JCTC Journal of Chemical Theory and Computation

SIMUFLEX: Algorithms and Tools for Simulation of the Conformation and Dynamics of Flexible Molecules and Nanoparticles in Dilute Solution

José García de la Torre,*,[II] José G. Hernández Cifre,[II] Álvaro Ortega,[II] Ricardo Rodríguez Schmidt,[II] Miguel X. Fernandes,[†] Horacio E. Pérez Sánchez,[‡] and R. Pamies[§]

Departamento de Química Física, Facultad de Química Universidad de Murcia, 30071 Murcia, Spain, Centro de Química da Madeira, Universidade da Madeira, 9000-390 Funchal, Portugal, Forschungszentrum Karlsruhe GmbH, Institut für Nanotechnologie, D-76021 Karlsruhe, Germany, and Department of Physical Chemistry, University of Oslo, Oslo, Norway

Received May 25, 2009

Abstract: A computer programs suite, SIMUFLEX, has been constructed for the calculation of solution properties of flexible macromolecules modeled as bead-and-connector models of arbitrary topology. The suite consists mainly of two independent programs, BROWFLEX that generates the macromolecular trajectory by using the Brownian dynamics technique and ANAFLEX that analyzes that trajectory to get solution properties of the macromolecule. In this paper, we describe theoretical aspects about the macromolecular model and the Brownian dynamics algorithm used and describe some of the numerous properties that can be evaluated. In order to provide examples of the application of the methodology, we present simulations of dynamic properties of DNA with length ranging from 10 to 10^5 base pairs. SIMUFLEX is able to run simulations with more or less coarse-grained models, thus enabling such multiple-scale studies.

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

1. Introduction

Solution properties of macromolecules (hydrodynamic coefficients, intrinsic viscosity, radiation scattering-related quantities...) shed light about their shape and conformation[1] and therefore are a primary source of information to predict their solution behavior. For example, the determination of such properties are of fundamental relevance when treating with biological macromolecules (DNA, proteins...), since their physiological functions are closely related to the solution conformations that they can adopt.[2] A powerful tool that helps to predict and understand macromolecular structure and dynamics is computational modeling and simulation.[3]

There exist well developed procedures based on bead modeling[4−8] or alternative approaches[9−12] to predict hydrodynamic properties of rigid macromolecules and nanoparticles. However, most synthetic polymers and many biological macromolecules are flexible and do not present a defined shape. Therefore, the development of computational procedures to predict the solution behavior of flexible and semiflexible macromolecules is of great interest. The large size typical of macromolecules and nanoparticles, and the long times typical of their dynamics, precludes usually the use of atomic-level models, and the conformational variability of flexible entities adds further complexity. Thus, the prediction of solution properties requires simplified schemes, based on more or less coarsely grained models. The classical bead-and-spring model of polymer physics, in which the model elements represents large pieces (subchains) of the long polymer chain,[13,14] is a very coarse grained model. Nowadays, the coarse-grained modeling concept is being

* Corresponding author e-mail: jgt@um.es.
[II] Facultad de Química Universidad de Murcia.
[†] Universidade da Madeira.
[‡] Forschungszentrum Karlsruhe GmbH, Institut für Nanotechnologie.
[§] University of Oslo.

10.1021/ct900269n CCC: $40.75 © 2009 American Chemical Society
Published on Web 08/24/2009

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

applied with more detail, with models whose elements represent, for instance, the repeating units - amino acid or nucleotide residues - of biomacromolecules.[15] On the other hand, useful schemes to build coarse-grained models have been recently developed.[16,17] Because of the widespread utilization - over the past two decades - of atomistic molecular dynamics simulations, there are many commercial and public domain tools for that purpose. However, for multiscale, coarse-graining simulation, one misses a wide availability of similar tools. Thus, we have intended[18] to develop computational methodologies where flexible macromolecules are represented at such a coarse-grained level as bead-and-connector models and to predict their solution behavior by simulation techniques as Monte Carlo (MC) and Brownian dynamics (BD).[18]

When the simulation of flexible entities is restricted to the prediction of conformational, equilibrium properties, and some overall hydrodynamic coefficients, Monte Carlo methods are applicable. In order to provide a tool for the MC simulation of quite general flexible bead-and-connector models, we recently published the public-domain program MONTEHYDRO,[19] that implements an importance sampling Monte Carlo procedure for the generation of random conformations of flexible structures, which includes the calculation of overall hydrodynamic properties in the so-called rigid-body treatment,[20−22] obtained as conformational averages over the values calculated for instantaneous conformations considered as rigid structures.[5,23]

However, to study dynamic aspects of flexible macromolecules in solution, such as relaxation processes and nonequilibrium behavior, it is necessary to solve the equation of motion that governs the macromolecular dynamics. This can be done by using molecular dynamics (MD) or Brownian dynamics (BD).[24] Because of the above-mentioned drawbacks, MD is not adequate for long time and size scales. BD is a numerical technique to solve the stochastic equation of motion that arises from considering the solvent as a continuum, thus eliminating the solvent degrees of freedom which allows for reaching longer times in the simulated physical system. In other words, BD simulations describe the Brownian motion of a collective of frictional elements, beads in our model, which can interact with each other through different potentials.

