquist, K.; Hazel, A.; Pavlova, A.; Lynch, D.L.; Reggio, P.H.; Wang, Y.; Gumbart, J.C. Accelerating Membrane Simulations with Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2019**, *15*, 4673–4686. [[CrossRef](#)] [[PubMed](#)]
32. Huber, T.; Torda, A.E.; van Gunsteren, W.F. Local elevation: A method for improving the searching properties of molecular dynamics simulation. *J. Comput. Aided Mol. Des.* **1994**, *8*, 695–708. [[CrossRef](#)]
33. Grubmüller, H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E* **1995**, *52*, 2893–2906. [[CrossRef](#)] [[PubMed](#)]
34. Miao, Y.; Sinko, W.; Pierce, L.; Bucher, D.; Walker, R.C.; McCammon, J.A. Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation. *J. Chem. Theory Comput.* **2014**, *10*, 2677–2689. [[CrossRef](#)] [[PubMed](#)]
35. Bekker, G.-J.; Fukuda, I.; Higo, J.; Kamiya, N. Mutual population-shift driven antibody-peptide binding elucidated by molecular dynamics simulations. *Sci. Rep.* **2020**, *10*, 1406. [[CrossRef](#)] [[PubMed](#)]
36. Kröger, L.C.; Kopp, W.A.; Döntgen, M.; Leonhard, K. Assessing Statistical Uncertainties of Rare Events in Reactive Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2017**, *13*, 3955–3960. [[CrossRef](#)]
37. Yang, L.; Liu, C.-W.; Shao, Q.; Zhang, J.; Gao, Y.Q. From Thermodynamics to Kinetics: Enhanced Sampling of Rare Events. *Acc. Chem. Res.* **2015**, *48*, 947–955. [[CrossRef](#)]
38. Darve, E.; Rodríguez-Gómez, D.; Pohorille, A. Adaptive biasing force method for scalar and vector free energy calculations. *J. Chem. Phys.* **2008**, *128*, 144120. [[CrossRef](#)]
39. Blanc, F.; Isabet, T.; Benisty, H.; Sweeney, H.L.; Cecchini, M.; Houdusse, A. An intermediate along the recovery stroke of myosin VI revealed by X-ray crystallography and molecular dynamics. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6213–6218. [[CrossRef](#)]
40. Granata, D.; Camilloni, C.; Vendruscolo, M.; Laio, A. Characterization of the free-energy landscapes of proteins by NMR-guided metadynamics. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6817–6822. [[CrossRef](#)]
41. Prinz, J.H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J.D.; Schütte, C.; Noe, F. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* **2011**, *134*, 174105. [[CrossRef](#)] [[PubMed](#)]
42. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **1933**, *24*, 417–441. [[CrossRef](#)]
43. Naritomi, Y.; Fuchigami, S. Slow dynamics in protein fluctuations revealed by time-structure based independent component analysis: The case of domain motions. *J. Chem. Phys.* **2011**, *134*, 065101. [[CrossRef](#)] [[PubMed](#)]
44. Scherer, M.K.; Trendelkamp-Schroer, B.; Paul, F.; Perez-Hernandez, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J.H.; Noe, F. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* **2015**, *11*, 5525–5542. [[CrossRef](#)] [[PubMed](#)]
45. Harrigan, M.P.; Sultan, M.M.; Hernandez, C.X.; Husic, B.E.; Eastman, P.; Schwantes, C.R.; Beauchamp, K.A.; McGibbon, R.T.; Pande, V.S. MSMBuilder: Statistical Models for Biomolecular Dynamics. *Biophys. J.* **2017**, *112*, 10–15. [[CrossRef](#)]

46. Deng, N.J.; Dai, W.; Levy, R.M. How kinetics within the unfolded state affects protein folding: An analysis based on markov state models and an ultra-long MD trajectory. *J. Phys. Chem. B* **2013**, *117*, 12787–12799. [[CrossRef](#)]
