Effect of DNA Superhelicity and Bound Proteins on Mechanistic Aspects of the Hin-Mediated and Fis-Enhanced Inversion

Jing Huang, Qing Zhang, and Tamar Schlick
Department of Chemistry and Courant Institute of Mathematical Sciences, New York University and the Howard Hughes Medical Institute, New York, New York 10012

ABSTRACT Using a recently developed inhomogeneous, macroscopic model for long DNA bound to proteins, we examine topological and geometric aspects of DNA/protein structures and dynamics on various stages of the Hin inversion pathway. This biological reaction involves exchange of DNA in a synaptic complex that brings together several DNA sites bound to Hin dimers as well as Fis enhancers. Brownian dynamics simulations in the millisecond timescale allow us to follow and analyze the DNA/protein dynamics trajectories and to examine the effects of DNA superhelicity and protein binding on various reaction steps. Analysis of the generated kinetic pathways helps explain mechanistic aspects regarding the process by which two or three protein-bound DNA sites come to close spatial proximity and show that how topological selectivity (two trapped supercoils), enhancer binding, and properties of supercoiled DNA play critical roles in regulating the inversion reaction. Specifically, a critical amount of DNA superhelicity (e.g., $|\sigma| > 0.02$) leads to an optimal interplay for the first reaction step—two-site juxtaposition—between large-scale random rearrangements of Hin-bound DNA and local slithering within branches of plectonemes. The three-site juxtaposition, the second step, is significantly accelerated by the presence of an enhancer protein that, due to severe local bending, also alters juxtaposition mechanisms, especially for superhelical density magnitude greater than around 0.04.

INTRODUCTION

Many biological processes involve DNA/protein interactions. The binding of proteins affects structural and dynamic properties of DNA and therefore directly impacts biological function (Lilley, 1995; Bustamante and Rivetti, 1996). Such protein binding effects are complicated in part because the biologically active DNAs are usually topologically closed or behave as such (via effective constraints on the DNA imposed by proteins). The supercoiling of DNA is regulated naturally by various enzymes like gyrase, topoisomerase, and others (Bates and Maxwell, 1993; Vinograd et al., 1965; Vologodskii and Cozzarelli, 1994; Watson et al., 1996; Gralla, 1996).

Extensive research has focused on the effect of protein binding on DNA. X-ray crystal structures and other atomic-level studies can investigate protein/DNA interactions in detail (Lilley, 1995), but current techniques limit the size of DNA being studied to relatively short segments. On the other hand, large supercoiled DNA systems with thousands of basepairs have been modeled and simulated by Monte Carlo methods (Vologodskii and Cozzarelli, 1994) and Brownian dynamics (Allison et al., 1989; Jian et al., 1998; Huang et al., 2001) to complement experiments such as electron microscopy and other suitable methods for large floppy molecules (Boles et al., 1990; Dubochet et al., 1992; Zuccheri et al., 2001). In the theoretical studies, DNA is usually represented by a discrete polymer-like chain model, which well reproduces equilibrium and dynamic properties of supercoiled DNA systems.

Submitted December 31, 2002, and accepted for publication April 18, 2003.

Address reprint requests to Tamar Schlick, Fax: 212-995-4152; E-mail: schlick@nyu.edu.

0006-3495/03/08/804/14 $2.00

Protein binding not only changes the local structure of DNA (as well known from high-resolution data) but also the global geometric and dynamic properties of the macromolecular system. To study these effects, we have developed (Huang and Schlick, 2002) a general ''mesoscale'' model and simulation procedure to model such protein/DNA systems. By a mesoscopic model, we mean one that incorporates local details as necessary and global characteristics where possible. In our case, the former involves an inhomogeneous treatment for the DNA/protein sites and the latter includes a polymer-level description for the free DNA segments. The goal in the model is to incorporate sufficient detail (or accuracy) but at the same time allow long simulations of large systems. Essential aspects of the inhomogeneous model are local inhomogeneous bending, electrostatic treatment via effective charges distributed on a macromolecular surface (based on available crystal structures and our DiSCO algorithm (Beard and Schlick, 2001a)), and an augmented Brownian dynamics protocol based on a nonidentical bead hydrodynamic description.

Our previous paper (Huang and Schlick, 2002) discussed details of the parameterization of the DNA/protein model in the Hin-mediated inversion system (Silverman and Simon, 1980; Johnson and Simon, 1985; Heichman and Johnson, 1990), which motivated our modeling and algorithmic work (see Fig. 1). That paper also presented analyses of simple two-site juxtaposition dynamics in the system and the effect of protein binding on supercoiled DNA dynamics. The model and simulation protocol are summarized in the section of ''DNA and protein models'' under Materials and Methods here. In this paper, we develop a set of analytic methods to numerically distinguish DNA/protein structures based on their topological properties and a protocol for modeling the *hix*-pairing intermediate complex (see the sections ''Topo-

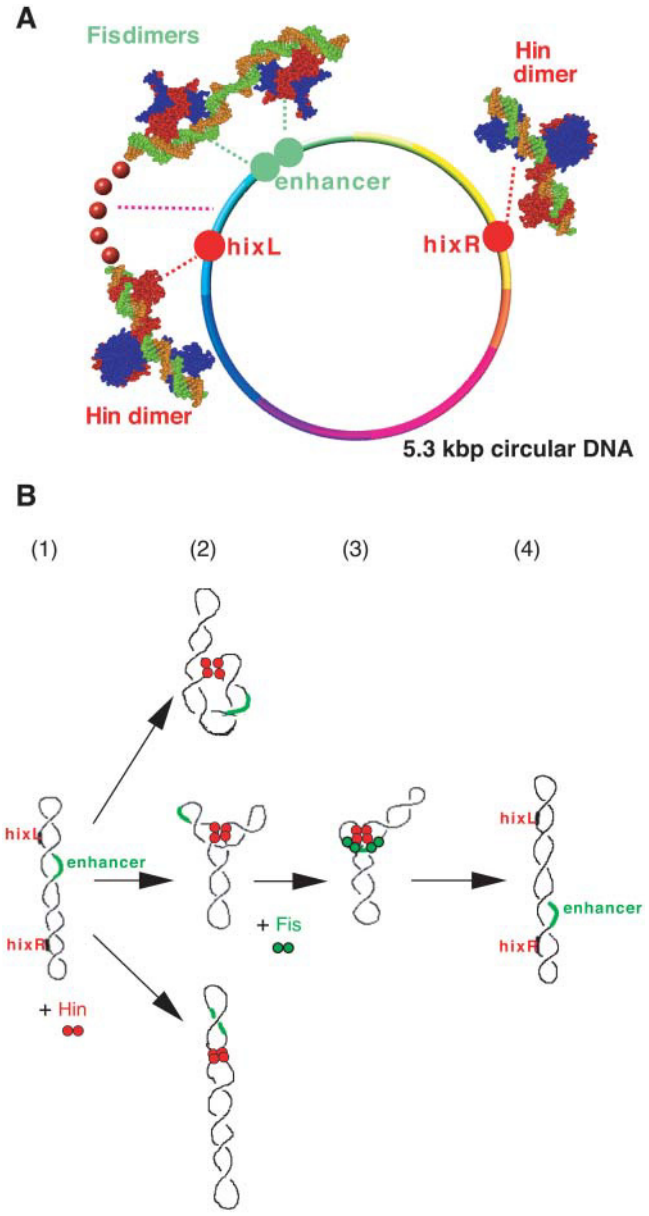

FIGURE 1 Macromolecular model and topological selectivity in a Hin-mediated inversion. (*A*) Macromolecular model used in the study of the Hin-mediated inversion. (*B*) Proposed pathway for the Hin-mediated inversion reaction (Merickel et al., 1998). For an intermediate structure with two Hin-bound *hix* sites juxtaposed, the Hin-Hin pairing must trap exactly two negative supercoils (*middle row* in step 2) to allow the following steps to successfully occur.

logical description of a juxtaposed structure'' and ''Modeling the *hix*-pairing complex'' under Materials and Methods and Figs. 2 and 3). Together with the protein/DNA model, we use these techniques to simulate different stages of the Hin inversion pathway over the millisecond time frame to analyze how large-scale DNA motions with bound proteins affect the reaction mechanism. In the Results and Discussions section, we first analyze simulations with and without the topological criteria for DNA as a function of DNA superhelical densities to systematically study the effects of topological selectivity on the DNA dynamics. The results help explain why it is more facile to align two sites with the correct topology (two trapped negative supercoils) in a naturally supercoiled DNA/protein system than in a system with less tight DNA supercoiling (smaller superhelical density $|\sigma|$). It is the combined contribution from large-scale random rearrangements and local slithering in the branched DNA that produces an optimal range of DNA superhelicity (e.g., $|\sigma| > 0.02$ for juxtaposition with two trapped negative supercoils). Analyzing this behavior, consistent with experimental observations (Lim and Simon, 1992), provides further insights into this supercoiling effect. We also examine behavior of various systems and compare the dynamics with and without the binding of the auxiliary Fis protein (factor for inversion stimulation) to the DNA substrate. Our results show that protein binding has profound effects on the DNA dynamics: Fis substantially distorts the system and crucially accelerates the three-site juxtaposition times, most markedly when the superhelical density magnitude $|\sigma|$ is greater than 0.04. We also estimate the juxtaposition rates for various proposed stages of the reaction pathway; though we cannot present accurate reaction rates with our mesoscale model —the rates of forming protein/DNA and protein/protein complexes are also unknown—our measurements based on global DNA dynamics suggest that the two-step pathway, instead of the direct three-site juxtaposition mechanism, is the dominant mechanism in the Hin inversion process.

MATERIALS AND METHODS

Full details of our general method to model and simulate large systems of proteins bound to supercoiled DNA molecules have been described (Huang and Schlick, 2002) and are summarized below. Essentially, the free DNA segments are modeled based on the worm-like chain and bead model, which is parameterized by average properties of supercoiled DNA (Jian, 1997; Huang et al., 2001). The DNA sites bound to proteins require a more elaborate treatment of inhomogeneous bending and electrostatic interactions, which necessitate changes in all aspects of the mechanical model, hydrodynamics formulation, and Brownian dynamics protocol. Here, we discuss the new methods as well as additional analysis tools we developed to analyze the inhomogeneous and topological properties of large DNA systems bound to proteins. The variables used in the model and their assigned values are tabulated in Tables 1 and 2. We also summarize how we model the *hix*-pairing intermediate complex: because we cannot afford an explicit atom-level description of hydrogen bonding and other interactions, an effective potential is introduced. These combined methods allow us to quantitatively track and analyze configurations generated in the dynamics simulations of DNA and proteins in a fashion consistent with the way biologists investigate these processes.