An essential aspect in the BD simulation is the inclusion of the so-called hydrodynamic interaction (HI) effect, which determines the solvent-mediated influence of the motion of every element of the model on the others. Our group[25−29] was among others[30−34] who pioneered the use of BD simulations including hydrodynamic interaction (HI) effects to predict dynamic properties of macromolecules in solution. As it is known from polymer hydrodynamic theory,[14,35] and confirmed by BD simulations,[28,36,37] the rigorous inclusion of the HI effect (avoiding approximations, like that of conformational preaverage) is essential for the accurate prediction of hydrodynamic properties results comparable to experiments. Nevertheless, BD simulations without inclusion of HI sample correctly the configurational space, so that some authors have proposed that BD could be used as a smart Monte Carlo method.[38] This adds a further utility to BD methodologies, providing an efficient way to obtain also equilibrium conformational properties.

Along our previous works we have been developing a BD simulation scheme that enables for the calculation of solution properties of flexible macromolecules with arbitrary complexity. Our procedures take into account fluctuating (nonpreaveraged) hydrodynamic interaction as well as the possibility of including different types of intramolecular potentials to represent excluded volume conditions (solvent quality) and electrostatic interactions. That computational scheme is implemented in a suite of public domain (freely available from our Web server, see below), named SIMUFLEX, which is presented in this paper. The suite consists mainly of two programs BROWFLEX and ANAFLEX. The program BROWFLEX generates a Brownian trajectory of a flexible bead-and-connector model with arbitrary connectivity, and the program ANAFLEX analyzes that trajectory to obtain several steady and time-dependent macromolecular quantities. In this way, many conformational and hydrodynamic solution properties, from single-valued coefficients to more complex experiments as well as different time correlation functions, can be straightforwardly evaluated from the Brownian trajectory. Furthermore, a most interesting feature of the BD technique is that it allows the simulation of the behavior of an individual molecule,[39,40] which is of great importance due to the emergence of single-molecule experimental techniques.[41] At this respect, SIMUFLEX is an useful tool to study single-molecule behavior of flexible macromolecules with arbitrary topology. Thus, our contribution joins those of other groups who have published Brownian dynamics simulation programs with different scope or structure[42,43] and focused on particular macromolecular systems (for instance the UHBD package[42] is appropriate for studying protein−protein association). On the other hand, the SIMUFLEX package was devised to treat with a variety of macromolecular models and physical situations, for example the presence of external agents, as well as to analyze easily an amount of macromolecular properties including a number of commonly employed correlation functions.

In this paper, we first describe some theoretical aspects of the modeling and BD simulation methodology implemented in SIMUFLEX. Then, we present several examples, all concerning the dynamics of DNA molecules in solution. In order to illustrate the multiscale possibilities of SIMUFLEX, the examples span a wide range of DNA sizes and cover both bulk-solution and single-molecule properties.

2. Models and Simulation Methods

In this section we specify the two main features in the simulated model. The first one corresponds to the mechanical or energetic features pertaining to the molecule itself and, eventually, its interaction with external agents (e.g., fields, walls, etc.). The second group of aspects comprise those relative to the motion of the molecular model in the viscous solvent, like viscous drag, hydrodynamic interaction, Brownian motion, etc., which are key factors for the construction of the simulation algorithm.

2.1. Mechanical Model: The Force Field. The simulation model is composed by what we generically call elements,

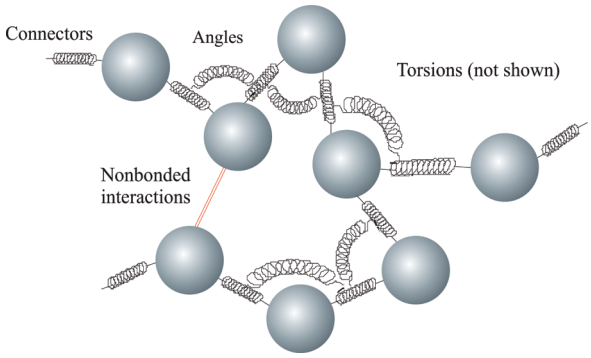

Figure 1. Pictorial representation of the generalized bead-and-spring model.

which will be later considered as spheres, or "beads", for the description of their hydrodynamic behavior. The N elements interact in a number of ways, which give rise to a potential energy that we decompose in several contributions

$$V^{(tot)} = \sum_{conn} V_{ij}^{(conn)} + \sum_{ang} V_{ijk}^{(ang)} + \sum_{tors} V_{ijkl}^{(tors)} + \sum_{EVpairs} V_{ij}^{(EVpair)} + \sum_{CHpairs} V_{ij}^{(CHpair)} + \sum_{i} V_{i}^{(E)} \quad (1)$$

The terms in eq 1 correspond to each kind of interaction, as described in the following paragraphs. A schematic overview of the model is displayed in Figure 1.

Primarily, the elements are joined by connectors, which describe the topology of the molecule or particle that is being modeled. A common case is that of linear chains, in which each element (except the terminal ones) is joined to its two neighbors. In general, an element may be joined to an arbitrary number of other elements. The sum extends over all the pairs of connected elements. Connectors behave mechanically as springs with an associated potential $V_{ij}^{(conn)}(l_{ij})$ that depends on the instantaneous distance between the two joined elements, equal to the length of the spring vector, $l_{ij} = |\mathbf{r}_{ij}| = |\mathbf{r}_j - \mathbf{r}_i|$, where $\mathbf{r}_i$ is the position vector of the i-th bead. Among others, BROWFLEX considers the following spring potential devised in a previous work to simulate dendrimer molecules[44]

$$V^{(conn)} = -\frac{1}{2}Hl_{\max}^2 \ln\left(\frac{l_{\max}^2 - l^2}{l_{\max}^2 - l_e^2}\right) - \frac{1}{2}Hl_{\max}l_e \ln\left[\frac{(l_{\max} + l)(l_{\max} - l_e)}{(l_{\max} - l)(l_{\max} + l_e)}\right] \quad (2)$$

