

## **BDflex: A method for efficient treatment of molecular flexibility in calculating protein-ligand binding rate constants from Brownian dynamics simulations**

Nicholas Greives and Huan-Xiang Zhou

Citation: *J. Chem. Phys.* **137**, 135105 (2012); doi: 10.1063/1.4756913

View online: <http://dx.doi.org/10.1063/1.4756913>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v137/i13>

Published by the [American Institute of Physics](#).

---

### **Additional information on J. Chem. Phys.**

Journal Homepage: <http://jcp.aip.org/>

Journal Information: [http://jcp.aip.org/about/about\\_the\\_journal](http://jcp.aip.org/about/about_the_journal)

Top downloads: [http://jcp.aip.org/features/most\\_downloaded](http://jcp.aip.org/features/most_downloaded)

Information for Authors: <http://jcp.aip.org/authors>

## **ADVERTISEMENT**



**AIP Advances**

Special Topic Section:  
**PHYSICS OF CANCER**

Why cancer? Why physics? [View Articles Now](#)

# BDflex: A method for efficient treatment of molecular flexibility in calculating protein-ligand binding rate constants from Brownian dynamics simulations

Nicholas Greives and Huan-Xiang Zhou<sup>a)</sup>

*Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306, USA*

(Received 9 July 2012; accepted 18 September 2012; published online 5 October 2012)

A method developed by Northrup *et al.* [J. Chem. Phys. **80**, 1517 (1984)] for calculating protein-ligand binding rate constants ( $k_a$ ) from Brownian dynamics (BD) simulations has been widely used for rigid molecules. Application to flexible molecules is limited by the formidable computational cost to treat conformational fluctuations during the long BD simulations necessary for  $k_a$  calculation. Here, we propose a new method called BDflex for  $k_a$  calculation that circumvents this problem. The basic idea is to separate the whole space into an outer region and an inner region, and formulate  $k_a$  as the product of  $k_E$  and  $\bar{\eta}_d$ , which are obtained by separately solving exterior and interior problems.  $k_E$  is the diffusion-controlled rate constant for the ligand in the outer region to reach the dividing surface between the outer and inner regions; in this exterior problem conformational fluctuations can be neglected.  $\bar{\eta}_d$  is the probability that the ligand, starting from the dividing surface, will react at the binding site rather than escape to infinity. The crucial step in reducing the determination of  $\bar{\eta}_d$  to a problem confined to the inner region is a radiation boundary condition imposed on the dividing surface; the reactivity on this boundary is proportional to  $k_E$ . By confining the ligand to the inner region and imposing the radiation boundary condition, we avoid multiple-crossing of the dividing surface before reaction at the binding site and hence dramatically cut down the total simulation time, making the treatment of conformational fluctuations affordable. BDflex is expected to have wide applications in problems where conformational fluctuations of the molecules are crucial for productive ligand binding, such as in cases where transient widening of a bottleneck allows the ligand to access the binding pocket, or the binding site is properly formed only after ligand entrance induces the closure of a lid. © 2012 American Institute of Physics. [<http://dx.doi.org/10.1063/1.4756913>]

## I. INTRODUCTION

Protein-ligand binding is a central part of many biochemical processes. The bimolecular association rate constant ( $k_a$ ) is essential for characterizing these processes. In 1984, Northrup, Allison, and McCammon<sup>1</sup> developed a method for calculating  $k_a$  from Brownian dynamics (BD) simulations. This method has inspired improvements<sup>2</sup> and alternative algorithms.<sup>3</sup> With some exceptions<sup>4</sup> among the many applications, the molecules are treated as rigid bodies undergoing translational (and possibly rotational) diffusion. Treating conformational fluctuations during the long BD simulations necessary for  $k_a$  calculation amounts to formidable computational cost. On the other hand, conformational fluctuations are necessary for productive ligand binding in many cases, such as when ligand access to the binding pocket requires transient widening of a bottleneck,<sup>5–8</sup> or when the binding site is properly formed only after ligand-induced lid closure.<sup>9–13</sup> In these cases, a rigid treatment precludes a realistic modeling of the binding events.<sup>8,14</sup> Here, we propose a new method, called BDflex, that allows conformational fluctuations to be efficiently treated in  $k_a$  calculation from BD simulations.

In the method of Northrup *et al.*,<sup>1</sup>  $k_a$  is formulated as

$$k_a = k_D(b)\eta(b), \quad (1.1)$$

where  $k_D(b)$  is the diffusion-controlled rate constant for the ligand to reach the spherical surface that has radius  $b$  and encloses the protein, and  $\eta(b)$  is the “capture” probability, i.e., the probability that the ligand, starting from the “ $b$ ” surface, will react at the binding site rather than escape to infinity.<sup>15</sup>  $k_D(b)$  can be found analytically (equal to  $4\pi Db$ , where  $D$  is the protein-ligand relative diffusion constant, in the absence of a protein-ligand interaction potential outside the  $b$  surface).  $\eta(b)$  has to be obtained from BD simulations. Because one cannot determine whether the ligand has escaped to infinity, an outer spherical surface is introduced, and absorption at this surface is used to replace the condition of escaping to infinity. A relation between the capture probability obtained under this altered outer boundary condition and  $\eta(b)$  allows the latter to be found.

Recently, we derived a more general form of Eq. (1.1).<sup>16</sup> Instead of a spherical surface, we find that, for any surface enclosing the protein, one has

$$k_a = - \oint_{\text{enclosing surface}} d\mathbf{s} \cdot \mathbf{J}_E(\mathbf{r})\eta(\mathbf{r}), \quad (1.2a)$$

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: hzhou4@fsu.edu.

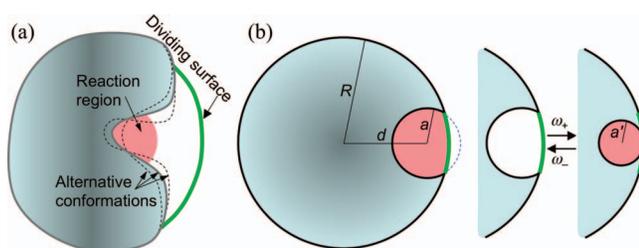


FIG. 1. Models for illustrating the basic idea and implementation of BDflex. (a) The dividing surface separating the inner region from the outer region. The inner region consists of the binding pocket and a buffer, with an inner boundary formed by the portion of the protein surface whose conformational dynamics is closely coupled to the diffusional motion of the incoming ligand. (b) A model system used to illustrate the implementation of BDflex. The binding pocket either is rigid or stochastically switches between an open, non-reactive conformation and a closed, reactive conformation.

where  $\mathbf{r}$  is the position vector of the ligand,  $\mathbf{J}_E(\mathbf{r})$  is the diffusive flux density of an exterior problem, and  $\eta(\mathbf{r})$  is the capture probability starting at position  $\mathbf{r}$ . Specializing the enclosing surface to  $r = b$  leads to Eq. (1.1). For the purpose of this paper, a surface enclosing just the protein molecule plus an “inner” region will be chosen (Fig. 1). The inner region is around the binding site and includes all the positions where translational/rotational diffusion of the ligand and conformational fluctuations of the molecules are strongly coupled. Most of the enclosing surface thus overlaps with the protein surface; the rest, defining the outside border of the inner region, will be referred to as the dividing surface between the inner region and the “outer” region. For the exterior problem, a reflecting boundary condition (i.e.,  $\mathbf{n} \cdot \mathbf{J}_E(\mathbf{r}) = 0$ ) applies to the part of the enclosing surface that overlaps with the protein surface, whereas an absorbing boundary condition (i.e., the pair distribution function = 0) applies to the dividing surface. Therefore, Eq. (1.2a) can be written as

$$k_a = - \int_{\text{dividing surface}} d\mathbf{s} \cdot \mathbf{J}_E(\mathbf{r}) \eta(\mathbf{r}). \quad (1.2b)$$

Equation (1.2b) is the starting point for the development of BDflex. The determination of  $\eta(\mathbf{r})$  requires considering both the inner and outer regions. The breakthrough of BDflex is to make the determination of  $\eta(\mathbf{r})$  as the solution to an interior problem, with the effect of the outer region captured by a boundary condition on the dividing surface. The calculation of  $k_a$  then involves separately solving the exterior and interior problems, and the treatment of conformational fluctuations is restricted to the interior problem, and can thus be efficient.

A number of ideas have been proposed to break the calculation of  $k_a$  into separate exterior and interior problems to obtain theoretical results for specific cases. We first achieved a breakup of the  $k_a$  calculation by introducing the approximation that the diffusive flux density is uniform over the dividing surface, leading to an interior problem with the equilibrium boundary value for the pair distribution function on the dividing surface.<sup>17</sup> The uniform-flux approximation was inspired by earlier work of Shoup *et al.*<sup>18</sup> and can be justified when the dividing surface spans a relatively small area. Using the same idea, we further treated the situation where a lid stochas-

tically opens and closes the entrance to the binding site. This situation will be referred to as “gated access.”

In recent work, we developed a model for the binding of a protein to a specific site on DNA that is facilitated by non-specific binding.<sup>16</sup> The solution of the model was based on Eq. (1.2a) and involved separately solving an exterior and an interior problem. The exterior problem deals with the diffusion-controlled nonspecific binding to the DNA surface, and the interior problem yields  $\eta(\mathbf{r})$  by modeling the one-dimensional search for the specific site on the DNA surface. The coupling of the latter problem to the outer region is accounted for by allowing for dissociation from the DNA surface, which serves to reduce  $\eta(\mathbf{r})$ . This formalism allowed us to treat the problem that the nonspecifically bound protein can switch between an active conformation (which can be recognized by the specific site) and an inactive conformation.

Very recently Berezhkovskii *et al.*<sup>19</sup> considered the case where the binding pocket is shaped like a tunnel. By assuming that the distribution function equilibrates quickly over each cross section of the tunnel, a radiation-type of boundary condition was derived for the interior problem of one-dimensional diffusion along the tunnel axis. We further extended this idea to treat models of conformational fluctuations, including gated access and gating binding pocket (i.e., ligand-induced lid closure).<sup>20</sup>

Here, we demonstrate that BDflex provides a general formalism for  $k_a$  calculation by separately solving the exterior and interior problems via BD simulations. This breakup allows for efficient treatment of conformational fluctuations. The rest of the paper is organized as follows. Section II presents the approximations that underlie the BDflex algorithm, followed by validation of these approximations in Sec. III on several model problems with conformational fluctuations. We describe the implementation of BDflex and illustrate its use in Sec. IV, and end with some concluding remarks in Sec. V.

