Reprinted from The Journal of Physical Chemistry, 1987, 91, 3624.
Copyright © 1987 by the American Chemical Society and reprinted by permission of the copyright owner.

Brownian Dynamics Simulation of Diffusion to Irregular Bodies

Kim Sharp,[†] Richard Fine,[‡] Klaus Schulten,[†§] and Barry Honig*[†]

Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, and Department of Biological Sciences, Columbia University, New York, New York 10027
(Received: December 22, 1986)

An extension of the Monte Carlo algorithm for solving the three-dimensional Smoluchowski diffusion equation (Brownian dynamics) for an arbitrarily shaped target particle with a nonsymmetric force field is described. The algorithm uses extensive table lookups to describe the molecular shape and force field in order to reduce computation during the simulation. This also allows the algorithm to be vectorized. Dynamic adjustment of the step size is used to handle rapidly changing and nonlinear force fields. The accuracy of the approximations introduced due to the discrete representation of the field and particle shape is assessed by comparison with several cases where analytical solutions are available. An application to the diffusion of a substrate to an enzyme, where the enzyme's shape and electrostatic field are believed to be important, is described. The electric fields calculated by using different Coulombic potentials do not account for the ionic strength dependence of the enzyme rate. However, the field calculated numerically from the Poisson–Boltzmann equation, which includes the effects of the dielectric boundary between the protein and the solvent, correctly reproduces the ionic strength dependence of the enzyme.

Introduction

Brownian dynamics is a powerful and flexible method for calculating observables, e.g., reaction yields, connected with diffusion-controlled association of molecules in one, two, or three dimensions.[1-4] In its most straightforward application, it is formally equivalent to solving the Smoluchowski diffusion–association equation. However, many factors can be included in these simulations for which direct or numerical solutions of the diffusion equation are difficult or impossible. These include reactive diffusion with partial reactivity on a boundary surface,[4] hydrodynamic interactions,[2] rotational diffusion, orientational constraints, and force fields.[5] With the appropriate constraints the Brownian dynamics method has been shown to converge rapidly. Results are obtained by running many trajectories and determining for each trajectory its contribution to the observables of interest. The accuracy of the method depends on the information content of the observables. For observables where the total yield of a fairly probable reaction is represented solely by one number whose value is close to 1, the method converges rapidly. In this case the simulation of a small number of trajectories is required. Observables for less probable events, or observables represented by continuous functions, e.g., distributions, which carry a very high information content, require a larger number of trajectories.

* To whom inquiries and reprint requests should be addressed.
† Department of Biochemistry and Molecular Biophysics.
‡ Department of Biological Sciences.
§ Permanent address: Physik Department, Technische Universitat Munchen, 8046 Garching, Federal Republic of Germany.

(1) Ermak, D. L. *J. Chem. Phys.* **1975**, *62*, 4189.
(2) Ermak, D. L.; McCammon, J. A. *J. Chem. Phys.* **1978**, *69*, 1532.
(3) Schulten, K.; Epstein, I. R. *J. Chem. Phys.* **1979**, *71*, 309.
(4) Lamm, G.; Schulten, K. *J. Chem. Phys.* **1983**, *78*, 2713.
(5) Northrup, S. H.; Allison, S. A.; McCammon, J. A. *J. Chem. Phys.* **1984**, *80*, 1517.

0022-3654/87/2091-3624$01.50/0 © 1987 American Chemical Society

metric force field, R_s in eq 1 and 2 can be replaced by the corrected radius, R_s^* [15]

$$1/R_s^* = \int_{R_s}^{\infty} \frac{\exp(U(r)/kT)}{4\pi r^2}\, dr \qquad (3)$$

where $U(r)$ is the centrosymmetric potential function describing the force field. Similarly R_e in (1) must be replaced by the equivalent expression for R_e^*. As an example, for a single point charge in a region of constant dielectric at zero ionic strength, the potential is given by

$$U(r) = q/\epsilon r \qquad (4)$$

where ϵ is the dielectric constant and q is the charge. Substituting (4) into (3) and integrating gives a general expression for any radius:

$$R^* = q/\epsilon(\exp(q/\epsilon R) - 1) \qquad (5)$$

The generalization of these expressions to situations with partially reactive surfaces has been given by Schulten and Schulten.[16]

Evolution of Trajectories. The new Cartesian coordinate of a trajectory, x_i', after a short time interval, τ, is calculated from the coordinate at the beginning of the time step, x_i, by the evolution equation

$$x_x' = x_i - \delta U/\delta x_i D\tau/kT + n_i D^{1/2}\tau^{1/2} \qquad \text{for } i = 1, 2, 3 \quad (6)$$

where k is Boltzmann's constant, T is the absolute temperature, $U(x_i)$ is the potential function, $-\delta U(x_i)/\delta x_i = f_i$, for $i = 1, 2, 3$, where f_i are the components of force in the three orthogonal directions, and n_i are three sets of independent normally distributed random numbers. The diffusion coefficient is assumed to be constant in this case (no hydrodynamic interactions). Equation 6 can be derived from Newton's third law and the Langevin equation by integration, using the fluctuation–dissipation theorem to relate the frictional drag on the diffusing particle to the Brownian forces which drive the diffusion.[1,3] It is valid for time steps short enough that the change in f_i during the time step is small. (Lamm and Schulten[4] have suggested less than 1% as a criterion.)