In eq 2 subscripts ij, that should be attached to $V^{(conn)}$, l, l_e, l_{max}, and H, are omitted for the sake of legibility. The equilibrium length l_e ($V^{(conn)}(l_e) = 0$), the maximum elongation l_{max}, and the force constant H are the three parameters of this general spring potential, which we call "hard-FENE" because it includes, as particular cases, several commonly used spring types. When $l_{max} \to \infty$ (in practice, a sufficiently large number), it reduces to $V^{(conn)} = 1/2H(l - l_e)^2$, proper of a Hookean (Fraenkel)[45] spring that is usually employed, with a large value of H, to represent stiff connectors with an equilibrium length l_e (the rms fluctuation in spring length, $(\langle l^2\rangle - \langle l\rangle^2)^{1/2} = H/(k_BT)$[25,46] is, for instance 10% of l_e for $H = 100k_BT/l_e^2$). Furthermore, with $l_{max} \to \infty$ and $l_e = 0$ we have $V^{(conn)} = 1/2Hl^2$, which is the potential associated with a Gaussian distribution of the spring length, with $\langle l^2\rangle = 3k_BT/H$ as used in the Rouse model[13] of linear polymer chains composed by Gaussian subchains. The Rouse model with a linear force and infinite extensibility is inappropriate when external agents, particularly strong flows, stretch the chain and the distribution is not Gaussian. For those cases, the most popular choice is the FENE (finitely extensible, nonlinear elastic; Warner)[47] spring, whose potential, $V^{(conn)} = -(1/2)Hl_{max}^2 \ln(1 - l^2/l_{max}^2)$ is a particular case of eq 2 for $l_e = 0$. For the spring potential, as for the other pairwise potentials depending on the distance between elements, the forces acting on the two elements are $\mathbf{F}_i = -\mathbf{F}_j = [dV(r_{ij})/dr_{ij}]\mathbf{r}_{ij}/r_{ij}$, where the derivative of the potential adopts a long but immediate expression (omitted) that allows an easy calculation of the forces.

The angles, α_{ijk}, between two neighbor springs joining beads i and j, and j and k, may have an associated potential $V_{ijk}^{(ang)}(\alpha_{ijk})$. A simple and useful potential for the bending angle is the quadratic form, $V(\alpha) = (1/2)Q(\alpha - \alpha_0)^2$, where α_0 is the equilibrium value of the angle, and Q is a bending force constant. Again, the subscripts ijk are omitted for brevity, but we recall that there may be specific values of the parameters for each angle in the model. In order to make the program applicable to chemical entities (real molecules), we have also included in SIMUFLEX torsional potentials associated with hindered internal rotation. If ij, jk, and kl, are three consecutive bonds, internal rotation around the jk bond can be represented by a potential $V_{ijkl}^{(tors)}(\phi_{ijkl})$, where ϕ_{ijkl} is the internal rotation angle. In the program we have included a variety of $V(\phi)$ functions, corresponding to the most frequent kinds of chemical bonds. The expressions for the forces associated with bending and internal rotations can be found in the literature.[24,26]

The force field includes two other kinds of pairwise intramolecular potentials. One kind, denoted in eq 1 as $V_{ij}^{(EVpair)}$ is usually for excluded-volume (EV) interactions, for which BROWFLEX considers various possibilities, including the useful and meaningful Lennard-Jones potential

$$V^{(EVpair)} = 4\varepsilon_{LJ}\left[\left(\frac{\sigma_{LJ}}{r}\right)^{12} - \left(\frac{\sigma_{LJ}}{r}\right)^{6}\right] \quad (3)$$

where ϵ_{LJ} and σ_{LJ} are the Lennard-Jones parameters, along with other forms, like hard-spheres, exponential or Gaussian repulsion, etc. The second kind of pairwise potentials indicated in eq 1 as $V_{ij}^{(CHpair)}$ is intended for any other interaction that would superimpose to the excluded volume effect. An example is the intramolecular electrostatic interaction between charged elements, which can be properly described by a screened Coulomb, Debye−Hückel potential

$$V^{(CHpair)} = (A/r)\exp(-\kappa r) \quad (4)$$

where A is related to the charge of the two elements and the dielectric constant of the solution, and κ, the screening parameter, is related to the ionic strength of the medium.[48] Again, in eqs 3 and 4 the subscripts ij are omitted, and each pair may have its own parameters.

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

Finally, $V_i^{(E)}$ indicates any interaction between any individual element and an external agent (or constrain), since BD is an adequate technique to simulate macromolecules in, for example, electric fields[49,50] and in biomembranes.[51] BROWFLEX includes various useful possibilities like interaction of a charged element with an external electric field; walls that the element cannot trespass; and anchorage, by means of a hard spring, of an element to a fixed point. In the BROWFLEX user guide there is a detail relation of the forces included in the force field of BROWFLEX and how they must be used to build the chain model.

