

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 10527–10537

www.rsc.org/pccp

PAPER

A mesoscopic model for protein enzymatic dynamics in solution

Carlos Echeverria,^a Yuichi Togashi,^b Alexander S. Mikhailov^c and Raymond Kapral^{*d}

Received 2nd January 2011, Accepted 1st March 2011

DOI: 10.1039/c1cp00003a

A multi-scale, coarse-grained description of protein conformational dynamics in a solvent is presented. The focus of the paper is on the description of the conformational motions that may accompany enzyme catalysis as the enzyme executes a catalytic cycle, starting with substrate binding and ending with product release and return to the original unbound enzyme. The protein is modeled by a network of beads representing amino acid residues, the solvent is described by multiparticle collision dynamics, and substrate binding and unbinding events are modeled stochastically by conformation-dependent transitions that modify the bonding in the network to correspond to the different binding states of the protein. The solvent dynamics is coupled to that of the protein and hydrodynamic interactions, which are important for the large-scale protein motions, are taken into account. The multi-scale model is used to study the dynamics of the adenylate kinase enzyme in solution. A potential function that describes the different binding and conformational states of the protein and accounts for partial unfolding during the catalytic cycle is constructed as a network built from elastic network and soft potential links. The conformational dynamics of the protein as it undergoes cyclic enzymatic dynamics, as well as its translational diffusion and orientational motion, are investigated using both multiparticle collision dynamics and dynamics that suppresses hydrodynamic coupling. Hydrodynamic interactions are found to have important effects on the large scale conformational motions of the protein and significantly affect the translational diffusion coefficients and orientational correlation times.

1. Introduction

Many biochemical processes rely on the catalytic activity of proteins acting as enzyme molecules since they greatly increase the rates of biochemical reactions. While enzymes operate by a variety of mechanisms, a common mode of operation involves large-scale conformational motions of protein domains in order to bring reactants into a configuration where reaction is possible. In such cases, the enzymatic cycle comprises the binding of one or more substrates to active sites in the protein, conformational changes induced by or accompanying the binding events, and reaction and product release which triggers return of the enzyme to its initial set of conformations in preparation for the next set of substrate binding events.¹

Enzymatic cycle times are typically long and the large-scale conformational motions are strongly influenced by both fluctuations and solvent interactions. Apart from specific solvent effects such as binding of water to certain parts of the protein and exclusion of water from the active site in order to enhance the efficiency of catalytic reactions, solvent hydrodynamical modes also play a role in the dynamics. Solvent collective modes can dissipate energy within the protein and can give rise to hydrodynamic coupling among remote parts of the protein, which changes the time scales on which large-scale protein motions occur.

A full description of such enzymatic processes should incorporate a molecular-level description of the enzyme, substrate binding, product formation and release processes, and a molecular-level description of the solvent in which the dynamics takes place. While molecular dynamics has proved to be a powerful simulation tool for the study of the conformational dynamics of proteins,^{2,3} it is very difficult to study many enzymatic cycles, or even a single cycle, at a fully microscopic level using molecular dynamics, although recent computational advances have allowed long time scale simulations of the folding of small proteins to be carried out.^{4,5} Consequently, in order to study the dynamics of large proteins in solution over many enzymatic cycles, one must rely on

^a Laboratorio de Investigación de Física Aplicada y Computacional, Universidad Nacional Experimental del Táchira, San Cristóbal 5001, Venezuela. E-mail: cecheve@unet.edu.ve

^b Cybermedia Center, Osaka University, 5-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan. E-mail: togashi@cmc.osaka-u.ac.jp

^c Abteilung Physikalische Chemie, Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany. E-mail: mikhailov@fhi-berlin.mpg.de

^d Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, ON M5S 3H6, Canada. E-mail: rkapral@chem.utoronto.ca

coarse-grain multi-scale methods.^{6–10} Features such as the large scale motions of the proteins and the influence of solvent collective motions on this dynamics can be described by models that neglect some aspects of the molecular detail.

In this article we present a mesoscopic dynamical model that can be used to investigate ligand-induced conformational changes in enzymatic cycles in solution. Coarse graining appears in several aspects of the model: both the protein and solvent are described at a coarse grained level, in addition, the time evolution of the entire system, protein plus solvent, is carried out by combining molecular dynamics (MD) for the protein and multiparticle collision (MPC) dynamics^{11,12} for the solvent. This hybrid MD-MPC scheme preserves the important conservation laws but allows for efficient simulation of the dynamics. The method and phenomena are illustrated by considering the dynamics of adenylate kinase in solution. This well-studied protein catalyzes the reversible conversion of ATP and AMP into two molecules of ADP and forms an essential part of many cellular processes. The protein undergoes large-scale domain motions in the course of a single catalytic cycle in order to bring the reactants into a configuration favorable for product formation and release; thus, its dynamics requires all components of our mesoscopic model.

In Section 2 we describe, in general terms, the ingredients of our mesoscopic model: protein potential, solvent dynamics and solvent coupling to the protein, and the description of substrate binding and unbinding events. Section 3 shows how a mesoscopic model for the enzymatic dynamics of adenylate kinase in solution can be constructed. The results of simulations of the model are presented in Section 4 while the conclusions of the study are given in Section 5.

2. Mesoscopic model of enzyme dynamics

2.1 Protein model

As in many other studies we adopt a coarse grained description of the protein where each of the N_p amino acid residues is represented by a single bead.¹³ The potential energy of the protein $V_p(\mathbf{R})$ is a function of the coordinates $\mathbf{R} = (\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{N_p})$ of the protein beads and we assume it may be represented by a sum of pair-wise interactions among the beads. Since we are interested in situations where the protein undergoes large conformational changes as a result of ligand binding and product release processes, $V_p(\mathbf{R})$ must be able to describe very different equilibrium conformations of the protein in various ligand bound and unbound states, as well as the transitions between these states.

In a full treatment of the enzymatic cycle, even at a coarse grained level of description, we must consider the binding of ligands to the enzyme and the subsequent formation of new bonds and breaking of other bonds as the enzyme–substrate complex evolves to a new conformational state. Similar considerations apply to the dissociation of the complex to form the enzyme and product. Rather than explicitly considering the dynamics of substrate and product molecules in these processes,^{14,15} we focus on the state of the protein with bound or unbound substrates or products. Such structure-based models have been used to study protein conformational

changes upon ligand binding and unbinding.^{16,17} A label, $\lambda = \lambda_1, \lambda_2, \dots, \lambda_{M_s}$, is used to indicate the potential function for the given binding state of the protein: $V_p(\mathbf{R}; \lambda)$, where we have supposed that there are M_s such states of binding. For binding state λ the protein will adopt an equilibrium conformation with configuration $\mathbf{R}_\lambda^0 = (\mathbf{R}_{1,\lambda}^0, \mathbf{R}_{2,\lambda}^0, \dots, \mathbf{R}_{N_p,\lambda}^0)$. The protein potential functions $V_p(\mathbf{R}; \lambda)$ may be constructed as elastic network models (ENM),¹⁸ Gō-like models,¹⁹ plastic network models¹⁶ or other models that account for large bond deformations or partial unfolding during the conformational evolution.^{17,20,21}

2.2 Solvent model and coupling to protein

The solvent is described at a mesoscopic level using MPC dynamics.^{11,12} In MPC dynamics, N_s solvent particles, representing coarse grained real molecules, free stream and undergo effective collisions at discrete time intervals τ , accounting for the effects of many real collisions during this time interval. The collisions are carried out by dividing the system into a grid of cells $\mathcal{V}_\xi$ and assigning rotation operators $\hat{\omega}_\xi$, chosen from some set of rotation operators, to each cell of the system at the time of collision. For example, the set of rotation operators may be taken to be the operators corresponding to rotations by a fixed angle about a randomly chosen axis, or those corresponding to a randomly chosen angle about a randomly chosen axis, *etc.* Particles within each cell “collide” with each other and the postcollision velocity of particle i in a cell $\mathcal{V}_\xi$ is given by

$$\mathbf{v}'_i = \mathbf{V}_\xi + \hat{\omega}_\xi(\mathbf{v}_i - \mathbf{V}_\xi), \quad (1)$$

where $\mathbf{V}_\xi$ is the center of mass velocity of particles in the cell. MPC dynamics satisfies the mass, momentum, and energy conservation laws. The dynamics is microcanonical and preserves phase-space volumes. Further descriptions of the MPC simulation scheme with applications can be found in recent reviews.^{22,23}

Next we consider how the coarse grained protein interacts with the mesoscopic solvent. Two coupling schemes have proven to be effective in modeling the system dynamics. Explicit solvent-bead intermolecular potentials can be introduced to couple the protein and solvent motions.¹² In such a case the free streaming step is replaced by molecular dynamics evolution for time intervals τ governed by the protein and protein–solvent intermolecular potentials. There are no solvent–solvent forces. These are accounted for by multiparticle collisions at the times τ as described above. This scheme allows specific attractive and repulsive (hydrophilic and hydrophobic) interactions with different protein beads to be taken into account. It also can describe exclusion of solvent from the interior of the protein or active site in certain protein conformations. The second scheme is simpler and computationally more efficient but cannot account for specific solvent–protein effects. In this scheme the protein beads that lie in a given MPC collision cell are simply included in the MPC collision step.²⁴ No protein–solvent forces need to be introduced or computed. Both of these MD-MPC dynamical coupling schemes conserve mass, momentum and energy. Consequently, hydrodynamic interactions, which play an important role in the large-scale protein conformational motions, are properly

taken into account. In this paper we adopt the second coupling scheme where explicit protein–solvent interactions are not considered.