47. Bowman, G.R.; Ensign, D.L.; Pande, V.S. Enhanced Modeling via Network Theory: Adaptive Sampling of Markov State Models. *J. Chem. Theory Comput.* **2010**, *6*, 787–794. [[CrossRef](#)]
48. Mao, Y.; Fisher, D.W.; Yang, S.; Keszycki, R.M.; Dong, H. Protein-protein interactions underlying the behavioral and psychological symptoms of dementia (BPSD) and Alzheimer's disease. *PLoS ONE* **2020**, *15*, e0226021. [[CrossRef](#)]
49. Brown, J.; Horrocks, M.H. A sticky situation: Aberrant protein-protein interactions in Parkinson's disease. *Semin. Cell Dev. Biol.* **2020**, *99*, 65–77. [[CrossRef](#)]
50. Albrecht, C.; Appert-Collin, A.; Bagnard, D.; Blaise, S.; Romier-Crouzet, B.; Efremov, R.G.; Sartelet, H.; Duca, L.; Maurice, P.; Bennisroune, A. Transmembrane Peptides as Inhibitors of Protein-Protein Interactions: An Efficient Strategy to Target Cancer Cells? *Front. Oncol.* **2020**, *10*. [[CrossRef](#)]
51. Porter, K.A.; Xia, B.; Beglov, D.; Bohnuud, T.; Alam, N.; Schueler-Furman, O.; Kozakov, D. ClusPro PeptiDock: Efficient global docking of peptide recognition motifs using FFT. *Bioinformatics* **2017**, *33*, 3299–3301. [[CrossRef](#)] [[PubMed](#)]
52. Wang, J.; Alekseenko, A.; Kozakov, D.; Miao, Y. Improved Modeling of Peptide-Protein Binding Through Global Docking and Accelerated Molecular Dynamics Simulations. *Front. Mol. Biosci.* **2019**, *6*, 112. [[CrossRef](#)] [[PubMed](#)]
53. Miao, Y.; Feher, V.A.; McCammon, J.A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J. Chem. Theory Comput.* **2015**, *11*, 3584–3595. [[CrossRef](#)]
54. Pan, A.C.; Jacobson, D.; Yatsenko, K.; Sriharan, D.; Weinreich, T.M.; Shaw, D.E. Atomic-level characterization of protein-protein association. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 4244. [[CrossRef](#)] [[PubMed](#)]
55. Marinari, E.; Parisi, G. Simulated Tempering: A New Monte Carlo Scheme. *Eur. Lett.* **1992**, *19*, 451–458. [[CrossRef](#)]
56. Wingbermhühle, S.; Schäfer, L.V. Capturing the Flexibility of a Protein-Ligand Complex: Binding Free Energies from Different Enhanced Sampling Techniques. *J. Chem. Theory Comput.* **2020**. [[CrossRef](#)]
57. Zwanzig, R.W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420–1426. [[CrossRef](#)]
58. Kirkwood, J.G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3*, 300–313. [[CrossRef](#)]
59. Aldeghi, M.; Heifetz, A.; Bodkin, M.J.; Knapp, S.; Biggin, P.C. Accurate calculation of the absolute free energy of binding for drug molecules. *Chem. Sci.* **2016**, *7*, 207–218. [[CrossRef](#)]
60. Lenselink, E.B.; Louvel, J.; Forti, A.F.; van Veldhoven, J.P.D.; de Vries, H.; Mulder-Krieger, T.; McRobb, F.M.; Negri, A.; Goose, J.; Abel, R.; et al. Predicting Binding Affinities for GPCR Ligands Using Free-Energy Perturbation. *ACS Omega* **2016**, *1*, 293–304. [[CrossRef](#)]
61. Abel, R.; Wang, L.; Harder, E.D.; Berne, B.J.; Friesner, R.A. Advancing Drug Discovery through Enhanced Free Energy Calculations. *Acc. Chem. Res.* **2017**, *50*, 1625–1632. [[CrossRef](#)] [[PubMed](#)]
62. Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lypyan, D.; Robinson, S.; Dahlgren, M.K.; Greenwood, J.; et al. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J. Am. Chem. Soc.* **2015**, *137*, 2695–2703. [[CrossRef](#)] [[PubMed](#)]
63. Lybrand, T.P.; McCammon, J.A.; Wipff, G. Theoretical calculation of relative binding affinity in host-guest systems. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 833–835. [[CrossRef](#)] [[PubMed](#)]