2.2. Hydrodynamics and Brownian Dynamics: The Algorithms. For the generation of Brownian trajectories of the mechanical molecular model in the viscous solvent, including hydrodynamic interaction effects, we propose the use of a procedure, based on the Ermak-McCammon[30,46] (E-M) algorithm, proposed by Iniesta and García de la Torre[29] (I-GT). In the E-M algorithm, the final position $\mathbf{r}_i$ of bead i after at time step Δt is calculated starting from its initial position $\mathbf{r}_i^0$, according to

$$\mathbf{r}_i = \mathbf{r}_i^0 + \frac{\Delta t}{k_B T}\sum_{j=1}^{N}\mathbf{D}_{ij}^0\cdot\mathbf{F}_j^0 + \Delta t\sum_{j=i}^{N}\left(\frac{\partial \mathbf{D}_{ij}}{\partial \mathbf{r}_j}\right)^0 + \mathbf{R}_i + (\Delta t)\mathbf{v}_i^0 \tag{5}$$

where $\mathbf{F}_j^0$ is the total force on bead i, $\mathbf{D}_{ij}$ is the 3×3 i,j-block of the $3N \times 3N$ grand diffusion tensor, $\mathbf{D}$, and $\mathbf{R}_i$ is a Gaussian random vector with zero mean and covariance

$$\langle \mathbf{R}_i\mathbf{R}_j\rangle = 2\mathbf{D}_{ij}\Delta t \quad i,j = 1, ..., N \tag{6}$$

If the dynamics takes place in a flowing solvent, we include in the E-M algorithm the term $\mathbf{v}_i^0$, which is the fluid velocity in the position of the bead i due to the flow field.

As the E-M is equivalent (without the Brownian drift term) to the first-order Euler algorithm for ordinary differential equations, Iniesta and García de la Torre proposed and algorithm inspired in the second-order Runge−Kutta procedure. In the I-GT procedure, each step is taken twice, in a predictor-corrector fashion. First, the predictor step is an E-M step, taking according to eq 5 that conducts to preliminary new bead positions $\mathbf{r}'_i$. Next, the forces, diffusion tensors, and their gradients are evaluated at these positions, and then the step is repeated, from the original initial position, taking the means of the quantities calculated before and after the predictor step (indicated with superscript ′); thus, the second corrector step is given by

$$\mathbf{r}_i = \mathbf{r}_i^0 + \frac{\Delta t}{k_B T}\frac{1}{2}\sum_{j=1}^{N}(\mathbf{D}_{ij}^0\cdot\mathbf{F}_j^0 + \mathbf{D'}_{ij}\cdot\mathbf{F'}_j) + \Delta t\frac{1}{2}\sum_{j=i}^{N}\left[\left(\frac{\partial \mathbf{D}_{ij}}{\partial \mathbf{r}_j}\right)^0 + \left(\frac{\partial \mathbf{D}_{ij}}{\partial \mathbf{r}_j}\right)'\right] + \mathbf{R}'_i + \Delta t\frac{1}{2}(\mathbf{v}_i^0 + \mathbf{v}'_i) \tag{7}$$

Although in the I-GT algorithm each step is taken twice, which amounts to duplicating the computing time per step, the time step Δt can be remarkably (over 1 order of magnitude) longer than in the first-order E-M procedure, so that the computing time needed to simulate a trajectory of a given duration is notably reduced. Several authors have commented on the advantages of the I-GT procedure.[52−55]

Fluctuating hydrodynamic interactions between beads can be accounted for by means of the Rotne-Prague-Yamakawa tensor,[56,57] valid when all elements or beads of the chain are equal size. García de la Torre and Bloomfield[58] extended that tensor to the case of nonequal elements. Using that HI tensor, the diffusion tensor that enters in the Brownian algorithm, $\mathbf{D}_{ij}$, reads

$$\mathbf{D}_{ij} = \frac{k_B T}{8\pi\eta_0 r_{ij}}\left[\mathbf{I} + \frac{\mathbf{r}_{ij}\mathbf{r}_{ij}}{r_{ij}^2} + \frac{\sigma_i^2 + \sigma_j^2}{r_{ij}^2}\left(\frac{1}{3}\mathbf{I} - \frac{\mathbf{r}_{ij}\mathbf{r}_{ij}}{r_{ij}^2}\right)\right] \tag{8}$$

where r_{ij} is the distance between beads i and j, and $\mathbf{I}$ is the unit tensor. If beads i and j overlap ($r_{ij} < \sigma_i + \sigma_j$), then

$$\mathbf{D}_{ij} = \frac{k_B T}{6\pi\eta_0\sigma}\left(1 - \frac{9}{32}\frac{r_{ij}}{\sigma}\right)\mathbf{I} + \frac{3}{32}\frac{\mathbf{r}_{ij}\mathbf{r}_{ij}}{r_{ij}\sigma} \tag{9}$$

where $\sigma = \sigma_i = \sigma_j$ if beads are equal size or $\sigma = (\sigma_i + \sigma_j)/2$ otherwise.[59]

Using this representation of the HI effect in the diffusion tensors instead of the original Oseen tensor, the gradient $\partial\mathbf{D}_{ij}/\partial\mathbf{r}_j$ terms in eqs 5 and 7 vanish and the simulation algorithm becomes simpler.

The most time-consuming process in BD with HI is the generation of the random displacement vectors, which require the calculation of a matrix $\mathbf{B}$ that satisfies $\mathbf{D} = \mathbf{B}\cdot\mathbf{B}^T$. For this purpose, McCammon and co-workers[30,46] used Cholesky decomposition, with computing time proportional to N^a with $a = 3$, and Fixman[32] proposed an alternative procedure, based on a Chebyshev polynomial approximation, that has been implemented by some authors,[60−62] with $a \approx 2$. It is clear that for sufficiently large N, the Fixman procedure will be more efficient, although (depending on details of the numerical implementation) for the moderate N employed in many instances, the procedure of McCammon may be faster. BROWFLEX will implement both methods, and a detailed comparison is to be published separately.

