

Review

Investigation of Structural Dynamics of Enzymes and Protonation States of Substrates Using Computational Tools

Chia-en A. Chang ^{*}, Yu-ming M. Huang, Leonard J. Mueller and Wanli You

Department of Chemistry, University of California, Riverside, CA 92521, USA;

yu-ming.huang@email.ucr.edu (Y.M.H.); leonard.mueller@ucr.edu (L.J.M.); wyou001@ucr.edu (W.Y.)

^{*} Correspondence: chiaenc@ucr.edu; Tel.: +1-951-827-7263

Academic Editor: David D. Boehr

Received: 12 April 2016; Accepted: 23 May 2016; Published: 31 May 2016

Abstract: This review discusses the use of molecular modeling tools, together with existing experimental findings, to provide a complete atomic-level description of enzyme dynamics and function. We focus on functionally relevant conformational dynamics of enzymes and the protonation states of substrates. The conformational fluctuations of enzymes usually play a crucial role in substrate recognition and catalysis. Protein dynamics can be altered by a tiny change in a molecular system such as different protonation states of various intermediates or by a significant perturbation such as a ligand association. Here we review recent advances in applying atomistic molecular dynamics (MD) simulations to investigate allosteric and network regulation of tryptophan synthase (TRPS) and protonation states of its intermediates and catalysis. In addition, we review studies using quantum mechanics/molecular mechanics (QM/MM) methods to investigate the protonation states of catalytic residues of β -Ketoacyl ACP synthase I (KasA). We also discuss modeling of large-scale protein motions for HIV-1 protease with coarse-grained Brownian dynamics (BD) simulations.

Keywords: force field; calculation; energy; substrate binding

1. Introduction

Conformational changes of enzymes are often related to regulating and creating an optimal environment for efficient chemical catalysis. The dynamics and conformational changes in enzymes may range from the fluctuation of side chains to large-scale protein motions. The former helps adjusting the environment for chemical reactions and the latter enables the protein to create a tertiary structure for substrate binding. The internal motions of proteins may serve as a “gate” in some systems, such as HIV-1 protease, which controls ligand–protein association [1]. In many enzymes or enzyme complexes, protein motions involve allosteric communication to coordinate the function and reactions, which may be intrinsic to many enzymes [2,3]. Residues in the active site directly involved in chemical catalysis may even change allosteric networks in different states during catalysis; example states are a ligand-free resting state and a substrate-bound working state.

In addition to conformational rearrangements of protein side-chains and substrates, protonation states of substrates and key catalytically important residues play critical roles in chemical reactions. For β -Ketoacyl ACP synthase I (KasA), the enzyme that involved in fatty acid synthesis, changes in protonation states on key catalytic residues at different steps are essential in enzyme activity [4,5]. The protonation states may significantly affect protein motions as well. With the cofactor pyridoxal-5'-phosphate (PLP), the bioactive form of vitamin B6, the protonation states of different sites on the coenzyme, such as the phenolic oxygen and PLP ring nitrogen in particular, are thought to be critical for establishing the efficient reaction pathway [6–9]. Moreover, changing a single proton location may interrupt the overall protein–substrate stability and restrain enzyme catalysis.

In this review, we discuss three well-known model systems, tryptophan synthase (TRPS) in Figure 1, β -Ketoacyl ACP synthase I (KasA) in Figure 2, and HIV-1 protease in Figure 3. The systems have been studied extensively for understanding large-scale protein motions, allosteric regulations and/or protonation states of key residues and ligands.

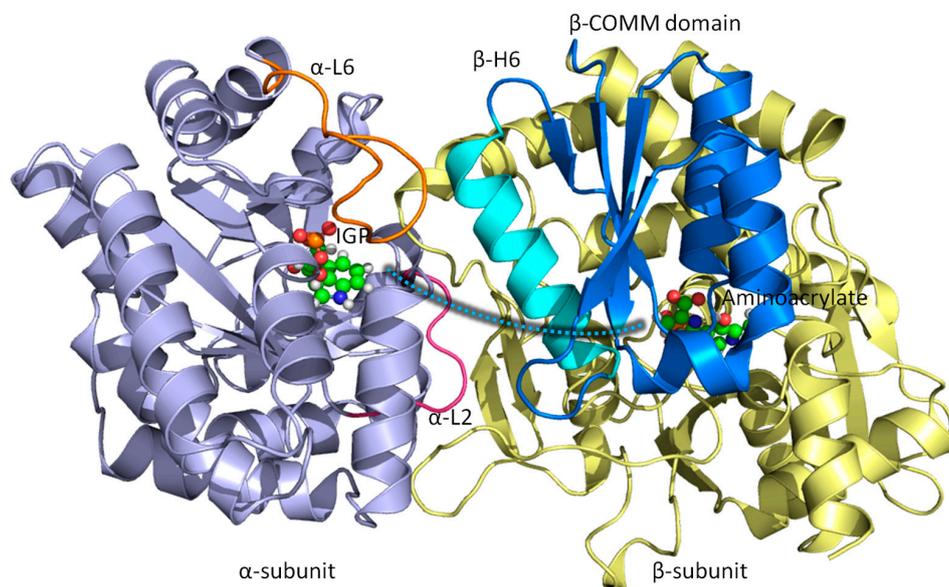


Figure 1. Overall structure and chemical reactions of tryptophan synthase (TRPS). TRPS is composed of an α - (purple) and β -subunit (yellow). The two ligands binding to each subunit are shown in bead representation. The open, partially closed, and fully closed conformations of the α -subunit are controlled by α -L2 (pink), α -L6 (orange), and β -Helix-6 of the communication domain (COMM) domain (cyan). The tunnel used to channel indole from the α -site to the β -site is marked as a cyan dashed shaded line.

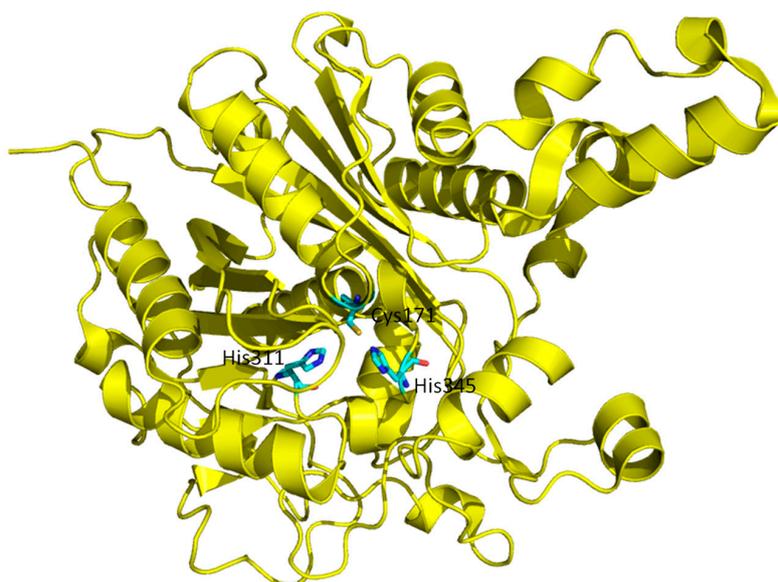


Figure 2. Overall structure of β -Ketoacyl ACP synthase I (KasA). Three residues in the active site, Cys171, His311 and His345, important for catalysis are highlighted in bond representation.

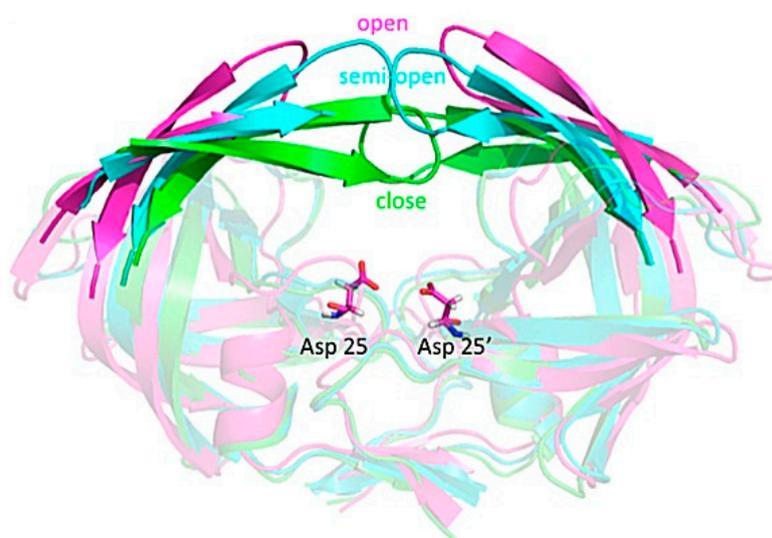
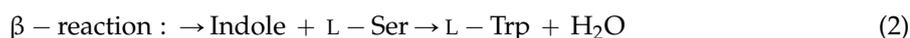


Figure 3. Structure of HIV-1 protease. Pink, cyan, and green indicate open, semi-open and closed flap conformation, respectively. The residues Asp-25 and Asp-25', highlighted in bond representation, are associated with chemical catalysis and their protonation states have been investigated.

Bacterial TRPS is a bifunctional tetrameric $\alpha\beta\beta\alpha$ enzyme complex that catalyzes the last two reactions in the biosynthesis of L-tryptophan (L-Trp). In its α -subunit, the enzyme catalyzes the α -reaction, which involves cleavage of indole-3-glycerol phosphate (IGP) to indole and glyceraldehydes-3-phosphate (G3P) [10]. Indole then diffuses in the channel within the enzyme complex and is condensed with serine at the β active site (β -reaction) (Figure 4). This pathway involves at least nine different intermediates—E(Ain), E(GD₁), E(Aex₁), E(Q₁), E(A-A), E(Q₂), E(Q₃), E(Aex₂), and E(GD₂)—formed from PLP and the reacting substrates (Figure 4) [11,12]. The $\alpha\beta$ -reaction involves the following combination:



TRPS complexes with E(Q₃) bound to the β -site; the complex readily reacts and is difficult to crystallize, so substrate analogues equivalent to the intermediates formed along the β -reaction pathway were widely studied. For example, the indole analogues indoline and 2-aminophenol (2AP) can rapidly react with E(A-A) to produce the long-lived indoline quinonoid E(Q)_{indoline}, and 2AP quinonoid E(Q)_{2AP}, respectively [13,14]. TRPS is a good model system for investigating PLP-dependent enzymatic mechanisms and for studying substrate channeling and allosteric networks that regulate protein functions.

β -Ketoacyl ACP synthase I (KasA) is a key enzyme in the survival of *Mycobacterium tuberculosis* (*M. tuberculosis*), the causative pathogen of tuberculosis [15]. It has been proven that KasA is important in the biosynthesis of mycolic acid, which is one of the building blocks of the cell wall in *M. tuberculosis*. Therefore, KasA is a promising drug target in tuberculosis treatment. The catalytic cycle of enzyme mechanism consists of three steps, acylation, decarboxylation and condensation (Figure 5). First, Cys171 of the active site attacks and binds with acetyl chain delivered by acyl carrier protein (ACP). ACP is subsequently eliminated, and His311 is deprotonated in the acylation step. Decarboxylation of another ACP in the active site generates an enolate, which is stabilized by His311 and His345. The enolate tautomerizes to a carbanion that attacks the carbonyl carbon of acylated Cys171, which finally forms an elongated acyl chain. However, the enzyme mechanism has been debated and discussed [5], due to the multiple possible protonation states of its catalytic residues, Cys171, His311 and His 345.