## II. THE CORE IDEA OF BDflex: SEPARATING INNER AND OUTER REGIONS

### A. General definition of the rate constant

Consider a ligand diffusing around a protein. The protein-ligand pair distribution function  $G(\mathbf{r})$  satisfies the steady-state Smoluchowski equation

$$\nabla \cdot \{ D e^{-\beta U_0(\mathbf{r})} \nabla [ e^{\beta U_0(\mathbf{r})} G(\mathbf{r}) ] \} = 0, \quad (2.1)$$

where  $U_0(\mathbf{r})$  is the protein-ligand interaction potential and  $\beta = 1/k_B T$ . The outer boundary condition is

$$G(\mathbf{r}) = 1 \text{ as } r \equiv |\mathbf{r}| \rightarrow \infty. \quad (2.2)$$

The diffusive flux density is

$$\mathbf{J}(\mathbf{r}) = - D e^{-\beta U_0(\mathbf{r})} \nabla [ e^{\beta U_0(\mathbf{r})} G(\mathbf{r}) ]. \quad (2.3)$$

Ligand binding can be modeled as reaction in a binding region, or as a radiation boundary condition over a binding site on the protein surface (Fig. 1(a))

$$\begin{aligned} -\mathbf{n} \cdot \mathbf{J}(\mathbf{r}) &\equiv \mathbf{n} \cdot D e^{-\beta U_0(\mathbf{r})} \nabla [ e^{\beta U_0(\mathbf{r})} G(\mathbf{r}) ] \\ &= \kappa_0 G(\mathbf{r}), \quad \mathbf{r} \in \text{binding site}, \end{aligned} \quad (2.4a)$$

where  $\kappa_0$ , referred to as reactivity, measures how quickly the ligand, once at the binding site, is captured. The rest of the protein surface is reflecting to the ligand

$$\mathbf{n} \cdot \mathbf{J}(\mathbf{r}) = 0, \quad \mathbf{r} \in \text{elsewhere on protein surface.} \quad (2.4b)$$

The protein-ligand binding rate constant is the total flux across any surface enclosing the protein molecule

$$k_a = - \oint_{\text{enclosing surface}} ds \mathbf{n} \cdot \mathbf{J}(\mathbf{r}), \quad (2.5a)$$

where  $ds$  is the surface area element and  $\mathbf{n}$  is the unit vector along the outward normal of the enclosing surface. If we specialize the enclosing surface to the protein surface, then

$$k_a = - \int_{\text{binding site}} ds \mathbf{n} \cdot \mathbf{J}(\mathbf{r}) = \int_{\text{binding site}} ds \kappa_0 G(\mathbf{r}). \quad (2.5b)$$

For later use we note that the pair distribution function  $G(\mathbf{r})$  and the capture probability  $\eta(\mathbf{r})$  are related via

$$G(\mathbf{r}) = e^{-\beta U_0(\mathbf{r})} [1 - \eta(\mathbf{r})]. \quad (2.6)$$

We are particularly interested in including conformational fluctuations during protein-ligand binding. It is straightforward to formally include conformational degrees of freedom, collectively represented by a variable,  $q$ , in the definition of the rate constant.<sup>21</sup> The pair distribution, now denoted as  $G(\mathbf{r}, q)$ , is governed by

$$\nabla \cdot \{ D e^{-\beta U(\mathbf{r}, q)} \nabla [ e^{\beta U(\mathbf{r}, q)} G(\mathbf{r}, q) ] \} + \mathcal{L}_q G(\mathbf{r}, q) = 0, \quad (2.7)$$

where the operator  $\mathcal{L}_q$  indicates how the conformational degrees of freedom evolve over time (e.g., continuous diffusion among conformations or rate-process type of transition between discrete conformational states). The outer boundary condition becomes

$$G(\mathbf{r}, q) = \frac{e^{-\beta U_\infty(q)}}{\int dq e^{-\beta U_\infty(q)}} \text{ as } r \rightarrow \infty, \quad (2.8)$$

where  $U_\infty(q)$  denotes the interaction potential at  $r = \infty$ . The rate constant is

$$k_a = - \oint_{\text{enclosing surface}} ds dq \mathbf{n} \cdot \mathbf{J}(\mathbf{r}, q). \quad (2.9)$$

The pair distribution function and the capture probability are now related via<sup>21</sup>

$$G(\mathbf{r}, q) = \frac{e^{-\beta U(\mathbf{r}, q)}}{\int dq e^{-\beta U(\mathbf{r}, q)}} [1 - \eta(\mathbf{r}, q)]. \quad (2.10)$$

## B. Formulation of the rate constant in terms of exterior and interior problems

Previously, we showed that Eq. (2.5a) defining the rate constant is identical to Eq. (1.2a).<sup>16</sup> Using our chosen enclosing surface, this identity leads to

$$k_a = - \int_{\text{dividing surface}} ds \mathbf{n} \cdot \mathbf{J}(\mathbf{r}), \quad (2.11a)$$

$$= - \int_{\text{dividing surface}} ds \mathbf{n} \cdot \mathbf{J}_E(\mathbf{r}) \eta(\mathbf{r}). \quad (2.11b)$$

To break the calculation of  $k_a$  into separate exterior and interior problems, we make the approximation

$$\begin{aligned} & -\mathbf{n} \cdot e^{\beta U_0(\mathbf{r})} \mathbf{J}(\mathbf{r}) \int_{\text{dividing surface}} ds e^{-\beta U_0(\mathbf{r})} \\ & \approx \eta(\mathbf{r}) \left[ - \int_{\text{dividing surface}} ds \mathbf{n} \cdot \mathbf{J}_E(\mathbf{r}) \right], \\ & \mathbf{r} \in \text{dividing surface.} \end{aligned} \quad (2.12)$$

Effectively, we assume that  $\mathbf{n} \cdot e^{\beta U_0(\mathbf{r})} \mathbf{J}(\mathbf{r})$  and  $\eta(\mathbf{r})$  are uniform over the dividing surface. The surface integral on the left-hand side defines a potential of mean force,  $V_d$

$$e^{-\beta V_d} = \frac{\int_{\text{dividing surface}} ds e^{-\beta U_0(\mathbf{r})}}{\sigma_d}, \quad (2.13)$$

where  $\sigma_d$  denotes the area of the dividing surface. The surface integral on the right-hand side of Eq. (2.12) can be recognized as the diffusion-controlled rate constant,  $k_E$ , for the ligand in the outer region to reach the dividing surface. Equation (2.12) then becomes

$$-\sigma_d e^{-\beta V_d} \mathbf{n} \cdot e^{\beta U_0(\mathbf{r})} \mathbf{J}(\mathbf{r}) \approx k_E \eta(\mathbf{r}), \quad \mathbf{r} \in \text{dividing surface.} \quad (2.14)$$

Using this approximation in Eq. (2.11a), we can express the rate constant as

$$k_a = k_E \bar{\eta}_d, \quad (2.15)$$

where  $\bar{\eta}_d$  denotes the average capture probability of the ligand starting on the dividing surface

$$\bar{\eta}_d = \frac{\int_{\text{dividing surface}} ds e^{-\beta U_0(\mathbf{r})} \eta(\mathbf{r})}{\int_{\text{dividing surface}} ds e^{-\beta U_0(\mathbf{r})}}. \quad (2.16)$$

Using Eq. (2.6), the relation between the pair distribution function and the capture probability, in Eq. (2.14) and rearranging, we find

$$-\mathbf{n} \cdot D \nabla \eta(\mathbf{r}) \approx \kappa_E \eta(\mathbf{r}), \quad \mathbf{r} \in \text{dividing surface,} \quad (2.17)$$

where we have defined

$$\kappa_E \equiv \frac{k_E}{\int_{\text{dividing surface}} ds e^{-\beta U_0(\mathbf{r})}}. \quad (2.18)$$

Equation (2.17) is a radiation boundary condition for the capture probability;  $\kappa_E$  appears as the reactivity. It allows  $\eta(\mathbf{r})$  to be obtained by solving a problem confined to the inner region.

In short,  $k_a$  can be calculated according to Eq. (2.15), by separately solving an exterior problem and an interior problem. The exterior problem, with an absorbing boundary condition on the dividing surface, yields  $k_E$ . The interior problem, with a radiation boundary condition on the dividing surface, yields the average capture probability  $\bar{\eta}_d$  starting from the dividing surface.

We end this subsection by noting the connection between Eq. (2.14) and the boundary condition obtained by Berezhkovskii *et al.*<sup>19</sup> These authors considered a tunnel-shaped inner region and made the approximation

$$G(\mathbf{r}) \approx \frac{e^{-\beta U_0(\mathbf{r})} g(x)}{\sigma(x) e^{-\beta V(x)}}, \quad (2.19)$$

where  $x$  is the coordinate along the tunnel axis,  $\sigma(x)$  is the cross-sectional area, and  $V(x)$  is the potential of mean force defined in analogy to Eq. (2.13). This form of  $G(\mathbf{r})$  ensures that  $\mathbf{n} \cdot e^{\beta U_0(\mathbf{r})} \mathbf{J}(\mathbf{r})$  and  $\eta(\mathbf{r})$  are uniform over the dividing surface (denoted by  $x = x_d$ ). So for the tunnel-shaped inner region considered by Berezhkovskii *et al.*, their boundary condition at steady state,

$$\begin{aligned} \mathcal{J}g(x_d) &\equiv D\sigma(x)e^{-\beta V(x)} \left. \frac{d}{dx} \frac{g(x)}{\sigma(x)e^{-\beta V(x)}} \right|_{x=x_d} \\ &\approx k_E \left[ 1 - \frac{g(x_d)}{\sigma(x_d)e^{-\beta V(x_d)}} \right], \end{aligned} \quad (2.20)$$

is equivalent to our Eq. (2.14).

### C. Extension to conformationally fluctuating molecules

We now consider ligand binding in the presence of conformational fluctuations. We assume that the influence of conformational fluctuations on ligand diffusion is restricted to a relatively small inner region around the binding site. In the outer region, the translational/rotational diffusion of the ligand is uncoupled from the conformational fluctuations. The interaction potential can be written as

$$U(\mathbf{r}, q) = U_0(\mathbf{r}) + U_\infty(q), \quad \mathbf{r} \in \text{outer region}, \quad (2.21)$$

where  $U_0(\mathbf{r}) \rightarrow 0$  as  $r \rightarrow \infty$ . Consequently,  $k_E$  can be obtained by assuming that conformational fluctuations are absent.

To enable the breakup of  $k_a$  calculation into separate exterior and interior problems, we assume that the radiation boundary condition of Eq. (2.17) holds not only for each position on the dividing surface but also for any  $q$  while the ligand is on the dividing surface

$$-\mathbf{n} \cdot D\nabla\eta(\mathbf{r}, q) \approx \kappa_E\eta(\mathbf{r}, q), \quad \mathbf{r} \in \text{dividing surface and any } q. \quad (2.22)$$

Using this in evaluating the surface integral of Eq. (2.9), we find that  $k_a$  can again be expressed as the product of  $k_E$  and  $\bar{\eta}_d$  [i.e., Eq. (2.15)]. Now the average capture probability is given by

$$\bar{\eta}_d = \frac{\int_{\text{dividing surface}} ds dq e^{-\beta U(\mathbf{r}, q)} \eta(\mathbf{r}, q)}{\int_{\text{dividing surface}} ds dq e^{-\beta U(\mathbf{r}, q)}}. \quad (2.23)$$

### D. The case of a large dividing surface

Equation (2.14) is based on the assumption that  $\mathbf{n} \cdot e^{\beta U_0(\mathbf{r})} \mathbf{J}(\mathbf{r})$  and  $\eta(\mathbf{r})$  are uniform over the dividing surface. This assumption can be justified when the dividing surface has a relatively small area, but may have to be modified when the dividing surface is large. We recently considered such a case, where a cylindrical dividing surface encloses a long DNA molecule.<sup>16</sup> We now re-derive the  $k_a$  results of this case by applying the formalism developed here.