64. Pearlman, D.A. A Comparison of Alternative Approaches to Free Energy Calculations. *J. Phys. Chem.* **1994**, *98*, 1487–1493. [[CrossRef](#)]
65. Boresch, S.; Karplus, M. The Role of Bonded Terms in Free Energy Simulations: 1. Theoretical Analysis. *J. Phys. Chem. A* **1999**, *103*, 103–118. [[CrossRef](#)]
66. Loeffler, H.H.; Bosisio, S.; Duarte Ramos Matos, G.; Suh, D.; Roux, B.; Mobley, D.L.; Michel, J. Reproducibility of Free Energy Calculations across Different Molecular Simulation Software Packages. *J. Chem. Theory Comput.* **2018**, *14*, 5567–5582. [[CrossRef](#)]
67. Gapsys, V.; Michielssens, S.; Seeliger, D.; de Groot, B.L. Accurate and Rigorous Prediction of the Changes in Protein Free Energies in a Large-Scale Mutation Scan. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 7364–7368. [[CrossRef](#)]

68. Plattner, N.; Doerr, S.; De Fabritiis, G.; Noe, F. Complete protein-protein association kinetics in atomic detail revealed by molecular dynamics simulations and Markov modelling. *Nat. Chem.* **2017**, *9*, 1005–1011. [[CrossRef](#)]
69. Woo, H.-J.; Roux, B. Calculation of absolute protein–ligand binding free energy from computer simulations. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6825–6830. [[CrossRef](#)]
70. Gumbart, J.C.; Roux, B.; Chipot, C. Efficient determination of protein-protein standard binding free energies from first principles. *J. Chem. Theory Comput.* **2013**, *9*. [[CrossRef](#)]
71. Suh, D.; Jo, S.; Jiang, W.; Chipot, C.; Roux, B. String Method for Protein-Protein Binding Free-Energy Calculations. *J. Chem. Theory Comput.* **2019**, *15*, 5829–5844. [[CrossRef](#)] [[PubMed](#)]
72. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, *10*, 449–461. [[CrossRef](#)] [[PubMed](#)]
73. Sabanés Zariquiey, F.; de Souza, J.V.; Bronowska, A.K. Cosolvent Analysis Toolkit (CAT): A robust hotspot identification platform for cosolvent simulations of proteins to expand the druggable proteome. *Sci. Rep.* **2019**, *9*, 19118. [[CrossRef](#)] [[PubMed](#)]
74. Ibarra, A.A.; Bartlett, G.J.; Hegedüs, Z.; Dutt, S.; Hobor, F.; Horner, K.A.; Hetherington, K.; Spence, K.; Nelson, A.; Edwards, T.A.; et al. Predicting and Experimentally Validating Hot-Spot Residues at Protein–Protein Interfaces. *ACS Chem. Biol.* **2019**, *14*, 2252–2263. [[CrossRef](#)]
75. Huang, D.; Wen, W.; Liu, X.; Li, Y.; Zhang, J.Z.H. Computational analysis of hot spots and binding mechanism in the PD-1/PD-L1 interaction. *RSC Adv.* **2019**, *9*, 14944–14956. [[CrossRef](#)]
76. Yan, Y.; Yang, M.; Ji, C.G.; Zhang, J.Z.H. Interaction Entropy for Computational Alanine Scanning. *J. Chem. Inf. Model.* **2017**, *57*, 1112–1122. [[CrossRef](#)]
77. Liu, X.; Peng, L.; Zhou, Y.; Zhang, Y.; Zhang, J.Z.H. Computational Alanine Scanning with Interaction Entropy for Protein–Ligand Binding Free Energies. *J. Chem. Theory Comput.* **2018**, *14*, 1772–1780. [[CrossRef](#)]
78. Bogan, A.A.; Thorn, K.S. Anatomy of hot spots in protein interfaces¹ Edited by J. Wells. *J. Mol. Biol.* **1998**, *280*, 1–9. [[CrossRef](#)]
79. Dias, D.M.; Ciulli, A. NMR approaches in structure-based lead discovery: Recent developments and new frontiers for targeting multi-protein complexes. *Prog. Biophys. Mol. Biol.* **2014**, *116*, 101–112. [[CrossRef](#)]
80. Allen, K.N.; Bellamacina, C.R.; Ding, X.; Jeffery, C.J.; Mattos, C.; Petsko, G.A.; Ringe, D. An Experimental Approach to Mapping the Binding Surfaces of Crystalline Proteins. *J. Phys. Chem.* **1996**, *100*, 2605–2611. [[CrossRef](#)]