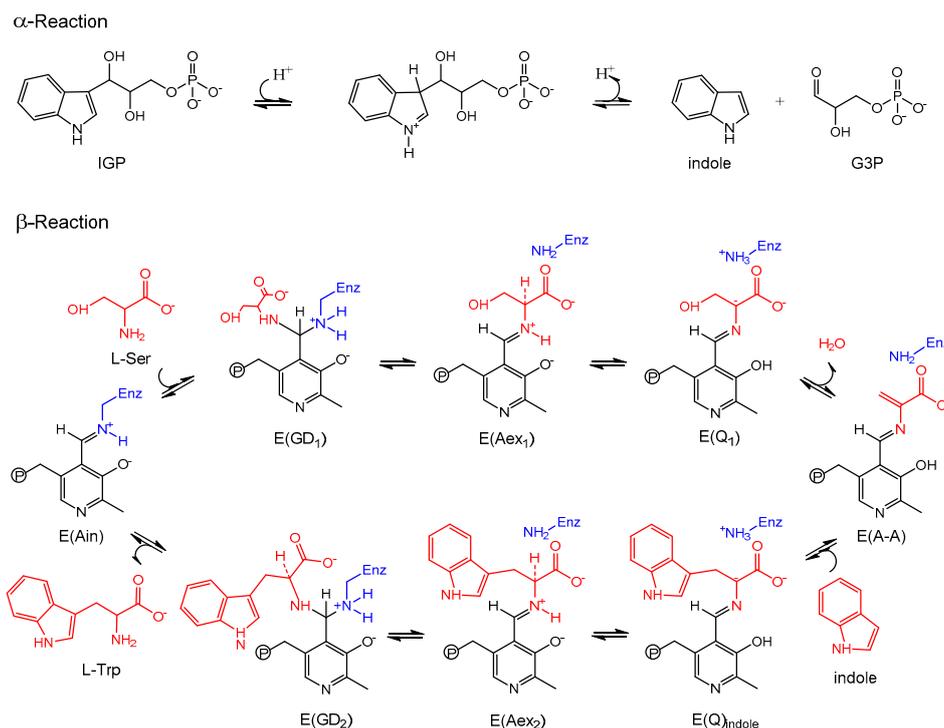


Figure 4. Overall chemical reactions in TRPS.

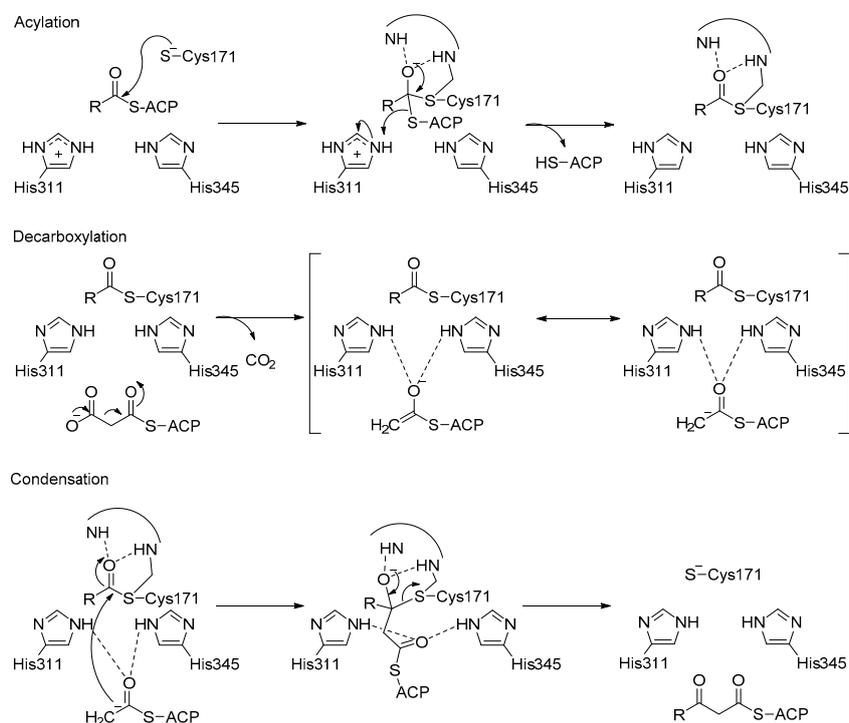


Figure 5. Catalytic mechanism of KasA.

HIV-1 protease is a dimeric aspartic protease that cleaves the premature polypeptide of HIV-1 and plays an essential role in viral replication [16–18]. Inhibitors targeting HIV-1 protease have been developed and represent among the major antiretroviral therapies for AIDS. Their development is considered a milestone success of structure-based drug design [19]. The motion in the flaps

of HIV-1 protease plays a critical role in substrate binding and catalysis. Because of the intrinsic flexibility of HIV-1 protease, the free enzyme contains various flap conformations, usually termed semi-open, open, and closed [20,21]. Most apo structures show a semi-open flap conformation; however, the flaps must open for binding to a natural substrate. The protonation states of the two catalytically important aspartates, Asp-25 and Asp-25', have been investigated to understand the catalytic mechanism and inhibitor recognition [22]. Because hydrogen atoms of Asp-25 and Asp-25' can contribute to an important hydrogen bond, the protonation states of the residues need to be considered in structure-based drug design.

Recently developed molecular modeling tools have advanced our knowledge of the atomistic details of enzyme dynamics and catalysis. Through combining experimental techniques such as NMR, single molecule spectroscopy and computational tools, one can not only unveil the protonation states of catalytically significant intermediates of enzymes, which serves as a key component to reveal the catalytic mechanism, but also examine large-amplitude conformational movements of proteins in greater details. In this review, we discuss the applications of molecular dynamics (MD) simulations to investigate atomistic information of TRPS, MD and multi-scale quantum mechanics/molecular mechanics (QM/MM) methods to investigate protonation states of catalytic key residues of TRPS and KasA, and Brownian dynamics (BD) simulations to model large-scale conformational changes in HIV-1 protease for substrate recognition.

2. Computational Tools

2.1. Atomistic Molecular Dynamics Simulations

MD simulations, first developed over 30 years ago [23,24], have advanced from a method to simulate movements of several hundreds of atoms to a widely used way to study the structure and dynamics of macromolecules such as proteins or nucleic acids. Simulation of systems with ~50,000–100,000 atoms are now routine. The simulated system can be represented with different levels of detail, among which atomistic representations can best reproduce the actual systems; coarse-grained representations are useful with large systems or long simulations [25].

In MD simulations, an initial model of the system is prepared from nuclear magnetic resonance (NMR), crystallographic, or built by homology-modeling when experimental data are not available. Different solvation models, including both explicit and implicit representations, can be used for the simulation [26–34]. Once the simulation system is built, forces acting on each atom can be obtained by deriving equations with parameters such as equilibrium bond length or angle, partial atomic charge, and van der Waals atomic radii (called “force-field” [35,36]) (Figure 6). Forces arising from interactions occur from both bonded and non-bonded atoms. Chemical bonds, atomic angles and improper angles are modeled by harmonic motions, and dihedral angles are modeled by using a sinusoidal function that approximates the energy differences between eclipsed and staggered conformations. Non-bonded forces include van der Waals interactions, modeled using the Lennard–Jones potential, and electrostatic interactions, modeled using Coulomb’s law.

Forces acting on individual atoms are then used for the calculation of accelerations and velocities with classical Newton’s law of motion. Therefore atom positions are updated after each time step. Here comes one principle challenge that limits the performance of atomistic MD simulation. To avoid possible collision, the simulation time is advanced, often by only 1 or 2 fs. The time step needs to be shorter than the timescale of bond stretching, the fastest motion in the molecule movements. Tens to hundreds of nanoseconds are currently standard simulation lengths for systems with ~100,000 atoms. However, the achievement of adequate sampling of conformational states may require at least microsecond-long simulations, which would cost much more computational time. While for coarse-grained simulation, which involves a more simplified representation of the system, we can use much larger time steps to greatly extend the length of simulations, in which case the long time length simulation is obtained at the cost of the accuracy including neglecting detailed atomic movements.

The good news is, with the rapid development of computational technology and algorithms, we have greatly improved the performance of atomistic MD simulations.

$$\begin{aligned}
 U(R) = & \sum_{\text{bonds}} k_r (r - r_{eq})^2 & \text{bond} & \text{---} \\
 & + \sum_{\text{angles}} k_\theta (\theta - \theta_{eq})^2 & \text{angle} & \text{---} \\
 & + \sum_{\text{dihedrals}} k_\phi (1 + \cos[n\phi - \gamma]) & \text{dihedral} & \text{---} \\
 & + \sum_{\text{impropers}} k_\omega (\omega - \omega_{eq})^2 & \text{improper} & \text{---} \\
 & + \sum_{i < j}^{\text{atoms}} \epsilon_{ij} \left[\left(\frac{r_m}{r_{ij}} \right)^{12} - 2 \left(\frac{r_m}{r_{ij}} \right)^6 \right] & \text{van der Waals} & \text{---} \\
 & + \sum_{i < j}^{\text{atoms}} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} & \text{electrostatic} & \text{---}
 \end{aligned}$$

Figure 6. Classical force fields used for MD simulations: **(Right)** potential energy terms in a force field; and **(Left)** energy function used to derive atomic forces for molecular movement. r is the bond length; θ is the atomic angle; ϕ is the dihedral angle; ω is the improper dihedral angle; r_{ij} is the distance in between atom i and j ; k_r , k_θ , k_ϕ , and k_ω are force constants; r_{eq} , θ_{eq} and ω_{eq} are equilibrium positions; the dihedral term is a periodic term characterized by a force constant (k_ω), multiplicity (n), and phase shift (γ); ϵ_{ij} is related to the Lennard–Jones well depth; r_m is the distance at which the potential reaches its minimum; q_i and q_j are the charges on the respective atoms; and ϵ_0 is the dielectric constant.

During the past decade, molecular dynamics simulations performing on computer clusters or supercomputers using hundreds of processors in parallel has become very common. With the Message Passing Interface (MPI), by using multiple processors for one calculation task simultaneously, we can largely reduce computation time [37]. The most popular simulation software packages (AMBER [38], CHARMM [39], GROMACS [40], NAMD [41] or TINKER [42]) have long been compatible with MPI. In recent years, the use of graphical processing unit (GPU) cards to accelerate calculations is a major breakthrough in computational simulation field. With the use of GPU cards, which include many arithmetic units working in parallel, MD simulations can be accelerated by tens of times, so a single PC with such a card has the power similar to that of a cluster of workstations with multiple processors. Many major MD codes have already been rewritten to incorporate GPUs [43,44]. Moreover, with the Anton machine that specifically designed for MD simulations [45,46], it is possible to study protein dynamics on a millisecond timescale [47].