Calculation of Forces. Inside the sphere R_s, space is divided up into a 64 × 64 × 64 array of cubic elements, with the coordinate origin inside the target particle (Figure 1). The scale is chosen so that the target particle occupies a fraction of the grid space no larger than 0.666. After each time step the location of the trajectory of the diffusing particle is determined. If it is outside the grid, the force is assumed to be either centrosymmetric and calculated analytically, or set to zero, depending on the particular simulation. Inside the grid, the trajectory is located by indexing into the grid. The force components are then obtained from precalculated tables by using this index.

Determining the Fate of Trajectories. If the radial coordinate of the trajectory is greater than the exit radius, R_e, the trajectory is terminated. If the trajectory lies within the grid, the grid index is used to determine the value of a status flag, which determines whether the trajectory continues to diffuse freely, collides with the molecule, or hits the target patch. The use of this grid and bit mapping to determine when the trajectory collides is equivalent to representing the target molecule as a collection of fine cubical elements or boxes. If the target patch is reached, the trajectory is terminated (complete reactivity). Partial reactivity can be incoporated into this scheme by using another random number to determine if the trajectory escapes or by weighting each trajectory by its survival probability,[4] but this is not considered in this work.

Treatment of Collisions. The flux normal to the nonreactive surface of the molecule must vanish (reflective boundary conditions). This condition is only obtained in Brownian dynamics for infinitely short time step, since with a finite step size there is a probability that the diffusing particle will end up inside the molecule at the end of a jump, a forbidden event.[6] However, the use of specular reflection to treat colliding trajectories has been shown to be a suitable approximation for the finite time steps needed in practical simulations.[6] In specular reflection the portion of the trajectory step inside the surface is reflected out into the solution about a plane tangent to the point of entry. In this case though, the additional computation that is required to calculate the position of the trajectory after a specular reflection from an irregular surface is probably not warranted since the surface is not described exactly but is represented as a collection of boxes. In addition to specular reflection, therefore, we also investigated the accuracy of two other, computationally simpler, collision schemes: backstepping and random backstepping. In backstepping, the trajectory that collides is simply moved back to its original position at the beginning of the time step. However, it is conceivable that when there is a strong attractive force normal to the surface, trajectories close to the surface will collide on nearly every time step. In this case backstepping would effectively immobilize the trajectory, suppressing any two-dimensional diffusion tangential to the surface. A way to avoid this potential problem is to use random backstepping: a backstep combined with a random displacement within the box containing the trajectory.

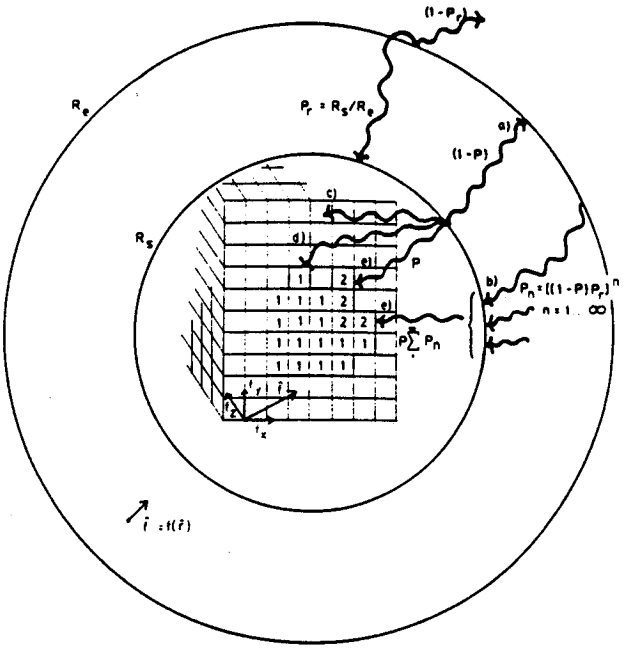

Figure 1. Schematic illustration of the method used to run irregular-body Brownian dynamic trajectories. R_s and R_e are the radii of the spheres where trajectories are initiated and terminated, respectively. Inside R_s the division of space into a grid is indicated. Blank boxes represent free diffusion space, 1 and 2 represent the nonreactive and reactive parts of the target molecule, respectively. The five classes of trajectories, with their associated probabilities are labeled as (a) exiting the diffusion space; (b) reentering the diffusion space at R_s after $n = 1-\infty$ previous exits and reentrances; (c) continuing to diffuse; (d) colliding with the surface; (e) reaching the target and reacting. The force vector, f, is calculated as a function $f(\mathbf{r})$ of the radial vector r outside the grid, or equal to f_i at the nearest grid point. Other symbols are defined in the text.