2.3 Stochastic MD-MPC dynamics

Given this specification of the basic interaction potentials and dynamics, the full enzymatic cycle can be simulated by stochastic MD-MPC dynamics which accounts for the changes in the protein potential functions when ligand binding and release events occur. As an example, consider the binding of a ligand to the protein which leads to a protein–ligand complex. As the ligand moves from the environment to the binding pocket of the protein, new bonds are formed and others are modified. This results in a new potential function that is different from that of the protein and ligand before the binding event. Our description of this binding process does not explicitly consider the dynamics of the ligand and instead focuses on the potential energy changes that result from such ligand binding. We model ligand binding as a stochastic process that changes the potential function to reflect the fact that the ligand is now bound to the protein. Similar considerations apply to the release of a ligand from the protein.

In more general terms, suppose that the enzymatic cycle, which starts with substrate binding and, after final product release, ends in the original unbound state, involves a sequence of M_s ligand binding–unbinding events. As noted above, the index λ is used to label the protein potential function when the protein is in a specified state of ligand binding. A transition from one state λ of ligand binding to another λ' , $\lambda \rightarrow \lambda'$, is assumed to occur with probability $p_{\lambda,\lambda'}(\mathbf{R})$ per unit time, which may depend on the instantaneous conformation $\mathbf{R}$ of the protein. We shall be concerned with systems that are far from equilibrium where substrates are fed into the systems and products are removed. In these circumstances the transition probabilities will not satisfy detailed balance. Equilibrium systems can also be considered in our general framework but we shall focus on systems that are far from equilibrium.

The full enzymatic dynamics is then as follows: given a conformational state $\mathbf{R}$ and binding state λ of the protein, the system evolves by MD-MPC dynamics governed by forces derived from the protein potential function $V_P(\mathbf{R};\lambda)$ and a solvent–protein potential if the explicit solvent–protein coupling scheme is used. In the course of this trajectory, transitions to other protein states λ' are attempted with probability $p_{\lambda,\lambda'}(\mathbf{R}(t))$. If such a protein state transition occurs, the protein potential function is changed to $V_P(\mathbf{R};\lambda')$ and MD-MPC dynamics continues.²⁵

Each stochastic trajectory provides a realization of the dynamics of a single protein as it executes catalytic cycles. As an illustration of this dynamical model, we now show how it can be used to explore various generic features, including the influence of hydrodynamic interactions, of the enzymatic dynamics of adenylate kinase.

3. Adenylate kinase

The adenylate kinase enzyme (AKE) catalyzes the Mg-dependent reversible conversion of ATP and AMP into two molecules of ADP,



This enzyme-catalyzed phosphotransfer regulates many cellular signaling processes such as hormone secretion, cell motility, the energetics of the cell cycle and DNA synthesis and repair. It controls AMP levels whose high or low values have been linked to human diseases such as diabetes, obesity and hypertrophic cardiomyopathy.²⁶

The enzyme consists of 214 amino acid residues and has three mobile domains: the CORE domain with 133 residues, the NMP domain with 38 residues and the LID domain with 43 residues (see Fig. 1). The LID domain binds ATP while the NMP domain binds AMP; both domains undergo large hinge motions in the course of the catalytic cycle. When no substrates are bound the enzyme exists in an open conformation where LID-CORE and NMP-CORE center of mass distances are large. Conversely, when both substrates are bound the protein exists in the fully closed conformation where these distances are small. Fluorescent resonance energy transfer (FRET) and NMR studies have indicated that the LID domain exhibits large fluctuations so that a range of LID-CORE distances, including distances comparable to those of the closed conformation, are visited.^{27–30} The phosphoryl transfer reaction to form the two ADP molecules occurs in the fully closed conformation where the solvent is excluded from the active site. There has been a considerable number of theoretical^{16,31–40} and experimental studies of this enzyme.^{27–30,41–47}

In addition to the fully open and fully closed conformations, there are other prominent intermediate conformational binding states: (LID-CORE closed, NMP-CORE open) and (LID-CORE open, NMP-CORE closed). Thus, there are four relevant states of the protein to consider, based on the open or closed configurations of the LID and NMP domains. We use the following notation to label these protein states: let χ_α be a variable that takes the value $\chi_\alpha = 1$ for an open conformation and $\chi_\alpha = 0$ for a closed conformation. The index $\alpha = \text{LC, NC}$ and the variables χ_{LC} and χ_{NC} are defined for the LID-CORE and NMP-CORE conformations, respectively. Using this notation the four binding states of the protein may be labeled by $\lambda = (\chi_{\text{LC}}, \chi_{\text{NC}})$. In particular, the state where both the LID-CORE and NMP-CORE domains are open may be labeled by $\lambda_1 = (1, 1)$, and the remaining states have the notation $\lambda_2 = (0, 1)$, $\lambda_3 = (0, 0)$ and $\lambda_4 = (1, 0)$. The mechanism for the enzymatic reaction will involve passage among these states.

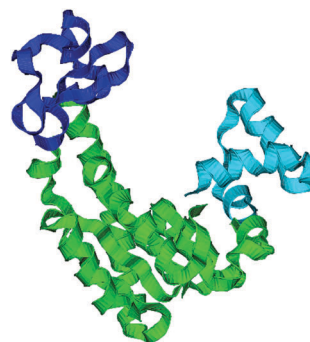


Fig. 1 Adenylate kinase domains: CORE (green), LID (blue) and NMP (light blue).

We now describe how the potential energy function for the coarse grained model of AKE that accounts for the different binding states can be constructed. The potential function incorporates elements that were used in earlier studies^{16,17} to construct a potential for AKE that can describe the conformational changes that occur.

3.1 AKE potential energy function

The coarse grained model of adenylate kinase consists of 214 beads that represent the amino acid residues in the protein. The protein potential energy functions $V_P(\mathbf{R};\lambda)$ were constructed from a network model where beads in the protein are connected by bonds. The building blocks of our potential function are the fully open conformation with no substrates bound, and the fully closed conformation with both ATP and AMP substrates bound. The equilibrium conformations corresponding to these two open (o) and closed (c) states, $\mathbf{R}_o^0$ and $\mathbf{R}_c^0$, respectively, are known from experimental data.^{48,49} In a network model, the equilibrium open conformation of the protein, $\mathbf{R}_o^0$, can be characterized by a set of equilibrium bond lengths, $\{R_{n,o}^b | n \in \mathbf{B}\}$, where $R_{n,o}^b$ denotes the equilibrium bond length of bond n for the protein in the fully open state and $\mathbf{B}$ denotes the set of bonds in the network. Similarly we can define the set $\{R_{n,c}^b | n \in \mathbf{B}\}$ for the fully closed state. Elastic network models are easily constructed for each of these two binding states. However, we cannot use these elastic network models directly for the enzymatic dynamics since there is partial unfolding of the LID and NMP domains in the course of the enzymatic cycle. A variety of potential models for AKE has been constructed.^{16,17,31,32} Our model is similar in spirit to the plastic network model¹⁶ and uses a combination of elastic network bonds and bonds described by soft potentials for the open and closed conformations, along with an interpolation function, to construct a network model for AKE.