Another strategy to improve of efficiency of MD is through enhanced MD simulations. Various enhanced MD simulation methods have been developed over the past decade, among which coarse grain is one of the early approaches [48–50]. Other approaches like hyperdynamics [51], accelerated MD [52], RaMD-db [53] and Gaussian accelerated MD (GaMD) [54] accelerate the simulation by raising the potential energy well to lower the energy barrier. Another group of enhanced MD methods improve sampling by employing additional forces to the region of interest, including steered MD [55,56], target MD [57] and self-guided Langevin dynamics (SGLD) [58,59]. There are also other important MD based simulation methods that can enhance sampling. LowModeMD [60] uses low frequency modes from normal mode analysis to guide MD simulation. Transition interface sampling [61] performs intensive sampling at the transition interface, and obtains the kinetics information from probability. Replica exchange molecular dynamics (REMD) [62,63] makes configurations at high temperatures available to the simulations at low temperatures and vice versa, leading to a robust ensemble sampling both low and high energy configurations. Both computational technologies and enhanced MD simulations have made great progress enhancing computational calculations, it is anticipated that future development will continue improving the efficiency and lead to a new chapter for MD simulation.

2.2. Quantum Mechanics/Molecular Mechanics (QM/MM) Method

Though MD is useful in simulating both local atomistic movements and large-scale protein motions, it lacks the ability to model chemical reactions that involve changes in electronic structures, such as bond breaking, bond forming and charge transfer. Quantum Mechanics (QM) would be the perfect method to describe the changes of electronic structures; however, the large computational demand for electronic structure calculation limits the use of QM methods to only small systems with several hundred atoms. To study large biomolecules, Warshel and Levitt first came up with a hybrid quantum mechanics/molecular mechanics (QM/MM) approach in 1976 [24], which is further developed and evaluated by Karplus by coupling semi-empirical QM methods to MM force field [64]. These three together won the 2013 Nobel Prize in Chemistry for “the development of multiscale models for complex chemical systems”.

The QM/MM methods combine the accuracy of QM and speed of MM; thus allowing for the study of chemical reactions in enzymes. The general idea of QM/MM methods is to divide a system into an inner QM region that covers substrates and residues in the enzyme active site, and an outer MM region that includes the rest part of the protein and solvent. If water molecules also play critical roles in catalysis, explicit water molecules in the active site are also included in the inner region. The inner region is treated by a highly accurate QM method, and the outer region is treated by an inexpensive MM method. Therefore, the total potential energy for the system is a sum of MM energy terms, QM energy terms and QM/MM coupling terms:

$$V_{\text{QM/MM}} = V_{\text{QM}}(\text{QM}) + V_{\text{MM}}(\text{MM}) + V_{\text{QM-MM}}(\text{QM} + \text{MM})$$

The QM/MM method is now widely used to model biomolecular systems [65–67], as well as other systems such as inorganic/organometallic [68,69], solid-state [70,71] and explicit solvent systems [72,73]. Moreover, combined with X-ray crystallography or NMR, QM/MM methods are also useful in the refinement of protein structures [74–77].

2.3. Combining Calculation, X-ray Crystallography and Solid-State NMR for Determining Protonation States

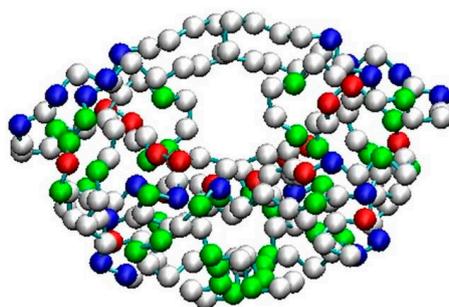
Using X-ray crystallography to determine high resolution protein structures is the first step to provide atomistic details for understanding enzyme function and mechanism. Even at high resolution, protonation states are difficult to be determined by solely X-ray crystallography. Therefore, we use the chemical shift in NMR spectroscopy as an extremely sensitive probe of the chemical environment to determine charge states and electrostatic fields at enzyme active sites. With NMR spectroscopy, isotropic chemical shifts can be used in two specific ways: (1) for directly reporting the chemical state of a probe atom to examine the protonation states or hybridization state at that site; or (2) analyzed for reporting the chemical and structural environment surrounding a probe nucleus. Notably, although solution NMR may not be applicable to proteins with more than several hundreds of residues, solid-state NMR can be used with large protein systems such as TRPS. Solid-state NMR spectra can be obtained on microcrystalline protein samples, prepared basically under the same conditions used to determine X-ray crystal structures [78]. Computational chemistry plays a key role in achieving a high level of structural details for determining protonation states. Typical X-ray and NMR experiments mostly show a protein structure in its static state, and molecular modeling tools are used to illustrate a dynamic view for protein motions and hydrogen bond networks. Although NMR chemical shifts are directly available from experiments, a proton may exchange between multiple ionizable sites. To more accurately predict the protonation states in the equilibrium, a proton can be placed in multiple sites, and *ab initio* calculations are then carried out for first-principle predictions of NMR chemical shifts. This calculation allows for further judging and ranking potential protonation site(s) based on their agreement with experiments [14].

2.4. Coarse-Grained Brownian Dynamics Simulations

Coarse-grained (CG) models for proteins have become popular in recent decades because of the need for modeling large-scale and/or long-time-scale protein motions such as protein folding or large-amplitude conformational fluctuations [79,80]. The CG model and BD simulations are also widely used approaches to investigate ligand–protein association processes. Although the association step is not discussed in this review, forming a substrate–enzyme complex is the very first step for any chemical reaction to happen in the active site of an enzyme. Typically a CG model uses one to six interacting centers (beads) to present each amino acid. A CG model such as a MARTINI CG force field can be a generic force field and the parameter set can be applied to most proteins [48]. Other CG parameters may be protein-dependent [81]. For example, Tozzini and co-authors developed a CG model for HIV-1 protease [82,83]. All the HIV-1 protease X-ray structures available in the Protein Data Bank were included in a statistical set, and the force-field parameters were derived by analyzing the set with the Boltzmann inversion procedure. The potential energy function is a sum of five kinds of interactions:

$$U = U_{\text{bond}} + U_{\text{angle}} + U_{\text{dih}} + U_{\text{vdw}} + U_{\text{elec}}$$

The representation of the one-bead model and the force-field functional forms are shown in Figure 7. The CG model has been implemented in the University of Houston Brownian Dynamics (UHBD) simulation package and the Reduced Molecular Dynamics (RedMD) simulation package [84,85].



Bonded interactions energy function

$$E_{\text{bond}} = \frac{1}{2} k_b (b - b_0)^2, E_{\text{morseb}} = E_0 \left[\left(1 - e^{-k(r_b - r_0)} \right)^2 - 1 \right]$$

$$E_{\text{angle}} = \frac{1}{2} k_a (\theta - \theta_0)^2 + \frac{1}{3} k_a' (\theta - \theta_0)^3 + \frac{1}{4} k_a'' (\theta - \theta_0)^4$$

$$E_{\text{dihedral}} = \frac{1}{2} k_d (\phi - \phi_0)^2$$

Non-bonded interactions energy function

$$E_{\text{intra}} = 0.20708 \left[\left(1 - e^{-0.70711(r_{ij} - 9.5)} \right)^2 - 1 \right]$$

$$E_{\text{inter}} = 0.5 \times \left[\left(\frac{r_i + r_j}{r_{ij}} \right)^8 - 1.5 \left(\frac{r_i + r_j}{r_{ij}} \right)^6 \right]$$

$$E_{\text{elec}} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}, \epsilon_0 = 4r_{ij}$$

Figure 7. Representation of HIV-1 protease in the coarse-grained model and energy function of the force field. **(Top)** Each bead represents a residue, and positively charged, negatively charged and polar residues are in blue, red and green, respectively. **(Bottom)** The potential energy function of the force field for BD simulations, where r is the bond length; θ is the atomic angle; ϕ is the dihedral angle; r_{ij} is the distance in between atom i and j ; k_b , k_a , k_d , and E_0 are force constants; and b_0 , θ_0 and ϕ_0 are equilibrium positions. r_i and r_j are the bead radius, and q_i and q_j are the charges on the respective bead. ϵ_0 is the dielectric constant. E_{morseb} is the Morse potential, a term for computing the bond energy between two consecutive beads. E_{intra} and E_{inter} are non-bonded interactions for $r_{ij} < 8 \text{ \AA}$ and $> 8 \text{ \AA}$, respective. Detailed parameters are in ref [83,86].

Because microseconds or longer simulation lengths may be required to sample and obtain reasonable statistics for large-scale protein motions, BD simulations are used to further accelerate the simulations. Using the BD algorithm, we solved the Langevin equation of internal motion in the overdamped limit. This Ermak–McCammon equation is then used to give a Brownian trajectory [87]. During the simulation, the time step dt may be set to 10–100 fs, which is significantly larger than with atomistic MD simulations, where time step is usually set to 1–2 fs. Hydrodynamic interactions are usually included in BD simulations, and this term is essential to simulate molecular systems under flow. However, calculating hydrodynamic interactions is time-consuming, and this term may be ignored if the hydrodynamic effects are negligible in modeling conformational fluctuations of a protein.

3. Examples of Modeling Enzymes and Substrates

3.1. TRPS: A Model System for Allosteric and Network Regulation in Enzyme Catalysis

Allosteric regulation is important in coordinating function and chemical reactions in many enzymes and protein complexes. A good model system that exhibits allosteric communication, synergistic regulation, and substrate channeling to enhance protein function is TRPS. As shown in Figure 8, TRPS shows synchronization of the α - and β -catalytic activities and conformational switching in its free and substrate bound states [88]. Atomistic MD simulations have been used to study the structure and dynamics of key residues for the α/β -dimer of TRPS and the motions of α -L6 and the communication (COMM) domain in TRPS with or without ligands. The MD simulations started from X-ray structures, the reference state. Because the flexible α -L6 loop is missing in all X-ray structures in the apo form, no open α -L6 loop structure is available experimentally. Therefore, the conformations of the open α -L6 were constructed by running MD simulations for a closed α -L6 loop conformation to open the loop [89,90]. With the absence of a ligand, the flexible α -L6 loop can sample open or partially closed conformations, but the loop tends to shift to fully closed conformations when its substrate is placed in the binding site. The fully closed conformations are induced by favorable ligand–protein interactions, which was not seen when the substrate is missing.