Choice of Time Step. Within the limits of validity of the integration of the Langevin equation, higher accuracy is obtained with smaller time steps, particularly for a rapidly changing force field, while faster convergence is obtained with larger time steps. The choice of time step is thus a compromise between these factors. The best compromise is to reduce the time step when the force change during a step is above some limit, reduce the time step when the trajectory is close to the surface of the molecule, and increase it when it is far from the surface and when the force is small. Higher accuracy prescriptions for choosing the time step which use information about the gradient of the force have been suggested.[4,17] However, in keeping with the general strategy of

(15) Debye, P. *Trans. Electrochem. Soc.* **1942**, *82*, 265.
(16) Schulten, Z.; Schulten, K. *J. Chem. Phys.* **1977**, *66*, 4616.
(17) van Gunsteren, W. F.; Berendsen, H. J. *Mol. Phys.* **1982**, *43*, 637.

reducing computation during the simulation to a minimum, we investigated the suitability of implementing a change of step size depending on the region in space. For this purpose we added to each grid cell a further index which points to a table of step sizes. This algorithm was combined with automatic scaling such that in regions where the force is very large, the maximum length of step, rather than the time, is kept below some required limit (typically 0.5 Å). This simpler approach, although not optimal, ensures rapid convergence.

Implementation. At the end of each time step, four possible situations arise for each trajectory: it continues diffusing; it collides with the nonreactive surface; it collides with the target and is terminated; it leaves the diffusion space and is terminated. All decisions about these different cases are implemented by table lookup and bit masking, in order to avoid all program branching. The algorithm is completely vectorized, and written in array processor control language for the Star ST-100 array processor. A vector of 1600 trajectories is run simultaneously, until the fates of at least 99% of them are determined. Generation of the large number of normally distributed random numbers required for Brownian dynamics is relatively time intensive, so 1.5 megaword tables are precalculated and stored in fast access memory.

Speed. Comparisons of the speed of Brownian dynamics algorithms are difficult, since the termination of the simulation depends on step size, the accuracy required, the size of the diffusion space, and the rate itself that one is trying to simulate. The program described here runs at about 20 megaflops, and executes about 1.5×10^5 steps/s, irrespective of the simulation conditions, how often the step size is changed, and how complicated the particle and field shape are. This represents a speed up of about 80 times over a VAX-780.

Simulation Conditions. Trajectories were run until enough hits were accumulated to give 5% error limits on rates, unless indicated otherwise in the tables. The maximum step size was usually limited to 0.5 Å close to the surface. The time step given in the results section is the basic time step when the trajectories are initiated at R_s; far from the surface it will be larger. Closer to the surface, and in regions of large forces, it will be smaller. The diffusion constant was chosen as 100 Å^2/ns, which is a typical value for a small molecule or ion.

Generation of Bit Maps. Bit maps of the spheres and spheres with recessed patches are generated by checking analytically whether each grid point is inside the object, and whether it is at the target patch. The appropriate flag value is then assigned to that grid point.

Simulation of Enzyme/Substrate Association. The bit map of the dimeric protein superoxide dismutase (SOD), a highly irregular object for which no analytical description of the surface shape is available, was generated from the Brookhaven Data Bank atomic coordinates.[18] The protein is modelled as a large number (2196) of spherical atoms, each with a radius equal to its van der Waals radius plus the radius of the superoxide ion substrate. It thus consists of a collection of intersecting spheres. Each grid point is checked for inclusion within the collision sphere of every atom. The target is defined as all the interior grid points that lie within the collision surface of the copper atoms of SOD. The scale is 1.5 Å per grid unit. The identity of the active site cleft atoms were taken from Tainer et al.[14]

The electric field was calculated numerically for every point of the $65 \times 65 \times 65$ grid for the constant and the two dielectric models for both zero and physiological ionic strength (0.15 M) as described by Klapper et al.[9]. This method involves assigning a value for the charge, dielectric constant, and Debye–Hückel parameter to every grid point, and solving the Poisson–Boltzmann equation by finite difference methods to obtain the potential at every grid point. For the two dielectric model, every point that lay inside the solvent accessible surface of the protein was assigned a dielectric constant of 2 and a Debye–Hückel parameter of zero (excluding solvent ions from the protein interior). Every point outside the protein was assigned a dielectric constant of 80, and a Debye–Hückel parameter of 0 (for zero ionic strength) or 1/(8 Å) (for physiological ionic strength). The uniform dielectric case was treated identically, except for an interior dielectric constant of 80. The charge distribution was determined by using from the Brookhaven data dase coordinates of SOD. For the distance-dependent dielectric model, the potential at each grid point, ϕ_i, was calculated by using

$$\phi_i = \Sigma q_j \exp(-\kappa r_{ij})/r_{ij}^2 \qquad (7)$$

where q_j is the charge at distance r_{ij} and κ is the Debye–Hückel parameter.