To build the mixed network model we assume that two beads in the network are connected by a bond n if their separation $r_n \leq 8.5$ Å. We determine all bonds in the fully open and closed conformations in this way and this defines the set of bonds $\mathbf{B} = \{n | n \text{ with } r_n < 8.5 \text{ Å}\}$. The distance value 8.5 Å is the parameter that limits the number of bonds in the $\mathbf{B}$ set. The bonds in this set are further classified as follows: the bonds are divided into two sets $\mathbf{B}^e$ and $\mathbf{B}^s$, where superscripts e and s indicate the type of bond, elastic and soft, respectively, which will be associated with a given type of potential energy of interaction. The bonds are classified according to the criteria,

$$\begin{aligned} n \in \mathbf{B}^e, & \text{ if } \{R_{n,o}^b \leq 8.5 \text{ Å and } R_{n,c}^b \leq 8.5 \text{ Å}\} \\ n \in \mathbf{B}^s, & \text{ if } \{R_{n,o}^b \leq 8.5 \text{ Å or } R_{n,c}^b \leq 8.5 \text{ Å}\} - \mathbf{B}^e \end{aligned} \quad (3)$$

We note that $\mathbf{B}^s$ is formed by union of all bonds with bead distances smaller than 8.5 Å in open or closed conformation, minus bonds in the $\mathbf{B}^e$ set. Since a change in the binding state is associated with a specific type of domain opening or closing, the bonds in the set $\mathbf{B}^s$ may be further classified as LID-CORE or NMP-CORE bonds. If a bond in this set connects a residue in the LID domain with a residue in the CORE domain it is

classified as a LID-CORE bond. A similar procedure was used to determine the NMP-CORE bonds.

We let $V_n^P(r_n; \chi_\alpha^n)$ to denote the potential energy function for bond n in the network, where we use a parameter χ_α^n in the bond potential function to determine the nature of that bond depending on the binding state. For example, if a bond n is a LID-CORE bond, then $\chi_\alpha^n = \chi_{LC}^n$. For those bonds in the $\mathbf{B}^e$ set, we use elastic network potentials since these bonds are common to both open and closed configurations and not highly strained. For bonds $n \in \mathbf{B}^s$, we use an attractive potential, similar to that proposed earlier.¹⁷ A detailed description of how these bond potential functions are constructed is given in the Appendix. Once the bond potential functions are specified, the total potential energy of the AKE network is given by

$$V_P(\mathbf{R}; \lambda) = \sum_{n=1}^{n_L} V_n^P(r_n; \chi_\alpha^n). \quad (4)$$

The sum is over all n_L bonds in the network. The protein binding state λ is determined from the values of χ_{LC} and χ_{NC} that enter in the specification of the bond potentials in the sum in eqn (4).

3.2 Transition rates

The transition rates between the different binding states depend on the instantaneous conformation of the protein. To account for the fact that binding and unbinding occur primarily from a certain set of open and closed conformational states, we assume that the AKE conformations can be characterized by the distances between the centers of mass of the LID and CORE domains, R_{LC}^{cm} , and NMP and CORE domains, R_{NC}^{cm} , and use these collective variables to determine when a binding or unbinding event occurs.^{31,32} Thus, the transition probability per unit time for a state transformation $\lambda \rightarrow \lambda'$ has the functional dependence $p_{\lambda,\lambda'}(R_{LC}^{\text{cm}}, R_{NC}^{\text{cm}})$ in our model. As noted earlier, we shall be concerned with systems that are far from equilibrium where substrates are fed into the system and products removed so that the transition probabilities will not satisfy detailed balance.

In order to investigate some aspects of AKE conformational dynamics in solution in the course of enzymatic catalysis, we consider two models for the transition probabilities. The first model, termed the sequential binding (SB) model, assumes a sequential passage through the different binding states. This model is suggested by recent investigations that describe the enzymatic cycle as a sequential series of domain closing and opening events upon ligand binding and product release.^{31,32} In particular, in this mechanism the cycle starts with a configuration where both the LID-CORE and NMP-CORE domains are in an open configuration. As ATP binds, the protein then passes to a state where the LID-CORE is closed but NMP-CORE is still open. After binding of AMP the NMP-CORE closes. As the ADP products are released, the LID-CORE first opens, followed by opening of the NMP-CORE, to yield the initial configuration. During these conformational changes partial unfolding of the LID and NMP domains occurs. The binding states visited in the enzymatic cycle described above can be indicated as $\lambda_1 \rightarrow \lambda_2 \rightarrow \lambda_3 \rightarrow \lambda_4 \rightarrow \lambda_1$. This enzymatic cycle is shown in Fig. 2.

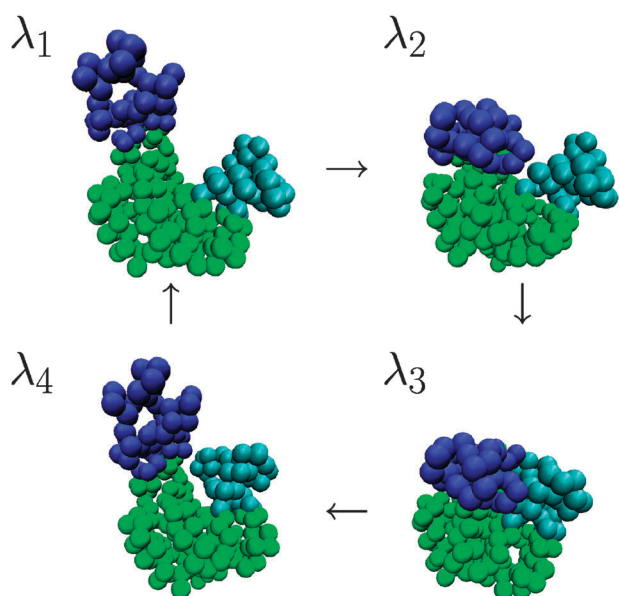


Fig. 2 Conformational changes of AKE from the fully open (top left) to fully closed (bottom right) conformations in the sequential binding mechanism. Domain color coding: CORE (green), LID (blue) and NMP (light blue).

To model this enzymatic cycle we compute the transition probabilities as follows. To determine how the transition probabilities depend on the protein conformation we compare the values of R_{LC}^{cm} and R_{NC}^{cm} to functions of their values in the fully open and closed conformations, $R_{LC,o}^{cm}$, $R_{NC,o}^{cm}$ and $R_{LC,c}^{cm}$, $R_{NC,c}^{cm}$, respectively, so that

$$\begin{aligned} p_{\lambda_1,\lambda_2}(\mathbf{R}) &= p_{oc}\theta(\delta_{LC} - R_{LC}^{cm})\theta(R_{NC}^{cm} - \delta_{NC}), \\ p_{\lambda_2,\lambda_3}(\mathbf{R}) &= p_{oc}\theta(\delta_{LC} - R_{LC}^{cm})\theta(\delta_{NC} - R_{NC}^{cm}), \\ p_{\lambda_3,\lambda_4}(\mathbf{R}) &= p_{co}\theta(R_{LC}^{cm} - \delta_{LC})\theta(\delta_{NC} - R_{NC}^{cm}), \\ p_{\lambda_4,\lambda_1}(\mathbf{R}) &= p_{co}\theta(R_{LC}^{cm} - \delta_{LC})\theta(R_{NC}^{cm} - \delta_{NC}). \end{aligned} \quad (5)$$

Here $\delta_{LC} = (R_{LC,o}^{cm} + R_{LC,c}^{cm})/2$ and $\delta_{NC} = (R_{NC,o}^{cm} + R_{NC,c}^{cm})/2$. The Heaviside θ functions account for the fact that ligand binding and unbinding events may exploit large fluctuations in the conformations. The probabilities per unit time, p_{oc} and p_{co} , account for factors other than the protein conformation that influence the binding rates; for instance, they determine how long the protein will remain in a given conformational state once the domain distance criteria are satisfied and incorporate the rates at which substrates encounter the active site of the protein. Other variants of these transition rates can be constructed to examine in further detail how protein conformations influence binding and unbinding rates.