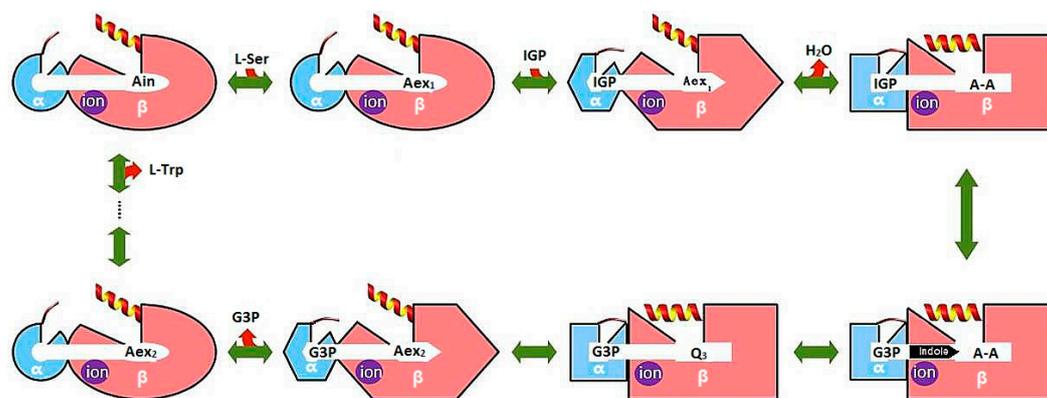


Figure 8. Cartoon representation of synergistic regulation in TRPS. The α - and β -subunit are represented by blue and pink, respectively. Different shapes represent different protein conformations, and α -L6 (thin line), and β -Helix-6 (helix) control open/closed of the active sites during different steps of chemical reactions.

Post-analysis of MD simulations also reveals the most common pattern of protein motion and the dynamic correlations. For example, principal component analysis shows that α -L6 and α -L2 loops and the COMM domain have a strong tendency to move in concert (Figure 1). This synergistic movement is important because α -L6 and α -L2 facilitate the formation of the fully closed conformation to catalyze the α -reaction. In the meantime, the COMM domain receives the signal from allosteric communication to induce the fully closed conformation of the β -site and thereby enhance the overall

rate of the α/β -reaction. The correlated motion that supports the experimental suggestions is shown in Figure 8, whereby a bound α -substrate induced a fully closed conformation of α -L6 and α -L2 loop and then further accelerated the COMM domain closure. A detailed cross-correlation matrix can be plotted with programs such as T-Analyst or Bio3D [91,92], which illustrates residue interactions during MD simulations. Correlations between pairs of residues are presented in cyan (correlated motion) or pink (anti-correlated motion) in Figure 9; correlated motions between α -L6 and α -L2 and the COMM domain are highlighted in squares. The COMM domain and α -L2 loop tend to move to the same direction (region I in Figure 9) and the motion of α -L6 is anti-correlated with that of the COMM domain (regions II and III).

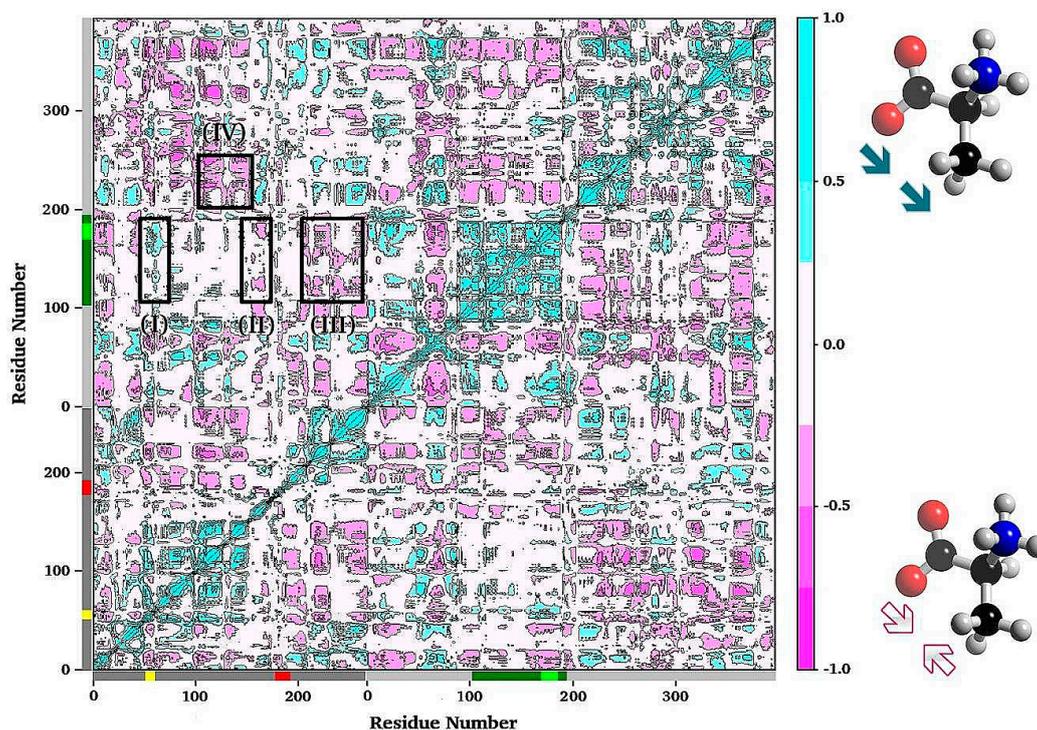


Figure 9. Plot of the residue–residue cross-correlation matrix. The plot was generated from a MD simulation of an apo TRPS by using the Bio3D program. Only C α atom of each residue was used in the analysis. Negative values (pink) indicate anti-correlated motions that C α atoms move along opposite directions. Positive values (cyan) represent correlated motions that C α atoms move along the same direction. Square I shows the correlations between α -L2 and the COMM domain. Squares II and III show that residues of the two edges of α -L6 have anti-correlation motions with the COMM domain. Another example for correlated motions between the α/β interfaces is shown in square IV.

In addition to large amplitude loops and domain motions to open or close the active sites, some residues can form networks of correlated motions for chemical catalysis. NMR chemical shift covariance analyses (CHESCA) and MD simulations are powerful tools to detect such amino acid networks when substrates or products are absent (resting state) and present (working state) in the active site [93–95]. We investigated TRPS in its resting and working states to identify the correlation networks from a cluster of residues. We found that a cluster of residues can form a strongly correlated network in its working state, but the same correlation does not exist in its resting state (Figure 10). Mechanistic studies suggested that Glu49 of α -subunit is a key residue for chemical reaction. NMR studies also indicated that Glu49 and a cluster of other residues formed a network motion in the working state for the α -subunit of TRPS from *Escherichia coli*. Figure 10 shows a cluster of residues that form a correlation network: Glu49, Leu48, Thr39, Val197, Gly98, Met101, Phe107, Ile151, Ile166, Ala167, Ser168, Gly170, Ala198, Ala205, Aal254, Val257 and Thr266. Post-analysis for a 100-ns MD simulation

for an isolated α -subunit of TRPS suggested that the correlated motions changes significantly between the working and resting state, which are consistent with the results from NMR experiments (Figure 10). In particular, the catalytic important Glu49 shows considerable correlated motions with residues identified in the above cluster when the substrate is in the bound form (working state) but not in the resting state. Notably, X-ray structures show that Glu49 can stay at different locations, so proper positioning of Glu49 may play a role in regulating enzyme activity. This observation is consistent with results from NMR and MD that Glu49 performs highly correlated motions with other residues. These studies of TRPS demonstrated that examining the networks for catalysis and synergistic allosteric regulation is critical for fully understanding the functions of an enzyme or enzyme complex. The information may be useful for engineering enzymes and for other applications.

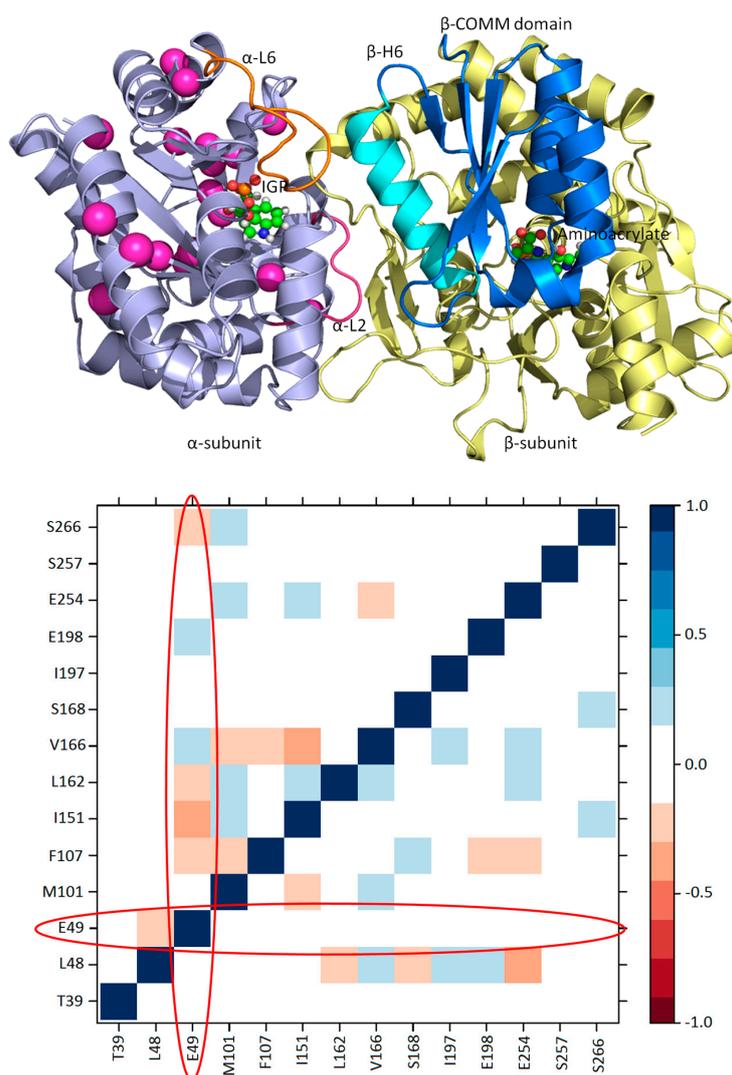


Figure 10. A cluster of residues of the TRPS α -subunit and their cross-correlation map: **(Top)** residues identified by CHESCA that show correlated motions in the working state are presented as pink spheres; and **(Bottom)** a cross-correlation map generated from a 100-ns MD simulation for the isolated α -subunit of TRPS from *Salmonella typhimurium*. The residues are shown in **(Top)**. The program T-Analyst was used to compute the side chain dihedral angles and their correlations with other residues. If a residue has more than one side chain dihedral angle, only the one closest to the protein backbone is used in the analysis. The lower triangle is for the ligand-free (resting) state, and the upper triangle is from the substrate-bound (working) state. The circles indicate the cross-correlations of Glu49 to other residues in the resting and working state, with no significant correlations with the residues in the resting state.