To ensure that the results were independent of the type of boundary conditions used in the calculations for the two dielectric model, simulations were run with field maps calculated either with zero boundary conditions or with Coulombic boundary conditions. The two rates agreed to within 8%. The monopole distribution was calculated by placing a charge of −4e at the center of the protein and using a uniform dielectric of 80 everywhere, reproducing the electrostatic field parameters of Allison and McCammon.[8]

Trajectories were initiated at random positions on a 60-Å sphere and terminated at an exit radius of 70 Å. The basic time step was 0.1 ps. Simulations run with different values for the starting radius, exit radius, and time step gave the same results, verifying that the calculated association rates are independent of these three simulation parameters.

Trapping of Trajectories. When simulating the association of superoxide with superoxide dismutase, an additional complication arises, since some of the protein's exposed charged side chains form small but highly attractive regions of positive potential in areas away from the target site. As a consequence, trajectories tend to accumulate in these regions during the simulation, and many remain trapped. This cannot of course happen in practice since only a finite number of ions can occupy any one region of space, so the other ions get displaced and escape. This artifact occurs because Brownian dynamics treats each trajectory independently of all others. To deal with this, simulations could be run long enough that the trapped trajectories escape with high probability during the simulation. However, this is very time consuming. A simpler procedure, which was found to give the same results, is to just increase the collision radii of the positively charge atoms somewhat.

TABLE I: Effect of Discrete Target Representation

target size, deg	reduced rate	
	simulated[a]	theoretical[b]
180	0.97 ± 0.05	1.00
90	0.69 ± 0.04	0.707
60	0.46 ± 0.04	0.465
45	0.35 ± 0.02	0.335
10 (9.5)[c]	0.053 ± 0.006	0.0615 (0.058)

[a] Simulated by using a time step of 1 ps, a sphere radius of 10 Å, a starting radius of 12.5 Å, and an exit radius of 30 Å. [b] Values calculated from eq 24 in Shoup et al.[19] [c] Figures in parentheses are based on the actual area fraction occupied by the target.

Results

Effect of Discrete Representation. The main objectives of this work are to extend the Brownian dynamics method to deal with irregularly shaped target bodies in noncentrosymmetric force fields. To obtain convergence with reasonable amounts of computer time it is necessary to use table lookups to represent the shape and force field. This discrete representation reduces computation drastically but inevitably introduces approximations. However, the errors introduced by these approximations can be tolerated if they are smaller than the increase in accuracy due to more realistic modelling of the shape and force field. To estimate the error due to the discrete representation of the shape, a 10-Å sphere was modeled as a collection of about 4000 1-Å cubes, with a circular target patch of half-angle θ. The simulated association rates are

(18) Tainer, J. A.; Getzoff, E. D.; Beem, K. M.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1982**, *160*, 181.

TABLE II: Effect of Discrete Force Representation

target size, deg	charge (e)	reduced rate[a] calcd	look-up	theory[b]
180	1	1.0 ± 0.05	1.02 ± 0.05	1.00
180	10	1.05 ± 0.05	1.01 ± 0.05	1.00
180[c]	−1	0.97 ± 0.05	0.95 ± 0.05	1.00
90	1	0.825 ± 0.05	0.836 ± 0.04	na[d]
90	−1	0.61 ± 0.04	0.61 ± 0.04	na
45	1	0.47 ± 0.03	0.45 ± 0.03	na
45	−1	0.28 ± 0.03	0.3 ± 0.03	na

[a] Simulated by using a time step of 1 ps, a sphere radius of 10 Å, a starting radius of 12.5 Å, and an exit radius of 50 Å. The reduced rate is calculated by using the sphere, starting and exit radii corrected for the centrosymmetric force field, using eq 5. [b] Calculated by using the corrected sphere radius. [c] Simulations used a time step of 0.3 ps. [d] Not available.

compared to analytically calculated rates (taken from Shoup, Lipari, and Szabo[19]) in Table I. The results are given as reduced rates, that is, as the rate normalized by the theoretical rate for a completely reactive sphere ($\theta = 180°$) of the same radius. All simulations are run to 5% accuracy and agree well with the analytical results, except for the 10° patch, in which case the simulated value is about 10% lower than predicted. This is a result of the discrete representation of the target patch. A patch of this size occupies 0.76% of the total surface area. When represented discretely in our simulation, the target occupies 0.68% of the total surface, which is equivalent to a circular patch of 9.5°. The theoretical rate for this target patch is also given in Table I, and better agreement is obtained. This illustrates the point that, when a small target patch is represented in a discrete fashion, it is probably more important to model the fractional area it occupies accurately, perhaps at the expense of a less accurate literal description in terms of its angular size or precise position.