DNA and protein models

Homogeneous DNA model and algorithms

Energy components and parameterization. The basis for our work on DNA supercoiling and macroscopic biopolymers, such as focusing on site juxtaposition (Huang et al., 2001) and chromatin fiber structure (Beard and

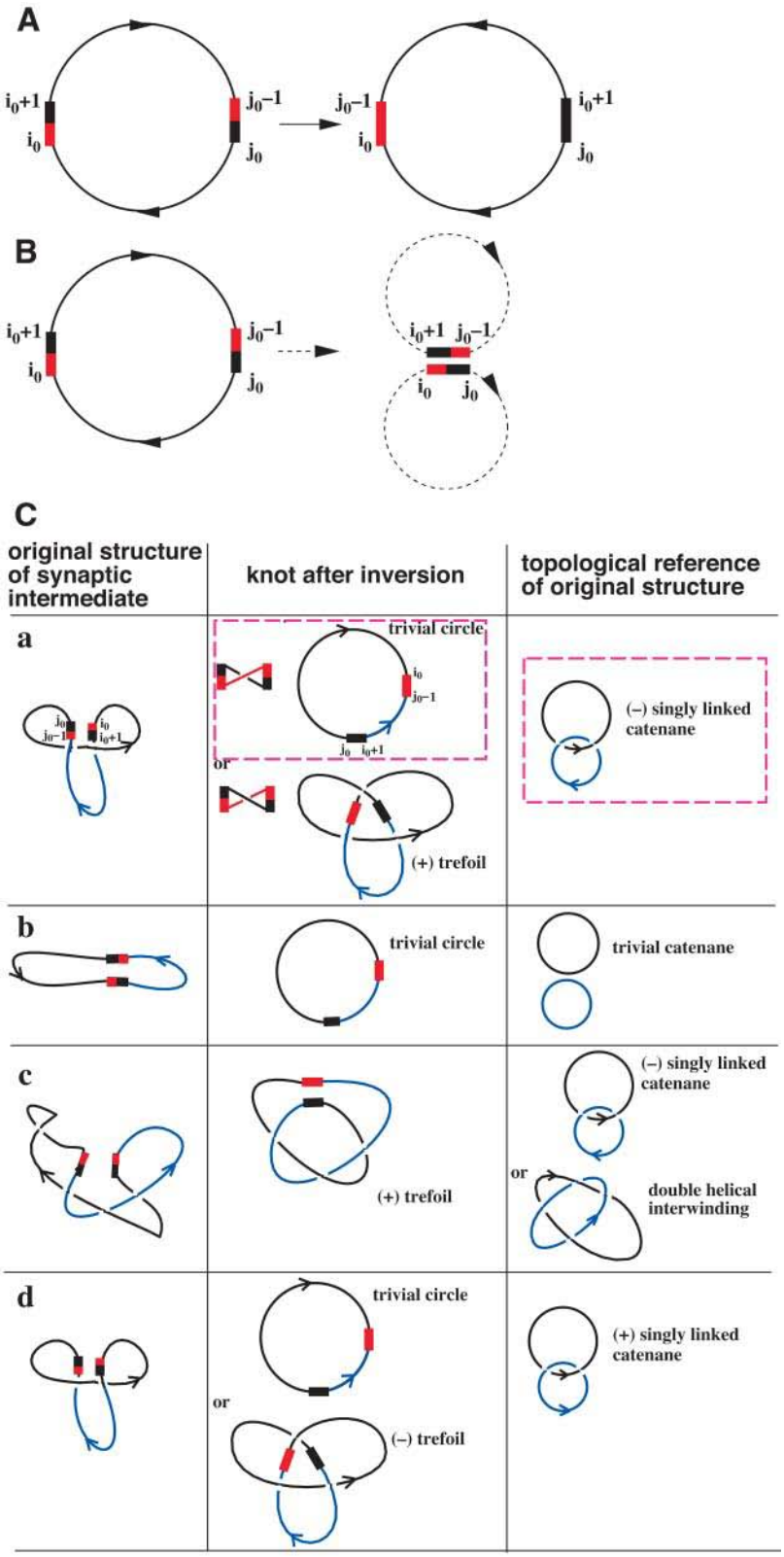

FIGURE 2 (*A*) Analysis of reaction. Illustration of the inversion result: The inversion results in a circular DNA with a part of DNA inverted. (*B*) A different result generated by a break-and-seal action different from the one in the real process (*A*). Topological status of this reference structure depends on how many supercoils are trapped in the juxtaposed structure (*C*). (*C*) Identifying the properly juxtaposition conformation for *hix*-pairing structure based on topological properties.

Schlick, 2001b), is the homogeneous worm-like chain/bead model of DNA, combining well-known polymer chain representations with the bead model required for defining hydrodynamic interactions. The homogeneous worm-like chain/bead model is suitable for average properties of DNA of mixed sequence. Based on the discrete worm-like chain developed by Allison and co-workers (Allison, 1986; Allison et al., 1989) and extended by Chirico and Langowski for supercoiled DNA (Chirico and Langowski, 1994; Chirico and Langowski, 1996), we first developed a carefully parameterized model

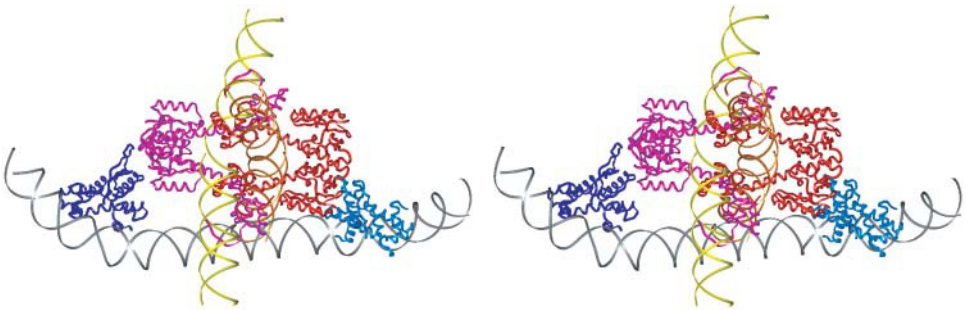

FIGURE 3 Stereo views of the molecular model of invertasome complex from Johnson's group. Two Fis dimers are bound to an enhancer sequence and two Hin dimers are complexed to *hix* sites.

and simulation procedure for studying the dynamic properties of long *linear* DNA (Jian, 1997; Jian et al., 1997). Accuracy and reliability were demonstrated with respect to all available equilibrium and dynamic properties as obtained from well-established Monte Carlo procedures and experimental data such as translational diffusion constants for various DNA lengths; predictive measurements for autocorrelation times, end-to-end distances, and rotational diffusion constants, as a function of length, were also reported (Jian, 1997; Jian et al., 1997). Extensions from linear to *circular* supercoiled DNA were subsequently reported (Jian et al., 1998), including improvements that extended the coverage from low salt to physiological salt conditions (Huang et al., 2001).

TABLE 1 List of symbols

Symbol	Definition
r_i	Coordinates of the ith vertex
$\{\alpha_{i,i+1},\beta_{i,i+1},\gamma_{i,i+1}\}$	Euler angles for the transformation from $\{\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i\}$ to $\{\mathbf{a}_{i+1}, \mathbf{b}_{i+1}, \mathbf{c}_{i+1}\}$
l_i	Segment length between ith and $i+1$th vertices
$\mathbf{r}_{ij}$	Displacement between the ith and jth vertices
E^b	Elastic bending potential
E^t	Elastic twisting potential
E^s	Elastic stretching potential
E^e	Electrostatic potential
E^v	Excluded volume potential
θ_i	Bending angle around the vertex i
ϕ_0	Homogeneous intrinsic twist in one model segment
Γ_i	Roll-like bending angle
Υ_i	Tilt-like bending angle
n^q	Number of charge points used to describe a complex or a segment
$\mathbf{D}^t$	Hydrodynamic diffusion tensor for nonidentical bead model
$\mathbf{F}^t,\mathbf{r}^t$	Collective force and position vectors at time t
$\mathbf{R}^t$	Random forces in Brownian dynamic algorithms at time t
E^{hh}	Virtual potential for *hix*-pairing
$r_{h_1h_2}$	Distance between the center of two Hin/DNA complexes
$\mathbf{r}^i_{ihin}$	Coordinates of the ith Hin/DNA complex
$\mathbf{r}_{i_{hin}}$	Coordinates of the midpoint of two paired Hin/DNA complexes
$\mathbf{r}^i_{ifis}$	Coordinates of the ith Fis binding site
$\{\mathbf{a}^i_{ihin}, \mathbf{b}^i_{ihin}, \mathbf{c}^i_{ihin}\}$	Local coordinate system on the vertex for the ith Hin/DNA complex
h^{hh}	Stretching potential for the virtual *hix-hix* complex
g^{hh}_a, g^{hh}_b	Bending potential for the virtual *hix-hix* complex
τ_{c1}	Time taken for two *hix* sites to juxtapose with exactly two negative supercoils trapped
τ'_{c1}	Time taken for two *hix* sites to juxtapose without any topological criteria
σ	DNA superhelical density
τ_{c2}	Time taken for the three-site juxtaposition of two *hix* sites and the enhancer sequence after the *hix*-pairing event
Lk	Linking number

Our homogeneous model represents a DNA molecule of size n Kuhn statistical lengths (or equivalently, $2n$ persistence lengths) as a chain of kn straight elastic segments of equilibrium length l_0. Each chain vertex i has a body-fixed coordinate frame (with corresponding Euler angles $\alpha_{i,i+1}$, $\beta_{i,i+1}$, $\gamma_{i,i+1}$) describing the transformation from local frame i to frame $i+1$.

The elastic energy contains terms for stretching (E^s), bending (E^b), torsion (E^t), and electrostatics (E^e), as follows:

$$E = E^s + E^b + E^t + E^e$$
$$= \frac{h}{2}\sum_{i=1}^{nk}(l_i - l_0)^2 + \frac{g}{2}\sum_{i=1}^{nk}(\theta_i)^2 + \frac{C}{2l_0}\sum_{i=1}^{nk}(\alpha_{i,i+1} + \gamma_{i,i+1} - \phi_0)^2 + \frac{\lambda^2 l_0^2}{n^{q^2}\epsilon}\sum_{i,j}^{N^q}\frac{\exp(-\kappa r_{ij})}{r_{ij}}. \quad (1)$$

Here h and g in E_s and E_b are stretching and bending constants, respectively; l_i and θ_i (or Euler angle $\beta_{i,I+1}$) are the length and angular displacement of the ith segment. We set $h = 100\,k_BT/l_0^2$, where k_BT is the Boltzmann factor, so that the "bond" variance of l_i is close to $l_0^2/100$ (Jian et al., 1998; Jian, 1997). The bending rigidity g (related to the bending persistence length (Schlick, 1995)) is defined so that the Kuhn statistical length corresponds to k rigid segments, where $k \geq 10$ (Vologodskii, 1992); typically, $k = 10$, $l_0 = 10$ nm, $g = 4.81k_BT$, and the Kuhn length is 100 nm. The torsional angle is computed between adjacent segments using the body-fixed coordinate frame attached to each vertex as $\phi_{i,i+1} = \alpha_{i,i+1} + \gamma_{i,i+1}$, and the reference angle ϕ_0 is the equilibrium twist between DNA segments. The torsional rigidity constant C is set to $C = 3 \times 10^{-19}$ erg-cm based on experimental data (see (Jian et al., 1998)). The superhelical density σ is set in a given simulation to the desired value by specifying the initial twist $\phi_{i,i+1}$ between each pair of segments i and $i+1$ (Jian, 1997).