3.2. TRPS: How Protonation States Affect Protein Dynamics and Catalysis

Determining the protonation states of ionizable groups in β -reaction intermediates is critical for understanding the mechanisms of chemical reactions in TRPS catalysis since the relocation of a single proton is enough to promote or initialize a chemical reaction in the enzyme active site. On each PLP-bound intermediate, the protonation states play a crucial role in enhancing or weakening the attractions between the enzyme and substrate at six different locations: the phosphoryl group (PG), pyridine nitrogen (PN), pyridoxyl oxygen (PO), Schiff Base (SB) nitrogen, and both carboxyl oxygens (COs) of the L-Ser substrate (Figure 11A). Combined studies of solid-state NMR and MD simulations have shown that the proton tends to stay at the SB nitrogen in the E(Ain) intermediate, whereas the proton transfers to PO in the E(A-A), E(Q)_{indoline}, and E(Q)_{2AP} state. Moreover, the locations of PG, PN and both COs are all deprotonated during most of the lifetime [96,97]. The preference of protonation states in the four β -intermediates is shown in Figure 11B.

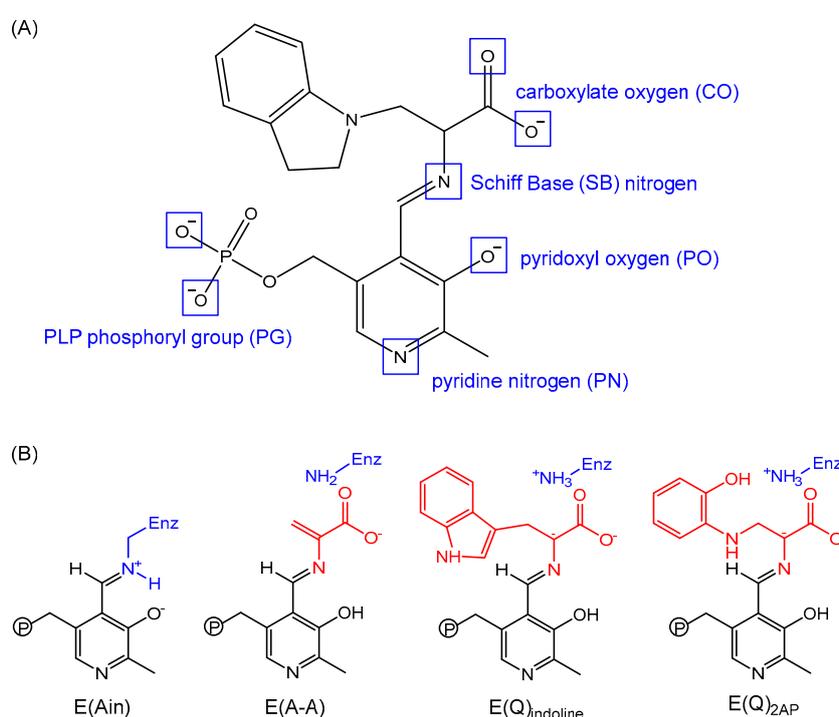


Figure 11. (A) Potential sites of protonation on an indoline quinonoid substrate; and (B) protonation states of four PLP substrates: E(Ain), E(A-A), E(Q)_{indoline} and E(Q)_{2AP}.

The catalysis of TRPS involves a series of proton transfers. Protonation or deprotonation of atoms at the catalytic site or the substrate directly affects structural stability, conformational changes and enzyme activity. For example, in the E(A-A) system, the proton at the PO helps stabilize the substrate in the binding site, so that the PG can form H-bonds with several nearby residues; the PN forms interactions with the β -hydroxyl group of the Ser377 side chain; and the CO atoms interact with the hydroxyl group of the Thr110 side chain and the backbone nitrogens of Gly111 and His115, which keeps the TRPS-E(A-A) complex in a closed conformation of the β -subunit, with no water molecules near the PN, PO, SB nitrogen, and COs. However, the simulations of the proton at the COs, SB nitrogen, PO and PG indicate that the weaker interactions between the COs and nearby residues result in the fluctuation of other residues around the substrate. As well, the H-bonds formed between the PO and Gln114 cause the opening of the β -site, so the solvent moves from bulk solution to the substrate binding site. Thus, changing the location with a single proton not only affects the stability of the local binding site but also alters the overall protein dynamics and molecular motions.

cycle, Cys171(−)/His311(ε) can change back to its initial neutral protonation state via zwitterionic state. By applying the QM/MM methods and FEP, a complete catalytic cycle has been obtained.

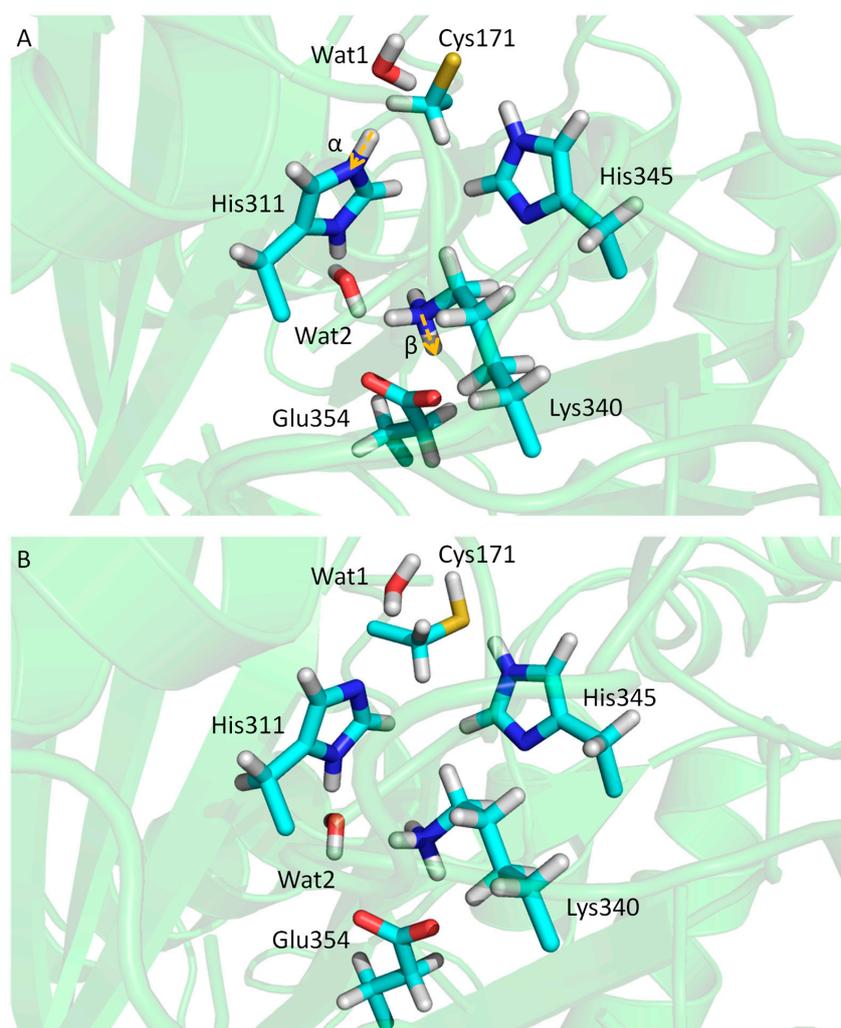


Figure 13. (A) Conformations of key residues of KasA in the resting state with protonation Cys171(−)/His311(+); and (B) conformations of key residues of KasA in the resting state with protonation Cys171/His311(δ).

3.4. Investigating Large-Scale Conformational Changes in Enzymes: HIV-1 Protease

The dynamics of HIV-1 protease flaps are essential for substrate binding and catalysis. Atomistic MD and coarse-grained BD have been used to study the flap motions, flap rearrangements and substrate/inhibitor binding processes. HIV-1 protease is a C₂-symmetric homodimer with the active site consisting of the residues Asp25/Asp25', Thr26/Thr26' and Gly27/Gly27'. The active site is gated by the two flaps that are extended beta hairpin loops and act as clamps to bind to a substrate or inhibitor. NMR studies suggested that the free HIV-1 protease has a substantial conformational change in the flap region, and the large-scale motion occurs on a micro- to millisecond (μs–ms) time scale [21]. Because HIV-1 protease is a well-studied system, hundreds of structures have been solved experimentally. Superimposing the experimental structures showed that most parts of the protein have similar conformations except for the flap regions. Three major conformations of HIV-1 protease have been widely discussed—fully-open, semi-open and closed conformation. Atomistic MD simulations suggested that the closed form is rare when a ligand is absent [108–110]. The semi-open conformations are the predominant form, and only <15% of apo HIV-1 protease appears with a fully

open conformation (Figure 14 top). Both MD and BD simulations revealed that the flap elbows act as “hinges” to control the movement of the flexible regions, whereas the beta hairpin loops move relatively rigidly. Multiple 20- μ s BD simulations with a coarse-grained model suggested that the average flap fully open and closed times are \sim 70 and 430 ns, respectively [83]. The orientation of flaps is called handedness. In addition to the large-scale conformational changes of the flaps, the handedness also varies along with the alternation of flaps motions from open to close. In general, semi-open and closed flap conformations are with semi-open and closed handedness, respectively (Figure 14 bottom).

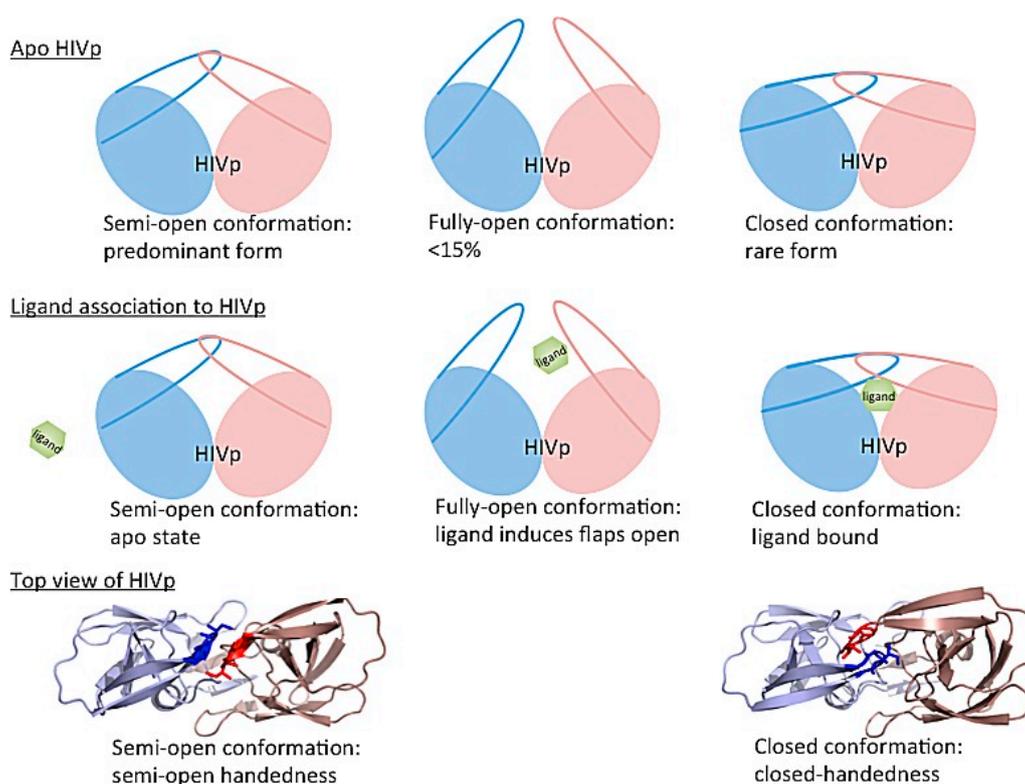


Figure 14. Cartoon representation of conformational rearrangements of HIV-1 protease. (**Top**) Three major HIV-1 protease conformations have been widely discussed. Semi-open conformations are the dominant form in the free state; however, fully-open and closed conformations are rare. (**middle**) A ligand binding to HIV-1 protease. The protein conformations change from semi-open, fully-open to closed forms. (**bottom**) The top view of HIV-1 protease shows that semi-open and closed conformations are with semi-open and closed handedness, respectively.