Calculations of electrostatic force fields in cases where there are regions of different dielectric constants and counterion screening can themselves be extremely difficult and time consuming.[9,20] Clearly it is impractical to calculate these forces during a Brownian simulation, except for simplified cases.[8] Again precalculating these fields and using look-up tables greatly reduces the total amount of computation. As a check on the errors introduced by discretizing the force field, simulations were run for both attractive and repulsive centrosymmetric fields resulting from a point charge at the center of a 10-Å sphere in a uniform dielectric constant of 80 with no counterion screening. The results of this simulation can be checked against the analytical values obtained from eq 3 and 4. The results are summarized in Table II. Simulations were run under two conditions: the forces were calculated analytically by using the instantaneous trajectory positions or were obtained by look-up from the table, taking the value at the nearest grid point. The reduced rates are calculated by using the force-corrected radii (eq 5) and are expressed as the rate normalized by that for a wholly reactive sphere with the same radius and charge. For the 180° case where theoretical values are available, the results from both methods are extremely good, whether the potential is attractive ($q = 1$) or repulsive ($q = -1$). Results for a stronger attractive potential ($q = 10$) were also accurate. For all targets sizes, the results are independent of the method of force calculation. Within the accuracy of simulations (5%), the approximation made by using the force at the nearest grid point, which can be no more than 0.5 Å from the true trajectory position, is small.

Effect of Time Step. Equation 6 is only strictly valid when the force is constant over the distance taken during the simulation step. The most practical way to assess the error due to this approximation is to vary the step size. The four simulations in Table II with 90° target patches were each run with time steps of 0.1, 0.3, and 1 ps. Within the accuracy of the simulations, no effect of step size was seen, for neither the calculated nor the look-up force determination. Lamm and Schulten have suggested an empirical limit of 1% for the maximum change in force over one time step.[4] Statistics on the change in force during a time step were collected during two of the above simulations. For a 1-ps time step, about 90% of the steps had a force change of less than 20%, while for 0.1 ps, 90% had less than a 5% change in force. Thus for these type of simulations, with 5% accuracy limits, a <5% change in force during any one time step is an acceptable criterion for the validity of eq 6. When table lookups are used, the force at the trajectory location can be estimated more accurately from the values at the surrounding eight grid points by using a trilinear interpolation scheme. However, the extra computation involved is not warranted at the 5% statistic level used here.

(19) Shoup, D.; Lipari, G.; Szabo, A. *Biophys. J.* **1981**, *36*, 697.
(20) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *157*, 671.

TABLE III: Association Rate with a Recessed Patch[a]

patch depth, Å	reduced rate theory[b]	simulation[c]	simulation
3	0.0296	0.0265 ± 0.0005	0.028 ± 0.003
15	0.0066	0.0047 ± 0.0004	0.0053 ± 0.0005

[a] Simulated with a time step of 0.1 ps, a starting radius of 50 Å, and an exit radius of 60 Å for a sphere of 30 Å with a circular target patch of 10°, recessed below the surface by the distance in column 1 (the model of Samson and Deutch[21]). [b] From Samson and Deutch. [c] From Allison, Northrup, and McCammon.[7] Figures are taken for the case where the trajectories were started at a radius of 31 Å.

TABLE IV: Effect of the Electrostatic Field on the Relative Association Rate of Superoxide to Superoxide Dismutase[a]

field map	ionic strength, mM 0	0.144
no field	0.017 ± 0.002	
monopole (−4e)	0.011 ± 0.001	
all charges, $\epsilon = 80$	0.16 ± 0.01	0.21 ± 0.01
all charges, $\epsilon = r$	0.0085 ± 0.09	0.025 ± 0.003
all charges, $\epsilon_{protein} = 2$, $\epsilon_{solution} = 80$	0.58 ± 0.03	0.41 ± 0.02

[a] Simulated with a time step of 0.1 ps, a starting radius of 60 Å and an exit radius of 70 Å, using field maps calculated by the method of Klapper et al.[9] The association rate is relative to that for the whole neutral protein (2.36×10^{11} M^{-1} s^{-1}).

Collision Scheme. For the 90° target patch with an attractive force and a time step of 1 ps, three different methods for dealing with the trajectory step when it collides with the nonreactive surface of the molecule were used (see Methods section). No effect on the association rate was found, indicating that the details of the trajectory behavior at the surface have little effect on the probability that a trajectory reaches the target. This does not mean that the detail consideration of boundary effects in previous works[4,6] would not be important in cases where the force is larger and varies more rapidly at the surface, or where accuracies of better than 1% are required. However, for irregular molecules where it is necessary to describe the global shape, unavoidable inaccuracies accumulated by simplified boundary conditions probably outweigh these errors.

The Recessed Patch Model. We next consider the model of a recessed circular target patch in a sphere. This model is of some interest it is an example of a more irregular shape that can be both treated analytically[21] and represented exactly in Brownian dynamics simulations.[7] It has also been used as a model for enzymes with an active site cleft. The results from a simulation using a discrete representation of a 10° patch recessed by 3 or 15 Å, with a scale of 1 Å per grid unit is shown in Table III. The rates are somewhat lower than the analytically calculated rates, particularly for the deeper patch, and in excellent agreement with the simulations which use an exact representation of the shape.