The electrostatic energy E^e of the DNA model is defined by Debye-Hückel point charges on the DNA segments. The variable υ is DNA's effective linear charge density at a given molar salt concentration of monovalent ions (c_S) with corresponding Debye length $1/\kappa$, ε is the dielectric constant of water, $N^q = n^q kn$ is the total number of charges, and r_{ij} is the distance between charges i and j. We set the number of charges per segment n^q so that it accurately reproduces a continuously charged DNA segment (Schlick et al., 1994) based on Stigter's pioneering contribution (Stigter, 1977); typically, $n^q = 5$ works for our highest monovalent salt concentration (Huang et al., 2001). Thus, the isotropic salt concentration is modeled by adjusting the salt-dependent variables λ and κ (Schlick et al., 1994); for a 0.2 M salt concentration, $\lambda = 40.9$ e/nm and $\kappa = 1.477$ nm^{-1} (Stigter, 1977). This "high-salt" modeling required the improvement ($n^q > 1$) described in (Huang et al., 2001). If DNA segments are closer than 2 nm, the electrostatic energy is replaced by the short-range exclusion (repulsion)

TABLE 2 Elastic, geometric, and electrostatic parameters used in the DNA/protein model

Parameter	Definition	Value
N	Number of the vertices in the circular DNA model	176 (5.28 kbp)
l_0	Equilibrium segment length	10 nm
l_e	Kuhn statistical length of DNA	100 nm
k	Number of elastic segment for each Kuhn statistical length	10
C	Twisting rigidity constant	3.0×10^{-19} erg·cm
k_B	Boltzmann's constant	1.38×10^{-23} J/K
T	Absolute temperature	298 K
g	Bending rigidity constant of average DNA	pk_BT/l_0
h	Stretching rigidity constant	$100\,kBT/l_0^2$
Δt	Time step for Brownian dynamic simulations	600 ps
d_0	Radial distance criterion for site juxtaposition	10 nm
c_S	Monovalent salt concentration	0.2 M
i_{hin}^1, i_{hin}^2	Vertices to which a Hin dimer can bind	1, 34
i_{fis}^1, i_{fis}^2	Vertices to which a Fis dimer can bind	5, 6
$\Gamma_{ihin}^0, \Upsilon_{ihin}^0$	Equilibrium tilt-like and roll-like bends on an *ihin* vertex by a Hin dimer	−18°, 0°
$\Gamma_{ifis}^0, \Upsilon_{ifis}^0$	Equilibrium tilt-like and roll-like bends on an *ifis* vertex by a Fis dimer	60°, 0°
$g_{ihin}^\Gamma, g_{ihin}^\Upsilon$	Computational tilt-like and roll-like bending rigidity for *ihin* vertices	g, g
$g_{ifis}^\Gamma, g_{ifis}^\Upsilon$	Computational tilt-like and roll-like bending rigidity for *ifis* vertices	3.0 g, 3.0 g
κ	Inverse Debye length (salt-dependent, here for 0.2 M)	1.477 nm^{-1}
λ	Effective linear charge density of double helix (here for 0.2 M)	40.9 e/nm
ϵ	Relative dielectric constant of aqueous medium	80
n_i^q	Number of effective charges for a DNA segment, a Hin-dimer/DNA complex, and a Fis-dimer/DNA complex, respectively	5, 200, 100
μ	Computational short-range repulsion force	35 pN
ρ_0	Hydrodynamic radius of a DNA segment of l_0	2.24 nm

energy E^v, necessary to prevent DNA strands from passing through each other (Huang et al., 2001), where:

$$E^v = -\sum_{i,j}^{N^q} \mu r_{ij}, \quad \text{if } r_{ij} < 2\,\text{nm}; \quad E^v = 0, \quad \text{if } r_{ij} \geq 2\,\text{nm}, \tag{2}$$

and $\mu = 35$ pN. The electrostatic energy is neglected between these adjacent segments because the bending rigidity already accounts for the local stiffness.

Hydrodynamic interactions and BD protocols. Hydrodynamic interactions between the DNA and the solvent are computed by placing virtual beads of radius ρ_0 at each vertex of the worm-like chain; these beads are only involved in hydrodynamic interactions and do not affect equilibrium properties of the model chain. When $k = 10$, a value of $\rho_0 = 2.24$ nm for the hydrodynamic radius reproduces experimental diffusion coefficients closely (Huang et al., 2001). We define the identical bead hydrodynamic interactions for protein-free DNA using the Rotne-Prager-Yamakawa diffusion tensor (Rotne and Prager, 1969; Yamakawa, 1970)—a second-order approximation for two identical beads diffusing in a Stokes fluid (see also Schlick (2002, Chapter 9) for details). Algorithmic improvements such as updating the hydrodynamic tensor less frequently than the position vector (e.g., every 10 steps here) led to a speedup of 4–7 times. With the current energy terms, we are able to increase the feasible time step size to around 600 ps (Jian, 1997; Jian et al., 1998).

For long-time simulations of DNA dynamics, we use the Iniesta-Garcia de la Torre second-order modification of the first-order Ermak-McCammon algorithm (Ermak and McCammon, 1978; Iniesta and Garcia de la Torre, 1990), which updates particle positions according to:

$$\mathbf{r}^{t+\Delta t} = \mathbf{r}^t + \left(\frac{\Delta t}{k_B T}\right)\mathbf{D}(\mathbf{r}^t)\cdot\mathbf{F}^t + \mathbf{R}^t. \tag{3}$$

Here Δt is the time step, $\mathbf{r}^t$ and $\mathbf{F}^t$ are the collective position and force vectors for the system at time t, $\mathbf{D}(\mathbf{r}^t)$, or $\mathbf{D}^t$ for short, is the configuration-dependent hydrodynamic diffusion tensor, and D_{ij} is the ijth entry of $\mathbf{D}^t$. The random-force vector $\mathbf{R}^t$ is a Gaussian white noise process (used to mimic thermal interactions with the solvent) related to the diffusion tensor $\mathbf{D}$ by the fluctuation/dissipation theorem (Kubo, 1966):

$$\langle(\mathbf{R}^t)(\mathbf{R}^{t'})^T\rangle = 2\Delta t\mathbf{D}^t, \quad \text{with } \langle\mathbf{R}^t\rangle = 0 \quad \text{and}$$
$$\langle(\mathbf{R}^t)(\mathbf{R}^{t'})^T\rangle = 0 \quad \text{for } t \neq t'. \tag{4}$$

The computations involved in the random-force vector are actually the most time-consuming aspect of the Brownian dynamics (BD) protocol. The traditional way, based on a Cholesky decomposition of $\mathbf{D}^t = \mathbf{L}\mathbf{L}^T$ (see (Dahlquist and Bjorck, 1974)) increases in computational time as the *cube* of the system size (three times the number of beads), since a Cholesky factorization of $\mathbf{D}^t$ is required at every step. The alternative approach based on Chebyshev polynomials proposed by Fixman (1986) only increases in complexity in theory with the *square* of the number of variables.

Our development and application of the Chebyshev approach for computing $\mathbf{R}$ in simulations of long DNA (see (Schlick et al., 2000, box on Page 49) and (Schlick, 2002, Chapter 13)) has demonstrated computational savings for large systems: for DNA longer than 10 kbp, the Chebyshev scheme reduces the simulation time by >50% (Schlick et al., 2000): 60 versus 120 days for 10 ms simulations of 12 kbp DNA on SGI Origin 2000, 300 MHz processors. In the current Hin system, we tested both Cholesky and Chebyshev approaches and found that the standard Cholesky approach is slightly faster for 5.28 kbp DNA. Therefore, all the simulations reported here use the Cholesky approach.

The time step in BD algorithms of DNA is generally of order 100 ps, much greater than the nanosecond time step used in all-atom molecular dynamics simulations (Schlick, 2001). Improvements of the hydrodynamic calculations and parallelization of the Brownian dynamics algorithm on SGI multiprocessors made possible the modeling of plasmids of 6–12 kbp for hundreds of milliseconds in days to a few weeks of computing (Schlick et al., 2000). These advances are essential to address site juxtaposition events in large systems at the physiological salt concentrations as used in the experiments (e.g., Benjamin et al., 1996; Oram et al., 1997; Sessions et al., 1997).

Inhomogeneous model of DNA bound to proteins

Our inhomogeneous model was developed with the motivation to study the Hin/Fis inversion reaction (see Fig. 1). Details are extensive and given in full in (Huang and Schlick, 2002; Huang, 2003). Here we only sketch some key features. Essentially, electrostatics interactions are evaluated using the DiSCO algorithm (Beard and Schlick, 2001a). Directional bending is formulated by using nonidentical beads with more complex potentials to further mimic the inhomogeneous effects caused by protein binding. Hydrodynamic interactions are treated with the nonidentical Oseen diffusion tensor (Rotne and Prager, 1969; Garcia de la Torre and Bloomfield, 1977). The resulting models thus account for basic elements of protein binding effects while remaining computationally tractable.

DiSCO's electrostatic approximation. The DiSCO (Discrete Surface Charge Optimization) electrostatic approach aims to model the proteins and protein/DNA sites in much greater detail than done in macroscopic models by bridging traditional descriptions of electrostatic interactions on the all-atom level with continuum solvation approximations on the macroscopic level (Beard and Schlick, 2001a, Zhang et al., 2003). On the basis of available crystal structures, DiSCO effectively defines charged surfaces using a discrete set of Debye-Hückel charges distributed on a virtual surface enclosing the macromolecule. The approximation relies on the linear behavior of the Poisson-Boltzmann equation in the far zone; thus contributions from a number of molecules may be superimposed, and the electrostatic potential, or equivalently the electrostatic field, may be quickly and efficiently approximated by the summation of contributions from the set of charges. The desired accuracy of the approximation is achieved by minimizing the difference between the Poisson-Boltzmann electrostatic field (solved using DelPhi (Rocchia et al., 2002)) and that produced by the linearized, Debye-Hückel approximation using our efficient truncated Newton optimization package (Schlick and Fogelson, 1992; Derreumaux et al., 1994; Xie and Schlick, 1999a,b). We use 200 and 100 effective charges to model Hin/DNA and Fis/DNA complexes, respectively (Huang and Schlick, 2002), parameterized using a program (Zhang et al., 2003) developed to construct irregular surfaces enclosing macromolecular systems.

Inhomogeneous bending. The inhomogeneous bending potentials rely on standard nucleic acid bending angles roll and tilt (Dickerson, 1989) formulations. Orthogonal bending directions Γ_i (tilt-like) and Υ_i (roll-like) describe the position of bead i with respect to bead $i-1$ in the body-fixed reference frame (Huang and Schlick, 2002). The energy of bending E_i^b as a function of Γ_i and Υ_i is:

$$E_i^b = \frac{g_i^\Gamma}{2}(\Gamma_i - \Gamma_i^0)^2 + \frac{g_i^\Upsilon}{2}(\Upsilon_i - \Upsilon_i^0)^2, \qquad (5)$$

where Γ_i^0 and Υ_i^0 are corresponding equilibrium bending values and g_i^Γ and g_i^Υ are the corresponding bending rigidities. Based on available structures and data for the Hin inversion system (Perkins-Balding et al., 1997; Pan et al., 1996; Feng et al., 1994; Safo et al., 1997), we set the equilibrium bending angles as: $\Gamma_{Hin}^0 = -18°$, $\Upsilon_{Hin}^0 = 0°$, $\Gamma_{Fis}^0 = 60°$, and $\Upsilon_{Fis}^0 = 0°$ (Fig. 1). Measurements of the bending rigidity of the Fis dimmer binding sites suggest that the DNA becomes more rigid; we approximate $g_i^\Gamma = g_i^\Upsilon = 3.0\,g$ for Eq. 5, where g is the homogeneous value used in Eq. 1.