In the apo protein, semi-open conformations, the predominant form of the flaps, still do not open widely enough for substrate association. Although the flaps can open spontaneously, the opening may be induced when a ligand is approaching to the enzyme, which results in an induced flap motion that may accelerate ligand binding [86,111,112]. After substrate/inhibitor binds to HIV-1 protease, the conformations of the flaps will switch from semi-open form with semi-open handedness to closed form with closed handedness (Figure 14 middle). Simulations complement existing observations from experiments and also provide more complete pictures of large-scale enzyme motions and residue rearrangements for ligand binding to help understand binding/catalysis mechanisms and for inhibitor design.

4. Outlook

Understanding enzyme dynamics and functions is important to building a complete picture of the catalytic mechanism and further help for enzyme design. Mechanistic studies of enzymes usually

begin with investigating X-ray structures, which provide atomistic details for the first examination of chemical reactions and enzyme functions. With recent advances in the synergistic combination of NMR, MD simulations and *ab initio* QM calculations, we can now specify all atomistic information, including the protonation states, of an enzyme system. MD simulations and NMR further reveal dynamic features important to the enzyme functions. Although less directly observable, the allosteric regulation and correlated motions of a residue network are also involved in controlling substrate binding and chemical catalysis. Various computer modeling techniques, such as coarse-grained BD simulations discussed in this review, have been applied to complement atomistic MD simulations to sample large-amplitude and/or long-time-scale protein motions. The field of enzymology has evolved, and describing enzyme functions and catalysis is no longer based solely on simple organic chemical interactions. In addition, we need to consider protein motions involving communication between various residues and multiple regions of an enzyme complex. With more powerful computation tools and resources, molecular simulations will more significantly contribute to our understanding of enzyme systems and assist in enzyme engineering in the future.

Acknowledgments: We thank support from the US National Science Foundation (MCB-1350401), US National Institute of Health (GM-109045 and GM-097569) and National Science Foundation (NSF) national super computer centers.

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

References

1. McCammon, J.A. Gated diffusion-controlled reactions. *BMC Biophys.* **2011**, *4*, 4. [[CrossRef](#)] [[PubMed](#)]
2. Csermely, P.; Sandhu, K.S.; Hazai, E.; Hoksza, Z.; Kiss, H.J.; Miozzo, F.; Veres, D.V.; Piazza, F.; Nussinov, R. Disordered proteins and network disorder in network descriptions of protein structure, dynamics and function: Hypotheses and a comprehensive review. *Curr. Protein Pept. Sci.* **2012**, *13*, 19–33. [[CrossRef](#)] [[PubMed](#)]
3. Tsai, C.J.; Del Sol, A.; Nussinov, R. Protein allostery, signal transmission and dynamics: A classification scheme of allosteric mechanisms. *Mol. Biosyst.* **2009**, *5*, 207–216. [[CrossRef](#)] [[PubMed](#)]
4. Kremer, L.; Dover, L.G.; Carrere, S.; Nampoothiri, K.M.; Lesjean, S.; Brown, A.K.; Brennan, P.J.; Minnikin, D.E.; Loch, C.; Besra, G.S. Mycolic acid biosynthesis and enzymic characterization of the beta-ketoacyl-ACP synthase A-condensing enzyme from *Mycobacterium tuberculosis*. *Biochem. J.* **2002**, *364*, 423–430. [[CrossRef](#)] [[PubMed](#)]
5. White, S.W.; Zheng, J.; Zhang, Y.M.; Rock. The structural biology of type II fatty acid biosynthesis. *Annu. Rev. Biochem.* **2005**, *74*, 791–831. [[CrossRef](#)] [[PubMed](#)]
6. Toney, M.D. Controlling reaction specificity in pyridoxal phosphate enzymes. *Biochim. Biophys Acta* **2011**, *1814*, 1407–1418. [[CrossRef](#)] [[PubMed](#)]
7. Crugeiras, J.; Rios, A.; Riveiros, E.; Richard, J.P. Substituent effects on electrophilic catalysis by the carbonyl group: Anatomy of the rate acceleration for PLP-catalyzed deprotonation of glycine. *J. Am. Chem. Soc.* **2011**, *133*, 3173–3183. [[CrossRef](#)] [[PubMed](#)]
8. Major, D.T.; Gao, J. A combined quantum mechanical and molecular mechanical study of the reaction mechanism and alpha-amino acidity in alanine racemase. *J. Am. Chem. Soc.* **2006**, *128*, 16345–16357. [[CrossRef](#)] [[PubMed](#)]
9. Major, D.T.; Nam, K.; Gao, J. Transition state stabilization and alpha-amino carbon acidity in alanine racemase. *J. Am. Chem. Soc.* **2006**, *128*, 8114–8115. [[CrossRef](#)] [[PubMed](#)]
10. Sachpatzidis, A.; Dealwis, C.; Lubetsky, J.B.; Liang, P.H.; Anderson, K.S.; Lolis, E. Crystallographic studies of phosphonate-based alpha-reaction transition-state analogues complexed to tryptophan synthase. *Biochemistry* **1999**, *38*, 12665–12674. [[CrossRef](#)] [[PubMed](#)]
11. Drewe, W.F.; Dunn, M.F. Detection and identification of intermediates in the reaction of L-serine with *escherichia-coli* tryptophan synthase via rapid-scanning ultraviolet visible spectroscopy. *Biochemistry* **1985**, *24*, 3977–3987. [[CrossRef](#)] [[PubMed](#)]

12. Drewe, W.F.; Dunn, M.F. Characterization of the reaction of L-serine and indole with Escherichia-coli tryptophan synthase via rapid-scanning ultraviolet visible spectroscopy. *Biochemistry* **1986**, *25*, 2494–2501. [[CrossRef](#)] [[PubMed](#)]
13. Ngo, H.; Kimmich, N.; Harris, R.; Niks, D.; Blumenstein, L.; Kulik, V.; Barends, T.R.; Schlichting, I.; Dunn, M.F. Allosteric regulation of substrate channeling in tryptophan synthase: Modulation of the L-serine reaction in stage i of the ss-reaction by alpha-site ligands. *Biochemistry* **2007**, *46*, 7740–7753. [[CrossRef](#)] [[PubMed](#)]
14. Lai, J.; Niks, D.; Wang, Y.; Domratcheva, T.; Barends, T.R.; Schwarz, F.; Olsen, R.A.; Elliott, D.W.; Fatmi, M.Q.; Chang, C.-E.A.; *et al.* X-ray and NMR crystallography in an enzyme active site: The indoline quinonoid intermediate in tryptophan synthase. *J. Am. Chem. Soc.* **2011**, *133*, 4–7. [[CrossRef](#)] [[PubMed](#)]
15. Bhatt, A.; Kremer, L.; Dai, A.Z.; Sacchettini, J.C.; Jacobs, W.R., Jr. Conditional depletion of KasA, a key enzyme of mycolic acid biosynthesis, leads to mycobacterial cell lysis. *J. Bacteriol.* **2005**, *187*, 7596–7606. [[CrossRef](#)] [[PubMed](#)]
16. Debouck, C.; Gorniak, J.G.; Strickler, J.E.; Meek, T.D.; Metcalf, B.W.; Rosenberg, M. Human immunodeficiency virus protease expressed in Escherichia coli exhibits autoprocessing and specific maturation of the gag precursor. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8903–8906. [[CrossRef](#)] [[PubMed](#)]
17. Kohl, N.E.; Emini, E.A.; Schleif, W.A.; Davis, L.J.; Heimbach, J.C.; Dixon, R.A.; Scolnick, E.M.; Sigal, I.S. Active human immunodeficiency virus protease is required for viral infectivity. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 4686–4690. [[CrossRef](#)] [[PubMed](#)]
18. Krausslich, H.G.; Ingraham, R.H.; Skoog, M.T.; Wimmer, E.; Pallai, P.V.; Carter, C.A. Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 807–811. [[CrossRef](#)] [[PubMed](#)]
19. Wlodawer, A.; Vondrasek, J. Inhibitors of HIV-1 protease: A major success of structure-assisted drug design. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 249–284. [[CrossRef](#)] [[PubMed](#)]
20. Ishima, R.; Freedberg, D.I.; Wang, Y.X.; Louis, J.M.; Torchia, D.A. Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function. *Structure (London, England 1993)* **1999**, *7*, 1047–1055. [[CrossRef](#)]
21. Freedberg, D.I.; Ishima, R.; Jacob, J.; Wang, Y.X.; Kustanovich, I.; Louis, J.M.; Torchia, D.A. Rapid structural fluctuations of the free HIV protease flaps in solution: Relationship to crystal structures and comparison with predictions of dynamics calculations. *Protein Sci. Publ. Protein Soc.* **2002**, *11*, 221–232. [[CrossRef](#)] [[PubMed](#)]
22. Adachi, M.; Ohhara, T.; Kurihara, K.; Tamada, T.; Honjo, E.; Okazaki, N.; Arai, S.; Shoyama, Y.; Kimura, K.; Matsumura, H.; *et al.* Structure of HIV-1 protease in complex with potent inhibitor KNI-272 determined by high-resolution X-ray and neutron crystallography. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 4641–4646. [[CrossRef](#)] [[PubMed](#)]
23. McCammon, J.A.; Gelin, B.R.; Karplus, M. Dynamics of folded proteins. *Nature* **1977**, *267*, 585–590. [[CrossRef](#)] [[PubMed](#)]
24. Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249. [[CrossRef](#)]
25. Orozco, M.; Orellana, L.; Hospital, A.; Naganathan, A.N.; Emperador, A.; Carrillo, O.; Gelpi, J.L. Coarse-grained representation of protein flexibility. Foundations, successes, and shortcomings. *Adv. Protein Chem. Struct. Biol.* **2011**, *85*, 183–215. [[PubMed](#)]
26. Anandkrishnan, R.; Drozdetski, A.; Walker, R.C.; Onufriev, A.V. Speed of conformational change: Comparing explicit and implicit solvent molecular dynamics simulations. *Biophys. J.* **2015**, *108*, 1153–1164. [[CrossRef](#)] [[PubMed](#)]
27. Bashford, D.; Case, D.A. Generalized born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129–152. [[CrossRef](#)] [[PubMed](#)]
28. Haberthur, U.; Cafilisch, A. Facts: Fast analytical continuum treatment of solvation. *J. Comput. Chem.* **2008**, *29*, 701–715. [[CrossRef](#)] [[PubMed](#)]
29. Kleinjung, J.; Fraternali, F. Design and application of implicit solvent models in biomolecular simulations. *Curr. Opin. Struct. Biol.* **2014**, *25*, 126–134. [[CrossRef](#)] [[PubMed](#)]
30. Lazaridis, T.; Karplus, M. Effective energy function for proteins in solution. *Proteins* **1999**, *35*, 133–152. [[CrossRef](#)]