This model for the transition probabilities assumes a specific order of cycling through the four binding states. Other mechanisms have been proposed for the enzymatic cycle.^{26,46} For comparison, we also consider another model, termed the random binding (RB) model, where the binding or unbinding of a ligand does not depend on the binding state of the other ligand. The transition probabilities for the RB model are then given by,

$$\begin{aligned} p_{\lambda_1,\lambda_2}(\mathbf{R}) &= p_{\lambda_4,\lambda_3}(\mathbf{R}) = p_{oc}\theta(\delta_{LC} - R_{LC}^{cm}), \\ p_{\lambda_2,\lambda_3}(\mathbf{R}) &= p_{\lambda_1,\lambda_4}(\mathbf{R}) = p_{oc}\theta(\delta_{NC} - R_{NC}^{cm}), \\ p_{\lambda_3,\lambda_4}(\mathbf{R}) &= p_{\lambda_2,\lambda_1}(\mathbf{R}) = p_{co}\theta(R_{LC}^{cm} - \delta_{LC}), \\ p_{\lambda_4,\lambda_1}(\mathbf{R}) &= p_{\lambda_3,\lambda_2}(\mathbf{R}) = p_{co}\theta(R_{NC}^{cm} - \delta_{NC}). \end{aligned} \quad (6)$$

In this model there is no regular order for the opening and closing of the domains.

4. Simulation of adenylate kinase dynamics

The simulations of AKE dynamics were carried out in a cubic box containing the protein and solvent. Periodic boundary conditions were employed. For the multiparticle collision events, the simulation box was divided into $(42)^3$ cubic cells with side a . In the MPC algorithm, velocities were rotated about randomly chosen axes by angles $\pm\pi/2$. Grid shifting was implemented in the MPC step of the dynamics.^{50,51} Dimensionless units were used in the calculations: lengths were measured in units of a , energy in units of ϵ , mass in units of m . The dimensionless mass of a solvent molecule was taken to be $m = 1$ while the mass of a protein bead was $m_b = 5$. In these units other system parameters are: elastic network bond force constant, $k = 40$; non-common soft bond energy parameter, $c = 0.4$; potential parameters, $\Delta = 1.0$ and $\Delta V = 0.0$; solvent average number density, $n_s = 7$; and reduced temperature $k_B T/\epsilon = 1/3$. Time was expressed in units of the multiparticle collision time τ and in this time unit the MD time step was $\Delta t = 0.002$. The probabilities that enter the transition rate expressions are $p_{oc} = 0.005$ and $p_{co} = 0.0005$.

All results will be reported in the dimensionless simulation units given above. In these units the LID-CORE and NMP-CORE equilibrium separations in the open and closed states are $R_{LC,o}^{cm} = 5.025$, $R_{LC,c}^{cm} = 3.49$, $R_{NC,o}^{cm} = 3.66$, $R_{NC,c}^{cm} = 3.06$. For reference, distances can be converted to Angstrom units by taking the cell size to be $a = 6 \text{ \AA}$, and in Angstrom units these distances are $R_{LC,o}^{cm} = 30.15 \text{ \AA}$, $R_{LC,c}^{cm} = 20.95 \text{ \AA}$, $R_{NC,o}^{cm} = 21.95 \text{ \AA}$, $R_{NC,c}^{cm} = 18.35 \text{ \AA}$, which correspond to the experimental values.

4.1 Catalytic cycles

The AKE dynamics with transition probabilities given by the SB model is shown in Fig. 3 for several catalytic cycles.⁵² In this figure we see that the enzyme passes between the fully open and fully closed conformations with average LID-CORE and NMP-CORE center of mass distances close to their experimental values. Large fluctuations occur in the domain distances in the course of the catalytic cycle. The SB transition probability rule in eqn (5) is constructed to exploit these fluctuations so that binding predominantly takes place when large fluctuations bring the system to a favorable configuration.

As the system executes catalytic cycles it also visits intermediate conformational states, with lower probability, where one domain is in an open configuration while the other is in a closed conformation. This can be seen in Fig. 4, which plots the steady state probability density of the LID-CORE and NMP-CORE distances, $P(R_{LC}^{cm}, R_{NC}^{cm})$. Enzyme conformations

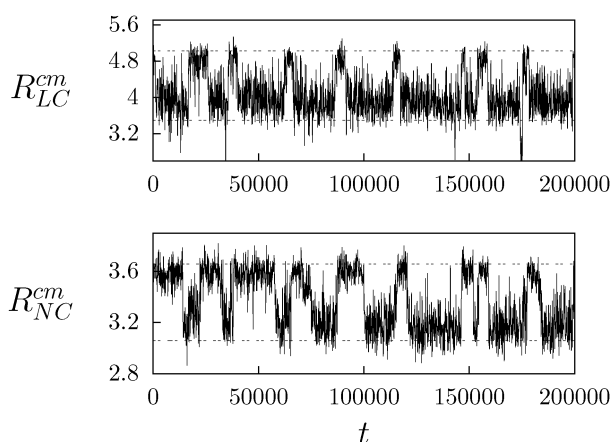


Fig. 3 Plots of the LID-CORE (top) and NMP-CORE (bottom) center of mass distances, R_{LC}^{cm} and R_{NC}^{cm} , as a function of time in τ units. The dashed lines are the crystallographic values of R_{LC}^{cm} (top) and R_{NC}^{cm} (bottom) for the open and closed configurations.

(LID (open), NMP (closed)), and *vice versa*, have probabilities that are lower than fully open or fully closed conformations. Also, the fully closed conformation is favored. The trajectories that underlie this plot explore the configuration space region indicated by blue color in the figure, similar to that in earlier studies.^{17,32} Our stochastic rule for substrate binding and unbinding, embodied in the forms of the transition probabilities in eqn (5), was designed to capture these general features of the sequential mechanism.

Further insight into the dynamics in the enzymatic cycle can be gained by examining the potential energy of the protein during its catalytic cycles. Fig. 5 shows a plot of the enzyme potential energy *versus* R_{LC}^{cm} and R_{NC}^{cm} . The potential energy variations in the protein reflect the conformational changes that accompany the ligand binding and product release events in the catalytic cycles. The figure, which was constructed from 200 realizations of the cyclic dynamics of the protein, shows that there are substantial fluctuations in the potential energy, whose magnitudes may be gauged by the dashed line in the figure. The protein potential energy magnitudes in this figure reflect the structure of the $P(R_{LC}^{cm}, R_{NC}^{cm})$ probability density

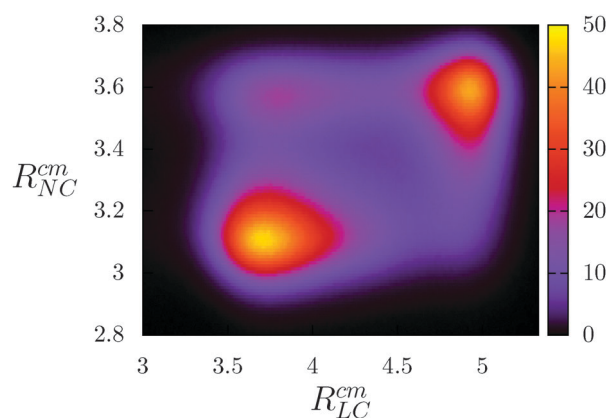


Fig. 4 The probability density $P(R_{LC}^{cm}, R_{NC}^{cm})$ *versus* R_{LC}^{cm} and R_{NC}^{cm} . The color scale is indicated in the box on the right side.

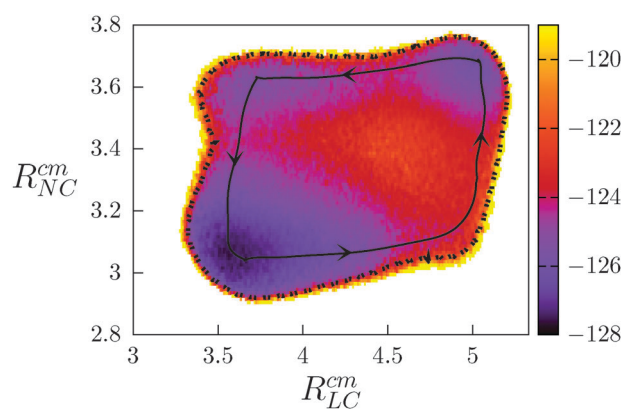


Fig. 5 Plot of the potential energy, $V(\mathbf{R}; \lambda)$, of the protein *versus* the LID-CORE and NMP-CORE center of mass distances, R_{LC}^{cm} and R_{NC}^{cm} , determined from an average over 200 realizations of the cyclic dynamics of the protein. The dashed line denotes values of the potential energy which correspond to 10% of the maximum value of the potential energy probability density. This dashed line serves to indicate the magnitude of the fluctuations in the protein potential energy in the course of enzyme catalytic cycles. The solid line with arrows shows the minimum energy path on this surface as the enzyme executes a SB cycle.

shown in Fig. 4. The minimum energy path, shown as a solid line, is consistent with the states visited in the SB mechanism.