81. Kimura, S.R.; Hu, H.P.; Ruvinsky, A.M.; Sherman, W.; Favia, A.D. Deciphering Cryptic Binding Sites on Proteins by Mixed-Solvent Molecular Dynamics. *J. Chem. Inf. Model.* **2017**, *57*, 1388–1401. [[CrossRef](#)] [[PubMed](#)]
82. Liu, J.; Nussinov, R. Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLoS Comput. Biol.* **2016**, *12*, e1004966. [[CrossRef](#)] [[PubMed](#)]
83. Guarnera, E.; Berezovsky, I.N. Allosteric sites: Remote control in regulation of protein activity. *Curr. Opin. Struct. Biol.* **2016**, *37*, 1–8. [[CrossRef](#)] [[PubMed](#)]
84. Wodak, S.J.; Paci, E.; Dokholyan, N.V.; Berezovsky, I.N.; Horovitz, A.; Li, J.; Hilser, V.J.; Bahar, I.; Karanicolas, J.; Stock, G.; et al. Allostery in Its Many Disguises: From Theory to Applications. *Structure* **2019**, *27*, 566–578. [[CrossRef](#)] [[PubMed](#)]
85. Guarnera, E.; Berezovsky, I.N. On the perturbation nature of allostery: Sites, mutations, and signal modulation. *Curr. Opin. Struct. Biol.* **2019**, *56*, 18–27. [[CrossRef](#)]
86. Zhang, Y.; Doruker, P.; Kaynak, B.; Zhang, S.; Krieger, J.; Li, H.; Bahar, I. Intrinsic dynamics is evolutionarily optimized to enable allosteric behavior. *Curr. Opin. Struct. Biol.* **2020**, *62*, 14–21. [[CrossRef](#)]
87. Korczynska, M.; Clark, M.J.; Valant, C.; Xu, J.; Moo, E.V.; Albold, S.; Weiss, D.R.; Torosyan, H.; Huang, W.; Kruse, A.C.; et al. Structure-based discovery of selective positive allosteric modulators of antagonists for the M2 muscarinic acetylcholine receptor. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E2419. [[CrossRef](#)]
88. Khajehali, E.; Valant, C.; Jörg, M.; Tobin, A.B.; Conn, P.J.; Lindsley, C.W.; Sexton, P.M.; Scammells, P.J.; Christopoulos, A. Probing the binding site of novel selective positive allosteric modulators at the M1 muscarinic acetylcholine receptor. *Biochem. Pharm.* **2018**, *154*, 243–254. [[CrossRef](#)]
89. Lückmann, M.; Trauelsen, M.; Bentsen, M.A.; Nissen, T.A.D.; Martins, J.; Fallah, Z.; Nygaard, M.M.; Papaleo, E.; Lindorff-Larsen, K.; Schwartz, T.W.; et al. Molecular dynamics-guided discovery of an ago-allosteric modulator for GPR40/FFAR1. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 7123. [[CrossRef](#)]

90. VanSchouwen, B.; Melacini, G. Cracking the allosteric code of NMR chemical shifts. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9407. [[CrossRef](#)]
91. Beglov, D.; Hall, D.R.; Wakefield, A.E.; Luo, L.; Allen, K.N.; Kozakov, D.; Whitty, A.; Vajda, S. Exploring the structural origins of cryptic sites on proteins. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E3416. [[CrossRef](#)] [[PubMed](#)]
92. Hollingsworth, S.A.; Kelly, B.; Valant, C.; Michaelis, J.A.; Mastromihalis, O.; Thompson, G.; Venkatakrishnan, A.J.; Hertig, S.; Scammells, P.J.; Sexton, P.M.; et al. Cryptic pocket formation underlies allosteric modulator selectivity at muscarinic GPCRs. *Nat. Commun.* **2019**, *10*, 3289. [[CrossRef](#)] [[PubMed](#)]
93. Oleinikovas, V.; Saladino, G.; Cossins, B.P.; Gervasio, F.L. Understanding Cryptic Pocket Formation in Protein Targets by Enhanced Sampling Simulations. *J. Am. Chem. Soc.* **2016**, *138*, 14257–14263. [[CrossRef](#)] [[PubMed](#)]