31. Luchko, T.; Gusarov, S.; Roe, D.R.; Simmerling, C.; Case, D.A.; Tuszynski, J.; Kovalenko, A. Three-dimensional molecular theory of solvation coupled with molecular dynamics in amber. *J. Chem. Theory Comput.* **2010**, *6*, 607–624. [[CrossRef](#)] [[PubMed](#)]
32. Orozco, M.; Luque, F.J. Theoretical methods for the description of the solvent effect in biomolecular systems. *Chem. Rev.* **2000**, *100*, 4187–4226. [[CrossRef](#)] [[PubMed](#)]
33. Roux, B.; Simonson, T. Implicit solvent models. *Biophys. Chem.* **1999**, *78*, 1–20. [[CrossRef](#)]
34. Vorobjev, Y.N. Advances in implicit models of water solvent to compute conformational free energy and molecular dynamics of proteins at constant ph. *Adv. Protein Chem. Struct. Biol.* **2011**, *85*, 281–322. [[PubMed](#)]
35. Rueda, M.; Ferrer-Costa, C.; Meyer, T.; Perez, A.; Camps, J.; Hospital, A.; Gelpi, J.L.; Orozco, M. A consensus view of protein dynamics. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 796–801. [[CrossRef](#)] [[PubMed](#)]
36. Perez, A.; Lankas, F.; Luque, F.J.; Orozco, M. Towards a molecular dynamics consensus view of b-DNA flexibility. *Nucleic Acids Res.* **2008**, *36*, 2379–2394. [[CrossRef](#)] [[PubMed](#)]
37. Larsson, P.; Hess, B.; Lindahl, E. Algorithm improvements for molecular dynamics simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 93–108. [[CrossRef](#)]
38. Case, D.A.; Babin, V.; Berryman, J.; Betz, R.M.; Cai, Q.; Cerutti, D.S.; Cheatham III, T.E.; Darden, T.A.; Duke, R.E.; Gohlke, H.; *et al.* *Amber 14*; University of California: San Francisco, CA, USA, 2014.
39. Brooks, B.R.; Brooks, C.L., 3rd; Mackerell, A.D., Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; *et al.* Charmm: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [[CrossRef](#)] [[PubMed](#)]
40. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447. [[CrossRef](#)] [[PubMed](#)]
41. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [[CrossRef](#)] [[PubMed](#)]
42. Ponder, J.W.; Richards, F.M. An efficient newton-like method for molecular mechanics energy minimization of large molecules. *J. Comput. Chem.* **1987**, *8*, 1016–1024. [[CrossRef](#)]
43. Harvey, M.J.; Giupponi, G.; Fabritiis, G.D. Acemd: Accelerating biomolecular dynamics in the microsecond time scale. *J. Chem. Theory Comput.* **2009**, *5*, 1632–1639. [[CrossRef](#)] [[PubMed](#)]
44. Eastman, P.; Friedrichs, M.S.; Chodera, J.D.; Radmer, R.J.; Brunts, C.M.; Ku, J.P.; Beauchamp, K.A.; Lane, T.J.; Wang, L.P.; Shukla, D.; *et al.* OpenMM 4: A reusable, extensible, hardware independent library for high performance molecular simulation. *J. Chem. Theory Comput.* **2013**, *9*, 461–469. [[CrossRef](#)] [[PubMed](#)]
45. Shaw, D.E.; Deneroff, M.M.; Dror, R.O.; Kuskin, J.S.; Larson, R.H.; Salmon, J.K.; Young, C.; Batson, B.; Bowers, K.J.; Chao, J.C.; *et al.* Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM* **2008**, *51*, 91–97. [[CrossRef](#)]
46. Shaw, D.E.; Grossman, J.P.; Bank, J.A.; Batson, B.; Butts, J.A.; Chao, J.C.; Deneroff, M.M.; Dror, R.O.; Even, A.; Fenton, C.H.; *et al.* Anton 2: Raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. In Proceedings of the SC14: International Conference for High Performance Computing, Networking, Storage and Analysis, New Orleans, LA, USA, 16–21 November 2014; pp. 41–53.
47. Shaw, D.E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R.O.; Eastwood, M.P.; Bank, J.A.; Jumper, J.M.; Salmon, J.K.; Shan, Y.; *et al.* Atomic-level characterization of the structural dynamics of proteins. *Science* **2010**, *330*, 341–346. [[CrossRef](#)] [[PubMed](#)]
48. Marrink, S.J.; Risselada, H.J.; Yefimov, S.; Tieleman, D.P.; de Vries, A.H. The martini force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824. [[CrossRef](#)] [[PubMed](#)]
49. Nagarajan, A.; Andersen, J.P.; Woolf, T.B. Coarse-grained simulations of transitions in the E2-to-E1 conformations for Ca ATPase (SERCA) show entropy-enthalpy compensation. *J. Mol. Biol.* **2012**, *422*, 575–593. [[CrossRef](#)] [[PubMed](#)]
50. Tozzini, V. Coarse-grained models for proteins. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144–150. [[CrossRef](#)] [[PubMed](#)]
51. Voter, A.F. Hyperdynamics: Accelerated molecular dynamics of infrequent events. *Phys. Rev. Lett.* **1997**, *78*, 3908–3911. [[CrossRef](#)]
52. 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](#)]

53. Doshi, U.; Hamelberg, D. Achieving rigorous accelerated conformational sampling in explicit solvent. *J. Phys. Chem. Lett.* **2014**, *5*, 1217–1224. [[CrossRef](#)] [[PubMed](#)]
54. 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](#)] [[PubMed](#)]
55. Leech, J.; Prins, J.F.; Hermans, J. Smd: Visual steering of molecular dynamics for protein design. *IEEE Comput. Sci. Eng.* **1996**, *3*, 38–45. [[CrossRef](#)]
56. Nicolini, P.; Frezzato, D.; Gellini, C.; Bizzarri, M.; Chelli, R. Toward quantitative estimates of binding affinities for protein-ligand systems involving large inhibitor compounds: A steered molecular dynamics simulation route. *J. Comput. Chem.* **2013**, *34*, 1561–1576. [[CrossRef](#)] [[PubMed](#)]
57. Schlitter, J.; Engels, M.; Kruger, P. Targeted molecular-dynamics—A new approach for searching pathways of conformational transitions. *J. Mol. Graph.* **1994**, *12*, 84–89. [[CrossRef](#)]
58. Wu, X.W.; Brooks, B.R. Self-guided langevin dynamics simulation method. *Chem. Phys. Lett.* **2003**, *381*, 512–518. [[CrossRef](#)]
59. Wu, X.; Brooks, B.R.; Vanden-Eijnden, E. Self-guided langevin dynamics via generalized langevin equation. *J. Comput. Chem.* **2016**, *37*, 595–601. [[CrossRef](#)] [[PubMed](#)]
60. Labute, P. Lowmodemd-implicit low-mode velocity filtering applied to conformational search of macrocycles and protein loops. *J. Chem. Inform. Model.* **2010**, *50*, 792–800. [[CrossRef](#)] [[PubMed](#)]
61. Swenson, D.W.H.; Bolhuis, P.G. A replica exchange transition interface sampling method with multiple interface sets for investigating networks of rare events. *J. Chem. Phys.* **2014**, *141*. [[CrossRef](#)] [[PubMed](#)]
62. Voter, A.F. Parallel replica method for dynamics of infrequent events. *Phys. Rev. B* **1998**, *57*, 13985–13988. [[CrossRef](#)]
63. Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151. [[CrossRef](#)]
64. Field, M.J.; Bash, P.A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700–733. [[CrossRef](#)]
65. Hu, H.; Yang, W. Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601. [[CrossRef](#)] [[PubMed](#)]
66. Senn, H.M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem. (Int. Ed. Engl.)* **2009**, *48*, 1198–1229. [[CrossRef](#)] [[PubMed](#)]
67. Van der Kamp, M.W.; Mulholland, A.J. Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* **2013**, *52*, 2708–2728. [[CrossRef](#)] [[PubMed](#)]
68. Chatwin, S.L.; Davidson, M.G.; Doherty, C.; Donald, S.M.; Jazzar, R.F.R.; Macgregor, S.A.; McIntyre, G.J.; Mahon, M.F.; Whittlesey, M.K. H–X bond activation via hydrogen transfer to hydride in ruthenium N-heterocyclic carbene complexes: Density functional and synthetic studies. *Organometallics* **2006**, *25*, 99–110. [[CrossRef](#)]
69. Genest, A.; Woiterski, A.; Krüger, S.; Shor, A.M.; Rösch, N. The imomm (integrated molecular orbitals/molecular mechanics) approach for ligand-stabilized metal clusters. Comparison to full density functional calculations for the model thiolate cluster $\text{Cu}_{13}(\text{SCH}_2\text{CH}_3)_8$. *J. Chem. Theory Comput.* **2006**, *2*, 47–58. [[CrossRef](#)] [[PubMed](#)]
70. Shiota, Y.; Suzuki, K.; Yoshizawa, K. QM/MM study on the catalytic mechanism of benzene hydroxylation over fe–zsm-5. *Organometallics* **2006**, *25*, 3118–3123. [[CrossRef](#)]
71. To, J.; Sherwood, P.; Sokol, A.A.; Bush, I.J.; Catlow, C.R.A.; van Dam, H.J.J.; French, S.A.; Guest, M.F. QM/MM modelling of the TS-1 catalyst using HPCx. *J. Mater. Chem.* **2006**, *16*, 1919–1926. [[CrossRef](#)]
72. Gunaydin, H.; Acevedo, O.; Jorgensen, W.L.; Houk, K.N. Computation of accurate activation barriers for methyl-transfer reactions of sulfonium and ammonium salts in aqueous solution. *J. Chem. Theory Comput.* **2007**, *3*, 1028–1035. [[CrossRef](#)] [[PubMed](#)]
73. Alexandrova, A.N.; Jorgensen, W.L. Why urea eliminates ammonia rather than hydrolyzes in aqueous solution. *J. Phys. Chem. B* **2007**, *111*, 720–730. [[CrossRef](#)] [[PubMed](#)]
74. Hsiao, Y.W.; Sanchez-Garcia, E.; Doerr, M.; Thiel, W. Quantum refinement of protein structures: Implementation and application to the red fluorescent protein dsred.M1. *J. Phys. Chem. B* **2010**, *114*, 15413–15423. [[CrossRef](#)] [[PubMed](#)]
75. Ryde, U. Combined quantum and molecular mechanics calculations on metalloproteins. *Curr. Opin. Chem. Biol.* **2003**, *7*, 136–142. [[CrossRef](#)]