For this case we use cylindrical coordinates  $(\rho, \phi, z)$ . Let the cylindrical dividing surface be represented by  $\rho = R$ ,  $0 \leq \phi < 2\pi$ ,  $-L \leq z \leq L$ , where  $L$  is the half-length of the

DNA. The area of the dividing surface is  $\sigma_d = 4\pi RL$ . We define a potential of mean force [cf. Eq. (2.13)]

$$e^{-\beta V_d(z)} = \frac{\int_{\rho=R} d\xi e^{-\beta U_0(\mathbf{r})}}{2\pi R}, \quad (2.24)$$

where  $d\xi = R d\phi$ . Here, to make use of Eq. (2.11b), we assume that the following are valid on the cylindrical dividing surface:

- (1)  $\mathbf{n} \cdot e^{\beta U_0(\mathbf{r})} \mathbf{J}(\mathbf{r})$  and  $\eta(\mathbf{r})$  depend only on  $z$ , not on  $\xi$ ;
- (2)  $\int_{\rho=R} d\xi \mathbf{n} \cdot \mathbf{J}_E(\mathbf{r})$  does not depend on  $z$ ;
- (3) Equation (2.11b) holds for every  $z$  after integration over  $\xi$ .

Then Eq. (2.14) is changed to

$$-\sigma_d e^{-\beta V_d(z)} \mathbf{n} \cdot e^{\beta U_0(\mathbf{r})} \mathbf{J}(\mathbf{r}) \approx k_E \eta(\mathbf{r}), \quad \mathbf{r} \in \text{dividing surface}. \quad (2.25)$$

Using this approximation in Eq. (2.11b), we again find that the rate constant can be expressed in the form of Eq. (2.15). Here, the average capture probability is

$$\bar{\eta}_d = \int_{\rho=R} dz d\xi e^{-\beta U_0(\mathbf{r})} \eta(\mathbf{r}) / \sigma_d e^{-\beta V_d(z)}, \quad (2.26a)$$

$$= \int_{-L}^L dz \eta_d(z) / 2L, \quad (2.26b)$$

where we have defined

$$\eta_d(z) = \frac{\int_{\rho=R} d\xi e^{-\beta U_0(\mathbf{r})} \eta(\mathbf{r})}{\int_{\rho=R} d\xi e^{-\beta U_0(\mathbf{r})}}. \quad (2.27)$$

Corresponding to Eq. (2.25), the radiation boundary condition for the capture probability on the dividing surface becomes

$$-D \left. \frac{\partial \eta(\mathbf{r})}{\partial \rho} \right|_{\rho=R} \approx \kappa_E(z) \eta_d(z), \quad (2.28)$$

where the reactivity becomes  $z$ -dependent,

$$\kappa_E(z) = \frac{k_E}{\sigma_d e^{-\beta V_d(z)}}. \quad (2.29)$$

We now focus on the interior problem. We model the DNA as a cylinder with radius  $R_-$ . The inner region is specified by  $R_- \leq \rho \leq R$ ,  $0 \leq \phi < 2\pi$ ,  $-L \leq z \leq L$ . In previous studies,<sup>16,22</sup> we considered the situation where the width of the inner region,  $R - R_- \equiv \varepsilon$ , is small. An approximation was proposed without justification. Here, we obtain this approximation using the present formalism.

The capture probability satisfies the backward Smoluchowski equation, which in cylindrical coordinates takes the form

$$\begin{aligned} \rho^{-1} \frac{\partial}{\partial \rho} \rho D e^{-\beta U_0(\mathbf{r})} \frac{\partial \eta(\mathbf{r})}{\partial \rho} + \frac{\partial}{\partial \xi} D e^{-\beta U_0(\mathbf{r})} \frac{\partial \eta(\mathbf{r})}{\partial \xi} \\ + \frac{\partial}{\partial z} D e^{-\beta U_0(\mathbf{r})} \frac{\partial \eta(\mathbf{r})}{\partial z} = 0. \end{aligned} \quad (2.30)$$

We already introduced the assumption that  $\eta(\mathbf{r})$  is independent of  $\xi$  on the dividing surface. Now given that the width of the inner region is small we can further assume that  $\eta(\mathbf{r})$  has

a weak dependence on  $\rho$  and on  $\xi$  in the inner region. Then integrating over  $\rho$  and on  $\xi$ , Eq. (2.30) becomes

$$\varepsilon^{-1} D \left. \frac{\partial \eta(\mathbf{r})}{\partial \rho} \right|_{\rho=R} + e^{\beta V_d(z)} \frac{\partial}{\partial z} D e^{-\beta V_d(z)} \frac{\partial \eta_d(z)}{\partial z} = 0. \quad (2.31)$$

Finally, using the radiation boundary condition of Eq. (2.28), we find

$$e^{\beta V_d(z)} \frac{\partial}{\partial z} D e^{-\beta V_d(z)} \frac{\partial \eta_d(z)}{\partial z} - \frac{k_E}{\varepsilon \sigma_d e^{-\beta V_d(z)}} \eta_d(z) = 0. \quad (2.32)$$

This is the approximation proposed in the recent work.<sup>16</sup> The interior problem now amounts to solving the one-dimensional diffusion problem on the DNA surface. Instead of a radiation boundary condition on the dividing surface, the coupling to the outer region is contained in the second term on the left-hand side of Eq. (2.32), which models the escape from the DNA surface to the outer region. Note that the rate of escape, given by the factor in front of  $\eta_d(z)$  in the second term and to be denoted by  $\nu$ , is related to the reactivity [Eq. (2.29)] on the dividing surface via

$$\nu = \frac{\kappa_E}{\varepsilon}. \quad (2.33)$$

### III. VALIDATION OF EQ. (2.22) IN MODELS OF FLUCTUATING MOLECULES

The approximation of Eq. (2.17), applicable to the case without conformational fluctuations, is related to similar approximations proposed previously,<sup>17,19</sup> and can be justified when the area of the dividing surface is relatively small. Equation (2.22) makes a stronger approximation for the case with conformational fluctuations, and is worth some scrutiny. We do so in this section.

We use analytical results derived previously<sup>20</sup> for three models of conformational fluctuations (Fig. 2) to test the accuracy of Eq. (2.22). We now summarize the common features of these models. All of them dealt with stochastic switching between two conformational states, referred to as active and inactive and denoted by subscripts ‘‘a’’ and ‘‘i,’’ respectively. However, the meanings of the conformational states are different among the three models, and will be explained below. The pair distribution functions,  $G_g(\mathbf{r})$ , where  $g = \text{‘‘a’’}$  and ‘‘i,’’ are governed by<sup>21</sup>

$$\begin{aligned} \nabla \cdot \{ D e^{-\beta U_a(\mathbf{r})} \nabla [ e^{\beta U_a(\mathbf{r})} G_a(\mathbf{r}) ] \} \\ - \omega_-(\mathbf{r}) G_a(\mathbf{r}) + \omega_+(\mathbf{r}) G_i(\mathbf{r}) = 0, \end{aligned} \quad (3.1a)$$

$$\begin{aligned} \nabla \cdot \{ D e^{-\beta U_i(\mathbf{r})} \nabla [ e^{\beta U_i(\mathbf{r})} G_i(\mathbf{r}) ] \} \\ + \omega_-(\mathbf{r}) G_a(\mathbf{r}) - \omega_+(\mathbf{r}) G_i(\mathbf{r}) = 0, \end{aligned} \quad (3.1b)$$

where  $\omega_{\pm}(\mathbf{r})$  are the transition rates between the two conformational states. They satisfy the following detailed balance condition:<sup>21,23</sup>

$$\frac{\omega_+(\mathbf{r})}{\omega_-(\mathbf{r})} = \frac{\omega_{\infty+}}{\omega_{\infty-}} e^{-\beta[U_a(\mathbf{r}) - U_i(\mathbf{r})]}, \quad (3.2)$$

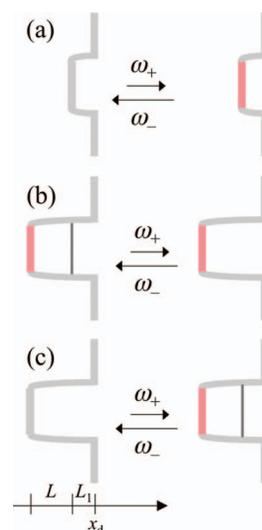


FIG. 2. Three models of conformational switching. (a) Gating binding site. The binding site switches between being a reflecting boundary to being a radiation boundary. (b) Gated access. A lid along the passageway to the binding site transiently opens and closes, allowing for and blocking the passage of the ligand. (c) Gating binding pocket. The binding site becomes reactive only when the ligand is behind the closed lid. Note that in each model the binding pocket has a cylindrical shape, and a slice along the cylindrical axis is shown. The  $x$  axis is along the cylindrical axis.

in which  $\omega_{\infty\pm}$  are the transition rates at  $r = \infty$  [where  $U_g(\mathbf{r}) = 0$ ]. In the outer region, ligand diffusion and conformational switching are assumed to be uncoupled. This means that, when  $\mathbf{r}$  is in the outer region,  $U_g(\mathbf{r})$  are independent of  $g$  and will be denoted as  $U_0(\mathbf{r})$ , and  $\omega_{\pm}(\mathbf{r})$  are independent of  $\mathbf{r}$  and will have the reference to  $\mathbf{r}$  dropped.<sup>21</sup> At infinite protein-ligand separation, the boundary values of  $G_g(\mathbf{r})$  are

$$G_g(\mathbf{r}) = p_g \text{ as } r \rightarrow \infty, \quad (3.3a)$$

where

$$p_a = \omega_+ / \omega; \quad p_i = \omega_- / \omega, \quad (3.3b)$$

with  $\omega = \omega_+ + \omega_-$ .

All the three models of conformational switching have a cylindrical inner region, with the binding site located at the bottom. The previous analytical results themselves were based on approximations. These generalize Eq. (2.19) to<sup>20</sup>

$$G_g(\mathbf{r}) \approx \frac{e^{-\beta U_g(\mathbf{r})} g_g(x)}{\sigma(x) e^{-\beta V_g(x)}}, \quad (3.4)$$

for the inner region. As with Eq. (2.19), these approximations can be well justified and the resulting  $k_a$  is expected to be accurate. Here, we refer to these results as ‘‘exact.’’ The governing equations for the inner region become

$$\begin{aligned} \nabla \cdot \{ D \sigma(x) e^{-\beta V_a(x)} \nabla [ g_a(x) / \sigma(x) e^{-\beta V_a(x)} ] \} \\ - \omega_-(x) g_a(x) + \omega_+(x) g_i(x) = 0, \end{aligned} \quad (3.5a)$$

$$\begin{aligned} \nabla \cdot \{ D \sigma(x) e^{-\beta V_i(x)} \nabla [ g_i(x) / \sigma(x) e^{-\beta V_i(x)} ] \} \\ + \omega_-(x) g_a(x) - \omega_+(x) g_i(x) = 0, \end{aligned} \quad (3.5b)$$

where

$$\omega_+(x) = \frac{\int_{\text{cross section at } x} ds \omega_+(\mathbf{r}) e^{-\beta U_i(\mathbf{r})}}{\sigma(x) e^{-\beta V_i(x)}}, \quad (3.6a)$$

$$\omega_-(x) = \frac{\int_{\text{cross section at } x} ds \omega_-(\mathbf{r}) e^{-\beta U_a(\mathbf{r})}}{\sigma(x) e^{-\beta V_a(x)}}. \quad (3.6b)$$

The new detailed balance relation is

$$\frac{\omega_+(x)}{\omega_-(x)} = \frac{\omega_{\infty+}}{\omega_{\infty-}} e^{-\beta[V_a(x) - V_i(x)]}. \quad (3.7)$$

Within each model, two types of conformational switching are considered. The first, referred to as indifferent switch, was studied in earlier work.<sup>17,24</sup> In this type, ligand diffusion and conformational switching are uncoupled everywhere, even in the inner region. The second, referred to as induced switch, was introduced recently to account for the fact that protein-ligand interactions in the binding pocket often favor the switching to the active conformation.<sup>16,21,23</sup> The boundary conditions on the dividing surface that led to the “exact”  $k_a$  results<sup>20</sup> were derived by assuming uncoupled ligand diffusion and conformational switching in the outer region and an infinitesimal layer around the dividing surface. Continuity of  $g_g(x)/\sigma(x)e^{-\beta V_g(x)}$  and the flux  $\mathcal{J}g_g(x) \equiv D\sigma(x)e^{-\beta V_g(x)}\nabla[g_g(x)/\sigma(x)e^{-\beta V_g(x)}]$  then took care of any transition to induced switch in the rest of the inner region.