Topological description of a juxtaposed structure

Many biological reactions possess essential topological selectivity mechanisms. For the Hin-inversion process, for example, only those *hix-hix* paired conformations with exactly two negative supercoils trapped can successfully proceed to form the final three-site juxtaposition complex and invert the sequence (Fig. 1 *B*) (Johnson and Simon, 1985). Our work focuses on the large-scale conformational rearrangements (steps 1–2 and 2–3 of Fig. 1 *B*) but we must consider how the various possible products arriving to step 3 affect the topological change (step 3–4). We analyze the possibilities in Fig. 2 (discussed below), where we introduce points i_0, i_0+1, j_0, j_0-1 to label the two *hix*-bound DNA sites.

The topology of a closed chain does not change during the large-scale conformational rearrangement—the DNA chain does not break or pass through itself in this process—(such as step 2—3 in Fig. 1 *B*). Therefore, a correct conformation with two juxtaposed *hix* sites at status 2 of Fig. 1 *B* must satisfy two criteria: the conformation must trap two negative supercoils, and its recombination product must be the trivial circle (status 4 in Fig. 1 *B*) instead of topologically more complicated knots. Our challenge is to develop a protocol of quantitative and programmable criteria based on topology theory to distinguish such conformations from others.

Fig. 2 *A* illustrates an inversion process. We use $i_0 \rightarrow i_0+1$ and $j_0 \rightarrow j_0-1$ to denote the two *hix* recombination sequences; the segment between them $(i_0+1 \rightarrow j_0-1)$ is the sequence to be inverted. In the Hin inversion reaction, the resulting conformation (Fig. 2 *A*, at *right*) has two new *hix* sites $i_0 \rightarrow j_0-1$ and $j_0 \rightarrow i_0+1$, with the sequence between in the direction $i_0+1 \rightarrow j_0-1$. The knot type of this resulting conformation reflects the geometric and topological properties of the initial structure (*middle column* in Fig. 2 *C*). We also consider another path to help analyze the correct intermediate structure of the reaction. Fig. 2 *B* shows that two closed chains would be produced if cross-linking would occur (sites i_0+1 to j_0-1 and i_0 to j_0). In general, these two closed chains are linked catenanes, and the type of linkage reflects a topological property of the original substrate (i.e., how many supercoils are trapped in a juxtaposed structure).

To analyze the supercoiled conformation at step 2 (Fig. 1 *B*), we show in Fig. 2 *C* (*middle* and *last* column) topologies of structures corresponding to the two cases of the DNA being inverted (as in Fig. 2 *A*) or broken and resealed (as in Fig. 2 *B*), respectively. Various topological products can result. For the correct intermediate structure in Fig. 2 *C* (*a*) with two trapped negative supercoils, both a trivial circle after inversion and a negatively single-linked catenane can result, both highlighted by the boxes.

Other juxtaposed conformations result in different knot/catenane products. Fig. 2 *C* (*b*) shows a simply juxtaposed conformation with no supercoils trapped. Its corresponding inversion produces a pair of nonlinked circles. Fig. 2 *C* (*c*) shows the case where three supercoils are trapped. The inverted conformation is now a positive trefoil, and two different catenanes result from the reaction—a negatively single-linked catenane or a catenane formed by double helical interwinding. Fig. 2 *C* (*d*) shows the conformation with two positive supercoils trapped. It is a mirror image of the conformation in example 2 *C* (*a*), so resulting knots and catenanes are also mirror images of those in that example. More complicated conformations are likely to produce knots and catenanes of higher complexity.

Having enumerated these scenarios, we can now develop a protocol to distinguish the correct conformation that would result in a trivial circle from a virtual negatively single-linked catenane. Using this method, we can automate tests for a particular juxtaposed conformation to occur. This allows us to filter specific juxtaposed conformations that can proceed as consistent with observations (Glasgow et al., 1989; Heichman and Johnson, 1990; Merickel et al., 1998) to the DNA strand exchange stage; alternative conformations can be rejected as an ''unsuccessful juxtaposition attempt''.

By definition, topological invariants do not change with geometric deformations that do not entail chain breaking; those conformations thus belong to the same isotopic class in topology. These mathematical descriptors can thus determine the knot/catenane type of a particular conformation. An ideal invariant should provide a unique result for each isotopic class. The Alexander polynomial, named after its creator, James Waddell Alexander II, is a polynomial of one variable, in the case of knot, and of two variables, in the case of catenanes. It can distinguish knots/catenanes based on the concept of linear color tests and it has shown to be effective computationally (Frank-Kamenetskii and Vologodskii, 1981; Michels and Wiegel, 1986). However, the Alexander polynomial cannot distinguish a knot/catenane from its mirror image, such as a positive trefoil and a negative trefoil. The linking number *Lk* (Bates and Maxwell, 1993), one of the simplest topological invariant, however, can distinguish between two mirror images:

$$Lk = \frac{1}{4\pi}\oint_{C_1}\oint_{C_2}\frac{(\mathbf{r}'(s)\times\mathbf{r}'(t))(\mathbf{r}(s)-\mathbf{r}(t))ds\,dt}{|\mathbf{r}(s)-\mathbf{r}(t)|^3}, \quad (6)$$

where $\mathbf{r}'(s) = d^2\mathbf{r}/ds^2$ and $\mathbf{r}'(t) = d^2\mathbf{r}/dt^2$ represents the curvature of the curve $\mathbf{r}(s)$ and $\mathbf{r}(t)$ along the curves C_1 and C_2, respectively, and s, t are two independent contour distances along C_1 and C_2.

We can identify the trivial circle by computing the value of the Alexander polynomial at two points as $I(-1)$ and $I(-2)$. For the trivial circle, $I(-1) = 1$ and $I(-2) = 1$. For both types of trefoil, $I(-1) = 3$, $I(-2) = 7$, and so on. We calculate the linking number of the reference structures (as in *last column* in Fig. 2 *C*) to distinguish mirror images, such as structure (*a*) ($Lk = -1$) versus structure (*d*) ($Lk = 1$). We can thus combine tests of the Alexander polynomial and the linking number to determine the unique topological state of any candidate DNA conformation.

Modeling the *hix*-pairing complex

A stable *hix*-pairing intermediate was characterized in experiments after two Hin dimers bound to two *hix* sites have properly juxtaposed (see Fig. 1 *B*). Part of our studies of the inversion process requires us to filter DNA/protein systems that have properly paired *hix* sites out of all possible conformations. Though the *hix*-pairing structure stabilized mainly by hydrogen bonds, our electrostatic model (Zhang et al., 2003) is not sufficiently detailed to describe this level of interaction as well as other extremely short-range (≤10 Å) electrostatic interactions. Thus, rigorously speaking, the *hix*-paired structure cannot be generated exactly in the dynamic simulations. To address this limitation, we introduce a computational potential to conserve the correct *hix*-pairing structure.

Johnson and co-workers have developed a molecular model of the *hix*-paired structure as well as its further pairing with the Fis-bound enhancer sequence based on chemical protection data (Merickel et al., 1998); their model is shown in Fig. 3. We see that two Hin dimers bend the two DNA strands by 18° and that these two DNA strands are perpendicular to each other. These two Hin/DNA complexes are properly aligned for stabilizing interactions and the centers of the two protein-bound sites are separated by ~50 Å, with the two Hin dimers located on opposite sides with respect to each other. We can conserve those features by the following potential:

$$E^{hh} = \frac{h^{hh}}{2}(r_{h_1h_2} - r^0_{h_1h_2})^2 + \frac{g^{hh}_a}{2}(\mathbf{a}^1_{ihin}\cdot\mathbf{a}^2_{ihin} + \mathbf{c}^1_{ihin}\cdot\mathbf{c}^2_{ihin})^2 + \frac{g^{hh}_b}{2}(\mathbf{b}^1_{ihin}\cdot\mathbf{b}^2_{ihin} + 1)^2, \quad (7)$$

$$r_{h_1h_2} = |\mathbf{r}^1_{ihin} - \mathbf{r}^2_{ihin}|, \quad (8)$$

in which $r_{h_1h_2}$ is the distance between the centers of two Hin/DNA complexes, and $\{\mathbf{a}^1_{ihin}, \mathbf{b}^1_{ihin}, \mathbf{c}^1_{ihin}\}$ and $\{\mathbf{a}^2_{ihin}\cdot\mathbf{b}^2_{ihin}, \mathbf{c}^2_{ihin}\}$ are local coordinates systems for two Hin/DNA complexes. The equilibrium value for $r_{h_1h_2}$ is set as $r^0_{h_1h_2} = 5.0$ nm, and the values for the constraining weights h^{hh}, g^{hh}_a, g^{hh}_b are set as $h^{hh} = h$, g^{hh}_a, $g^{hh}_b = 5\,g$, where h and g are the stretching and bending rigidity constants we used to model DNA with average sequences (Huang et al., 2001; Huang and Schlick, 2002). These computational values were found to maintain stable *hix*-paired structures while DNA stretching motion are the fastest motion (defined by h); this allows us to use the same time step for the simulations as described previously ($\Delta t = 600$ ps). We have tested those parameters on several simulations. We found the *hix*-pairing potential, together with our excluded volume potential (Huang and Schlick, 2002), to work well in terms of conserving the major features of the *hix*-pairing local structure. However, we noted an increasing number of segment passing attempts (10^{-2}–10^{-4} per simulation step, $\Delta t = 600$ ps) in the simulations of the *hix*-paired structure with Fis dimers included; this is because the Fis dimers are only 120 bp away from the *hix*-pairing structure, and thus multiple DNA strands as well as four DNA/protein regions make the local environment sterically crowded. To prevent such segment passing events and conserve the DNA topology, we use the writhing number and the Alexander polynomial (Jian et al., 1998; Huang et al., 2001).

RESULTS AND DISCUSSION

The generally accepted pathway of the Hin inversion—a topologically sensitive site-specific recombination reaction—involves two major steps of DNA global conformational changes (Fig. 1 *B*) (Berg and Howe, 1989; Lilley, 1995; Heichman and Johnson, 1990): a Hin dimer binds to each of the *hix* sites and the two sites juxtapose to form a stable paired *hix* structure; the paired *hix* structure interacts with the Fis-bound enhancer to form a nucleoprotein complex called ''invertasome'' (see Fig. 3). Finally, the two DNA strands exchange to produce the final inversion of the sequence (see Fig. 2 *C* (*a*)).

Intriguingly, experiments show clear differences in the dependence of the formation of paired *hix* sites and of the formation of the invertasome on the superhelical density σ of the DNA substrate (Fig. 4) (Lim and Simon, 1992). Namely, *hix*-pairing (*square data points* in Fig. 4) requires a relatively low degree of negative supercoiling, whereas formation of the invertasome (*circular data points* in Fig. 4) as well as the final inversion reaction, require a much higher degree of DNA superhelicity.