As in most dynamic simulation techniques, the time step Δt must be sufficiently small so that the forces do not change much in the step. When hard springs and other strong interactions are present in the model, this requires quite short steps. However, the fluctuations in hydrodynamic interaction are much slower than those interactions, and during such short time steps the change in the diffusion tensor is quite small. Then, in an efficient strategy,[43,63] the $\mathbf{D}$ tensor is not calculated at each time step; instead, it is kept fixed for a block of (say, 5−50) consecutive time steps, during which the same $\mathbf{B}$ is used.

As indicated above, BROWFLEX includes also the simulation of Brownian dynamics in a flowing solvent, which allows the prediction of rheological properties and single-molecule phenomena in flows.[37,39,64] In a homogeneous flow, the velocity of a fluid element can be written as

$$\mathbf{v}_i^0 = \mathbf{G}\cdot\mathbf{r}_i \tag{10}$$

Table 1. Velocity Gradient Tensors for Different Types of Flows

tensor	simple shear	uniaxial elongational	planar elongational
veloc. grad., **G**	$\begin{pmatrix} 0 & \dot{\gamma} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} \dot{\varepsilon} & 0 & 0 \\ 0 & -\dot{\varepsilon}/2 & 0 \\ 0 & 0 & -\dot{\varepsilon}/2 \end{pmatrix}$	$\begin{pmatrix} \dot{\varepsilon} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\dot{\varepsilon} \end{pmatrix}$

where $\mathbf{r}_i$ is the position vector of the fluid element, and **G** is the velocity gradient tensor that characterizes the flow. Table 1 gives the expressions[65,66] of that tensor for three common type of flows which, among others, are included in our program.

3. BROWFLEX, the Simulation Program

As commented on above, BROWFLEX is the program devised to perform both equilibrium and nonequilibrium BD simulations of bead-and-spring chains with any topology and with the possibility of selecting among several interaction potentials associated with connectors, angles, torsions, and nonbonded beads.

The information required to control the simulation is organized in several input files having a simple, clear format, so that long input files that are required in some situations can be written by other user-supplied ancillary programs. Thus, as other available software from our lab, BROWFLEX is a data-file driven program, and it is not necessary to have any script for running the simulations. The main input data file just contains the collection of names of both the other input files and the output files, as appreciated in Figure 2A. One of the output files will provide a run-time simulation report and the other one will contain the trajectory, i.e. the Cartesian coordinates defining the macromolecular conformations along the time.

Three compulsory input files are those containing (i) the initial conformation (initcfile.txt) that consists just of a list with the beads Cartesian coordinates; (ii) information on molecular features as number of beads and their radii, connectivity, and parameters of the forces (moleculefile.txt); and (iii) information on simulation features as its duration, the sample size (number of molecules), the time step value, and the type of algorithm used (brownfile.txt). In Figure 2, we show two examples of moleculefile.txt, one for a 12 base pairs double-helical DNA model (Figure 2B), and another one for a 471 base pairs bent DNA model (Figure 2C) (point lines indicate that content is larger but it was omitted to save space). As appreciated, the molecular file is formed by several blocks of information that allow for defining individually the different components of the model. Thus, we have a list with the beads hydrodynamic radii, next a list of bonds where the indices of the two connected beads and the connector force parameters are supplied, next the list of bending interactions with the indices of the three beads involved and the bending force parameters, then a block for torsions that in this case are not present, and finally the list of excluded volume interactions with the indices of the pair of beads not involved in bonds or angles and the excluded volume force parameters.

4. Analysis of Trajectories and Calculation of Properties. ANAFLEX, the Analysis Program

The other program that forms part of the SIMUFLEX suite, named ANAFLEX, was designed to analyze the trajectories generated by BROWFLEX. Separating the generation and the analysis of trajectories has the obvious advantages of speeding up the trajectory generation and allowing for analyzing the trajectory in as many ways as desired.

(A)

```
bentDNA471bp-log.txt          !log file
bentDNA471bp-tra.txt          !trajectory file
moleculefile.txt              !molecular file
initcfile.txt                 !initial file
-                             !no flow file
-                             !no electric field
-                             !no wall file
-                             !no special file
brownfile.txt                 !brownian file
```

(B)

```
20.                 !temperature (Celsius)
0.01                !solvent viscosity
12000.              !molecular weight
DNA 20 base pairs   !title
     40             !number of beads
 3.5E-8             !bead radius
........
    166             !number of connectors
.....................................
 7  10   1   2   1623.9   1.9127E-7 !middle connector
 8  10   1   2   1623.9   1.3584E-7
 9  10   1   2   1623.9   7.0538E-8
10  11   1   2   1623.9   7.0538E-8
10  12   1   2   1623.9   1.3584E-7
10  13   1   2   1623.9   1.9127E-7
10  29   1   2   1623.9   1.9322E-7
10  30   1   2   1623.9   2.0000E-7
10  31   1   2   1623.9   1.9322E-7
.....................................
      0             !number of bending angles
      0             !number of torsions
      0             !number of nonbonded interactions
```

(C)

```
20.                       !temperature (Celsius)
0.01                      !solvent viscosity
2.8E+5                    !molecular weight
Bent DNA 471 base pairs   !title
     11                   !number of beads
26.7E-8                   !bead radius
........
     10             !number of connectors
 1   2   1   2   1.580   160.0E-8     !1st connector
.....................................
10  11   1   2   1.580   160.0E-8     !last connector
      9             !number of bending angles
 1   2   3   1   2   0.000   7.39E-14 !1st angle
.....................................
 5   6   7   1   2   2.356   7.39E-14 !middle angle
.....................................
 9  10  11   1   2   0.000   7.39E-14 !last angle
      0             !number of torsions
      0             !number of nonbonded interactions
```

Figure 2. Examples of two of the user-supplied input data files for BROWFLEX: (A) main input file and (B,C) molecular input files for the 12 base pairs double-helical DNA model and for the 471 base pairs bent DNA model (point lines indicate that content is larger but it was omitted to save space).