### A. Gating binding site

In this model (Fig. 2(a)), it is the binding site that switches between a reactive conformation and a non-reactive conformation. The former allows for ligand binding, with the binding site (at  $x = x_b$ ) serving as a radiation boundary

$$\mathcal{J}g_a(x_b) = \kappa_0 g_a(x_b). \quad (3.8a)$$

In the latter conformation, the binding site is a reflecting boundary

$$\mathcal{J}g_i(x_b) = 0. \quad (3.8b)$$

The boundary condition on the dividing surface (i.e.,  $x = x_d$ ), based on Eq. (3.4) and assuming uncoupled ligand diffusion and conformational switching in the outer region and an infinitesimal layer around the dividing surface, is<sup>20</sup>

$$\mathcal{J}g(x_d) = k_E \left[ 1 - \frac{g(x_d)}{\sigma(x_d) e^{-\beta V(x_d)}} \right], \quad (3.9a)$$

$$\mathcal{J}f(x_d) = -\omega \hat{k}_E(\omega) \frac{f(x_d)}{\sigma(x_d) e^{-\beta V(x_d)}}. \quad (3.9b)$$

Here,  $g(x)$  and  $f(x)$ , defined for  $x$  in an infinitesimal interval centered at  $x_d$ , are linear combinations of  $g_a(x)$  and  $g_i(x)$

$$g(x) = g_a(x) + g_i(x), \quad (3.10a)$$

$$f(x) = p_i g_a(x) - p_a g_i(x), \quad (3.10b)$$

and  $\hat{k}_E(s)$  is the Laplace transform of the time-dependent rate coefficient of the exterior problem [the steady-state value  $k_E = \lim_{s \rightarrow 0} s \hat{k}_E(s)$ ].

BDflex gives a different boundary condition on the dividing surface. According to Eq. (2.22), this is

$$-\frac{d\eta_g(x)}{dx} \Big|_{x=x_d} = \frac{k_E}{\sigma(x_d) e^{-\beta V(x_d)}} \eta_g(x_d), \quad (3.11)$$

where the capture probability  $\eta_g(x)$  is related to the pair distribution function  $\eta_g(x)$  via

$$g_g(x) = p_g \sigma(x) e^{-\beta V_g(x)} [1 - \eta_g(x)]. \quad (3.12)$$

Note that the capture probabilities (or, equivalently, the pair distribution functions) for the two conformational states are uncoupled in Eq. (3.11), but coupled in Eqs. (3.9). The uncoupling offers a significant practical advantage in treating the boundary condition during BD simulations (see Sec. IV). The boundary condition of Eqs. (3.9) would be very cumbersome to implement in BD simulations. The boundary condition of Eq. (3.11) nevertheless is closely related to that of Eqs. (3.9). To see this, we express Eq. (3.11) in terms of the pair distribution functions

$$\mathcal{J}g_g(x_d) = k_E \left[ p_g - \frac{g_g(x_d)}{\sigma(x_d) e^{-\beta V(x_d)}} \right]. \quad (3.13a)$$

Using the combinations of Eq. (3.10), we further transform the boundary condition to

$$\mathcal{J}g(x_d) = k_E \left[ 1 - \frac{g(x_d)}{\sigma(x_d) e^{-\beta V(x_d)}} \right], \quad (3.13b)$$

$$\mathcal{J}f(x_d) = -k_E \frac{f(x_d)}{\sigma(x_d) e^{-\beta V(x_d)}}. \quad (3.13c)$$

We now see that the BDflex boundary condition involves the further approximation of replacing  $\omega \hat{k}_E(\omega)$  by  $k_E$  in the boundary condition for the function  $f(x)$ . This replacement is strictly valid only in the limit of slow conformational transition (i.e.,  $\omega \rightarrow 0$ ). However, in the limit of fast conformational transition (i.e.,  $\omega \rightarrow \infty$ ), equilibration between the conformational states makes  $f(x) \rightarrow 0$ , and the rate constant is then solely determined by  $g(x)$ . Therefore, the BDflex boundary condition is correct in both the slow and fast limits of conformational transition, and is expected to be accurate overall. This expectation is confirmed by the results below.

Under indifferent switch, with a constant potential  $V_1$  for both conformational states in the inner region having constant cross section  $\sigma$ , the rate constant calculated based on the boundary condition of Eqs. (3.9) is<sup>20</sup>

$$\frac{1}{k_a} = \frac{1}{k_E} + \frac{1}{p_a \kappa_0 \sigma e^{-\beta V_1}} + \frac{L_1}{D \sigma e^{-\beta V_1}} + \frac{p_i \nu \coth(\nu L_1) + \omega \hat{k}_E(\omega) / D \sigma e^{-\beta V_1}}{p_a \omega \hat{k}_E(\omega) \nu \coth(\nu L_1) + \omega \sigma e^{-\beta V_1}}, \quad (3.14)$$

where  $L_1 = x_d - x_b$  and  $\nu = (\omega/D)^{1/2}$ . In Subsection III B, we will use  $k_{\text{GBS}}(L_1)$  to denote this rate constant, for binding

to a gating binding site separated by a distance  $L_1$  from the dividing surface. As noted above, the BDFlex result involves replacing  $\omega \hat{k}_E(\omega)$  by  $k_E$

$$\frac{1}{k_a} \approx \frac{1}{k_E} + \frac{1}{p_a \kappa_0 \sigma e^{-\beta V_1}} + \frac{L_1}{D \sigma e^{-\beta V_1}} + \frac{p_i \nu \coth(\nu L_1) + k_E / D \sigma e^{-\beta V_1}}{p_a k_E \nu \coth(\nu L_1) + \omega \sigma e^{-\beta V_1}}, \quad (3.15)$$

When  $\omega \rightarrow 0$ , both Eqs. (3.14) and (3.15) predict

$$\frac{1}{k_a} \rightarrow \frac{1}{p_a} \left( \frac{1}{k_E} + \frac{1}{\kappa_0 \sigma e^{-\beta V_1}} + \frac{L_1}{D \sigma e^{-\beta V_1}} \right), \quad (3.16a)$$

which gives  $k_a$  as the product of the rate constant when the binding site stays reactive and the equilibrium probability for the reactive state. In the opposite limit  $\omega \rightarrow \infty$ , both

Eqs. (3.14) and (3.15) predict

$$\frac{1}{k_a} \rightarrow \frac{1}{k_E} + \frac{1}{p_a \kappa_0 \sigma e^{-\beta V_1}} + \frac{L_1}{D \sigma e^{-\beta V_1}}, \quad (3.16b)$$

which gives  $k_a$  as the rate constant when the binding site stays reactive but has reactivity  $p_a \kappa_0$ . Figure 3 shows that the BDFlex result for  $k_a$  agrees with the “exact” result over the entire range of conformational transition rates. BDFlex does break down when  $L_1 \rightarrow 0$ ; for the set of parameters chosen the discrepancy becomes noticeable at  $L_1/a = 0.1$  ( $a$ : radius of the circular cross section of the inner region). So in implementing BDFlex the dividing surface should be separated from the gating binding site by at least a small buffer region.

Under induced switch, with a constant potential  $V_g$  for conformational state  $g$  and transition rates  $\omega_{I\pm}$  between the states while the ligand is in the inner region, the rate constant calculated based on the boundary condition of Eqs. (3.9) is<sup>20</sup>

$$\frac{1}{k_a} = \frac{1}{k_E} + \frac{1}{p_{1a} \kappa_0 \sigma e^{-\beta V_{\text{eff}}}} + \frac{L_1}{D \sigma e^{-\beta V_{\text{eff}}}} + \frac{p_i p_{1i} \left( 1 + \frac{\Delta p^2}{p_{1i}^2} \right) \nu_1 \coth(\nu_1 L_1) + \frac{p_a p_i}{p_{1a} p_{1i}} \frac{\omega \hat{k}_E(\omega)}{D \sigma e^{-\beta V_{\text{eff}}}} + \frac{2 \Delta p}{p_{1i}} \frac{\nu_1}{\sinh(\nu_1 L_1)}}{(p_a p_i / p_{1a} p_{1i}) \omega \hat{k}_E(\omega) \nu_1 \coth(\nu_1 L_1) + \omega_1 \sigma e^{-\beta V_{\text{eff}}}}, \quad (3.17)$$

where  $\omega_1 = \omega_{1+} + \omega_{1-}$ ,  $p_{1a} = \omega_{1+} / \omega_1$ ,  $p_{1i} = \omega_{1-} / \omega_1$ ,  $\Delta p = p_{1a} - p_a$ ,  $\nu_1 = (\omega_1 / D)^{1/2}$ , and

$$e^{-\beta V_{\text{eff}}} = p_a e^{-\beta V_a} + p_i e^{-\beta V_i}. \quad (3.18)$$

Again the BDFlex result is obtained by replacing  $\omega \hat{k}_E(\omega)$  with  $k_E$

$$\frac{1}{k_a} = \frac{1}{k_E} + \frac{1}{p_{1a} \kappa_0 \sigma e^{-\beta V_{\text{eff}}}} + \frac{L_1}{D \sigma e^{-\beta V_{\text{eff}}}} + \frac{p_i p_{1i} \left( 1 + \frac{\Delta p^2}{p_{1i}^2} \right) \nu_1 \coth(\nu_1 L_1) + \frac{p_a p_i}{p_{1a} p_{1i}} \frac{k_E}{D \sigma e^{-\beta V_{\text{eff}}}} + \frac{2 \Delta p}{p_{1i}} \frac{\nu_1}{\sinh(\nu_1 L_1)}}{(p_a p_i / p_{1a} p_{1i}) k_E \nu_1 \coth(\nu_1 L_1) + \omega_1 \sigma e^{-\beta V_{\text{eff}}}}. \quad (3.19)$$

When  $\omega \rightarrow 0$ , both Eqs. (3.17) and (3.19) predict

$$\frac{1}{k_a} \rightarrow \frac{1}{p_a} \left( \frac{1}{k_E} + \frac{1}{\kappa_0 \sigma e^{-\beta V_a}} + \frac{L_1}{D \sigma e^{-\beta V_a}} \right), \quad (3.20a)$$

which gives  $k_a$  as the product of the rate constant when the binding site stays reactive and the equilibrium probability for the reactive state. In the opposite limit  $\omega \rightarrow \infty$ , both Eqs. (3.17) and (3.19) predict

$$\frac{1}{k_a} \rightarrow \frac{1}{k_E} + \frac{1}{p_{1a} \kappa_0 \sigma e^{-\beta V_{\text{eff}}}} + \frac{L_1}{D \sigma e^{-\beta V_{\text{eff}}}}, \quad (3.20b)$$

which gives  $k_a$  as the rate constant when the binding site stays active but has reactivity  $p_{1a} \kappa_0$  and the interaction potential is  $V_{\text{eff}}$ . As noted previously,<sup>16,20,21</sup> relative to indifferent switch, the rate constant under induced switch has a reduced difference between the slow and fast limits of conformational tran-

sition, and the shift from the slow limit to the fast limit occurs at a lower conformational transition rate. Both effects serve to make the BDFlex result more accurate over intermediate conformational transition rates. Figure 3 shows that the BDFlex result for  $k_a$  agrees with the “exact” result over the entire range of conformational transition rates, even at  $L_1/a = 0.1$ .