Diffusion of Superoxide to Superoxide Dismutase. Three dielectric models were used to calculate the field of SOD due to the entire charge distribution: a uniform dielectric constant of 80 (used previously by Allison and McCammon[5] in similar studies); a distance-dependent dielectric, which is commonly used

(21) Samson, R.; Deutch, J. M. *J. Chem. Phys.* **1978**, *68*, 285.

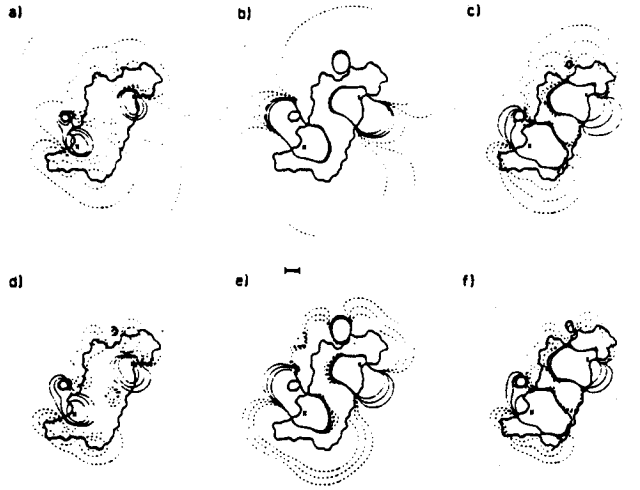

Figure 2. Potential maps of superoxide dismutase for three different dielectric models: uniform dielectric constant of 80 (a, d); distance-dependent dielectric $\epsilon = r$ (in Å) (b, e); solvent dielectric constant of 80, protein dielectric constant of 2 (c, f). Maps were calculated for an ionic strength of 0 mM (a, b, c), and 144 mM (d, e, f). Negative contours are dashed. Contour levels are at 0.5, 1, and 2 kT/e. The slice is taken through both active site copper atoms, marked by crosses. The protein surface is indicated by the bold solid line. The scale bar is 5 Å.

in molecular dynamics simulations for potentials and is the model used by Getzoff et al. in displaying the field around the SOD active site;[22] and finally a two-dielectric model, where the protein is considered as a low dielectric cavity in a high dielectric medium.[9] The potential maps for these three models are shown in Figure 2 for a slice through the two active site copper atoms. Results from the simulations based on these maps and for the monopole distribution are given in Table IV. The rates are normalized by the rate of association with the whole molecule in the absence of the field. This rate is 2.36×10^{11}, which we note is close the value of 2.27×10^{11} for a neutral 30-Å sphere, used by McCammon and co-workers as an idealized model for SOD.[7,8,23]

Considering the results at zero ionic strength first, the uniform dielectric model gives an enhancement of 16 over the monopole. This is a much greater increase than for the spherical model (1.4-fold[23]), and results from the inclusion of the actual protein shape, which gives a more accurate cleft and target location with respect to the protein charges. In fact it can be seen from Figure 2a that, using a 30-Å sphere to define the protein surface, the entire positive contour emanating from the active site lies beneath the surface, where it can play no role in guiding the substrate. This fact might explain the small increase in rate for the spherical model. The distance-dependent model gives a lower rate than the monopole case, since the lower effective dielectric constant for points close to the charges enhances the negative portion of the field, resulting in a barrier of about $2kT$ to the diffusing particle entering the active site.[9] The two-dielectric model gives a fiftyfold increase over the monopole case, giving a rate that is in fact a good fraction of that for the whole, neutral protein. This is a consequence of the cleft potential, resulting from focusing of field lines into the active site, away from the low-dielectric region of the protein. This causes the positive region of potential to extend from the protein much more, effectively providing a large target area for the diffusing substrate.[9]

An important experimental observation that provides evidence for the steering effect of the electric field is the decrease in enzyme rate with increasing ionic strength, in the range 0–150 mM.[13] The effect of ionic strengh on the association rate for the three different electrostatic models is shown in Table III. The striking feature is that only the two-dielectric model shows a decrease in association rate with increasing ionic strength. (Although we only show results for two concentrations, the calculated rate for this model decreases monotonically over the entire ionic strength range 0–0.45 M, in agreement with experiment.) Moreover, the size of the decrease at 144 mM is 30%, very similar to the measured percent decrease in enzyme rate.[13] The opposite behavior of the two other dielectric models can be understood qualitatively from examination of the electric field maps in Figure 2. For the two-dielectric model, the dominant effect is the decrease in effective target area with increasing ionic strength, which reduces the rate. For the other two models, however, the main effect of the increased shielding at higher ionic strengths is to reduce the extent of the negative potential areas that provide a barrier to the association. This is most pronounced for the distance-dependent case, which shows a threefold increase in rate, but this also occurs for the uniform dielectric model. This effect would account for the incorrect dependence seen in this ionic strength range in the previous simulations of Allison and McCammon, in which a uniform dielectric constant of 80 was used.[8]