Kinetic questions regarding the role of DNA supercoiling are the focus of our modeling work. We investigate in turn two specific dynamic stages in the Hin inversion reaction: site juxtaposition of two *hix* sites, and three-site juxtaposition with Fis after formation of *hix*-pairing complex. Both steps are affected by large-scale motions of supercoiled DNA, and an understanding of some aspects of these complex dynamic processes can add insights that may be difficult to obtain by regular experimental approaches.

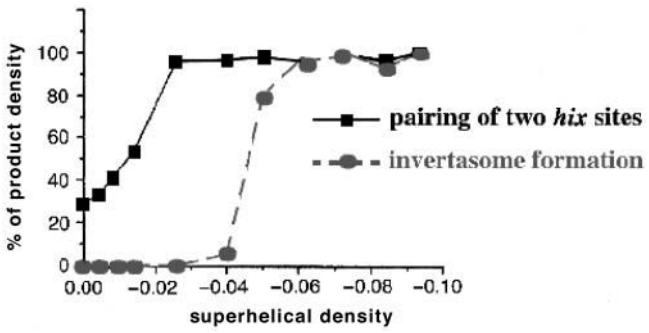

FIGURE 4 Dependence of two different stages in the Hin-mediated inversion on the average superhelical density σ (Lim and Simon, 1992). Reaction products are quantitatively measured using a densitometer by a published method to calculate the number of inversions (Bruist and Simon, 1984). Stage-specific band formation is represented as a percentage of that formed at physiological density.

Site juxtaposition of two *hix* sites

Site juxtaposition times with and without topological criteria

To simulate the juxtaposition process preceding the formation of the *hix-hix* pairing, we model Hin-dimers on both *hix* sites of a randomly selected, equilibrated DNA substrate and let it relax for a short time (10,000 steps of 600 ps each, or 6 μs). Fis dimers are not added in this stage because the *hix*-pairing experiments in vitro (Lim and Simon, 1992) (*square data points* in Fig. 4) were conducted with Hin but not Fis in the buffer to identify the formation of the paired *hix* complex. We thus start simulations and monitor each trajectory until two *hix* sites are properly juxtaposed (i.e., within 10 nm of each other) with exactly two negative supercoils trapped; we call this first juxtaposition time τ_{c1}. Technical details of how to identify such a conformation are discussed under Materials and Methods (''Topological description of a juxtaposed structure''). For comparison, we also record the first juxtaposition times τ'_{c1} based on only the distance criteria ($|\mathbf{r}^1_{hin} - \mathbf{r}^2_{hin}| < d_0$) (i.e., no filtering geometric criteria) to study the effect of the topological selectivity on the juxtaposition kinetics. For each superhelical density value σ, we perform 20 independent simulations of length 10 ms each to sample the site juxtaposition times as well as other dynamic properties.

Fig. 5 *A* clearly shows the effects of the topological selectivity as a function of DNA superhelicity. For superhelical densities from -0.02 to -0.06, site juxtaposition times corresponding to structures with two negative trapped supercoils are nearly five times longer than site juxtaposition times without this topological criterion. If the DNA is nearly relaxed (superhelical density is close to $\sigma \sim 0$), it becomes much more difficult to produce a juxtaposed conformation with the proper topology; this is shown by the much larger $\langle \tau_{c1} \rangle$ value (by two orders of magnitude).

We can explain this dependence based on conformational properties of the supercoiled DNA. When a DNA molecule is negatively supercoiled, its more tightly interwound conformation facilitates significantly the trapping of two negative supercoils in the juxtaposed conformation. However, the exact rate of site juxtaposition with proper topology depends on many other factors. Our studies of free supercoiled DNA dynamics (Huang et al., 2001) reveal that DNA superhelicity strongly affects the autocorrelation patterns in site juxtaposition events. For DNA with $\sigma = 0$, site juxtaposition events are less correlated due to random collisions (Huang et al., 2001; Huang and Schlick, 2002). With the increase of $|\sigma|$, slithering becomes the dominant juxtaposition mechanism and makes site juxtaposition events strongly correlated (Huang et al., 2001; Huang and Schlick, 2002). These two complex opposing and competing mechanisms explain why equilibrium juxtaposition probabilities may be 100-fold different whereas site juxtaposition times may be similar. (Slithering scales much slower with DNA length than random collisions (Huang et al., 2001; Huang and Schlick, 2002).) In our simulations of DNA bound to proteins, we also expect the site juxtaposition mechanism to be strongly affected by slithering and branching motions in a supercoiled DNA, while random, three-dimensional movements dominate in relaxed DNA. The interplay between these two mechanisms—random, large-scale DNA rearrangements and local slithering in branched DNA structures—produces the plateau in Fig. 5 *A*, with an optimal range of σ between -0.02 and -0.06.

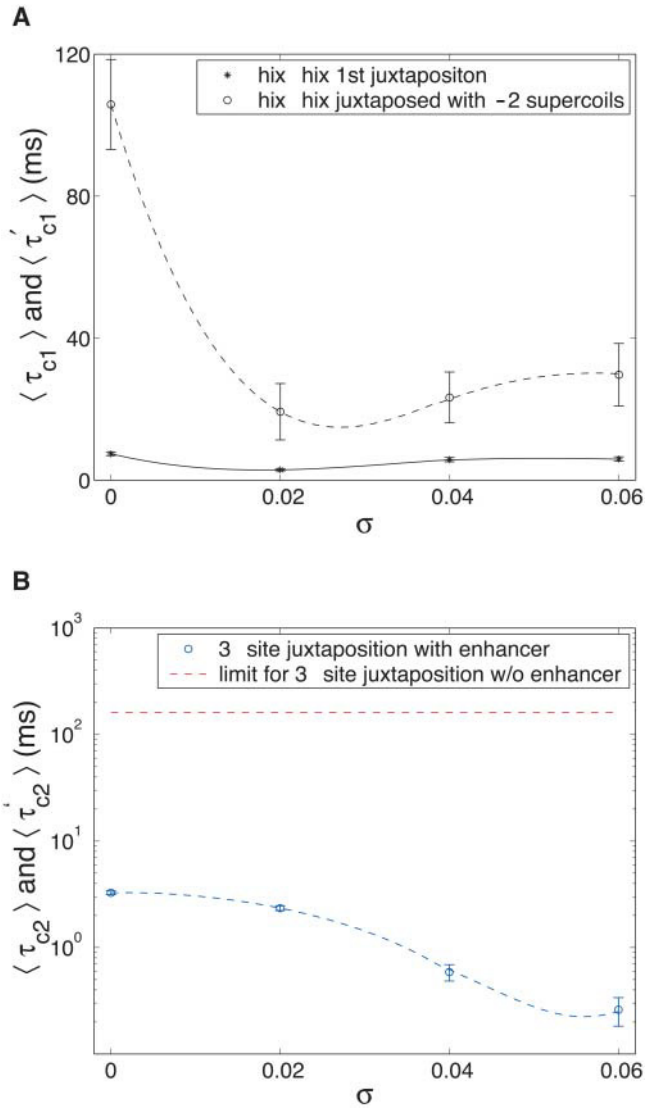

FIGURE 5 Computation of juxtaposition times. (*A*) Dependence for site juxtaposition times of two *hix* sites on $|\sigma|$. Circles show the juxtaposition times for two *hix* sites with less than d_0 separation while having two negative supercoils trapped. Asterisks show average site juxtaposition times without the topological criteria. The values for $N, i^1_{hix}, i^2_{hix}, i^1_{fis}, i^2_{fis}$, and c_S are 176, 1, 34, 5, 6, and 0.2 M, respectively. The simulation systems include two bound Hin dimers but no Fis dimers. (*B*) Effect of the enhancer: three-site juxtaposition times after the *hix*-pairing with and without Fis-dimer binding. The values $N, i^1_{hix}, i^2_{hix}, i^1_{fis}, i^2_{fis}$, and c_S are 176, 1, 34, 5, 6, and 0.2 M, respectively. Circles refer to values of DNA systems properly bound to the Hin dimers and Fis dimers. The dashed straight line refers to the lower limit of the reaction times estimated by simulations without the Fis binding.

Dynamics

We further investigate the dynamic details of simulations for the *hix*-pairing processes to dissect the roles of the DNA superhelicity, the salt environment, and the topological specificity on the reaction process.

Fig. 6 shows Brownian dynamic snapshots from a typical trajectory simulating the dynamics of a DNA substrate with *hix* sites bound to two Hin dimers. Since the Hin dimers bend DNA only slightly, both *hix* sites do not have a strong tendency to occupy superhelical ends or any specific regions of supercoiled conformations. We clearly observe branch formation and deletion events within the millisecond timescale, similar to the dynamics of free, supercoiled DNA. Branching motions are especially important in the *hix*-pairing dynamics because a branch point is a natural position for sites to juxtapose with two trapped negative supercoils, as shown in the conformation at 29.2 ms of this trajectory. Since positions of branch points move randomly on the 2D-like branched supercoiled structure, such juxtapositions with two negative supercoils occur on the timescale of a 2D-like random walk. We also observe juxtaposed conformations without the proper topology; these result from either random collisions between two superhelical branches (usually more than two negative supercoils are trapped) or from juxtapositions on two strands of the same superhelical branch (zero supercoil trapped). For example, the conformation at 18.8 ms traps 10 negative supercoils instead of two.

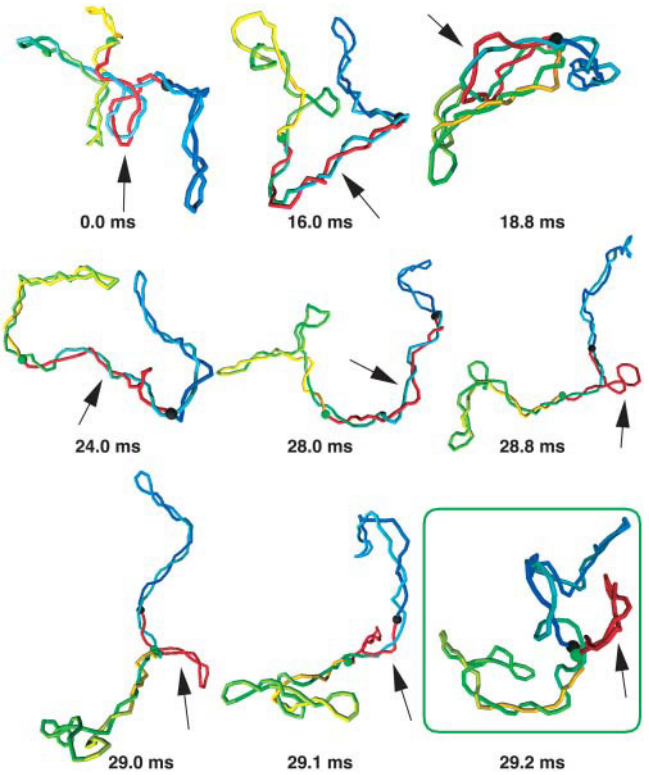

FIGURE 6 BD snapshots from a trajectory for the *hix*-pairing process in Hin-mediated inversion. The two *hix* sites ($i^1_{\text{hin}} = 10, i^2_{\text{hin}} = 43$) are gray and black spheres. Arrows show the sequence between two *hix* sites. Hin-dimers/DNA complexes are incorporated in both *hix* sites ($\sigma = -0.06$).