## B. Gated access

In the second model of conformational switching, a lid stochastically opens and closes the entrance to the binding site (Fig. 2(b)). The active and inactive conformations correspond to the open and closed lid, respectively. The boundary condition for  $g_a(x)$  at the location of the lid (where  $x = x_1$ ), based on Eq. (3.4) and assuming that ligand diffusion and conformational switching are uncoupled all the way to an infinitesimal

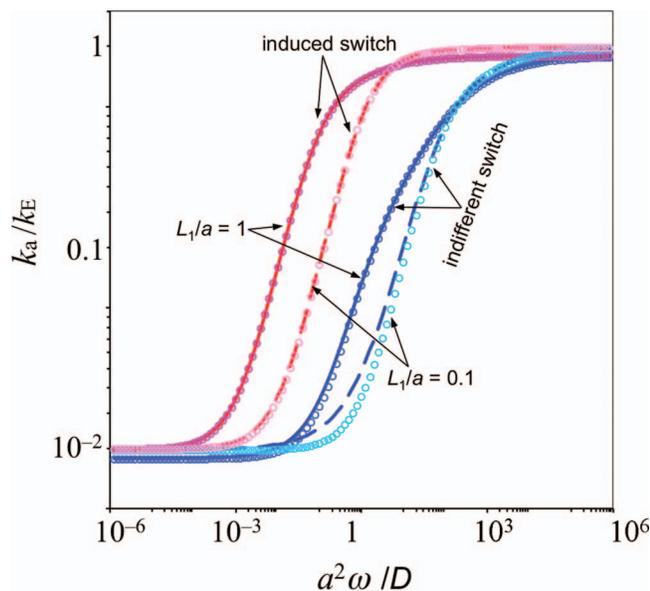


FIG. 3. Comparison of BDflex and “exact” results for the gating binding-site model. Curves represent the “exact” results and symbols represent the BDflex results. Results for two values of the distance,  $L_1$ , between the binding site and the dividing surface are shown.  $k_E = 4Da$  and  $\kappa_0 = \infty$ . For induced switch,  $\omega_+/ \omega_- = 0.01$ ,  $\omega_{I+}/ \omega_{I-} = 10$ ,  $\omega_{I-} = \omega_-$ , and  $V_1 = 0$ . Correspondingly,  $e^{-\beta V_{\text{eff}}} = 11/1.01$ . For indifferent switch,  $\omega_+/ \omega_- = 0.01$  and  $V_1 = V_{\text{eff}}$ .

layer beyond  $x = x_1$ , is<sup>20</sup>

$$\mathcal{J}g_a(x_1) = k_{\text{GBS}}(L_1) \left( 1 - \frac{g_a(x_1)}{p_a \sigma(x_1) e^{-\beta V(x_1)}} \right), \quad (3.21a)$$

where  $k_{\text{GBS}}(L_1)$  is the rate constant for binding to a gating binding site, with  $\kappa_0 = \infty$ , located at a distance  $L_1 = x_d - x_1$  from the dividing surface [see Eq. (3.14)]. For  $g_i(x)$  the boundary condition at  $x = x_1$  is reflecting

$$\mathcal{J}g_i(x_1) = 0. \quad (3.21b)$$

The binding site serves as a radiation boundary for both  $g_a(x)$  and  $g_i(x)$

$$\mathcal{J}g_a(x_b) = \kappa_0 g_a(x_b), \quad (3.22a)$$

$$\mathcal{J}g_i(x_b) = \kappa_0 g_i(x_b). \quad (3.22b)$$

Under indifferent switch, with a constant potential  $V_0$  for both conformational states when the ligand is in the binding pocket (i.e.,  $x_b \leq x \leq x_1$ ), the rate constant calculated based on the boundary condition of Eqs. (3.21) is<sup>20</sup>

$$\frac{1}{k_a} = \frac{1}{k_{\text{GBS}}(L_1)} + \frac{1}{\kappa_0 \sigma e^{-\beta V_0}} + \frac{L}{D \sigma e^{-\beta V_0}} + \frac{p_i v \coth(\nu L) + k_0/D \sigma e^{-\beta V_0}}{p_a k_0 \nu \coth(\nu L) + \omega \sigma e^{-\beta V_0}}, \quad (3.23a)$$

where  $L = x_1 - x_b$  and  $k_0 = \kappa_0 \sigma e^{-\beta V_0}$ . Note the resemblance of Eq. (3.15) to Eq. (3.23a). In the former problem, the reactivity stays at  $k_E/\sigma e^{-\beta V_1}$  at one end of the binding pocket and switches between  $\kappa_0$  and 0 at the other end [see Eqs. (3.13a) and (3.8), respectively]; in the latter problem, the reactivity stays at  $\kappa_0$  at one end of the binding pocket and switches

between  $k_{\text{GBS}}(L_1)/p_a \sigma e^{-\beta V_0}$  and 0 at the other end [see Eqs. (3.22) and (3.21), respectively]. If the potential takes a constant value  $V_1$  for  $x_1 \leq x \leq x_d$ , then  $k_{\text{GBS}}(L_1)$  is given by Eq. (3.14) with  $\kappa_0 = \infty$ . Consequently,

$$\frac{1}{k_a} = \frac{1}{k_E} + \frac{L_1}{D \sigma e^{-\beta V_1}} + \frac{1}{\kappa_0 \sigma e^{-\beta V_0}} + \frac{L}{D \sigma e^{-\beta V_0}} + \frac{p_i v \coth(\nu L_1) + \omega \hat{k}_E(\omega)/D \sigma e^{-\beta V_1}}{p_a \omega \hat{k}_E(\omega) \nu \coth(\nu L_1) + \omega \sigma e^{-\beta V_1}} + \frac{p_i v \coth(\nu L) + k_0/D \sigma e^{-\beta V_0}}{p_a k_0 \nu \coth(\nu L) + \omega \sigma e^{-\beta V_0}}. \quad (3.23b)$$

This result could have been derived by using: (1) the boundary condition of Eqs. (3.9) at  $x = x_d$ ; (2) the continuity conditions for  $g_a(x)e^{\beta V_a(x)}$  and  $\mathcal{J}g_a(x)$  and reflecting boundary condition for  $g_i(x)$  at  $x = x_1$ ; and (3) the radiation boundary conditions for  $g_a(x)$  and  $g_i(x)$  at  $x = x_b$ .

We now present the rate constant calculated according to BDflex. This would use the three sets of conditions just listed, except that Eqs. (3.9) at  $x = x_d$  are replaced by Eqs. (3.13b) and (3.13c). More specifically,  $\omega \hat{k}_E(\omega)$  is replaced by  $k_E$ . Therefore, the BDflex result is

$$\frac{1}{k_a} = \frac{1}{k_E} + \frac{L_1}{D \sigma e^{-\beta V_1}} + \frac{1}{\kappa_0 \sigma e^{-\beta V_0}} + \frac{L}{D \sigma e^{-\beta V_0}} + \frac{p_i v \coth(\nu L_1) + k_E/D \sigma e^{-\beta V_1}}{p_a k_E \nu \coth(\nu L_1) + \omega \sigma e^{-\beta V_1}} + \frac{p_i v \coth(\nu L) + k_0/D \sigma e^{-\beta V_0}}{p_a k_0 \nu \coth(\nu L) + \omega \sigma e^{-\beta V_0}}. \quad (3.24)$$

When  $\omega \rightarrow \infty$ , both Eqs. (3.23b) and (3.24) predict

$$\frac{1}{k_a} \rightarrow \frac{1}{k_E} + \frac{L_1}{D \sigma e^{-\beta V_1}} + \frac{1}{\kappa_0 \sigma e^{-\beta V_0}} + \frac{L}{D \sigma e^{-\beta V_0}}, \quad (3.25a)$$

which gives  $k_a$  as the rate constant when the lid stays open. In the opposite limit  $\omega \rightarrow 0$ , both Eqs. (3.23b) and (3.24) predict

$$\frac{1}{k_a} \rightarrow \frac{1}{p_a} \left( \frac{1}{k_E} + \frac{L_1}{D \sigma e^{-\beta V_1}} + \frac{1}{\kappa_0 \sigma e^{-\beta V_0}} + \frac{L}{D \sigma e^{-\beta V_0}} \right), \quad (3.25b)$$

which gives  $k_a$  as the product of the rate constant when the lid stays open and the equilibrium probability for the lid being open. Figure 4 shows the good agreement between the BDflex “exact” results over the entire range of conformational transition rates. BDflex still breaks down when  $L_1 \rightarrow 0$  (not shown); so again the dividing surface should be separated by a buffer from the region (the lid in the present case) undergoing conformational transition to ensure accuracy of BDflex.

Next we consider induced switch, with a constant potential  $V_g$  for conformational state  $g$  and transition rates  $\omega_{I\pm}$  between the states when the ligand is in the binding pocket. The interaction of the ligand with the binding pocket favors the lid open state (i.e.,  $p_{Ia}/p_{Ii} \gg p_a/p_i$ ). The problem can again be solved by using the three sets of conditions at  $x = x_d$ ,  $x_1$ , and

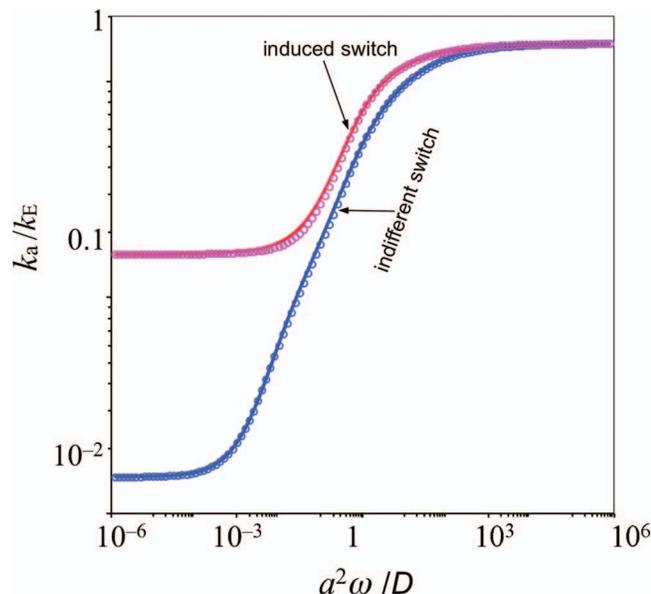


FIG. 4. Comparison of BDflex and “exact” results for the gated access model.  $L_1/a = 1$  and  $L/a = 5$ , where  $L_1$  is the distance between the lid and the dividing surface and  $L$  is the distance between the binding site and the lid.  $k_E = 4Da$  and  $\kappa_0 = 0.01D/a$ . The gating parameters for induced switch are the same as in Fig. 3; in addition  $V_1 = V_{\text{eff}}$ . For indifferent switch,  $\omega_{\pm}/\omega = 0.01$  and  $V_0 = V_1 = V_{\text{eff}}$ .