Due to the discrete representation of the protein shape, and the relatively small size of the copper, there is some uncertainty in representing the target precisely. The effect of possible errors arising from this were examined by varying the target area by over a factor of 30. This did not alter the rate by more than about 10%. This indicates that, once the trajectory gets close to the copper, the probability of hitting the copper is extremely high, making the simulations insensitive to detailed features of the target and active site topography. The largest target size considered corresponds approximately to the mouth of the active site cleft (about 600 Å^2), which is roughly delineated by the $2kT$ contour in Figure 2c. Since the rate for the smallest target size examined (16 Å^2) is about 7% less than for this case, the probability for escape from the active site cleft at the $2kT$ contour level can be estimated as about equal to this figure. This compares with an escape probability of over 90% in the absence of the electric field (results not shown). The low escape probability indicates that the $2kT$ potential contour can effectively be considered as the target area, as was suggested earlier.[9] The area of this contour is about 2000 Å^2, which represents about 20% of the total protein surface area. This also confirms that most collisions with the active site channel are productive, as originally suggested by Getzoff et al.[22]

Discussion

The aim of this work is to develop a general strategy for simulating diffusion to an irregular particle in a nonsymmetric force field. A flexible and rapidly converging algorithm for irregular-body Brownian dynamics is developed by using a number of different computational strategies. The particle shape and the inner part of the force field are represented on a grid. The data that a trajectory needs to locate itself with respect to any shaped surface and target, to determine the force however complicated its spatial variation, to choose the step size, and to determine its fate, are all precalculated for every point in the grid. During simulations, this information is accessed by table lookup using the trajectory location as an index. The force at the trajectory is approximated by that at the nearest grid point, which requires no interpolation calculation. Collisions are handled by backstepping, which also requires no further calculation. Decisions about the fate of trajectories are handled by lookup tables, eliminating program branching, and allowing complete vectorization. Random numbers are generated in tables, rather than by calculation during simulations. Finally, the program is implemented on an array processor, which allows parallel processing, and all tables are stored in rapid access memory.

For the size of grids used here, the resolution in representing shape is limited to about 3–5%, so simulations are only run to 5% statistics for comparison with the analytical cases, although greater precision can be obtained with reasonable length runs.

For cases where the predominant concern is to represent the overall features of the particle shape and force field, describing the large scale spatial variations is more important than representing local topography or small scale changes in force accurately.

(22) Getzoff, E. D.; Tainer, J. A.; Weiner, P. K.; Kollman, P. A.; Richardson, J. S.; Richardson, D. C. *Nature* **1983**, *306*, 28.
(23) Allison, S. A.; Northrup, S. H.; McCammon, J. A. *Biophys. J.* **1986**, *49*, 167.

This is particularly true in the Brownian dynamics approach to solving the diffusion equation, since the observable of interest is the association rate, not the concentration distribution throughout space. Thus these type of simulations would be expected to be fairly robust with respect to small errors in modelling local conditions. This is borne out by the results of the first set of simulations. Within the simulation uncertainties, the rate is independent of the method used for handling nonreactive collisions with the surface. The discrete representation of the target particle has no effect, providing the fractional surface are occupied by the target is accurately reproduced. Similarly, at 5% statistics, even the relatively simple representation of the local force by that at the nearest grid point does not reduce the accuracy of the simulations significantly.

The results of varying the time step suggest an empirical criterion for constancy of force during a simulation time step to be that 90% or more of the steps should result in less than a 5% force change. For the type of electrostatic fields considered here, this can be achieved with a basic time step of 0.1 ps, combined with dynamic reduction of step size near the particle surface. This criterion is somewhat less stringent than previously suggested,[3] allowing larger time steps, which results in more rapid convergence.

By using realistic representations of the shape of the protein superoxide dismutase, and its field, the experimentally observed monotonic decrease in enzyme rate when the ionic strength is increased from zero to physiological[13] is obtained. This is not the case for simpler models, which represent the protein as a sphere, or which do not include explicitly the dielectric discontinuity at the protein surface when calculating the field.[8] With these simpler models, although the association rate begins to decrease at high ionic strength (>0.1 M), it is always higher than at 0 M. Another difference is that the large enhancement in association rate due to the field is not seen in simpler models. Although previous work[8,10,23] has shown that by including more realistic charge distributions the association rate is increased compared to that for the monopole distribution, it is still lower than or comparable to that in the absence of the field.

Representation of the actual shape of the protein also allows the target to be defined explicitly as the exposed surface of the active site copper atoms. With spherical models of the protein there is a serious difficulty in accurately representing both the size and location of the target. For example, there is no a priori reason to believe that the 10° patch on a sphere used in previous work consitutes a realistic representation of the active site.

In addition the association rates calculated here confirm the conclusions drawn in our previous work based on qualitative analysis of the potential maps alone.[9] These conclusions were based on an intuitively appealing argument which considered the effect of two features of the maps: the effective size of the target area, and the minimum barrier height that must be crossed by the diffusing substrate in order to reach the target. Two other results also justify this type of interpretation. Firstly, the rate of association with the copper is increased over thirtyfold by the field. Clearly the attractive part of the field is very important. Secondly, the chance of hitting the target once the substrate has reached the mouth of the active site cleft, at about the $2kT$ contour level, is over 90%. It therefore makes sense to talk about a large effective target area in this context.