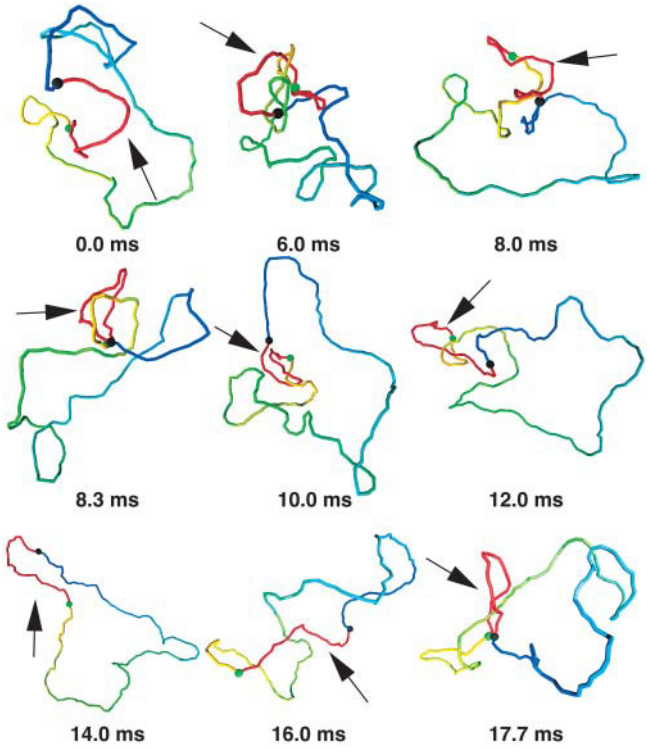

FIGURE 7 BD snapshots from a trajectory for the *hix*-pairing process in Hin-mediated inversion. The two *hix* sites ($i^1_{\text{hin}} = 10, i^2_{\text{hin}} = 43$) are shown colored gray and black. Arrows show the sequence between two *hix* sites. Hin-dimers/DNA complexes are incorporated in both *hix* sites ($\sigma = 0$).

Fig. 7 shows a typical dynamics trajectory of the same DNA substrate for relaxed DNA ($\sigma = 0$). Comparing with Fig. 6, we note that conformations are far less correlated. The conformation at 8.3 ms shows juxtaposition with no supercoils trapped and the conformation at 17.7 ms reveals a correct juxtaposed structure. The correct juxtaposed structure is thus achieved by a three-dimensional random walk process.

We next track the time evolution of the DNA sites near the two *hix* sites to investigate detailed dynamic aspects of the juxtaposition mechanism. Fig. 8 shows typical results from simulations of DNA at both $\sigma = 0$ and $\sigma = -0.06$. A circle at a particular time indicates that one of the Hin-bound DNA sites (*open circles* for i^1_{hin} and *filled circles* for i^2_{hin}) comes close to another DNA site, whose location (i.e., the index number of the bead) is plotted as the y axis. In the bottom of the figure, two reference miniplots show typical patterns for the two juxtaposition mechanisms of random collisions and slithering (data from our studies of supercoiled DNA dynamics (Huang et al., 2001)). We notice that random collision patterns dominate the plot of $\sigma = 0$ (*top*), whereas continuous slithering motions appear frequently as broad

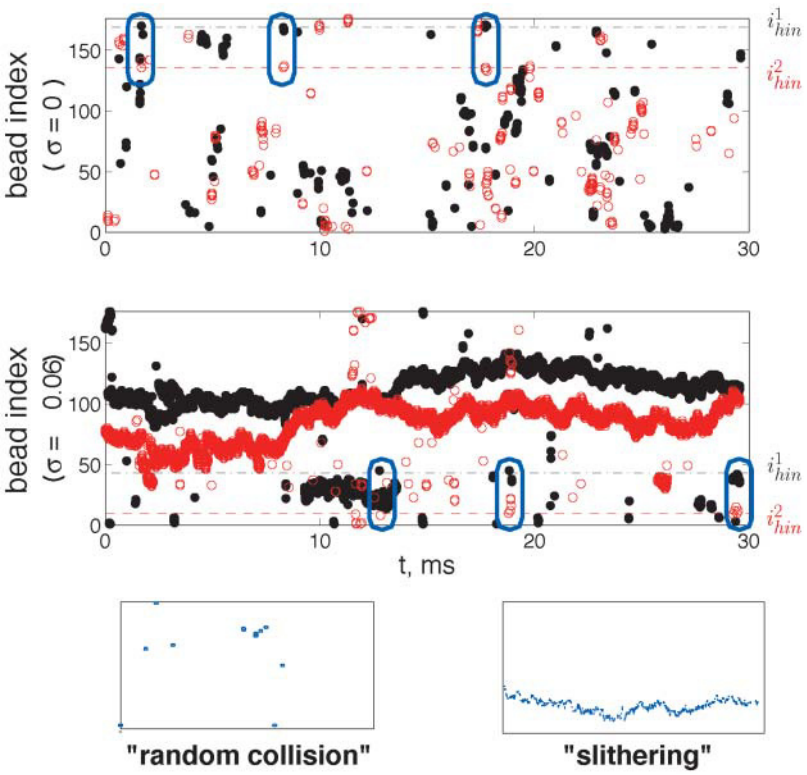

FIGURE 8 The DNA segments near two *hix* sites as a function of time. We plot a site if the distance between that site and a *hix* site is <10 nm. In the upper panel, sites near $i^1_{\text{hin}} = 10$ in the simulation of a DNA with $\sigma = 0$ are shown in black and those near $i^2_{\text{hin}} = 43$ are colored gray. In the bottom panel, sites near $i^1_{\text{hin}} = 136$ with $\sigma = -0.06$ are in black and those near $i^2_{\text{hin}} = 169$ are in gray. Juxtaposition events of two *hix* sites are highlighted by ovals. The two miniplots at the bottom illustrate the patterns expected for two dynamic mechanisms, random collision and slithering (Huang et al., 2001).

bands for physiological superhelical densities (*middle plot*). There we also notice branch formation and the slithering through branch points—as short bands and clusters—and random collisions between two superhelical branches as isolated circles.

Besides analyzing the DNA dynamic mechanism, such figures also provide direct information about juxtaposition events between specific sites. In the plots of Fig. 8, an intersection between lines and circles indicates a juxtaposition event between two *hix* sites (i^1_{hin} and i^2_{hin}), highlighted by ovals. For example, in the middle plot, we can find two simultaneous intersections with clusters (i.e., new branch formations) at 29.2 ms, indicating a juxtaposition event near the branch point (as shown in Fig. 6). Examination of our trajectories reveals consistency with our previous analysis: different dynamic mechanisms operate for DNA substrates with different amount of DNA superhelicity; in the range of from $\sigma = -0.02$ to -0.06, juxtaposition near the branch points results from global rearrangements of DNA branches coupled to local slithering along the branches; this is the dominant mechanism for properly juxtaposing two Hin-bound *hix* sites.

We further examine the square data points of Fig. 4 describing the experimental pairing of two *hix* sites as a function of σ. To form a stable *hix*-pairing structure requires two kinetic steps: juxtaposition of two *hix* sites with two negative supercoils trapped, as we simulated, and the short-range reaction between two Hin/DNA complexes to form a stable intermediate structure. The second part involves formation of hydrogen bonding and requires very accurate electrostatic description, beyond the capability of our model. Therefore, we cannot quantitatively compare the experimentally measured rate (in Fig. 4) with our simulated juxtaposition times with the proper topology (*open circles* in Fig. 5 *A*). (Note that in the experimental results of Fig. 4, the product density as the percentage of the physiological density is measured as a function of time in the experiments. These data are used to derive the kinetic reaction rates instead of equilibrium properties (Lim and Simon, 1992).) We do, however, observe a similar plateau in both curves appearing around $\sigma = -0.02$: both reactions are substantially slowed down for relaxed DNA and rates do not dramatically change for supercoiled DNA in the wide range from $\sigma = -0.02$ to -0.06.

Three-site juxtaposition after forming *hix*-pairing complex

After forming the *hix*-pairing complex with two trapped negative supercoils, the DNA conformation continues to rearrange until the enhancer sequence can directly interact with the paired *hix* sites and form the invertasome. In our millisecond simulations, we can ignore the possibility of dissociation of the *hix*-pairing structures, because experiments (Heichman and Johnson, 1990) suggest that the *hix*-paired structure is stable for hours at room temperature.

Fig. 3 shows the molecular model of invertasome complex built by R. C. Johnson and co-workers (Merickel et al.,

1998). A distinctive structural feature of the enhancer sequence is that the two local bends caused by the binding of two Fis dimers form a phased structure with double kinks. To investigate the dynamics of the three-site juxtaposition (i.e., two *hix* sites and one enhancer site containing two proximal Fix-binding sequences) as well as effects of the Fis-bound enhancer sequence, we now analyze simulations of the three-site juxtapositions with and without Fis under various conditions.

Site juxtaposition times

We first select DNA conformations from our previous simulations with two Hin-dimer bound *hix* sites paired and two trapped negative supercoils. We conserve the *hix*-pairing structure with the proper geometry as observed in the molecular model using methodology described separately (see ''Modeling the *hix*-pairing complex'' under Materials and Methods). Our substrate DNA has both Hin-dimer and Fis-dimer bound, and simulations are continued until the three sites juxtapose. This three-site juxtaposition is defined by satisfying two distance criteria:

$$|\mathbf{r}_{i_{hin}} - \mathbf{r}^1_{ifis}| < d_0 \quad \text{and} \quad |\mathbf{r}_{i_{hin}} - \mathbf{r}^2_{ifis}| < d_0, \tag{9}$$

in which $\mathbf{r}^1_{ifis}$ and $\mathbf{r}^2_{ifis}$ are coordinates for the enhancer sequence, and $\mathbf{r}_{i_{hin}} = 0.5(\mathbf{r}^1_{i_{hin}} + \mathbf{r}^2_{i_{hin}})$ is the midpoint of two paired *hix* sites. For each different superhelical density σ, multiple trajectories are performed from various starting conformations.

Fig. 5 *B* shows the three-site juxtaposition times τ_{c2} (*open circles*). We see that DNA superhelicity affects strongly this process: the decrease of $|\sigma|$ significantly slows down the process by an order of magnitude. Note that the distance between the enhancer sequence and one of the *hix* site is only 120 bp in our model (110 bp in experiments), shorter than DNA's persistence length, and this juxtaposition is expected to be slow because a higher free-energy is required to form a small loop. Interestingly, a higher magnitude of DNA superhelicity ($|\sigma| > 0.04$) is required here compared with the previous juxtaposition process between two sites (Fig. 5).

Effects of the Fis enhancer

According to experimental observations, the enhancer sequence with Fis bound is required for the efficient Hin inversion both in vitro (Johnson et al., 1986) and in vivo (Johnson et al., 1988) and can accelerate the reaction by more than 150-fold (Johnson et al., 1986).