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

Actually, ANAFLEX analyzes the Brownian trajectories in a number of ways:

• One of the analysis modes consists of the evaluation, as averages over a trajectory of a molecule simulated at equilibrium (steady-state) conditions, of overall properties, either conformational, like the radius of gyration, or hydrodynamic coefficients, in the above-mentioned rigid-body Monte Carlo approach (RBMC),[20−22,67] such as intrinsic viscosity or diffusion coefficient.

• In another mode, the trajectories of a number of molecules can be analyzed obtaining the averages, over the sample, of the properties as a function of time, thus predicting the time evolution of bulk solution properties upon the cessation of external agents (e.g., electrical or flow fields). Of course, it is possible to follow the evolution of each single molecule in order to characterize molecular individualism in single-molecule properties.[39,41,68,69]

• The equilibrium BD trajectories can be also analyzed to study the translational, rotational, and internal dynamics of rigid and flexible particles through the calculation of various time correlation functions, $C(t) = \langle F(t_0,t_0 + t)\rangle_{t_0}$, where $F(t_0,t_0 + t)$ is a quantity that depends on macromolecular conformation at time t_0 and at a later time $t_0 + t$, averaging over all possible choices of the initial time t_0. Some relevant correlation functions computed by ANAFLEX are as follows:

Translational Correlation.

$$C_{trans}(t) = \langle [\mathbf{r}_{cm}(t_0) - \mathbf{r}_{cm}(t_0 + t)]^2 \rangle_{t_0} \tag{11}$$

That is the Einstein equation for the center of mass (cm) mean-squared displacement, where the quantity $\mathbf{r}_{cm}$ is the position vector of the center of mass. From a linear fit of $C_{trans}(t)$, the translational diffusion coefficient can be obtained.

Correlation of Any Interelement Vector, Including the End-to-End Vector for Linear Topology.

$$C_{ij}(t) = \langle \mathbf{r}_{ij}(t_0) \cdot \mathbf{r}_{ij}(t_0 + t) \rangle_{t_0} \tag{12}$$

In this case, the correlated quantity is the scalar product of the value of some characteristic vector, $\mathbf{r}_{ij}$, defined between elements i and j of the macromolecular model, at time t_0 times its value at time $t_0 + t$. Similar correlations can be carried out for linear combinations of the $\mathbf{r}_{ij}$'s, such as those involved in the Rouse modes[13,14] of flexible polymer chains. Thus, the longest relaxation time of the chain can be computed from the decay of the correlation function of the first Rouse mode.[28,70] For the specific case of a linear chain, the information on the relaxation time is also contained in the correlation function of the end-to-end vector $\mathbf{r}_{1N}$.

***P*$_2$ Function of Some Characteristic Vector.** The internal dynamics of a flexible macromolecule can be characterized by the Brownian reorientation of some unitary vector defined inside the molecule, **u**. In such a case, the correlated quantity will be the angle θ (indeed its cosine) formed by two successive orientations of the vector when a time t has elapsed, i.e. the scalar product of **u** at two times separated by t. The correlation functions so defined are the Legendre polynomials. Particularly, a quite common correlation function, involved in the time-dependence of various observable properties (transient electric or flow birefringence and deformation, NMR relaxation, etc.)[30,71] is the second Legendre polynomial, P_2

$$\langle P_2(t)\rangle = \frac{3\langle [\mathbf{u}(t_0) \cdot \mathbf{u}(t_0 + t)]^2\rangle - 1}{2} = \frac{3\langle cos^2\theta(t_0, t_0 + t)\rangle_{t_0} - 1}{2} \tag{13}$$

P_2 decay is usually fitted to a multiexponential in order to obtain relaxation times associated with macromolecular internal dynamics. In the case of P_2-related electro-optic properties (birefringence, dichroism...) of rigid macromolecules, theory predicts that a set of up to five reorientational relaxation times can be found.[72,73]

DDLS Correlation Function. Another interesting function based on a second Legendre polynomial is the depolarized dynamic light scattering correlation function, C_{DDLS}[74,75] (also related to the electric birefringence decay that is proportional to the birefringence decay),[76,77] from which information on rigid-body rotation and internal dynamics can be extracted

$$C_{DDLS}(t) = \frac{1}{(N-1)^2} \sum_{j=1}^{N-1} \sum_{j'=1}^{N-1} \langle P_2[\mathbf{u}_j(t_0) \cdot \mathbf{u}_{j'}(t_0 + t)]\rangle_{t_0} \tag{14}$$

As observed, C_{DDLS} is related to P_2 and the scalar product $\mathbf{u}_j(t_0) \cdot \mathbf{u}_{j'}(t_0 + t)$ that is the cosine of the angle subtended by the connector vector j at instant t_0 and the connector vector j' at instant $t_0 + t$.

DLS Correlation Function. The (polarized) dynamic light scattering correlation function, C_{DLS}, may allow the simultaneous determination of the translational diffusion coefficient and quantities related to the internal dynamics of flexible particles[28,74,75] and the macromolecular global size

$$C_{DLS}(t,q) = \frac{1}{N^2}\left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} e^{-i\mathbf{q} \cdot [\mathbf{r}_i(t_0) - \mathbf{r}_j(t_0+t)]} \right\rangle_{t_0} \tag{15}$$

where the modulus of the scattering vector, $|\mathbf{q}| = (4\pi/\lambda)\sin(\theta/2)$, is determined by the wavelength of light, λ, and the scattering angle, θ.