94. Hart, K.M.; Moeder, K.E.; Ho, C.M.W.; Zimmerman, M.I.; Frederick, T.E.; Bowman, G.R. Designing small molecules to target cryptic pockets yields both positive and negative allosteric modulators. *PLoS ONE* **2017**, *12*, e0178678. [[CrossRef](#)] [[PubMed](#)]
95. Comitani, F.; Gervasio, F.L. Exploring Cryptic Pockets Formation in Targets of Pharmaceutical Interest with SWISH. *J. Chem. Theory Comput.* **2018**, *14*, 3321–3331. [[CrossRef](#)]
96. Ghanakota, P.; Carlson, H.A. Moving Beyond Active-Site Detection: MixMD Applied to Allosteric Systems. *J. Phys. Chem. B* **2016**, *120*, 8685–8695. [[CrossRef](#)]
97. Graham, S.E.; Leja, N.; Carlson, H.A. MixMD Probeview: Robust Binding Site Prediction from Cosolvent Simulations. *J. Chem. Inf. Model.* **2018**, *58*, 1426–1433. [[CrossRef](#)]
98. Alvarez-Garcia, D.; Barril, X. Molecular Simulations with Solvent Competition Quantify Water Displaceability and Provide Accurate Interaction Maps of Protein Binding Sites. *J. Med. Chem.* **2014**, *57*, 8530–8539. [[CrossRef](#)]
99. Faller, C.E.; Raman, E.P.; MacKerell, A.D., Jr.; Guvench, O. Site Identification by Ligand Competitive Saturation (SILCS) simulations for fragment-based drug design. *Methods Mol. Biol.* **2015**, *1289*, 75–87. [[CrossRef](#)]
100. Fischer, A.; Smieško, M. Allosteric Binding Sites On Nuclear Receptors: Focus On Drug Efficacy and Selectivity. *Int. J. Mol. Sci.* **2020**, *21*, 534. [[CrossRef](#)]
101. Martinez-Rosell, G.; Lovera, S.; Sands, Z.A.; De Fabritiis, G. PlayMolecule CrypticScout: Predicting Protein Cryptic Sites Using Mixed-Solvent Molecular Simulations. *J. Chem. Inf. Model.* **2020**, *60*, 2314–2324. [[CrossRef](#)] [[PubMed](#)]
102. Buch, I.; Harvey, M.J.; Giorgino, T.; Anderson, D.P.; De Fabritiis, G. High-Throughput All-Atom Molecular Dynamics Simulations Using Distributed Computing. *J. Chem. Inf. Model.* **2010**, *50*, 397–403. [[CrossRef](#)] [[PubMed](#)]
103. Lefèvre, F.; Rémy, M.-H.; Masson, J.-M. Alanine-stretch scanning mutagenesis: A simple and efficient method to probe protein structure and function. *Nucleic Acids Res.* **1997**, *25*, 447–448. [[CrossRef](#)] [[PubMed](#)]
104. Ning, J.; Li, R.; Ren, J.; Zhangsun, D.; Zhu, X.; Wu, Y.; Luo, S. Alanine-Scanning Mutagenesis of α -Conotoxin GI Reveals the Residues Crucial for Activity at the Muscle Acetylcholine Receptor. *Mar. Drugs* **2018**, *16*, 507. [[CrossRef](#)]
105. Zhang, Y.; Raudah, S.; Teo, H.; Teo, G.W.S.; Fan, R.; Sun, X.; Orner, B.P. Alanine-shaving Mutagenesis to Determine Key Interfacial Residues Governing the Assembly of a Nano-cage Maxi-ferritin. *J. Biol. Chem.* **2010**, *285*, 12078–12086. [[CrossRef](#)] [[PubMed](#)]
106. Boukharta, L.; Gutiérrez-de-Terán, H.; Aqvist, J. Computational prediction of alanine scanning and ligand binding energetics in G-protein coupled receptors. *PLoS Comput. Biol.* **2014**, *10*, e1003585. [[CrossRef](#)] [[PubMed](#)]
107. Qiu, L.; Yan, Y.; Sun, Z.; Song, J.; Zhang, J.Z.H. Interaction entropy for computational alanine scanning in protein–protein binding. *Wires Comput. Mol. Sci.* **2018**, *8*, e1342. [[CrossRef](#)]
108. Barlow, K.A.; Ó Conchúir, S.; Thompson, S.; Suresh, P.; Lucas, J.E.; Heinonen, M.; Kortemme, T. Flex ddG: Rosetta Ensemble-Based Estimation of Changes in Protein–Protein Binding Affinity upon Mutation. *J. Phys. Chem. B* **2018**, *122*, 5389–5399. [[CrossRef](#)]
109. Smith, C.A.; Kortemme, T. Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. *J. Mol. Biol.* **2008**, *380*, 742–756. [[CrossRef](#)]