4.2 Influence of hydrodynamic interactions

In solution, the motion of the protein beads perturbs the local solvent velocity field. These disturbances propagate through the fluid by collective solvent viscous modes giving rise to long-range hydrodynamic interactions among the beads. Since MPC dynamics conserves mass, momentum and energy, these hydrodynamic interactions are accurately described. Simulations of polymer collapse,^{53–56} colloidal suspensions⁵⁷ and molecular machine dynamics¹⁵ using MPC dynamics have indicated the importance of such long-range hydrodynamic interactions for these systems.

In order to investigate the influence of hydrodynamic interactions on the catalytic cycles of AKE, we performed a series of simulations where such interactions were absent. Simulations without hydrodynamic interactions are easily carried out in the context of our formalism by replacing MPC dynamics with an alternative collision rule that destroys solvent correlations.⁵⁸ In this rule the explicit solvent particles are replaced by a heat bath. Each protein bead is coupled to an effective solvent momentum, $\mathbf{P}$, which is chosen from a Maxwell-Boltzmann distribution with variance $m\gamma k_B T$, where γ is the average number of solvent particles per cell. The center of mass velocity of a cell containing a protein bead with velocity $\mathbf{v}_b$ used in the MPC collision step is then given by $\mathbf{V}_\xi = (m_b \mathbf{v}_b + \mathbf{P}) / (m\gamma + m_b)$. Since momentum is no longer conserved hydrodynamic interactions are destroyed and the protein beads act as independent sources of friction. This effective solvent without hydrodynamic interactions has properties similar to the MPC solvent, *e.g.*, the solvent viscosity, which depends weakly on hydrodynamic interactions, is similar in both solvents.⁵⁸ The time series of the LID-CORE and NMP-CORE distances are plotted in Fig. 6 for systems where

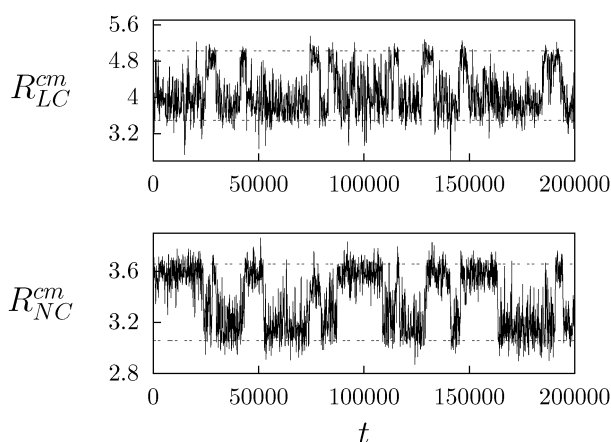


Fig. 6 Plots of the LID-CORE (top) and NMP-CORE (bottom) center of mass distances, R_{LC}^{cm} and R_{NC}^{cm} , as function of time for AKE dynamics without hydrodynamic interactions. The dashed lines are the crystallographic values of R_{LC}^{cm} (top) and R_{NC}^{cm} (bottom) for the open and closed configurations.

hydrodynamic interactions are suppressed. Visual comparison of these time series with those in Fig. 3 shows that the enzymatic cycle times are longer when hydrodynamic interactions are not taken into account. This is expected since each protein bead acts as an independent source of friction when hydrodynamic interactions are ignored and the friction that a domain experiences as it opens or closes is approximately equal to the number of beads in the domain times the bead friction. However, if hydrodynamic interactions are present the entire domain experiences a friction that is proportional to its linear size, and this leads to a much smaller friction on the domain. Consequently, domain motions in solution are more rapid. In addition, when hydrodynamic interactions are absent, the slower time scale of domain motions can lead to differences in the binding and unbinding rates since these depend on the instantaneous conformational state of the protein. The histograms of the cycle times for MPC dynamics and dynamics without hydrodynamic interactions are shown in Fig. 7. Analysis of the data shows that the average cycle time, τ_C , for MPC dynamics is $\tau_C = 15\,000 \pm 300$, while it is $\tau_C^{NH} = 18\,300 \pm 400$ when hydrodynamic interactions are suppressed. (Here and henceforth the errors refer to $\pm$ one standard deviation.) While there is a modest increase in the cycle time when hydrodynamic interactions are suppressed ($\tau_C^{NH}/\tau_C \approx 1.2$), the histograms show a broader distribution with a longer tail of long cycle times than the MPC results. The τ_C values reflect both the change in time scales of protein conformational changes and the changes in the frequencies of transition events.

Another more detailed measure of the influence of hydrodynamic interactions is provided by the decay of the autocorrelation of the LID-CORE or NMP-CORE fluctuations. These correlation functions focus explicitly on the time scale of the domain motions, which occur during the catalytic cycle. If we define the domain distance fluctuations by $\delta R_x^{cm} = R_x^{cm} - \langle R_x^{cm} \rangle$, with $x = LC, NC$, then the time autocorrelation function of this quantity is,

$$C_x(t) = \langle \delta R_x^{cm}(t) \delta R_x^{cm}(0) \rangle, \quad (7)$$

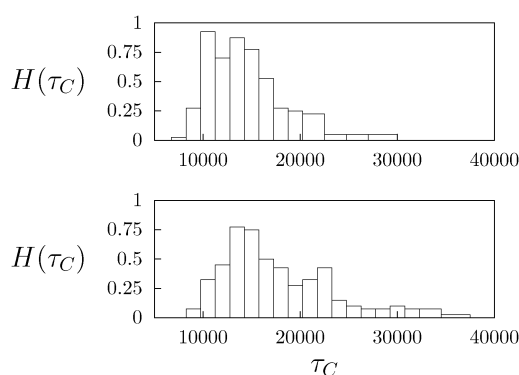


Fig. 7 Histograms $H(\tau_C)$ of the cycle times for MPC dynamics (top) and dynamics without hydrodynamic coupling (bottom).

where the angle brackets signify a time average and an average over realizations for the different types of protein dynamics. In Fig. 8 we see that the decay of $C_{LC}(t)$ for the MPC dynamics (solid line) occurs on a much shorter time scale than when hydrodynamic interactions are absent (dashed line). When the data for $C_x(t)$ are fit to an exponential form, $C_x(t) = C_x(0)e^{-t/\tau_x}$, we find the results for τ_{LC} shown in Table 1. The τ_{LC} values in the table are for the enzyme undergoing catalytic cycles by the SB and RB mechanisms and also for the enzyme in its four binding states. The ratio τ_{LC}^{NH}/τ_{LC} for the enzyme undergoing catalytic cycles has values similar to that for the ratio of catalytic cycle times discussed above. However, when this ratio is computed for the enzyme in specific binding states, we see from the table that hydrodynamic effects are considerably larger. These results reflect the different natures of the LID-CORE fluctuations in the various conformational binding states of the protein. Analogous results were obtained for the NMP-CORE distance correlation times, τ_{NC} .

We can conclude that simple Langevin models of the dynamics that ignore hydrodynamic interactions will not be able to capture all important aspects of the enzyme catalytic dynamics. Langevin models that incorporate a space dependent friction tensor using the Oseen approximation or other

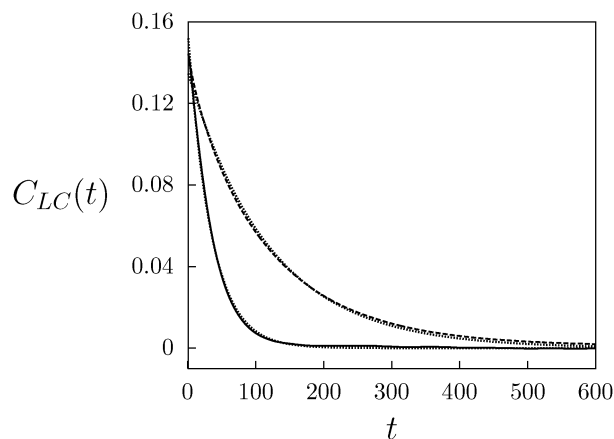


Fig. 8 Plot of the autocorrelation function of the LID-CORE fluctuations as a function of time: MPC dynamics (solid line) and dynamics without hydrodynamic interactions (dashed line) for the protein in the binding state $\lambda_3 = (1,0)$; i.e., LID-CORE open and NMP-CORE closed. The fits to exponential decay are also shown as (nearly indistinguishable) dotted lines.