Fig. 5 *B* shows this enhancement by Fis on the three-site juxtaposition rates (*open circles*) compared with measurements with Hin-bound sites only (the *dashed line* for the lower limit estimate). This behavior can be explained by the severe local bending (Huang and Schlick, 2002) induced by the two Fis dimers and its altering of the local electrostatic and hydrodynamic properties. How can such local changes affect DNA global dynamics and juxtaposition kinetics? All our five simulations of the system without Fis dimers (and two paired *hix* sites) for each $|\sigma|$ value failed to produce three-site juxtapositions within 40 ms (3 months CPU time per trajectory) (Huang, 2003). This time thus provides only a lower bound for the actual three-site juxtaposition times (dashed, straight line in Fig. 5 *B*).

Why does Fis play such a crucial role in the three-site juxtaposition and hence in invertasome formation? Let us first examine the local distortion of the Fis dimers. The large bend $\sim 60^\circ$ toward the minor groove in each Fis-dimer bound site certainly lowers the free-energy barrier to form a 120 bp loop required in the three-site juxtaposition. From the molecular model of the three-site juxtaposed invertasome (Fig. 3), we find many spatial restrictions as a consequence of the requisite topology. Since strand crossing is not allowed during juxtaposition, the two negative supercoils trapped in the *hix*-pairing events must be conserved. The strong bending caused by Fis can substantially alleviate the steric conflicts by allowing the two Fis dimer to bind the enhancer sequence properly. The Fis-dimer binding also induces significant changes of the local electrostatic fields, as we know from our DiSCO modeling (see Fig. 1 *A*). The negative charges of the DNA polyelectrolyte are further screened near physiological monovalent salt environments ($c_S = 0.2$ M). Since our studies of two-site juxtaposition indicate that the protein binding effect on electrostatic interactions is only significant after sites become very close, that is in the range of 7.5–15 nm (Huang and Schlick, 2002), we expect a similarly subtle effect for three-site juxtaposition kinetics.

Dynamics

We now inspect many trajectories simulating the three-site juxtaposition for various superhelical densities. Fig. 9 *A* shows a series of snapshots from a trajectory with $\sigma = -0.06$. In the first conformation (same as final in Fig. 6), *hix*-pairing with two trapped negative supercoils occurs at a branch point where the sequence between the two Hin-bound sites forms a superhelical branch. As discussed for the *hix*-pairing study, juxtapositions at branch points are favored since they naturally trap two negative supercoils. After forming the *hix*-pairing intermediate, we find that that the branch point is conserved in the dynamics, such as shown in Fig. 9 *A*. The three-site juxtaposition, therefore, happens as the short branch randomly changes until a proper juxtaposition forms.

For comparison, we show in Fig. 9 *B* several snapshots from a trajectory of the three-site juxtaposition with $\sigma = 0$. The DNA conformations are irregular, but the two trapped negative supercoils produce a short branch for the sequence between *hix* sites. Similarly, the final three-site juxtaposition is achieved mainly through random conformational changes.

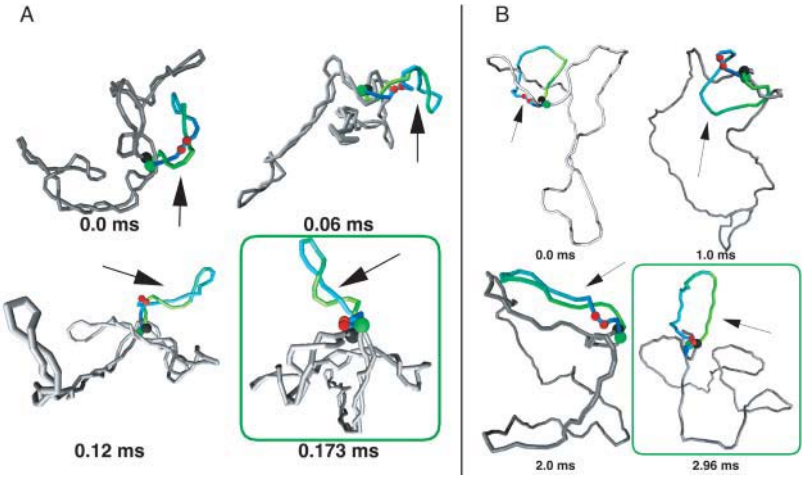

FIGURE 9 BD snapshots of three-site juxtaposition after pairing of the two *hix* sites in Hin-mediated inversion process. (*A*) The two *hix* sites ($i^1_{\rm hin} = 10, i^2_{\rm hin} = 43$) are shown colored gray and black. The sequence between two *hix* sites is indicated by arrows. Two Hin dimers bind to both *hix* sites and form a stable *hix*-pairing structure. Two Fis dimers bind on the enhancer sequence. The values $N, i^1_{\rm hix}, i^2_{\rm hix}, i^1_{\rm fis1}, i^2_{\rm fis2}, \sigma$, and $c_{\rm S}$ are set to 176, 1, 34, 5, 6, -0.06, and 0.2 M respectively. (*B*) Parameters are the same as in *A* except that $\sigma = 0$.

The movements, however, tend to be more open and flexible in more relaxed DNA conformations (lower $|\sigma|$).

Fig. 10 shows a trajectory from the same starting conformation as in Fig. 9 *A*, except that the enhancer sequence is not bound to two Fis dimers. The enhancer sequence is usually more straight, behaving as all other protein-free DNA segments in the model. The branched structure is conserved, but forming the three-site juxtaposition conformation is much more difficult: no successful juxtaposition over 40 ms, in contrast to a timescale of 1 ms where Fis is present. By analyzing the successful three-site juxtaposition structures (such as the two in Fig. 9, *A* and *B*), we find that a proper three-site juxtaposition always involves severe bending of the DNA to allow all three DNA strands as well as four DNA/protein complexes to fit into a small space. The DNA bending introduced by the Fis dimers reduces the free energy required for the three-site juxtaposition and makes the enhancer site less negatively charged, which also favors the crowded final structure. The sum of those effects explains why the binding of two Fis dimers on the enhancer sequence is so critical to the three-site juxtaposition reaction.

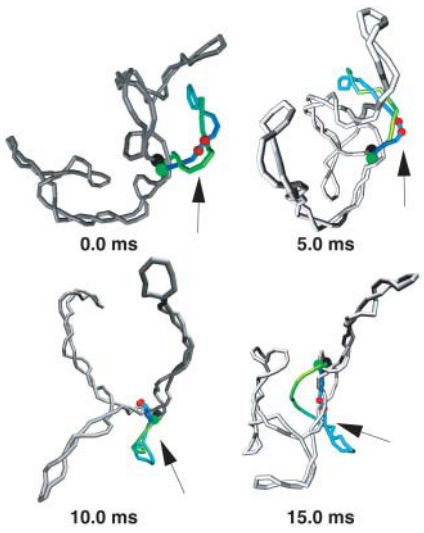

FIGURE 10 BD snapshots of the three-site juxtaposition after pairing two *hix* sites without the Fis binding. All parameters are the same as in Fig. 9 except that there is no Fis dimer binding on sites of $i_{\rm fis1}$ and $i_{\rm fis2}$.

Feasibility of other pathways

Three-site juxtaposition times are defined as the average time required for three sites, two *hix* sites and the enhancer sequence, to direct juxtapose without forming the *hix*-pairing structure and thus be sufficiently close for proteins to interact and form the invertasome complex. If the average three-site juxtaposition time is compatible to the sum of juxtaposition times ($\tau_{\rm c1} + \tau_{\rm c2}$) in the proposed pathway (Fig. 1 *B*), the direct pathway without forming the *hix*-pairing intermediate must be considered as an important competitor to the proposed mechanism of the two-step juxtaposition pathway.

In the previous sections, we only report the two-site juxtaposition time for *hix* sites $\tau_{\rm c1}$ when the enhancer Fis does not present, which corresponds to the case in vitro (Lim and Simon, 1992). However, in vivo, both Hin and Fis exist. We are not able to achieve acceptable statistics on the simulations of two-site juxtaposition for *hix* sites with both Hin-dimer and Fis-dimer properly bound mainly because of the computational time limit, but based on our preliminary data, we can estimate that the two-site juxtaposition times of *hix* sites, with both Hin and Fis or with only the Hin, are in the same timescale.

We have simulated many trajectories of DNA with both *hix* sites and the enhancer sequence bound to proper protein dimers to investigate the rates and probabilities of the direct juxtaposition of three sites. Those trajectories are of lengths ranging from 20 ms to 50 ms and required many months of computing times. For most of those trajectories, we did not observe a single three-site juxtaposition; statistics are thus too poor to provide quantitative measurements on direct three-site juxtaposition rates. Still, the rarity of successful three-site juxtaposed structures suggests the much slower

time for this single-step juxtaposition mechanism as opposed to the two-step juxtaposition pathway. The two-step mechanism is thus expected to be the dominant mechanism in real systems.

CONCLUSION

We have studied the global dynamics in the Hin-mediated inversion process by using the computational model and simulation protocols for supercoiled DNA molecules with bound proteins as previously developed. The resulting mesoscale model integrates atomic detail where essential with a macroscopic representation at the polymer level to permit long time trajectories of a large system (thousands of DNA basepairs). Our previous study (Huang and Schlick, 2002) validated the model by reproducing translational diffusion coefficients and radius of gyrations as measured by light-scattering experiments and computed by Monte Carlo simulations. Here we also developed required numerical protocols to distinguish among various juxtaposition conformations based on topology and to judge whether an intermediate structure is acceptable in the real biological pathway; these allowed us to quantitatively identify conformations ''with exactly two negative supercoils trapped'' as described in the literature.

We have performed Brownian dynamics simulations to mimic various dynamic steps involved in the proposed multistep inversion pathway (Fig. 1). By systematically comparing the simulation results under different conditions, we could delineate the effect of topological selectivity on the dependence of the reaction rates on DNA's superhelicity as observed experimentally, and study the effect of Fis binding on reaction rates, to help understand the importance of the Fis factor. We found that DNA superhelicity plays important regulatory roles in the different stages of the multistep pathway: the juxtaposition of two *hix* sites is markedly enhanced when $|\sigma| > 0.02$ whereas the subsequent juxtaposition between the *hix*-pair and the Fis-bound enhancer sequence requires a higher amount of DNA superhelicity ($|\sigma| > 0.04$), as observed experimentally (Lim and Simon, 1992). Our computed juxtaposition rates show how much more accelerated the two-step inversion pathway is compared to the alternative single-step pathway, and explain how local electrostatics and bend deformations caused by Fis work together with global supercoiling geometric features to accelerate juxtaposition in Hin- and Fis-bound DNA.

The methods developed here to model, simulate, and analyze large DNA/protein systems are general and potentially useful for applications involving large DNA molecules (thousands of basepairs) with bound proteins in solution, such as resolvase DNA and integrase DNA systems (Lilley, 1995; Glasgow et al., 1989). We hope to apply a similar approach to study interesting mechanisms of Tn5 transposase and related mobile DNAs.

We thank Dr. R. C. Johnson for providing the atomic structures of Hin-dimers/DNA and Fis-dimers/DNA complexes. We also thank the reviewer and handling editor for excellent critiques.

Support by the National Institutes of Health (GM 55164 to T.S.) is gratefully acknowledged.

T. Schlick is an investigator for the Howard Hughes Medical Institute.

REFERENCES

Allison, S. A. 1986. Brownian dynamics simulation of wormlike chains. Fluorescence depolarization and depolarized light scattering. *Macromolecules.* 19:118–124.