110. McIntosh-Smith, S.; Price, J.; Sessions, R.B.; Ibarra, A.A. High performance in silico virtual drug screening on many-core processors. *Int. J. High. Perform. Comput. Appl.* **2015**, *29*, 119–134. [[CrossRef](#)]

111. Pires, D.E.V.; Ascher, D.B.; Blundell, T.L. mCSM: Predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* **2013**, *30*, 335–342. [[CrossRef](#)] [[PubMed](#)]
112. Rodrigues, C.H.M.; Myung, Y.; Pires, D.E.V.; Ascher, D.B. mCSM-PPI2: Predicting the effects of mutations on protein–protein interactions. *Nucleic Acids Res.* **2019**, *47*, W338–W344. [[CrossRef](#)] [[PubMed](#)]
113. Jankauskaitė, J.; Jiménez-García, B.; Dapkūnas, J.; Fernández-Recio, J.; Moal, I.H. SKEMPI 2.0: An updated benchmark of changes in protein–protein binding energy, kinetics and thermodynamics upon mutation. *Bioinformatics* **2018**, *35*, 462–469. [[CrossRef](#)] [[PubMed](#)]
114. Kumar, M.D.; Bava, K.A.; Gromiha, M.M.; Prabaharan, P.; Kitajima, K.; Uedaira, H.; Sarai, A. ProTherm and ProNIT: Thermodynamic databases for proteins and protein-nucleic acid interactions. *Nucleic Acids Res.* **2006**, *34*, D204–D206. [[CrossRef](#)] [[PubMed](#)]
115. Duan, L.; Liu, X.; Zhang, J.Z.H. Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein–Ligand Binding Free Energy. *J. Am. Chem. Soc.* **2016**, *138*, 5722–5728. [[CrossRef](#)] [[PubMed](#)]
116. Lee, T.S.; Cerutti, D.S.; Mermelstein, D.; Lin, C.; LeGrand, S.; Giese, T.J.; Roitberg, A.; Case, D.A.; Walker, R.C.; York, D.M. GPU-Accelerated Molecular Dynamics and Free Energy Methods in Amber18: Performance Enhancements and New Features. *J. Chem. Inf. Model.* **2018**, *58*, 2043–2050. [[CrossRef](#)]
117. Durrant, J.D.; McCammon, J.A. Molecular dynamics simulations and drug discovery. *BMC Biol.* **2011**, *9*, 71. [[CrossRef](#)]
118. van Gunsteren, W.F.; Daura, X.; Hansen, N.; Mark, A.E.; Oostenbrink, C.; Riniker, S.; Smith, L.J. Validation of Molecular Simulation: An Overview of Issues. *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 884–902. [[CrossRef](#)]
119. Singhal, N.; Pande, V.S. Error analysis and efficient sampling in Markovian state models for molecular dynamics. *J. Chem. Phys.* **2005**, *123*, 204909. [[CrossRef](#)]
120. Pastor, R.W.; Brooks, B.R.; Szabo, A. An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* **1988**, *65*, 1409–1419. [[CrossRef](#)]
121. Zhang, W.; Wu, C.; Duan, Y. Convergence of replica exchange molecular dynamics. *J. Chem. Phys.* **2005**, *123*, 154105. [[CrossRef](#)] [[PubMed](#)]
122. Torshin, I.Y.; Namiot, V.A.; Esipova, N.G.; Tumanyan, V.G. Numeric analysis of reversibility of classic movement equations and constructive criteria of estimating quality of molecular dynamic simulations. *J. Biomol. Struct. Dyn.* **2020**, 1–11. [[CrossRef](#)] [[PubMed](#)]
123. Melcr, J.; Piquemal, J.-P. Accurate Biomolecular Simulations Account for Electronic Polarization. *Front. Mol. Biosci.* **2019**, *6*. [[CrossRef](#)] [[PubMed](#)]
124. Gao, Y.; Li, Y.; Mou, L.; Lin, B.; Zhang, J.Z.H.; Mei, Y. Correct folding of an α -helix and a β -hairpin using a polarized 2D torsional potential. *Sci. Rep.* **2015**, *5*, 10359. [[CrossRef](#)]
125. Cisneros, G.A.; Babin, V.; Sagui, C. Electrostatics Interactions in Classical Simulations. In *Biomolecular Simulations: Methods and Protocols*; Monticelli, L., Salonen, E., Eds.; Humana Press: Totowa, NJ, USA, 2013; pp. 243–270.