Table 1 LC domain correlation times

State	$\tau_{LC} \times 10^{-2}$	$\tau_{LC}^{NH} \times 10^{-2}$	τ_{LC}^{NH}/τ_{LC}
SB	45 ± 3	60 ± 3	1.33
RB	32 ± 3	39 ± 6	1.22
λ_1	3.3 ± 0.4	4.7 ± 0.8	1.42
λ_2	0.51 ± 0.01	1.40 ± 0.02	2.74
λ_3	0.20 ± 0.01	0.51 ± 0.02	3.53
λ_4	0.34 ± 0.01	1.20 ± 0.01	2.55

approximate forms will be able to capture a portion of these effects but are not quantitatively accurate.¹⁵

4.3 Comparisons with random binding mechanism

It is interesting to compare the catalytic dynamics of AKE when it visits the four binding states in a regular order, $\lambda_1 \rightarrow \lambda_2 \rightarrow \lambda_3 \rightarrow \lambda_4 \rightarrow \lambda_1$, with the dynamics when the opening or closing of a domain does not depend on the state of the other domain. This leads to an irregular pattern of binding states. We have carried out simulations of the AKE dynamics using the transition probabilities in eqn (6). Fig. 9 shows histograms $H(R_{LC}^{cm})$ and $H(R_{NC}^{cm})$ of the LID-CORE and NMP-CORE distances, R_{LC}^{cm} and R_{NC}^{cm} determined from the enzymatic dynamics. The SB and RB models are compared in the figure, and there are differences in the histograms for these two models. In particular, the closed conformations are more highly favored when there is random binding than when there is sequential binding. This feature can be understood from the fact that in the RB model there are additional routes to reach

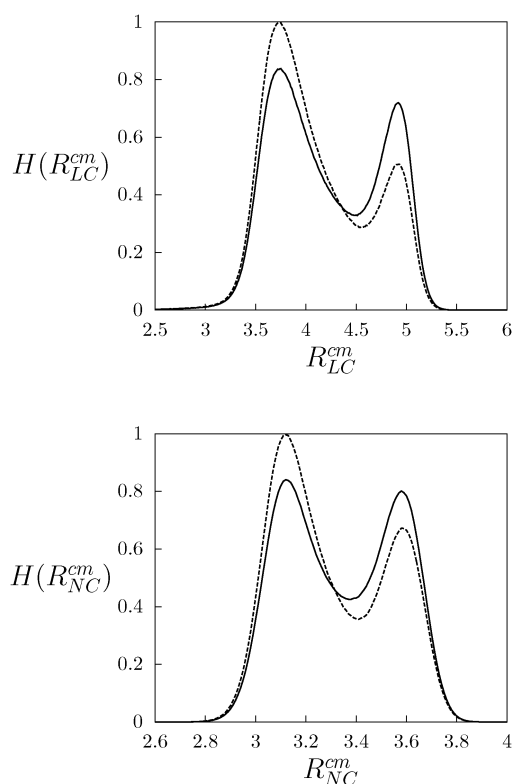


Fig. 9 Histograms $H(R_{LC}^{cm})$ (top) and $H(R_{NC}^{cm})$ (bottom) of the LID-CORE and NMP-CORE distances: SB model (solid line), RB model (dashed line) determined from MPC dynamics.

the more stable closed conformation, since the unidirectional cyclic nature of the SB mechanism is destroyed.

4.4 Transport properties

The transport properties of AKE, such as its center of mass diffusion coefficient and orientational relaxation, are influenced both by the conformational dynamics of the protein as it executes catalytic cycles and hydrodynamic interactions with the solvent. The diffusion coefficient, D , was estimated from the slope of the mean square displacement (MSD) versus time t , $MSD(t) = \langle |\mathbf{R}_P^{cm}(t) - \mathbf{R}_P^{cm}(0)|^2 \rangle = 6Dt$, where $\mathbf{R}_P^{cm}$ is the center of mass of the protein. A crude estimate of the diffusion coefficient can be obtained by assuming that AKE is a sphere with an effective radius R_P . The diffusion coefficient may then be written approximately as a sum of microscopic and hydrodynamic contributions,⁵⁹ $D \approx D_m + D_h$, where D_m is a microscopic contribution and D_h is a hydrodynamic contribution. Using expressions for D obtained for a network of beads,^{60,61} we may write $D_m = D_0/N_b = k_B T/N_b \zeta_0$, where D_0 is the single bead diffusion coefficient,^{22,23} N_b is the number of beads in AKE and ζ_0 is the single bead friction coefficient, and $D_h = k_B T/\zeta_h$ with $\zeta_h = 6\pi\eta R_P$. Since D_m scales as R_P^{-3} and D_h scales as R_P^{-1} , the diffusion coefficient will be dominated by the hydrodynamic contribution. Assuming Oseen interactions, the effective radius is given by conformation-dependent factor times the average radius of gyration.⁶⁰

In Table 2 we list several properties of AKE, including the diffusion coefficients, the average radius of gyration $\overline{R_g}$ and the ratio of $\overline{R_g}$ to the effective radius R_P . As expected from protein size considerations, the diffusion coefficients for the SB and RB models are approximately the same, D for the fully open conformation has the smallest value while the fully closed conformation has the largest value of D . Consistent with these D values, the average radius of gyration, $\overline{R_g}$, is approximately the same for the SB and RB mechanisms. This is reasonable since in both cases the protein visits all four conformational binding states.⁶² The results for the fully open and closed protein conformations are also in accord with expectations: $\overline{R_g}$ is smaller for the closed conformation than for the open conformation. The values of R_P used to compute the ratio $\overline{R_g}/R_P$ were determined from the approximate expression for D , taking the measured values of D and the viscosity η from the known analytical expressions for this transport coefficient for MPC dynamics.^{63–65} We see that the values of R_P are similar to $\overline{R_g}$ and the ratio $\overline{R_g}/R_P$ is roughly constant. We have also computed the diffusion coefficients for dynamics without hydrodynamic interactions (D^{NH}) and the table gives the values of ratio D^{NH}/D . Hydrodynamic interactions have a substantial effect on these values, with D^{NH} more than a factor of ten times smaller than D . Also, while the D values reflect the characteristic sizes of the protein in the different

Table 2 Diffusion coefficient and protein radius

State	$D \times 10^3$	D^{NH}/D	$\overline{R_g}$	$\overline{R_g}/R_P$
SB	4.9 ± 0.4	0.08 ± 0.02	2.93	0.98
RB	5.2 ± 0.5	0.08 ± 0.02	2.90	1.05
λ_1	4.2 ± 0.3	0.10 ± 0.02	3.14	0.89
λ_3	5.3 ± 0.4	0.06 ± 0.04	2.81	1.02

conformations, the values of D^{NH} only depend on the number of beads. They are approximately the same for the different conformational states and in accord with the approximate expression for D_{m} .

The rotational motion of the protein is also strongly influenced by hydrodynamic interactions. To characterize the rotational dynamics of the protein we consider a vector r in the CORE domain directed between beads corresponding to residues 4 and 111. The CORE domain beads do not undergo large relative motions when the LID and NMP domains open or close, so the dynamics of this vector provides information on the overall rotational motion of the protein. The rotational dynamics may be characterized by the decay of the autocorrelation function,

$$C_{\theta}(t) = \langle \mathbf{r}(t) \cdot \mathbf{r}(0) \rangle. \quad (8)$$

This correlation function is plotted in Fig. 10 for MPC dynamics and dynamics with hydrodynamic interactions suppressed. The correlation function exhibits exponential decay to a good approximation and the orientational correlation times were determined by fits to the form, $C_{\theta}(t) = C_{\theta}(0)\exp(-t/\tau_{\theta})$. The results are presented in Table 3. Trends similar to those seen in the translational diffusion data are also apparent in the orientational correlation times. Without hydrodynamic interactions, orientational correlation times are more than five times longer than those for MPC dynamics. In general, for MPC dynamics that includes solvent correlated motions, one would expect this time correlation function, after an initial exponential decay, to exhibit a power law long time tail due to coupling of the protein orientational and solvent velocity collective motions. For linear molecules this orientational correlation function is predicted to

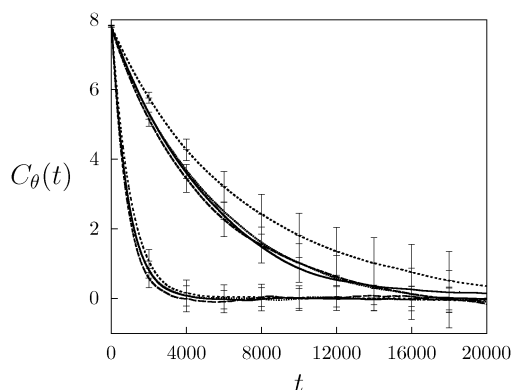


Fig. 10 Plot of the orientation autocorrelation function $C_{\theta}(t)$ versus time for MPC dynamics (rapidly decaying set of curves) and dynamics without hydrodynamic interactions (slowly decaying set of curves): (solid lines) SB model; (points) RB model; (short dashed lines) fully open conformation; (long dashed lines) fully closed conformation. The error bars denote $\pm$ one standard deviation.