Thus, dynamic coefficients (as translational diffusion coefficient) and time properties (as relaxation times) can be obtained by linear, polynomial, or multiexponential fits of different time correlation functions. Those fits are also carried out by ANAFLEX. In particular, ANAFLEX uses routines adapted from the program DISCRETE[78,79] to make multiexponential fits. It is well-known that fitting multiexponential functions with three or more components is an ill-posed problem. Simpler situations are those of rigid and symmetric particles or weakly bending rods. In any case one can hope to extract the longest or a mean relaxation time.[73]

As BROWFLEX, the ANAFLEX program is driven by simple data files. Figure 3 is an example of the only and simple input file for ANAFLEX. The three first lines are the names of different output files, then it comes a sequence of numeric codes or "flags" to inform ANAFLEX about the

```
bentDNA471bp-res.txt     !results file
bentDNA471bp-log.txt     !log file
bentDNA471bp-sum.txt     !summary file
10                       !sampling frequency
2                        !topology (linear case)
0                        !instant prop.(1-yes)
1                        !analysis mode (steady)
1,6                      !number of properties
18,21,0,0,0              !code for properties
bentDNA471bp-tra.txt     !trajectory file
```

Figure 3. Example of the user-supplied input data file for ANAFLEX.

properties to be calculated and the type of analysis to be performed, and finally the name of the trajectory file is supplied.

5. High-Performance Simulation in Multicore Platforms

In many instances, BD simulation problems are suitable for high-performance computing in multicore servers or clusters, because they may involve somehow independent simulations. The case is trivial when simulation is carried out for each of a sample of many molecules, as in the above-mentioned studies of time-dependency of bulk properties or single molecule behavior. Steady-state averages can be calculated either from a very long trajectory of one molecule or as mean values of the averages of the values computed for a number of molecules, and something similar happens when computing correlation functions $C(t)$. One can either obtain $C(t)$ from a single, very long simulation or determine the function for a number of independent trajectories, averaging $C(t)$ for each time t. The multiple trajectories can be made practically independent, if they are sufficiently long, by changing the sequence of random numbers - in practice, changing the seed of the sequential generator. Truly independent trajectories are those starting from different, initial conformations that would be generated a priori, for instance by Monte Carlo procedures, or even with an inexpensive BD simulation without HI.

In order to take advantage of multicore computers and clusters (even a computer with two *Quad* processor has eight cores), we have set up a scheme to run such multimolecule simulations, based on two ancillary tools. First, Multi-BROWFLEX "clones" what would be the files for a single simulation, producing multiple copies changing either the seed of the random numbers or the file containing the initial coordinates. It also generates a batch file for all the execution that is submitted to a load-balancing manager such as Sun Grid Engine. The outcome consists of multiple results files, one for each trajectory. Then, there is another tool, Multi-ANAFLEX, which is in charge of collecting and reading all those files, producing the final results as the proper averages over those of each molecule.

6. Examples: Multiscale Simulations of DNA

In order to demonstrate the usefulness and versatility of the mechanical model, BD algorithms and other methodologies implemented in the BROWFLEX suite, we have chosen a well-known and most relevant macromolecule, DNA, in sizes ranging from ~10 through ~10^5 base pairs. With a convenient, more or less coarse-grained bead-and-connector model, simulations of the dynamics in such multiple scales are possible employing the same methodology. Next, we show how BROWFLEX works by applying it to study the dynamics of several DNA models, always in experimentally observable situations.

6.1. A Double-Helical Model. There are macromolecular solution properties related to the local dynamics and structure of the molecule. In case of B-DNA this implies the convenience of modeling the double-helix properly. It is clear that atomic level simulations are quite expensive in CPU time. Then a suitable mesoscale model is built by considering the nucleotides in each strand as repetitive units. That model, proposed years ago by Horta and García de la Torre,[80,81] contains the characteristic double-helix parameters: number of base pairs, pitch, phase angle, helix diameter, etc. This is a highly valuable model to represent short fragments of DNA. Our group has already employed that model, where for simplicity beads lie only on the outside, in order to study hydrodynamic properties of double-helical DNA[63] (later, in the spirit of this kind of mesoscale simulations, similar models have been employed by other workers).[82,83] If larger DNA fragments were to be simulated, possible interpenetration of double helices can be avoided by using the hard-sphere excluded volume included in our program or alternatively some other DNA model.[84,85]

In this model, nucleotides are modeled by beads all of them with the same hydrodynamic radius, σ. Then, the number of beads will be $N = 2N_{bp}$, where N_{bp} is the number of base pairs (bp). In addition, some degree of flexibility is incorporated by using hard Hookean, elastic connectors (Fraenkel springs, defined in paragraph after eq 2). For the sake of minimizing the amount of interactions but keeping the double-helical shape and the stiffness at short scale, we found it adequate to connect each bead i to

1. its first neighbors along its strand (beads $i \pm 1$), which keeps connectivity and bond equilibrium length.

2. its second neighbors along its strand (beads $i \pm 2$), which accounts for bending interactions.

3. its third neighbors along its strand (beads $i \pm 3$), which accounts for torsional interactions.

4. its counterpart in the other strand (bead $i + N_{bp}$), which accounts for interactions between nucleotides forming the base pair.

5. the first neighbors of its counterpart in the other strand (beads $i + (N_{bp} \pm 1)$), which is necessary in order to keep the strands together.

Figure 4 shows this model (in the straight, equilibrium conformation) displaying all the connectors involving one of the innermost beads (indeed, these are the connectors whose data have been kept in Figure 2B).