126. Inakollu, V.S.S.; Geerke, D.P.; Rowley, C.N.; Yu, H. Polarizable force fields: What do they add in biomolecular simulations? *Curr. Opin. Struct. Biol.* **2020**, *61*, 182–190. [[CrossRef](#)]
127. Wang, J.; Cieplak, P.; Luo, R.; Duan, Y. Development of Polarizable Gaussian Model for Molecular Mechanical Calculations I: Atomic Polarizability Parameterization To Reproduce ab Initio Anisotropy. *J. Chem. Theory Comput.* **2019**, *15*, 1146–1158. [[CrossRef](#)]
128. Lemkul, J.A.; Huang, J.; Roux, B.; MacKerell, A.D. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **2016**, *116*, 4983–5013. [[CrossRef](#)]
129. Patel, S.; Brooks, C.L., 3rd. CHARMM fluctuating charge force field for proteins: I parameterization and application to bulk organic liquid simulations. *J. Comput. Chem.* **2004**, *25*, 1–15. [[CrossRef](#)]
130. Ponder, J.W.; Wu, C.; Ren, P.; Pande, V.S.; Chodera, J.D.; Schnieders, M.J.; Haque, I.; Mobley, D.L.; Lambrecht, D.S.; DiStasio, R.A.; et al. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549–2564. [[CrossRef](#)]
131. Liu, Z.; Timmermann, J.; Reuter, K.; Scheurer, C. Benchmarks and Dielectric Constants for Reparameterized OPLS and Polarizable Force Field Models of Chlorinated Hydrocarbons. *J. Phys. Chem. B* **2018**, *122*, 770–779. [[CrossRef](#)]

132. Das, A.K.; Demerdash, O.N.; Head-Gordon, T. Improvements to the AMOEBA Force Field by Introducing Anisotropic Atomic Polarizability of the Water Molecule. *J. Chem. Theory Comput.* **2018**, *14*, 6722–6733. [[CrossRef](#)] [[PubMed](#)]
133. Cieplak, P.; Caldwell, J.; Kollman, P. Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: Aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases. *J. Comput. Chem.* **2001**, *22*, 1048–1057. [[CrossRef](#)]
134. Wang, Z.-X.; Zhang, W.; Wu, C.; Lei, H.; Cieplak, P.; Duan, Y. Strike a balance: Optimization of backbone torsion parameters of AMBER polarizable force field for simulations of proteins and peptides. *J. Comput. Chem.* **2006**, *27*, 781–790. [[CrossRef](#)] [[PubMed](#)]
135. Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J.W.; Ren, P. Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046–4063. [[CrossRef](#)]
136. Wang, J.; Cieplak, P.; Li, J.; Cai, Q.; Hsieh, M.-J.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations. 4. van der Waals Parametrization. *J. Phys. Chem. B* **2012**, *116*, 7088–7101. [[CrossRef](#)]
137. Wang, J.; Cieplak, P.; Cai, Q.; Hsieh, M.-J.; Wang, J.; Duan, Y.; Luo, R. Development of polarizable models for molecular mechanical calculations. 3. Polarizable water models conforming to Thole polarization screening schemes. *J. Phys. Chem. B* **2012**, *116*, 7999–8008. [[CrossRef](#)]
138. Rackers, J.A.; Wang, Q.; Liu, C.; Piquemal, J.-P.; Ren, P.; Ponder, J.W. An optimized charge penetration model for use with the AMOEBA force field. *Phys. Chem.* **2017**, *19*, 276–291. [[CrossRef](#)]
139. Liu, C.; Qi, R.; Wang, Q.; Piquemal, J.P.; Ren, P. Capturing Many-Body Interactions with Classical Dipole Induction Models. *J. Chem. Theory Comput.* **2017**, *13*, 2751–2761. [[CrossRef](#)]