Table 3 Orientational correlation times

State	τ_{θ}	$\tau_{\theta}^{\text{NH}}$	$\tau_{\theta}^{\text{NH}}/\tau_{\theta}$
SB	878 ± 9	4800 ± 900	5.47 ± 1.02
RB	875 ± 6	4900 ± 300	5.60 ± 0.34
λ_1	1003 ± 3	6800 ± 200	6.78 ± 0.20
λ_3	784 ± 4	4800 ± 300	6.12 ± 0.38

decay as $C_{\theta}(t) \approx t^{-7/2}$ with a small amplitude.⁶⁶ Our data are described by exponential decay within our statistical uncertainty, suggesting that the power law contribution is small.

5. Conclusions

The particle-based mesoscopic description of protein dynamics during enzymatic catalysis presented in this paper combines a number of different coarse graining strategies. The protein is described at a coarse grained level as a network of beads representing amino acid residues; such protein models are in common use. In addition, binding and release of ligands are modeled by changes in the nature of the network bonds instead of explicitly considering the details of the binding and unbinding events including the ligands explicitly. The solvent is described at a mesoscopic level using MPC dynamics which coarse grains the effects of many collisions into single multiparticle collision events. Inclusion of the solvent at the coarse grained particle level has a number of advantages. In particular, hydrodynamic interactions among the protein beads are taken into account without having to specify approximate forms of space dependent friction tensors, and the algorithms for simulating the dynamics are simpler than those for Langevin equation methods with space dependent friction tensors. Such hydrodynamic interactions were shown to play an important role in the dynamics of proteins executing catalytic cycles.

Certain aspects of the description are easily generalized. For example the stochastic rules that specify binding and unbinding of ligands without explicitly considering these species can be replaced by more microscopic descriptions that explicitly include these species and account for their diffusion through the MPC solvent and binding with the enzyme.¹⁴ The simple coupling of the solvent to the enzyme considered in this paper, where the protein beads are included in the MPC collision step, can be replaced by specific protein bead-solvent interactions that model hydrophobic or hydrophilic portions of the protein. In this way exclusion of solvent from parts of the protein, in particular in the binding pocket where reactions take place, can be modeled. Consequently, coarse-grained models of protein enzymatic dynamics of the sort described in this paper and their generalizations provide a way to study solvent effects on the dynamics of proteins acting as catalysts for biochemical reactions.

Appendix: AKE potential function

For those bonds in the $\mathbf{B}^{\text{e}}$ set, we use elastic network potentials since these bonds are common to both open and closed configurations and not highly strained. The potential energy takes the form

$$V_n(r_n | R_{n,\lambda}^{\text{b}}) = \frac{1}{2}k(r_n - R_{n,\lambda}^{\text{b}})^2, \text{ if } n \in \mathbf{B}^{\text{e}} \quad (9)$$

where r_n is the internuclear distance of bond n , k is the force constant that is common to all elastic network bonds and $\lambda = \lambda_1 \equiv o, \lambda_3 \equiv c$.

For bonds $n \in \mathbf{B}^{\text{s}}$, we use an attractive potential, similar to that proposed earlier,¹⁷ which has been used successfully to model nonlocal native attractive interactions in proteins. In

our implementation, it depends on the parameter χ_z that enters in the specification of the binding state λ of the protein. Since a change in the binding state is associated with a specific type of domain opening or closing, the bonds in the set $\mathbf{B}^s$ may be further classified as LID-CORE bonds or NMP-CORE bonds. We use a parameter χ_z^n in the bond potential functions to determine the nature of that bond depending on the binding state. For example, if a bond n is a LID-CORE bond, then eqn (10)–(12) below are applied with $\chi_z^n = \chi_{LC}^n$.

For bonds in the $\mathbf{B}^s$ set from the fully closed configuration we have

$$V_n(r_n | R_{n,c}^b, \chi_z^n) = c\varepsilon \begin{cases} 5 \left(\frac{R_{n,c}^b}{r_n} \right)^{12} - 6 \left(\frac{R_{n,c}^b}{r_n} \right)^{10} + 1 & \chi_z^n = 0 \\ 0, & \chi_z^n = 1. \end{cases} \quad (10)$$

For bonds in the $\mathbf{B}^s$ set from the fully open configuration we take

$$V_n(r_n | R_{n,o}^b, \chi_z^n) = c\varepsilon \min \left\{ \chi_z^n, 5 \left(\frac{R_{n,o}^b}{r_n} \right)^{12} - 6 \left(\frac{R_{n,o}^b}{r_n} \right)^{10} + 1 \right\}, \quad (11)$$

if $r_n > \sqrt{5/6} R_{n,c}^b$, and

$$V_n(r_n | R_{n,o}^b, \chi_z^n) = c\varepsilon \begin{cases} 5 \left(\frac{R_{n,c}^b}{r_n} \right)^{12} - 6 \left(\frac{R_{n,c}^b}{r_n} \right)^{10} + 1, & \chi_z^n = 1 \\ 0, & \chi_z^n = 0 \end{cases} \quad (12)$$

if $r_n \leq \sqrt{5/6} R_{n,c}^b$. Here ε is the common energy parameter used later in the specification of dimensionless units and the parameter c is a constant used to gauge the strength of the interactions.

Finally, knowing the bond potentials for the open and closed configurations, we need to be able to interpolate between these two conformations in order to represent all intermediate bonding states of the protein in the enzymatic cycle. For this purpose we use the empirical valence bond combinations rule to obtain the final bond potentials:¹⁶

$$V_n^p(r_n; \chi_z^n) = \frac{1}{2} (V_n(r_n | R_{n,o}^b, \chi_z^n) + (V_n(r_n | R_{n,c}^b, \chi_z^n) + \Delta V) - \frac{1}{2} \{ (V_n(r_n | R_{n,o}^b, \chi_z^n) - V_n(r_n | R_{n,c}^b, \chi_z^n) - \Delta V)^2 + 4\Delta^2 \}^{1/2}), \quad (13)$$

although there are other methods for combining such potentials.^{68,69} The parameter ΔV is introduced to vary the relative stability of the two states and the constant Δ determines the magnitude of the energy barrier separating the states.

Acknowledgements

The work of R.K. was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada and the Alexander von Humboldt Foundation.

C.E. acknowledges support from Decanato de Investigación, Universidad Nacional Experimental del Táchira, San Cristóbal, Venezuela. Computations were performed on the GPC supercomputer at the SciNet HPC Consortium.⁶⁷ SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund-Research Excellence; and the University of Toronto.