Since the use of a grid representation of space limits the accuracy with which fine details can be modelled, the fact that the fate of the substrate trajectory is effectively determined once it reaches the $2kT$ contour level has another important consequence. It makes the simulation relatively insensitive to the local details of the electrostatic field, and the local trajectory behavior in the active site close to the copper, and hence insensitive to inaccuracies in these factors. In addition the accuracy of the simulation is not reduced by the imprecise representation of the target and its surroundings.

Apart from the discretization of space, a number of approximations and simplifications must be made in order perform the protein simulations. We now consider the possible errors arising from these.

Firstly, the field outside the grid clearly cannot be represented with the same detail as that inside. For the protein simulations, the potential was set to zero outside the grid, effectively truncating the force field. Since far from the protein the monopole contribution to the potential dominates, an estimate of the effect of this truncation error on the simulated association rates can be made using eq 3. The edge of the grid is about 50 Å from the center of the protein. With a net monopole charge of −4e for the protein, and an ionic strength of 144 mM, numerical integration of eq 5 for a centrosymmetric potential gives a correction of less than an angstrom, or less than 2%, for this radius. No appreciable correction occurs for ionic strengths down to 1 mM. What this means is that under these conditions the diffusion rate to the grid surface from infinity is essentially the same whether the field is present or not. At zero ionic strength the radius correction is about 27%. Inclusion of the higher order multipole terms will reduce the size of this correction by reducing the repulsive effect of the net negative charge. Since absolute zero ionic strength conditions can never actually be obtained in practice, due to the presence of protein counterions, buffer etc., the effect of truncating the field outside the grid in this case is probably also small. Although truncation of the fields does not affect the conclusions reached in this work, better representations of the field should be used if required. Two better approximations are to use either a centrosymmetric monopole field or to use a continuous dielectric approximation for calculating the potential outside the grid.

A second approximation is the use of the linearized Poisson–Boltzmann equation used to calculate the field. This significantly overestimates potentials above $1kT$. This effect will be considerable within the active site, where the potential is very large close to the copper. However, since the probability of hitting the target after reaching the $2kT$ contour level is over 90%, the simulation is probably not that sensitive to the overestimate of the potential further inside the active site.

No attempt has been made to include desolvation of the anion as it proceeds through the narrow part of the active site channel. This effect would reduce the steepness of the potential gradient in the active site. It is likely therefore that the probability of escape from the channel is larger than we calculate. Another unknown is the superoxide diffusion constant. The value used, 100 Å^2/ns, is typical for similar sized anions, although, depending on the solvation state of the ion, this may vary by about a factor of 2.[24] The diffusion constant is also assumed to be uniform, particularly within the active site. This may not be true, since the protein-bound water within the cleft may result in slower diffusion. Again, however, the high probability of the ion hitting the copper after having reached the mouth of the active site is unlikely to be much reduced due to either changes in the diffusion constant or desolvation effects. The superoxide may get to the copper faster or slower, but the rate is calculated only from the fraction of trajectories that reach the copper compared to those that do not.

Finally, we have neglected the rotational diffusion of the protein in these simulations. However, the rotational diffusion coeff. of a macromolecule the size of SOD is of the order of 0.06 rad/ns.[25] For a rapidly diffusing particle such as superoxide, with a translational diffusion coefficient around 100 Å^2/ns, the effect of rotational diffusion would be expected to be small.[19]

Simulations including both the nonlinear Poisson–Boltzmann equation and rotational diffusion are currently underway in order to check these conclusions.

In summary, the good agreement with analytical results at 5%, the robustness of the simulations with respect to the time step and surface representation and the ability to use computationally very simple procedures for the reflective boundary condition and force determination mean that this procedure for irregular-body Brownian dynamics can be made to converge rapidly and accurately for any arbitrary shape and field. Representation of very

(24) Conway, B. E. *Ionic Hydration in Chemistry and Biophysics*: Elsevier: Amsterdam, 1981; Chapter 4.
(25) Tanford, C. *Physical Chemistry of Macromolecules*: Wiley: New York, 1961; Chapter 6.

irregular, real shapes, such as enzymes, is straightforward. Complex electrostatic fields can also be represented easily. Using this method, we have demonstrated for the first time that a significant enhancement of the association of superoxide to superoxide dismutase is provided by this field and that this enhancement is decreased as the ionic strength is increased from zero to physiological ionic strength. The relative decrease is very similar to the experimentally determined decrease in enzyme rate.[13] Morover, this work shows that the correct ionic strength behavior is only obtained when both the shape of the enzyme and the dielectric discontinuity at the protein surface are included in the simulations.

Acknowledgment. This work was supported by grant No. GM-30518 (NIH), N00014-86-K-0483 (ONR), RR00442 (NIH), and AC02-72 (DE).

Registry No. SOD, 9054-89-1; $O_2^{\cdot-}$, 11062-77-4.