140. Bedrov, D.; Piquemal, J.-P.; Borodin, O.; MacKerell, A.D.; Roux, B.; Schröder, C. Molecular Dynamics Simulations of Ionic Liquids and Electrolytes Using Polarizable Force Fields. *Chem. Rev.* **2019**, *119*, 7940–7995. [[CrossRef](#)]
141. Zhang, D.; Lazim, R.; Yip, Y.M. Incorporating Polarizability of Backbone Hydrogen Bonds Improved Folding of Short α -Helical Peptides. *Biophys. J.* **2019**, *117*, 2079–2086. [[CrossRef](#)]
142. Duan, L.; Feng, G.; Wang, X.; Wang, L.; Zhang, Q. Effect of electrostatic polarization and bridging water on CDK2–ligand binding affinities calculated using a highly efficient interaction entropy method. *Phys. Chem.* **2017**, *19*, 10140–10152. [[CrossRef](#)] [[PubMed](#)]
143. Wang, A.; Zhang, Z.; Li, G. Higher Accuracy Achieved in the Simulations of Protein Structure Refinement, Protein Folding, and Intrinsically Disordered Proteins Using Polarizable Force Fields. *J. Phys. Chem. Lett.* **2018**, *9*, 7110–7116. [[CrossRef](#)] [[PubMed](#)]
144. Hazel, A.J.; Walters, E.T.; Rowley, C.N.; Gumbart, J.C. Folding free energy landscapes of β -sheets with non-polarizable and polarizable CHARMM force fields. *J. Chem. Phys.* **2018**, *149*, 072317. [[CrossRef](#)] [[PubMed](#)]
145. Peng, X.; Zhang, Y.; Chu, H.; Li, Y.; Zhang, D.; Cao, L.; Li, G. Accurate Evaluation of Ion Conductivity of the Gramicidin A Channel Using a Polarizable Force Field without Any Corrections. *J. Chem. Theory Comput.* **2016**, *12*, 2973–2982. [[CrossRef](#)] [[PubMed](#)]
146. Sun, R.-N.; Gong, H. Simulating the Activation of Voltage Sensing Domain for a Voltage-Gated Sodium Channel Using Polarizable Force Field. *J. Phys. Chem. Lett.* **2017**, *8*, 901–908. [[CrossRef](#)]
147. El Khoury, L.; Célerse, F.; Lagardère, L.; Jolly, L.-H.; Derat, E.; Hobaika, Z.; Maroun, R.G.; Ren, P.; Bouaziz, S.; Gresh, N.; et al. Reconciling NMR Structures of the HIV-1 Nucleocapsid Protein NCp7 Using Extensive Polarizable Force Field Free-Energy Simulations. *J. Chem. Theory Comput.* **2020**, *16*, 2013–2020. [[CrossRef](#)]
148. Yu, L.; Li, D.-W.; Brüschweiler, R. Balanced Amino-Acid-Specific Molecular Dynamics Force Field for the Realistic Simulation of Both Folded and Disordered Proteins. *J. Chem. Theory Comput.* **2020**, *16*, 1311–1318. [[CrossRef](#)]
149. Piana, S.; Robustelli, P.; Tan, D.; Chen, S.; Shaw, D.E. Development of a Force Field for the Simulation of Single-Chain Proteins and Protein–Protein Complexes. *J. Chem. Theory Comput.* **2020**, *16*, 2494–2507. [[CrossRef](#)]
150. Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C. Machine Learning for Molecular Simulation. *Annu. Rev. Phys. Chem.* **2020**, *71*, 361–390. [[CrossRef](#)]

151. Chmiela, S.; Sauceda, H.E.; Müller, K.-R.; Tkatchenko, A. Towards exact molecular dynamics simulations with machine-learned force fields. *Nat. Commun.* **2018**, *9*, 3887. [[CrossRef](#)]
152. Wehmeyer, C.; Noé, F. Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics. *J. Chem. Phys.* **2018**, *148*, 241703. [[CrossRef](#)] [[PubMed](#)]
153. Schneider, E.; Dai, L.; Topper, R.Q.; Drechsel-Grau, C.; Tuckerman, M.E. Stochastic Neural Network Approach for Learning High-Dimensional Free Energy Surfaces. *Phys. Rev. Lett.* **2017**, *119*, 150601. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).