76. Marti, S.; Moliner, V.; Tunon, I. Improving the QM/MM description of chemical processes: A dual level strategy to explore the potential energy surface in very large systems. *J. Chem. Theory Comput.* **2005**, *1*, 1008–1016. [[CrossRef](#)] [[PubMed](#)]
77. Hsiao, Y.W.; Drakenberg, T.; Ryde, U. NMR structure determination of proteins supplemented by quantum chemical calculations: Detailed structure of the Ca²⁺ sites in the EGF34 fragment of protein S. *J. Biomol. NMR* **2005**, *31*, 97–114. [[CrossRef](#)] [[PubMed](#)]
78. Mueller, L.J.; Dunn, M.F. NMR crystallography of enzyme active sites: Probing chemically detailed, three-dimensional structure in tryptophan synthase. *Acc. Chem. Res.* **2013**, *46*, 2008–2017. [[CrossRef](#)] [[PubMed](#)]
79. Clementi, C. Coarse-grained models of protein folding: Toy models or predictive tools? *Curr. Opin. Struct. Biol.* **2008**, *18*, 10–15. [[CrossRef](#)] [[PubMed](#)]
80. Gorfe, A.A.; Chang, C.E.A.; Ivanov, I.; McCammon, J.A. Dynamics of the acetylcholinesterase tetramer. *Biophys. J.* **2008**, *94*, 1144–1154. [[CrossRef](#)] [[PubMed](#)]
81. Leonarski, F.; Trovato, F.; Tozzini, V.; Les, A.; Trylska, J. Evolutionary algorithm in the optimization of a coarse-grained force field. *J. Chem. Theory Comput.* **2013**, *9*, 4874–4889. [[CrossRef](#)] [[PubMed](#)]
82. Tozzini, V.; McCammon, J.A. A coarse grained model for the dynamics of flap opening in HIV-1 protease. *Chem. Phys. Lett.* **2005**, *413*, 123–128. [[CrossRef](#)]
83. Chang, C.E.; Shen, T.; Trylska, J.; Tozzini, V.; McCammon, J.A. Gated binding of ligands to HIV-1 protease: Brownian dynamics simulations in a coarse-grained model. *Biophys. J.* **2006**, *90*, 3880–3885. [[CrossRef](#)] [[PubMed](#)]
84. Gorecki, A.; Szypowski, M.; Dlugosz, M.; Trylska, J. Redmd-reduced molecular dynamics package. *J. Comput. Chem.* **2009**, *30*, 2364–2373. [[CrossRef](#)] [[PubMed](#)]
85. Davis, M.E.; Madura, J.D.; Luty, B.A.; McCammon, J.A. Electrostatics and diffusion of molecules in solution: Simulations with the university of houston brownian dynamics program. *Comput. Phys. Commun.* **1991**, *62*, 187–197. [[CrossRef](#)]
86. Chang, C.-E.A.; Trylska, J.; Tozzini, V.; Andrew McCammon, J. Binding pathways of ligands to HIV-1 protease: Coarse-grained and atomistic simulations. *Chem. Biol. Drug Des.* **2007**, *69*, 5–13. [[CrossRef](#)] [[PubMed](#)]
87. Ermak, D.L.; McCammon, J.A. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* **1978**, *69*, 1352–1360. [[CrossRef](#)]
88. Dunn, M.F. Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase holoenzyme complex. *Arch. Biochem. Biophys.* **2012**, *519*, 154–166. [[CrossRef](#)] [[PubMed](#)]
89. Fatmi, M.Q.; Ai, R.; Chang, C.-E.A. Synergistic regulation and ligand-induced conformational changes of tryptophan synthase. *Biochemistry* **2009**, *48*, 9921–9931. [[CrossRef](#)] [[PubMed](#)]
90. Fatmi, M.Q.; Chang, C.-E.A. The role of oligomerization and cooperative regulation in protein function: The case of tryptophan synthase. *PLoS Comput. Biol.* **2010**, *6*, e1000994. [[CrossRef](#)] [[PubMed](#)]
91. Ai, R.; Qaiser Fatmi, M.; Chang, C.-E.A. T-analyst: A program for efficient analysis of protein conformational changes by torsion angles. *J. Comput. Aided Mol. Des.* **2010**, *24*, 819–827. [[CrossRef](#)] [[PubMed](#)]
92. Grant, B.J.; Rodrigues, A.P.; ElSawy, K.M.; McCammon, J.A.; Caves, L.S. Bio3d: An R package for the comparative analysis of protein structures. *Bioinformatics (Oxford, England)* **2006**, *22*, 2695–2696. [[CrossRef](#)] [[PubMed](#)]
93. Axe, J.M.; Yezdimer, E.M.; O'Rourke, K.F.; Kerstetter, N.E.; You, W.; Chang, C.-E.A.; Boehr, D.D. Amino acid networks in a (beta/alpha)₈ barrel enzyme change during catalytic turnover. *J. Am. Chem. Soc.* **2014**, *136*, 6818–6821. [[CrossRef](#)] [[PubMed](#)]
94. Axe, J.M.; Boehr, D.D. Long-range interactions in the alpha subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling. *J. Mol. Biol.* **2013**, *425*, 1527–1545. [[CrossRef](#)] [[PubMed](#)]
95. Axe, J.M.; O'Rourke, K.F.; Kerstetter, N.E.; Yezdimer, E.M.; Chan, Y.M.; Chasin, A.; Boehr, D.D. Severing of a hydrogen bond disrupts amino acid networks in the catalytically active state of the alpha subunit of tryptophan synthase. *Protein Sci. A Publ. Protein Soc.* **2015**, *24*, 484–494. [[CrossRef](#)] [[PubMed](#)]
96. Caulkins, B.G.; Bastin, B.; Yang, C.; Neubauer, T.J.; Young, R.P.; Hilario, E.; Huang, Y.-M.M.; Chang, C.-E.A.; Fan, L.; Dunn, M.F.; *et al.* Protonation states of the tryptophan synthase internal aldimine active site from solid-state NMR spectroscopy: Direct observation of the protonated schiff base linkage to pyridoxal-5'-phosphate. *J. Am. Chem. Soc.* **2014**, *136*, 12824–12827. [[CrossRef](#)] [[PubMed](#)]

97. Huang, Y.-M.M.; You, W.; Caulkins, B.G.; Dunn, M.F.; Mueller, L.J.; Chang, C.-E.A. Protonation states and catalysis: Molecular dynamics studies of intermediates in tryptophan synthase. *Protein Sci. A Publ. Protein Soc.* **2016**, *25*, 166–183. [[CrossRef](#)] [[PubMed](#)]
98. Schnackerz, K.D.; Andi, B.; Cook, P.F. P-31 NMR spectroscopy senses the microenvironment of the 5'-phosphate group of enzyme-bound pyridoxal 5'-phosphate. *Biochim. Biophys. Acta Proteins Proteom.* **2011**, *1814*, 1447–1458. [[CrossRef](#)] [[PubMed](#)]
99. Eliot, A.C.; Kirsch, J.F. Pyridoxal phosphate enzymes: Mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem.* **2004**, *73*, 383–415. [[CrossRef](#)] [[PubMed](#)]
100. Fogle, E.J.; Liu, W.S.; Woon, S.T.; Keller, J.W.; Toney, M.D. Role of Q52 in catalysis of decarboxylation and transamination in dialkylglycine decarboxylase. *Biochemistry* **2005**, *44*, 16392–16404. [[CrossRef](#)] [[PubMed](#)]
101. Sharif, S.; Fogle, E.; Toney, M.D.; Denisov, G.S.; Shenderovich, I.G.; Buntkowsky, G.; Tolstoy, P.M.; Huot, M.C.; Limbach, H.-H. NMR localization of protons in critical enzyme hydrogen bonds. *J. Am. Chem. Soc.* **2007**, *129*, 9558–9559. [[CrossRef](#)] [[PubMed](#)]
102. Griswold, W.R.; Toney, M.D. Role of the pyridine nitrogen in pyridoxal 5'-phosphate catalysis: Activity of three classes of PLP enzymes reconstituted with deazapyridoxal 5'-phosphate. *J. Am. Chem. Soc.* **2011**, *133*, 14823–14830. [[CrossRef](#)] [[PubMed](#)]
103. Shaw, J.P.; Petsko, G.A.; Ringe, D. Determination of the structure of alanine racemase from bacillus stearothermophilus at 1.9-angstrom resolution. *Biochemistry* **1997**, *36*, 1329–1342. [[CrossRef](#)] [[PubMed](#)]
104. Tai, C.H.; Cook, P.F. Pyridoxal 5'-phosphate dependent alpha,beta-elimination reactions: Mechanism of O-acetylserine sulfhydrylase. *Acc. Chem. Res.* **2001**, *34*, 49–59. [[CrossRef](#)] [[PubMed](#)]
105. Lee, W.; Luckner, S.R.; Kisker, C.; Tonge, P.J.; Engels, B. Elucidation of the protonation states of the catalytic residues in mtKasA: Implications for inhibitor design. *Biochemistry* **2011**, *50*, 5743–5756. [[CrossRef](#)] [[PubMed](#)]
106. Lee, W.; Engels, B. Clarification on the decarboxylation mechanism in KasA based on the protonation state of key residues in the acyl-enzyme state. *J. Phys. Chem. B* **2013**, *117*, 8095–8104. [[CrossRef](#)] [[PubMed](#)]
107. Lee, W.; Engels, B. The protonation state of catalytic residues in the resting state of KasA revisited: Detailed mechanism for the activation of KasA by its own substrate. *Biochemistry* **2014**, *53*, 919–931. [[CrossRef](#)] [[PubMed](#)]
108. Karthik, S.; Senapati, S. Dynamic flaps in HIV-1 protease adopt unique ordering at different stages in the catalytic cycle. *Proteins* **2011**, *79*, 1830–1840. [[CrossRef](#)] [[PubMed](#)]
109. Deng, N.J.; Zheng, W.; Gallicchio, E.; Levy, R.M. Insights into the dynamics of HIV-1 protease: A kinetic network model constructed from atomistic simulations. *J. Am. Chem. Soc.* **2011**, *133*, 9387–9394. [[CrossRef](#)] [[PubMed](#)]
110. Hornak, V.; Okur, A.; Rizzo, R.C.; Simmerling, C. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 915–920. [[CrossRef](#)] [[PubMed](#)]
111. Huang, Y.M.; Kang, M.; Chang, C.-E.A. Switches of hydrogen bonds during ligand-protein association processes determine binding kinetics. *J. Mol. Recognit. JMR* **2014**, *27*, 537–548. [[CrossRef](#)] [[PubMed](#)]
112. Kang, M.; Roberts, C.; Cheng, Y.; Chang, C.-E.A. Gating and intermolecular interactions in ligand-protein association: Coarse-grained modeling of HIV-1 protease. *J. Chem. Theory Comput.* **2011**, *7*, 3438–3446. [[CrossRef](#)] [[PubMed](#)]