$x_b$ . The “exact” result for  $k_a$  is given by

$$\frac{1}{k_a} = \frac{1}{k_E} + \frac{L_1}{D\sigma e^{-\beta V_1}} + \frac{1}{\kappa_0\sigma e^{-\beta V_{\text{eff}}}} + \frac{L}{D\sigma e^{-\beta V_{\text{eff}}}} + \frac{p_i}{p_a} \frac{\nu \coth(\nu L_1) + \omega \hat{k}_E(\omega)/D\sigma e^{-\beta V_1}}{\omega \hat{k}_E(\omega)\nu \coth(\nu L_1) + \omega\sigma e^{-\beta V_1}} + \frac{p_{\text{li}}}{p_{\text{la}}} \frac{\nu_1 \coth(\nu_1 L) + \kappa_0/D}{\kappa_0\sigma e^{-\beta V_{\text{eff}}}\nu_1 \coth(\nu_1 L) + \omega_1\sigma e^{-\beta V_{\text{eff}}}}. \quad (3.26)$$

This is nearly identical to Eq. (3.23b), except that here  $V_{\text{eff}}$ ,  $p_{\text{la}}$ , and  $p_{\text{li}}$  in the binding pocket substitute for  $V_0$ ,  $p_a$ , and  $p_i$ , respectively. The BDflex result is obtained by replacing  $\omega \hat{k}_E(\omega)$  with  $k_E$  in Eq. (3.26)

$$\frac{1}{k_a} = \frac{1}{k_E} + \frac{L_1}{D\sigma e^{-\beta V_1}} + \frac{1}{\kappa_0\sigma e^{-\beta V_{\text{eff}}}} + \frac{L}{D\sigma e^{-\beta V_{\text{eff}}}} + \frac{p_i}{p_a} \frac{\nu \coth(\nu L_1) + k_E/D\sigma e^{-\beta V_1}}{k_E\nu \coth(\nu L_1) + \omega\sigma e^{-\beta V_1}} + \frac{p_{\text{li}}}{p_{\text{la}}} \frac{\nu_1 \coth(\nu_1 L) + \kappa_0/D}{\kappa_0\sigma e^{-\beta V_{\text{eff}}}\nu_1 \coth(\nu_1 L) + \omega_1\sigma e^{-\beta V_{\text{eff}}}}. \quad (3.27)$$

When  $\omega \rightarrow 0$ , both Eqs. (3.26) and (3.27) predict

$$\frac{1}{k_a} \rightarrow \frac{1}{p_a} \left( \frac{1}{k_E} + \frac{L_1}{D\sigma e^{-\beta V_1}} + \frac{1}{\kappa_0\sigma e^{-\beta V_a}} + \frac{L}{D\sigma e^{-\beta V_a}} \right), \quad (3.28a)$$

which gives  $k_a$  as the product of the rate constant when the lid stays open (with the potential being  $V_1$  for  $x_1 < x < x_d$  and  $V_a$  for  $x_b < x < x_1$ ) and the equilibrium probability for the lid being open. In the opposite limit  $\omega \rightarrow \infty$ , both Eqs. (3.26)

and (3.27) predict

$$\frac{1}{k_a} \rightarrow \frac{1}{k_E} + \frac{L_1}{D\sigma e^{-\beta V_1}} + \frac{1}{\kappa_0\sigma e^{-\beta V_{\text{eff}}}} + \frac{L}{D\sigma e^{-\beta V_{\text{eff}}}}, \quad (3.28b)$$

which gives  $k_a$  as the rate constant when the lid stays open, with the potential being  $V_1$  for  $x_1 < x < x_d$  and  $V_{\text{eff}}$  for  $x_b < x < x_1$ . Figure 4 shows the good agreement between the BDflex and exact results for  $k_a$  in the gated access case under induced switch. Also note that, similar to the gating binding-site case, induced switch has the advantage over indifferent switch in that the rate constant has a reduced difference between the slow and fast limits of conformational transition and the shift from the slow limit to the fast limit occurs at a lower conformational transition rate.

### C. Gating binding pocket

The third model of conformational switching also has a lid at the entrance to the binding pocket, but after the ligand is inside, the lid must be closed for the binding site to be reactive (Fig. 2(c)). The active conformation of the binding pocket has the lid closed and the binding site reactive, whereas the inactive conformation of the binding pocket has the lid open and the binding site non-reactive.<sup>25</sup> The boundary conditions at the binding site become

$$\mathcal{J}g_a(x_b) = \kappa_0 g_a(x_b), \quad (3.29a)$$

$$\mathcal{J}g_i(x_b) = 0. \quad (3.29b)$$

Like the gated access case, the problem can be solved by using three sets of conditions at  $x = x_d$ ,  $x_1$ , and  $x_b$ . At  $x = x_d$ , the “exact” boundary condition is given by Eqs. (3.9). At  $x = x_1$ ,  $g_a(x)$  encounters a reflecting boundary but  $g_i(x)e^{\beta V_i(x)}$  and  $\mathcal{J}g_i(x)$  are continuous. At  $x = x_b$ , the boundary conditions for  $g_a(x)$  and  $g_i(x)$  are given by Eqs. (3.29).

Under indifferent switch, with transition rates  $\omega_{\pm}$  between the inactive and active conformations regardless of where the ligand is, the “exact” rate constant is given by

$$\frac{1}{k_a} = \frac{1}{k_E} + \frac{L_1}{D\sigma e^{-\beta V_1}} + \frac{1}{p_a\kappa_0\sigma e^{-\beta V_0}} + \frac{L}{D\sigma e^{-\beta V_0}} + \frac{p_a}{p_i} \frac{\nu \coth(\nu L_1) + \omega \hat{k}_E(\omega)/D\sigma e^{-\beta V_1}}{\omega \hat{k}_E(\omega)\nu \coth(\nu L_1) + \omega\sigma e^{-\beta V_1}} + \left( \frac{p_i}{p_a} + \frac{p_a}{p_i} \right) \frac{\nu \coth(\nu L)}{\omega\sigma e^{-\beta V_0}} + \frac{2\nu}{\omega\sigma e^{-\beta V_0} \sinh(\nu L)}. \quad (3.30)$$

The BDflex result is obtained by replacing  $\omega \hat{k}_E(\omega)$  with  $k_E$

$$\frac{1}{k_a} = \frac{1}{k_E} + \frac{L_1}{D\sigma e^{-\beta V_1}} + \frac{1}{p_a\kappa_0\sigma e^{-\beta V_0}} + \frac{L}{D\sigma e^{-\beta V_0}} + \frac{p_a}{p_i} \frac{\nu \coth(\nu L_1) + k_E/D\sigma e^{-\beta V_1}}{k_E\nu \coth(\nu L_1) + \omega\sigma e^{-\beta V_1}} + \left( \frac{p_i}{p_a} + \frac{p_a}{p_i} \right) \frac{\nu \coth(\nu L)}{\omega\sigma e^{-\beta V_0}} + \frac{2\nu}{\omega\sigma e^{-\beta V_0} \sinh(\nu L)}. \quad (3.31)$$

When  $\omega \rightarrow 0$ , both Eqs. (3.26) and (3.27) predict  $k_a \rightarrow 0$ . The fast-transition limit is

$$\frac{1}{k_a} \rightarrow \frac{1}{k_E} + \frac{L_1}{D\sigma e^{-\beta V_1}} + \frac{1}{p_{1a}\kappa_0\sigma e^{-\beta V_0}} + \frac{L}{D\sigma e^{-\beta V_0}}. \quad (3.32)$$

The results under induced switch can be obtained by substituting  $V_{\text{eff}}$ ,  $p_{1a}$ , and  $p_{1i}$  in the binding pocket for  $V_0$ ,  $p_a$ , and  $p_i$ , respectively. The “exact” result for  $k_a$  is given by

$$\begin{aligned} \frac{1}{k_a} = & \frac{1}{k_E} + \frac{L_1}{D\sigma e^{-\beta V_1}} + \frac{1}{p_{1a}\kappa_0\sigma e^{-\beta V_{\text{eff}}}} + \frac{L}{D\sigma e^{-\beta V_{\text{eff}}}} \\ & + \frac{p_a v \coth(vL_1) + \omega \hat{k}_E(\omega)/D\sigma e^{-\beta V_1}}{p_i \omega \hat{k}_E(\omega) v \coth(vL_1) + \omega \sigma e^{-\beta V_1}} \\ & + \left( \frac{p_{1i}}{p_{1a}} + \frac{p_{1a}}{p_{1i}} \right) \frac{v_1 \coth(v_1 L)}{\omega_1 \sigma e^{-\beta V_{\text{eff}}}} + \frac{2v_1}{\omega_1 \sigma e^{-\beta V_{\text{eff}}} \sinh(v_1 L)}. \end{aligned} \quad (3.33)$$

The BDflex result is

$$\begin{aligned} \frac{1}{k_a} = & \frac{1}{k_E} + \frac{L_1}{D\sigma e^{-\beta V_1}} + \frac{1}{p_{1a}\kappa_0\sigma e^{-\beta V_{\text{eff}}}} + \frac{L}{D\sigma e^{-\beta V_{\text{eff}}}} \\ & + \frac{p_a v \coth(vL_1) + k_E/D\sigma e^{-\beta V_1}}{p_i k_E v \coth(vL_1) + \omega \sigma e^{-\beta V_1}} \\ & + \left( \frac{p_{1i}}{p_{1a}} + \frac{p_{1a}}{p_{1i}} \right) \frac{v_1 \coth(v_1 L)}{\omega_1 \sigma e^{-\beta V_{\text{eff}}}} + \frac{2v_1}{\omega_1 \sigma e^{-\beta V_{\text{eff}}} \sinh(v_1 L)}. \end{aligned} \quad (3.34)$$

Figure 5 shows that, for both indifferent switch and induced switch, the BD results for  $k_a$  agree well with the exact results. Induced switch results in higher  $k_a$  and a shift to the fast limit of conformational transition at a lower transition rate.

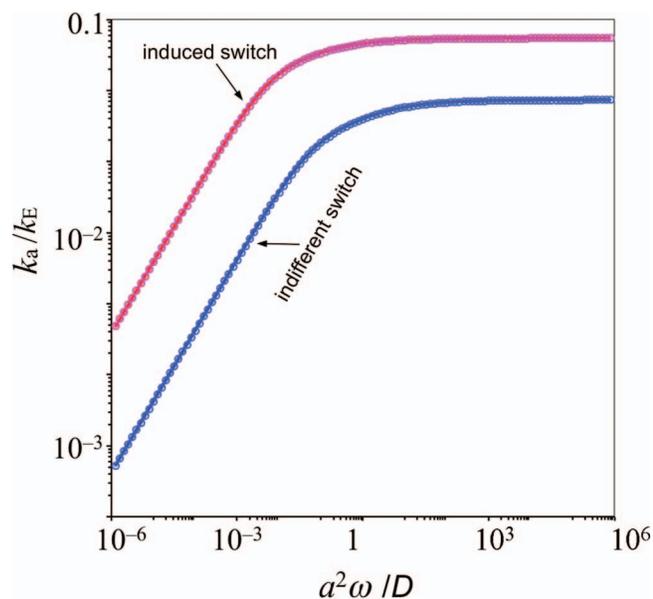


FIG. 5. Comparison of BDflex and “exact” results for the gating binding-pocket model. The parameters are the same as in Fig. 4, except  $\kappa_0 = D/a$ .