References

- 1 B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular Biology of the Cell*, Garland, New York, 4th edn, 1994.
- 2 J. A. McCammon and M. Karplus, *Annu. Rev. Phys. Chem. Soc.*, 1980, **131**, 24–45.
- 3 S. A. Adcock and J. A. McCammon, *Chem. Rev.*, 2006, **106**, 1589–1615.
- 4 D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan and W. Wriggers, *Science*, 2010, **330**, 341–346.
- 5 V. A. Voelz, G. R. Bowman, K. Beauchamp and V. S. Pande, *J. Am. Chem. Soc.*, 2010, **132**, 1526–1528.
- 6 I. Bahar and A. J. Rader, *Curr. Opin. Struct. Biol.*, 2010, **15**, 586–592.
- 7 G. A. Voth, *Coarse-graining of Condensed Phase and Biomolecular Systems*, CRC Press, New York, Boca Raton, 2008.
- 8 S. O. Nielsen, C. F. Lopez, G. Srinivas and M. L. Klein, *J. Phys.: Condens. Matter*, 2004, **16**, R481–R512.
- 9 M. Venturoli, M. M. Sperotto, M. Kraneburg and B. Smit, *J. Phys.: Condens. Matter*, 2006, **437**, 1–54.
- 10 H. Flechsig and A. S. Mikhailov, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 20863–20869.
- 11 A. Malevanets and R. Kapral, *J. Chem. Phys.*, 1999, **110**, 8605–8613.
- 12 A. Malevanets and R. Kapral, *J. Chem. Phys.*, 2000, **112**, 7260–7269.
- 13 V. Tozzini and J. A. McCammon, *Coarse-graining of Condensed Phase and Biomolecular Systems*, CRC Press, New York, 1st edn, 2008.
- 14 More refined models, where the diffusive, binding and unbinding dynamics of substrates and products are explicitly taken into account and explicitly solvent-bead interactions are included, can be constructed. Such a model was constructed to study the enzymatic dynamics of phosphoglycerate kinase [P. Inder, J. M. Schofield and R. Kapral, unpublished].
- 15 A. Cressman, Y. Togashi, A. S. Mikhailov and R. Kapral, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2008, **77**, 050901(R).
- 16 P. Maragakis and M. Karplus, *J. Mol. Biol.*, 2005, **352**, 807–822.
- 17 K. Okazaki, N. Koga, S. Takada, J. Onuchic and P. Wolynes, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 11844–11849.
- 18 M. M. Tirion, *Phys. Rev. Lett.*, 1996, **77**, 1905–1908.
- 19 N. Gō, *Annu. Rev. Biophys. Bioeng.*, 1983, **12**, 182–210.
- 20 O. Miyashita, J. N. Onuchic and P. G. Wolynes, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 12570–12575.
- 21 J.-W. Chu and G. A. Voth, *Biophys. J.*, 2007, **93**, 3860–3871.
- 22 R. Kapral, *Adv. Chem. Phys.*, 2008, **140**, 89–146.
- 23 G. Gompper, T. Ihle, D. M. Kroll and R. G. Winkler, *Adv. Polym. Sci.*, 2009, **221**, 1–87.
- 24 A. Malevanets and J. M. Yeomans, *Europhys. Lett.*, 2000, **52**, 231–237.
- 25 A similar model was constructed for a nanoscale swimmer whose dynamics was described deterministically in the overdamped limit [T. Sakaue, R. Kapral and A. S. Mikhailov, *Eur. Phys. J. B*, 2010, **75**, 381–387].
- 26 P. Dzeja and A. Terzic, *Int. J. Mol. Sci.*, 2009, **10**, 1729–1772.
- 27 J. A. Hanson, K. Duderstadt, L. P. Watkins, S. Bhattacharyya, J. Brokaw, J.-W. Chu and H. Yang, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 18055–18060.
- 28 Y. Santosoa, C. M. Joyceb, O. Potapovab, L. L. Restea, J. Hohlbeina, J. P. Torellaa, N. D. F. Grindleyb and

- A. N. Kapanidisa, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 715–720.
- 29 A. Korkut and W. A. Hendrickson, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 15673–15678.
- 30 J. Aden and M. Wolf-Watz, *J. Am. Chem. Soc.*, 2007, **129**, 14003–14012.
- 31 P. Whitford, O. Miyashita, Y. Levy and J. Onuchic, *J. Mol. Biol.*, 2007, **366**, 1661–1671.
- 32 P. Whitford, S. Gosavi and J. Onuchic, *J. Biol. Chem.*, 2008, **283**, 2042–2048.
- 33 K. Okazaki and S. Takada, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 11182–11187.
- 34 H. Krishnamurthy, H. Lou, A. Kimple, C. Vieille and R. I. Cukier, *Proteins: Struct., Funct., Bioinf.*, 2005, **58**, 88–100.
- 35 H. Lou and R. I. Cukier, *J. Phys. Chem. B*, 2006, **110**, 24121–24137.
- 36 P. Kern, R. M. Brunne and G. Folkers, *J. Comput.-Aided Mol. Des.*, 1994, **8**, 367–388.
- 37 K. Arora and C. L. Brooks, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 18496–18501.
- 38 C. Snow, G. Qi and S. Hayward, *Proteins: Struct., Funct., Bioinf.*, 2007, **67**, 325–337.
- 39 N. A. Temiz, E. Meirovitch and I. Bahar, *Proteins: Struct., Funct., Bioinf.*, 2004, **57**, 468–480.
- 40 O. Beckstein, E. J. Denning, J. R. Perilla and T. B. Woolf, *J. Mol. Biol.*, 2009, **394**, 160–176.
- 41 Q. Ruan, K. Ruan, C. Balny, M. Glaser and W. W. Mantulin, *Biol. Chem.*, 2001, **40**, 14706–14714.
- 42 D. C. Fry, S. A. Kuby and A. S. Mildvan, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 907–911.
- 43 E. Bae and G. N. Phillips, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 2132–2137.
- 44 M. A. Sinev, E. V. Sineva, V. Ittah and E. Haas, *Biochemistry*, 1996, **35**, 6425–6435.
- 45 C. R. Sanders, G. Tian and M. D. Tsai, *Biochemistry*, 1989, **28**, 9028–9043.
- 46 X. R. Sheng, X. Li and X. M. Pan, *J. Biol. Chem.*, 1999, **274**, 22238–22242.
- 47 C. Vonnrhein, G. J. Schlaunderer and G. E. Schulz, *Structure (London)*, 1995, **3**, 483–490.
- 48 C. W. Mueller, G. J. Schlaunderer, J. Reinstein and G. E. Schulz, *Structure (Cambridge, MA, U.S.)*, 1996, **4**, 147–156 (PDB ID: 4AKE).
- 49 C. W. Mueller and G. E. Schulz, *J. Mol. Biol.*, 1992, **224**, 159–177 (PDB ID: 1AKE).
- 50 T. Ihle and D. M. Kroll, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 2001, **63**, 020201.
- 51 T. Ihle and D. M. Kroll, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 2003, **67**, 066705.
- 52 For reference, a single enzymatic cycle with an average duration of 15000 τ in a solvent of approximately 5×10^5 particles takes approximately 4.4 hours on a single 2.53 GHz Intel Xeon CPU E5540.
- 53 N. Kikuchi, A. Gent and J. M. Yeomans, *Eur. Phys. J. E: Soft Matter Biol. Phys.*, 2002, **9**, 63–66.
- 54 N. Kikuchi, J. F. Ryder, C. M. Pooley and J. M. Yeomans, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2005, **71**, 061804.
- 55 S. H. Lee and R. Kapral, *J. Chem. Phys.*, 2006, **124**, 214901.
- 56 C. Echeverria and R. Kapral, *J. Chem. Phys.*, 2010, **132**, 104902.
- 57 J. T. Padding and A. A. Louis, *Phys. Rev. Lett.*, 2004, **93**, 220601.
- 58 M. Ripoll, R. G. Winkler and G. Gompper, *Eur. Phys. J. E: Soft Matter Biol. Phys.*, 2007, **23**, 349–354.
- 59 J. T. Hynes, R. Kapral and M. Weinberg, *J. Chem. Phys.*, 1979, **70**, 1456–1467.
- 60 B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Modern Theory of Polymer Solutions*, Garland, New York, 1st edn, 1994.
- 61 K. Mussawisade, M. Ripoll, R. G. Winkler and G. Gompper, *J. Chem. Phys.*, 2005, **123**, 144905.
- 62 The cyclic dynamics in the SB mechanism is asymmetric since the forward and reverse cycles are not equivalent. In principle, this implies that the AKE protein executing such dynamics could “swim” [E. Purcell, *Am. J. Phys.*, 1977, **45**, 3–11.]. Such an effect could manifest itself in enhanced diffusion for the protein. However, the fact that the SB and RB mechanisms yield nearly identical diffusion coefficients indicates that if such an effect operates it is very small.
- 63 T. Ihle, E. Tüzel and D. M. Kroll, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2005, **72**, 046707.
- 64 N. Kikuchi, C. M. Pooley, J. F. Ryder and J. M. Yeomans, *J. Chem. Phys.*, 2003, **119**, 6388–6396.
- 65 C. M. Pooley and J. M. Yeomans, *J. Phys. Chem. B*, 2005, **109**, 6505–6513.
- 66 F. Garisto and R. Kapral, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1974, **10**, 309–318.
- 67 C. Loken, *et al.*, *J. Phys.: Conf. Ser.*, 2010, **256**, 012026.
- 68 Y.-G. Chen and G. Hummer, *J. Am. Chem. Soc.*, 2007, **129**, 2414–2415.
- 69 W. Zheng, C. A. Brooks and G. Hummer, *Proteins: Struct., Funct., Bioinf.*, 2007, **69**, 43–57.