Allison, S. A., R. Austin, and M. Hogan. 1989. Bending and twisting dynamics of short linear DNAs. Analysis of the triplet anisotropy decay of a 209 base pair fragment by Brownian simulation. *J. Chem. Phys.* 90:3843–3854.

Bates, A. D., and A. Maxwell, editors. 1993. DNA Topology. Oxford University Press, New York.

Beard, D., and T. Schlick. 2001a. Modeling salt-mediated electrostatics of macromolecules: the algorithm DiSCO (Discrete Surface Charge Optimization) and its application to the nucleosome. *Biopolymers.* 58:106–115.

Beard, D. A., and T. Schlick. 2001b. Computational modeling predicts the structure and dynamics of chromatin fiber. *Structure.* 9:105–114.

Benjamin, K. R., A. P. Abola, R. R. Kanaar, and N. R. Cozzarelli. 1996. Contributions of supercoiling to Tn3 resolvase and phage Mu Gin Site-specific recombination. *J. Mol. Biol.* 256:50–65.

Berg, D. E., and M. M. Howe, editors. 1989. Mobile DNA. American Society for Microbiology, Washington, DC.

Boles, T., J. White, and N. R. Cozzarelli. 1990. Structure of plectonemically supercoiled DNA. *J. Mol. Biol.* 213:931–951.

Bruist, M. F., and M. I. Simon. 1984. Phase variation and the Hin protein: in vivo activity measurements, protein overproduction, and purification. *J. Bacteriol.* 159:71–79.

Bustamante, C., and C. Rivetti. 1996. Visualizing protein-nucleic acid interactions on a large scale with the scanning force microscope. *Annu. Rev. Bioph. Biom.* 25:395–429.

Chirico, G., and J. Langowski. 1994. Kinetics of DNA supercoiling studied by Brownian dynamics simulation. *Biopolymers.* 34:415–433.

Chirico, G., and J. Langowski. 1996. Brownian dynamics simulations of supercoiled DNA with bent sequences. *Biophys. J.* 71:955–971.

Dahlquist, G., and A. Bjorck. 1974. Numerical Methods. Prentice-Hall, Englewood Cliffs, NJ.

Derreumaux, P., G. Zhang, B. Brooks, and T. Schlick. 1994. A truncated-Newton method adapted for CHARMM and biomolecular applications. *J. Comp. Chem.* 15:532–552.

Dickerson, R. E. 1989. Definitions and nomenclature of nucleic acid structure components. *J. Mol. Biol.* 208:1797–1803.

Dubochet, J., M. Adrian, I. Dustin, P. Furrer, and A. Stasiak. 1992. Cryoelectron microscopy of DNA molecules in solution. *Methods Enzymol.* 211:507–518.

Ermak, D. L., and J. A. McCammon. 1978. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* 69:1352–1360.

Feng, J. A., R. C. Johnson, and R. E. Dickerson. 1994. Hin recombinase bound to DNA: the origin of specificity in major and minor groove interactions. *Science.* 263:348–355.

Fixman, M. 1986. Construction of Langevin forces in the simulation of hydrodynamic interaction. *Macromolecules.* 19:1204–1207.

Frank-Kamenetskii, M. D., and A. V. Vologodskii. 1981. Topological aspects of the physics of polymers: the theory and its biophysical applications. *Sov. Phys. Usp.* 24:679–696.

Garcia de la Torre, J., and V. A. Bloomfield. 1977. Hydrodynamics of macromolecular complexes. I. Translation. *Biopolymers.* 16:1747–1763.

Glasgow, A. C., K. T. Hughes, and M. I. Simon. 1989. Bacterial DNA inversion systems. In Mobile DNA. D. Berg and M. Howe, editors. American Society for Microbiology, Washington, DC. 637–659.

Gralla, J. D. 1996. Activation and repression of E. coli promoters. *Curr. Opin. Genet. Dev.* 6:526–530.

Heichman, K. A., and R. C. Johnson. 1990. The Hin invertasome: protein-mediated joining of distant recombination. *Science*. 249:511–517.

Huang, J. 2003. Macroscopic modeling and dynamic simulations of supercoiled DNA and bound proteins. PhD thesis. New York University, New York.

Huang, J., and T. Schlick. 2002. Macroscopic modeling and simulation of supercoiled DNA with bound proteins. *J. Chem. Phys.* 117:8573–8586.

Huang, J., T. Schlick, and A. V. Vologodskii. 2001. Dynamics of site juxtaposition in supercoiled DNA. *Proc. Natl. Acad. Sci. USA*. 98:968–973.

Iniesta, A., and J. Garcia de la Torre. 1990. A 2nd-order algorithm for the simulation of the Brownian dynamics of macromolecular models. *J. Chem. Phys.* 92:2015–2018.

Jian, H. 1997. A combined wormlike-chain and bead model for dynamic simulations of long DNA. PhD thesis. New York University, New York.

Jian, H., T. Schlick, and A. V. Vologodskii. 1998. Internal motion of supercoiled DNA: Brownian dynamics simulations of site juxtaposition. *J. Mol. Biol.* 284:287–296.

Jian, H., A. V. Vologodskii, and T. Schlick. 1997. A combined wormlike-chain and bead model for dynamic simulations of long linear DNA. *J. Comp. Phys.* 136:168–179.

Johnson, R. C., C. A. Ball, D. Pfeffer, and M. I. Simon. 1988. Isolation of the gene encoding the Hin recombinational enhancer binding protein. *Proc. Natl. Acad. Sci. USA*. 85:3484–3488.

Johnson, R. C., M. F. Bruist, and M. I. Simon. 1986. Host protein requirements for in vitro site-specific DNA inversion. *Cell*. 46:531–539.

Johnson, R. C., and M. I. Simon. 1985. Hin-mediated site-specific recombination requires two 26 bp recombination sites and a 60 bp recombinational enhancer. *Cell*. 41:781–791.

Kubo, R. 1966. The fluctuation-dissipation theorem. *Rep. Prog. Phys.* 29: 255–284.

Lilley, D. M. J., editor. 1995. DNA-Protein: Structural Interactions. Oxford University Press, New York.

Lim, H. M., and M. I. Simon. 1992. The role of negative supercoiling in Hin-mediated site-specific recombination. *J. Biol. Chem.* 5:11176–11182.

Merickel, S. K., M. J. Haykinson, and R. C. Johnson. 1998. Communication between Hin recombinase and Fis regulatory subunits during coordinate activation of Hin-catalyzed site-specific DNA inversion. *Genes Dev.* 12:2803–2816.

Michels, J. P. J., and F. W. Wiegel. 1986. On the topology of a polymer ring. *Proc. R. Soc. London A.* 403:269–284.

Oram, M., J. F. Marko, and S. E. Halford. 1997. Communications between distant sites on supercoiled DNA from non-exponential kinetics for DNA synapsis by resolvase. *J. Mol. Biol.* 270:396–412.

Pan, C. Q., S. E. Finkel, S. E. Cramton, J. A. Feng, D. S. Sigman, and R. C. Johnson. 1996. Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. *J. Mol. Biol.* 264:675–695.

Perkins-Balding, D., D. P. Dias, and A. C. Glasgow. 1997. Location, degree, and direction of DNA bending associated with the Hin recombinational enhancer sequence and Fis-enhancer complex. *J. Bacteriol.* 179:4747–4753.

Rocchia, W., S. Sridharan, A. Nicholls, E. Alexov, A. Chiabrera, and B. Honig. 2002. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. *J. Comp. Chem.* 23:128–137.

Rotne, J., and S. Prager. 1969. Variational treatment of hydrodynamic interaction in polymers. *J. Chem. Phys.* 50:4831–4837.

Safo, M. K., W. Z. Yang, L. Corselli, S. E. Cramton, H. S. Yuan, and R. C. Johnson. 1997. The transactivation region of the Fis protein that controls site-specific DNA inversion contains extended mobile beta-hairpin arms. *EMBO J.* 16:6860–6873.

Schlick, T. 1995. Modeling superhelical DNA: recent analytical and dynamic approaches. *Curr. Opin. Struct. Biol.* 5:245–262.

Schlick, T. 2001. Time-trimming tricks for dynamic simulations: splitting force updates to reduce computational work. *Structure*. 9:R45–R53.

Schlick, T. 2002. Molecular Modeling: An Interdisciplinary Guide. Springer-Verlag, New York.

Schlick, T., D. Beard, J. Huang, D. Strahs, and X. Qian. 2000. Computational challenges in simulating large DNA over long times. *IEEE Comput. Sci. Eng.* 2:38–51 (Special Issue on Computational Chemistry).

Schlick, T., and A. Fogelson. 1992. TNPACK—a truncated Newton minimization package for large-scale problems. I. Algorithms and usage. *ACM Trans. Math. Softw.* 18:46–70.

Schlick, T., B. Li, and W. K. Olson. 1994. The influence of salt on the structure and energetics of supercoiled DNA. *Biophys. J.* 67:2146–2166.

Sessions, R. B., M. Oram, M. D. Szczelkun, and S. E. Halford. 1997. Random walk models for DNA synapsis by resolvase. *J. Mol. Biol.* 270:413–425.

Silverman, M., and M. I. Simon. 1980. Phase variation: genetic analysis of switching mutants. *Cell.* 19:845–854.

Stigter, D. 1977. Interactions of highly charged colloidal cylinders with applications to double-stranded. *Biopolymers.* 16:1435–1448.

Vinograd, J., J. Lebowitz, R. Radloff, R. Watson, and P. Laipis. 1965. The twisted circular form of polyoma viral DNA. *Proc. Natl. Acad. Sci. USA.* 53:1104–1111.

Vologodskii, A. V. 1992. Topology and Physics of Circular DNA. CRC Press, Boca Raton, FL.

Vologodskii, A. V., and N. R. Cozzarelli. 1994. Conformational and thermodynamic properties of supercoiled DNA. *Annu. Rev. Biophys. Biomol. Struct.* 23:609–643.

Watson, M. A., M. R. Boocock, and W. M. Stark. 1996. Rate and selectivity of synapsis of res recombination sites by Tn3 resolvase. *J. Mol. Biol.* 257:317–329.

Xie, D., and T. Schlick. 1999a. Efficient implementation of the truncated Newton method for large-scale chemistry applications. *SIAM J. Opt.* 10:132–154.

Xie, D., and T. Schlick. 1999b. Remark on the updated truncated Newton minimization package, Algorithm 702. *ACM Trans. Math. Softw.* 25: 108–122.

Yamakawa, H. 1970. Transport properties of polymer chains in dilute solution: hydrodynamic interactions. *J. Chem. Phys.* 53:436–443.

Zhang, Q., D. A. Beard, and T. Schlick. 2003. Constructing irregular surfaces to enclose macromolecular complexes for mesoscale modeling using the discrete surface charge optimization (DiSCO) algorithm. *J. Comp. Chem.* In press.

Zuccheri, G., A. Scipioni, V. Cavaliere, G. Gargiulo, P. De Santis, and B. Samori. 2001. Mapping the intrinsic curvature and flexibility along the DNA chain. *Proc. Natl. Acad. Sci. USA.* 98:3074–3079.