#### IV. IMPLEMENTATION OF BDflex

We now illustrate the implementation of BDflex on a model system, consisting of a spherical-shaped protein (radius  $R$ ) with a buried binding pocket and a pointlike ligand (Fig. 1(b)). The binding pocket is defined by a second, smaller sphere, centered a distance  $d$  from the center of the protein sphere; the binding pocket includes only the portion of the second sphere that is inside the protein sphere. For convenience we define the center of the protein sphere as the origin of a coordinate system, and the center of the binding-pocket sphere as lying on the  $x$  axis. The ligand cannot penetrate the protein surface. Inside the binding pocket the ligand can bind to the protein with a rate constant  $\gamma$ . The protein either has a rigid structure or can undergo conformational switching. In the latter case, the radius of the binding pocket stochastically switches, with rates  $\omega_{\pm}$ , between two values,  $a$  and  $a'$ ; binding can occur only when the binding pockets adopts the second size. Both indifferent switch and induced switch are studied. We implement both the 1990 algorithm<sup>3</sup> for  $k_a$  calculation and BDflex to demonstrate the latter’s considerable savings in computational time. All times reported below are in units of  $R^2/2D$ , which has a typical value  $\sim 2$  ns.

##### A. Implementation procedure

We first summarize the 1990 algorithm. This is also the algorithm that we use to solve the exterior problem of BDflex. In addition, the implementation of the interior problem of BDflex shares many ingredients with the 1990 algorithm. In that algorithm, ligand trajectories are started in the reaction region and then propagated up to a cutoff time ( $t_{\text{cut}}$ ). Whenever the ligand is inside the reaction region it is allowed to form the product (with rate constant  $\gamma$ ). If production formation occurs, then the trajectory is terminated. The lifetime of that trajectory is recorded. If production formation does not occur before  $t_{\text{cut}}$ , then the lifetime of the trajectory is recorded as  $>t_{\text{cut}}$ . For these lifetimes, the fraction of surviving trajectories at times up to  $t_{\text{cut}}$  is calculated. The survival fraction,  $S(t)$ , is related to the time-dependent rate coefficient  $k(t)$  via

$$k(t) = k(0)S(t). \quad (4.1)$$

The initial value of the rate coefficient is

$$k(0) = p_a V_{\text{rr}} \left\langle e^{-\beta U_a(\mathbf{r})} \right\rangle_{\text{rr}} \gamma, \quad (4.2)$$

where  $V_{\text{rr}}$  is the volume of the reaction region. Note that in the case of the protein undergoing conformational switching, the reaction region is the binding pocket with radius  $a'$ . The average Boltzmann factor is present if the protein-ligand interaction potential is not zero when the ligand is in the reaction region and the protein is in the reactive conformation. The long-time limit of  $k(t)$  is the rate constant  $k_a$ . This is obtained by fitting the long-time portion of  $k(t)$  to the known asymptotic expansion<sup>26</sup>

$$k(t) = k_a [1 + k_a/4\pi D(\pi Dt)^{1/2} + \dots]. \quad (4.3)$$

To ensure conformity to the asymptotic behavior,  $t_{\text{cut}}$  has to be sufficiently large.

The 1990 algorithm works both when the protein has a rigid structure and when it undergoes conformational switching. However, the latter case may require enormous computational time, which is the motivation for the present work. For the model problem of Fig. 1(b), this is not a problem and we can use the 1990 algorithm for benchmarking. Some details of the implementation here are as follows. The propagation of ligand trajectories follows the Ermak-McCammon algorithm.<sup>27</sup> The reflecting boundary condition on the protein surface is treated by placing the ligand back to its prior position. The timesteps are  $10^{-5}$  when inside a sphere with radius  $R_1 = 1.1R$  and  $10^{-5} + 10^{-2}(r - R_1)^2/2D$  otherwise. In the case of the protein undergoing conformational switching, when the binding pocket switches from the large size to the smaller size, the ligand could be found in a forbidden region. In that case the ligand is moved toward the center of binding pocket to just inside it.

The protein-ligand interaction potential is absent in the case of a rigid binding pocket and in the case of indifferent switch between an open (with radius  $a$ ) but non-reactive conformation and a closed (with radius  $a'$ ) but reactive conformation. However, under induced switch, the ligand in the reaction region should interact favorably with the protein in the reactive conformation. We thus assign a negative value,  $U_0$ , to  $U_a(\mathbf{r})$  for  $\mathbf{r}$  in the reactive region. Around  $r = R$ , the interaction potential makes a smooth but sharp transition to zero<sup>21</sup>

$$U_a(\mathbf{r}) = -\frac{U_0}{2} \left( \tanh \frac{r - R}{L} - 1 \right). \quad (4.4)$$

We choose  $L/R = 0.005$ . The potential for the non-reactive conformation,  $U_i(\mathbf{r})$ , is zero both in the binding pocket and in the outside.

BDflex consists of two sets of BD simulations, solving the exterior and interior problems, respectively. We use different choices of the dividing surface to check whether and how the calculated  $k_a$  depends on the choice of the dividing surface. The inner most choice is the mouth of the binding pocket (green solid curve in Fig. 1(b)); a choice that is slightly displaced outward is the exposed cap of the binding-pocket sphere (dashed curve in Fig. 1(b)); other choices include the exposed surface of a third sphere, with various radii and centered at various positions on the  $x$  axis, that intersects the protein sphere.

As already mentioned, the exterior simulations follow the 1990 algorithm to yield  $k_E$ ; here the protein is treated as rigid. The boundary condition on the dividing surface should be absorbing. However, the 1990 algorithm cannot directly handle the absorbing boundary condition. Instead, we turn the boundary condition into the radiation type, and then let the reactivity  $\kappa$  approach infinity. Actually the 1990 algorithm cannot directly handle the radiation boundary either. The way out<sup>3</sup> is to grow the dividing surface outward into a thin spherical shell, with thickness  $\varepsilon$ ; the original dividing surface is turned into a reflecting boundary but the thin shell is turned into a reaction region. With a reaction rate equal to  $\kappa/\varepsilon$ , it can be shown that this treatment approximates the radiation boundary condition well when  $\varepsilon$  is small [see Eq. (2.33)]. Note that now the outer region has its entire inner boundary, defined by the dividing surface plus the protein surface connected

to it, satisfying the reflecting condition, which ensures that the ligand is restricted to the outer region. We use a similar treatment for implementing the radiation boundary condition, Eq. (2.22), of the interior problem (see below). Hereafter, we refer to the thin shell bordering the dividing surface as the sink region, to distinguish from the reaction region modeling ligand binding. With ligand trajectories started in the sink region and allowed to react with rate  $\kappa/\varepsilon$ , we obtain the rate constant  $k$  for a series of  $\kappa$  values. The desired  $k_E$  is obtained from fitting the dependence of  $k$  on  $\kappa$  according to<sup>26</sup>

$$\frac{1}{k} = \frac{1}{V_{\text{sr}}\kappa/\varepsilon} + \frac{1}{k_E}, \quad (4.5)$$

where  $V_{\text{sr}}$  is the volume of the sink region (an average Boltzmann factor would be present should the protein-ligand interaction potential be not zero in the sink region).

In the interior simulations, as mentioned above we transform the radiation boundary condition on the dividing surface [Eq. (2.22)] into the sink region with reaction rate  $\kappa_E/\varepsilon$ , where  $\kappa_E = k_E/\sigma_d$  [see Eq. (2.18)]. To confine the ligand to the inner region, we make the dividing surface invisible but the remaining surface of the sink region reflecting. Ligand trajectories are started in the sink region and are propagated until reaction occurs either in the sink region or in the reaction region. The fraction of trajectories that end up reacting in the reaction region is  $\bar{\eta}_d$ . Finally, the product  $k_E\bar{\eta}_d$  gives the overall rate constant  $k_a$ .

Note that the same radiation boundary condition, Eq. (2.22), is valid for any conformational state. Therefore, its implementation during the BD simulations is the same even in the presence of conformational fluctuations. A more rigorous formulation of the boundary condition would involve coupling of different conformational states [see, e.g., Eqs. (3.9)]. That type of boundary condition would be far more cumbersome to implement.

In the 1990 algorithm, the ligand crosses the dividing surface multiple times before reaction at the binding site (or the trajectory is terminated by exceeding the  $t_{\text{cut}}$  threshold). By confining the ligand to the inner region and then imposing the radiation boundary condition, BDflex avoids such multiple-crossing of the dividing surface. This results in considerable savings in computational time, as shown below.

## B. Results

In Fig. 6(a), we compare  $k_a$  values obtained by the 1990 algorithm and by BDflex for a rigid binding pocket with a range  $\gamma$ , which is the rate of product formation inside the binding pocket. We implemented BDflex using two dividing surfaces (Fig. 1(b)), resulting in  $k_E = 1.01DR$  and  $1.36DR$ , respectively. The final results for  $k_a$  are very similar and both are close to the results by the 1990 algorithm. Other dividing surfaces also give very similar results (not shown).

Relative to  $k_a$  by the 1990 algorithm,  $\bar{\eta}_d$  converges with a smaller number of trajectories. The  $k_a$  values shown in Fig. 6(a) are obtained using 6000 trajectories each by the 1990 algorithm but using 4000 trajectories for BDflex. Moreover, the total simulation time per trajectory is significantly reduced in the BDflex simulations, except for very high  $\gamma$ , under which most of the trajectories in the 1990 algorithm

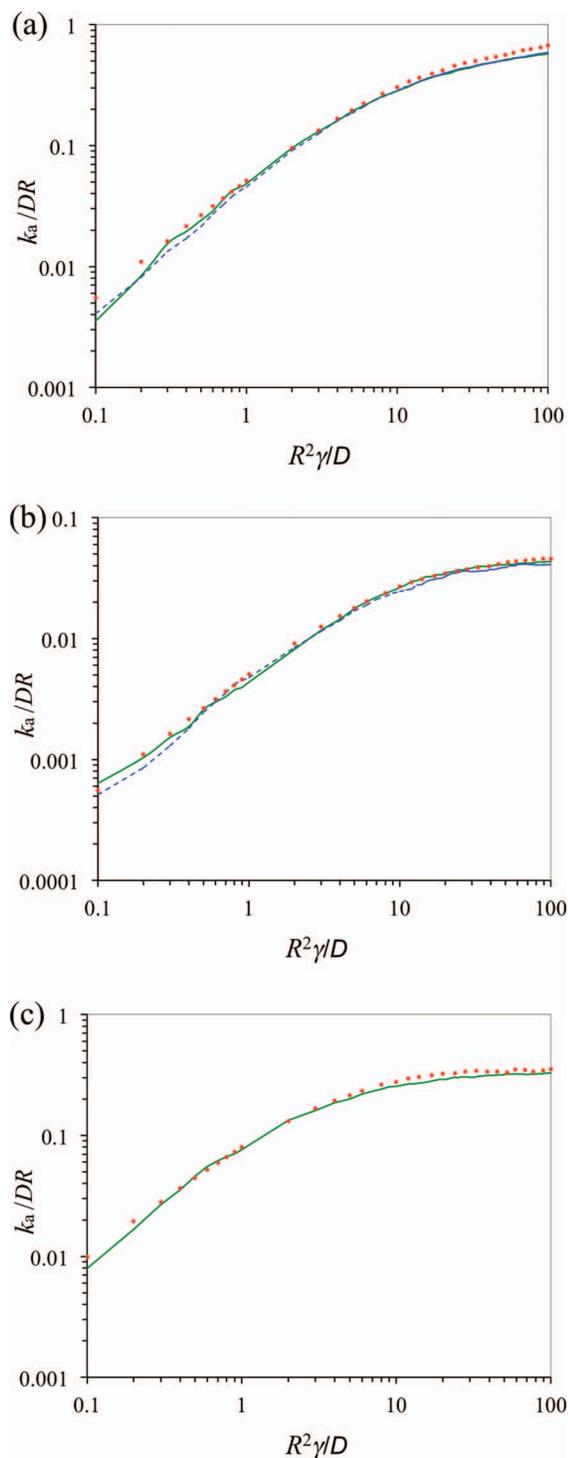


FIG. 6. Comparison of simulation results according to BDflex and to the 1990 algorithm. (a) Rigid binding pocket.  $a/R = 0.25$ ;  $d/R = 0.86$ . (b) Binding pocket undergoing indifferent switch. The binding pocket switches between an open (with radius  $a$ ) but non-reactive state and a closed (with radius  $a' = R - d = 0.14R$ ) but reactive state, with  $\omega_+ = \omega_- = 10D/R^2$ . The solid and dashed curves are BDflex results using two alternative dividing surfaces (Fig. 1(b)); symbols are results by the 1990 algorithm. For  $R^2\gamma/D \leq 10$ , the number of trajectories in the BDflex interior simulations is increased tenfold to 40 000, to demonstrate that long simulations for these more challenging cases can be afforded by BDflex. (c) Binding pocket undergoing induced switch. The rates of conformation switch are  $\omega_+ = D/R^2$  and  $\omega_- = 10D/R^2$  when the ligand is far away but change to  $\omega_+ = 100D/R^2$  and  $\omega_- = 10D/R^2$  when the ligand is inside the binding pocket. Correspondingly,  $e^{-\beta U_0} = 100$ . The solid curve displays BDflex results using the solid dividing surface in Fig. 1(b); symbols are results by the 1990 algorithm.

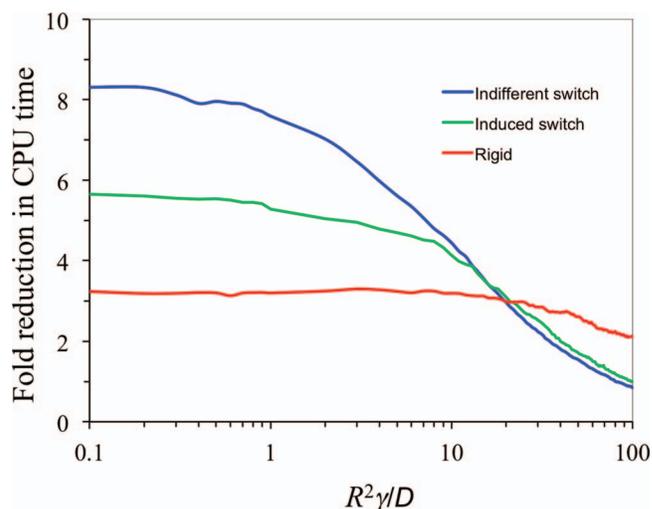


FIG. 7. The ratio of CPU times of BDflex and the 1990 algorithm. In each of the three cases studied, CPU times are from a total of 4000 trajectories for BDflex and 6000 trajectories for the 1990 algorithm.

terminate before they leave the reaction region. The factor by which the CPU time of the BDflex results over that by the 1990 algorithm is shown in Fig. 7. Threefold speedup is obtained.

In Fig. 6(b), we display  $k_a$  values for the binding pocket under indifferent switch, between an open but non-reactive conformation and a closed but reactive conformation (Fig. 1(b)), similar to the gating binding-pocket model of the preceding section. Again the BDflex results agree very well with results obtained by the 1990 algorithm. Relative to the rigid binding pocket, product formation in the present case becomes a rarer event; hence  $k_a$  is reduced in magnitude. Here, the saving in CPU time by BDflex becomes even more significant, with speedup as much as eightfold (Fig. 7). So BDflex becomes powerful especially for problems that are challenging to the traditional approaches.

In Fig. 6(c), we display  $k_a$  values for the binding pocket under induced switch. Again the BDflex results agree very well with results obtained by the 1990 algorithm. Relative to the indifferent switch case, the magnitude of  $k_a$  is higher, suggesting that a protein can use adaption of its conformation to the incoming ligand to speed up the binding process. Here again there is a significant saving in CPU time by BDflex, with speedup of  $\sim 5.5$ -fold (Fig. 7).

While we do not show explicit results for the saving in CPU time by BDflex over the 1984 algorithm of Northrup *et al.*,<sup>1</sup> a comparison of the algorithm designs makes it clear that BDflex should be significantly faster. In essence both algorithms find  $k_a$  by breaking it into the form  $k_a = k_E \eta$  [see Eqs. (1.1) and (2.15)]. We choose a dividing surface that is as close to the binding pocket as possible, such that  $k_E$  approximates  $k_a$  as much as possible and,  $\eta$ , the probability for a ligand starting on the dividing surface to react in the binding pocket, instead of escape to infinity, is as high as possible. A higher  $\eta$  makes it easier and hence requires less number of trajectories to determine. In contrast, the Northrup *et al.* algorithm uses a spherical dividing surface, so  $k_E$  is generally much higher than  $k_a$ , and corresponding  $\eta$  is very small,

especially when ligand binding is accompanied by protein conformational change, the case that motivated the present paper. The radiation boundary condition that we introduce here on the dividing surface so as to confine the motion of the ligand to the inner region may provide an additional saving in computational time.

## V. CONCLUDING REMARKS

We have developed a new method, BDflex, for efficiently treating conformational fluctuations in calculating protein-ligand binding rate constants ( $k_a$ ) from Brownian dynamics simulations. The core idea of BDflex, breaking up the calculation of  $k_a$  into separate exterior and interior problems, can be traced to theoretical developments over the years.<sup>16,17,19,20</sup> Connecting the exterior and interior problems is a critical link, and here we are able to treat this link via a radiation boundary condition that can be conveniently implemented in Brownian dynamics simulations, even in the presence of conformational fluctuations.

The radiation boundary condition on the dividing surface of the outer and inner regions is an approximation and we have demonstrated its accuracy on models of conformational fluctuations. Moreover, we have illustrated the implementation of BDflex on a model system and compared to the implementation of a previous algorithm<sup>3</sup> for calculating  $k_a$  from BD simulations. The latter algorithm does not invoke any approximation. Here, we show that BDflex yields the same  $k_a$  values at considerably reduced computational cost. The ligand in the inner region can either react in the binding site or be absorbed on the dividing surface. The avoidance of crossing the dividing surface multiple times before binding to the protein accounts for the savings in computational time.

The implementation of BDflex illustrated here has demonstrated its accuracy and expected considerable savings in computational time. This implementation did not rely on anything specific to the chosen model system and can be easily ported to more realistic models of protein-ligand systems. In such models, proteins and ligands are represented with more molecular details, albeit still at a coarse-grained level for practical purposes. Such models, after careful parameterization, allow both overall translational and rotational diffusion and internal conformational fluctuations to be modeled by Brownian dynamics simulations.<sup>28</sup> Implementation of BDflex within such more realistic models of protein-ligand systems is underway.

## ACKNOWLEDGMENTS

This study was supported by Grant No. GM58187 from the National Institutes of Health (NIH).

- <sup>1</sup>S. H. Northrup, S. A. Allison, and J. A. McCammon, *J. Chem. Phys.* **80**, 1517 (1984).
- <sup>2</sup>B. A. Luty, J. A. McCammon, and H. X. Zhou, *J. Chem. Phys.* **97**, 5682 (1984).
- <sup>3</sup>H.-X. Zhou, *J. Phys. Chem.* **94**, 8794 (1990).
- <sup>4</sup>R. C. Wade, B. A. Luty, E. Demchuk, J. D. Madura, M. E. Davis, J. M. Briggs, and J. A. McCammon, *Nat. Struct. Biol.* **1**, 65 (1994).
- <sup>5</sup>M. Harel, I. Schalk, L. Ehret-Sabatier, F. Bouet, M. Goeldner, C. Hirth, P. H. Axelsen, I. Silman, and J. L. Sussman, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9031 (1993).
- <sup>6</sup>H.-X. Zhou, S. T. Wlodek, and J. A. McCammon, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9280 (1998).
- <sup>7</sup>R. J. Russell, L. F. Haire, D. J. Stevens, P. J. Collins, Y. P. Lin, G. M. Blackburn, A. J. Hay, S. J. Gamblin, and J. J. Skehel, *Nature (London)* **443**, 45 (2006).
- <sup>8</sup>J. C. Sung, A. W. Van Wynsberghe, R. E. Amaro, W. W. Li, and J. A. McCammon, *J. Am. Chem. Soc.* **132**, 2883 (2010).
- <sup>9</sup>D. Joseph, G. A. Petsko, and M. Karplus, *Science* **249**, 1425 (1990).
- <sup>10</sup>X. Chen, A. A. Antson, M. Yang, P. Li, C. Baumann, E. J. Dodson, G. G. Dodson, and P. Gollnick, *J. Mol. Biol.* **289**, 1003 (1999).
- <sup>11</sup>B. G. Miller, A. M. Hassell, R. Wolfenden, M. V. Milburn, and S. A. Short, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2011 (2000).
- <sup>12</sup>K. Melcher, L.-M. Ng, X. E. Zhou, F.-F. Soon, Y. Xu, K. M. Suino-Powell, S.-Y. Park, J. J. Weiner, H. Fujii, V. Chinnusamy, A. Kovach, J. Li, Y. Wang, J. Li, F. C. Peterson, D. R. Jensen, E.-L. Yong, B. F. Volkman, S. R. Cutler, J.-K. Zhu, and H. E. Xu, *Nature (London)* **462**, 602 (2009).
- <sup>13</sup>H. Unal, R. Jagannathan, M. B. Bhat, and S. S. Karnik, *J. Biol. Chem.* **285**, 16341 (2010).
- <sup>14</sup>H.-X. Zhou, J. M. Briggs, and J. A. McCammon, *J. Am. Chem. Soc.* **118**, 13069 (1996).
- <sup>15</sup>Strict validity of Eq. (1.1) requires that the b surface be positioned far away from the protein such that both the protein-ligand interaction potential and the diffusive flux are spherically symmetric [H.-X. Zhou, *J. Chem. Phys.* **92**, 3092 (1990)].
- <sup>16</sup>H.-X. Zhou, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8651 (2011).
- <sup>17</sup>H.-X. Zhou, *J. Chem. Phys.* **108**, 8146 (1998).
- <sup>18</sup>D. Shoup, G. Lipari, and A. Szabo, *Biophys. J.* **36**, 697 (1981).
- <sup>19</sup>A. M. Berezhkovskii, A. Szabo, and H. X. Zhou, *J. Chem. Phys.* **135**, 075103 (2011).
- <sup>20</sup>J. L. Barreda and H. X. Zhou, *J. Chem. Phys.* **135**, 145101 (2011).
- <sup>21</sup>L. Cai and H. X. Zhou, *J. Chem. Phys.* **134**, 105101 (2011).
- <sup>22</sup>H. X. Zhou, *Biophys. J.* **88**, 1608 (2005).
- <sup>23</sup>H.-X. Zhou, *Biophys. J.* **98**, L15 (2010).
- <sup>24</sup>A. Szabo, D. Shoup, S. H. Northrup, and J. A. McCammon, *J. Chem. Phys.* **77**, 4484 (1982).
- <sup>25</sup>Note that the correspondence here between active/inactive conformation and closed/open lid is opposite to that in the gated access case.
- <sup>26</sup>H. X. Zhou and A. Szabo, *Biophys. J.* **71**, 2440 (1996).
- <sup>27</sup>D. L. Ermak and J. A. McCammon, *J. Chem. Phys.* **69**, 1352 (1978).
- <sup>28</sup>M. Kang, C. Roberts, Y. Cheng, and C.-e. A. Chang, *J. Chem. Theory Comput.* **7**, 3438 (2011).