As described in ref 86, we performed BD simulations by using program BROWFLEX for double-helical DNA models representing oligonucleotides with a different number of base pairs N_{bp}, all of them with the following features: helix radius, $r = 10$ Å, pitch $= 3.4$ Å, phase angle either $\phi = 180°$ (a symmetrical helix, Figure 4) or $\phi = 120°$ (a nonsymmetrical helix, more akin to the Watson−Crick structure), $H =$

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

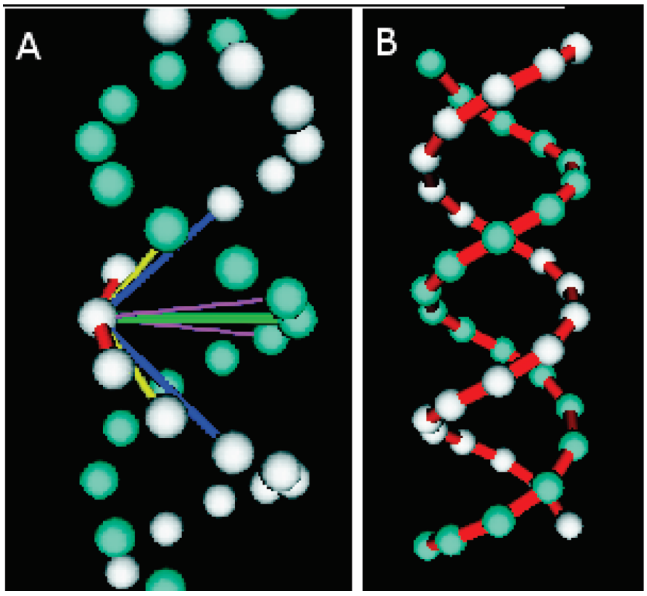

Figure 4. Double-helical model for DNA (A) showing the connectors supported by one bead and (B) showing only the connectors between neighbors in each strand.

Table 2. Diffusion Coefficients and Orientational Relaxation Times of DNA Oligonucleotides Obtained via Brownian Dynamics Simulations[a]

N_{bp}	simul. $\phi = 180^0$	simul. $\phi = 120^0$	exptl
	$D_t \times 10^7$(cm^2 s^{-1})		
8	16.3 ± 1.8	16.7 ± 1.6	15.3
12	13.2 ± 0.4	13.4 ± 1.2	13.4
20	10.6 ± 0.3	10.8 ± 0.5	10.9
	τ_{ee} (ns)		
8	3.3 ± 0.2	3.9 ± 0.5	3.2
12	6.1 ± 0.7	7.8 ± 0.8	6.4
20	16.7 ± 1.7	19.0 ± 1.3	16.2

[a] Comparison to experimental values obtained by Eimer and Pecora.[87]

$200k_BT/b_1^2$, where b_1 is the equilibrium length of the connector binding bead i to its first neighbor in its own strand, T = 293 K, $\eta_s = 0.01$ poise, hydrodynamic bead radius $\sigma = 3.5$ Å, and equilibrium spring length $b_1 = 7.0$ Å. The trajectories generated by the Brownian dynamics simulation were analyzed with program ANAFLEX in order to compute the translational diffusion coefficient, D_t, and the rotational relaxation time of the end-to-end vector, τ_{ee}. Within the statistical uncertainty of the simulations, the results for the two choices of ϕ are identical. Table 2 shows a comparison between the values of those dynamic properties obtained by simulation and the experimental values obtained by Eimer and Pecora.[87] As appreciated, the agreement is quite good. It is noteworthy that D_t and τ_{ee} are indeed quite close to the predictions of a rigid-body hydrodynamic calculation, using the HYDRO++ program[8] for the straight equilibrium conformation, demonstrating, thanks to BD simulations, the validity of the RBMC treatment for overall properties of quite stiff molecules. However, the relaxation time for the $\langle P_2(t)\rangle$ function for a vector perpendicular to the helical axis differs remarkably from the rigid-body prediction. While bending is scarcely noticeable in such short oligonucleotides, torsion of the helix, which influences the diffusivity of such transversal vector, is much more influential (for more details, see refs 63 and 86).

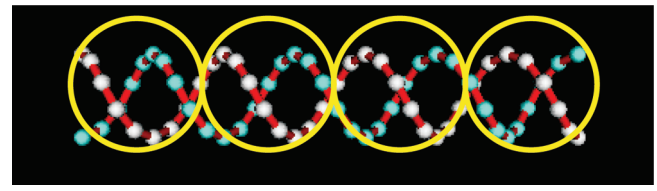

Figure 5. Sketch of a nearly touching beads model where each bead embraces a piece of the DNA double helix.

Mesoscale models, with one, two, or a few elements per amino acid residue[15,88] are now being considered as an alternative to atomistic simulation to predict dynamics of natively unfolded (intrinsically disordered) proteins or that of the folding process. Such models, and the important, recently demonstrated,[89] HI effects in protein dynamics, can be very adequately covered with our methodology.

6.2. A String of Touching Beads. A higher level in the coarse-graining procedure results from considering the macromolecular chain from a global perspective for which the precise local structure is not relevant and assuming that its flexibility is more or less uniformly distributed along its backbone. These assumptions give rise to the well-known wormlike chain model (or Kratky−Porod chain). In such a model, the double-helical structure of B-DNA is not explicitly considered. The wormlike chain can be represented with the generalized bead-and-spring model by setting constant the connector lengths, as in the case of a freely jointed chain, and allowing for the bond angles to fluctuate around their equilibrium values (0^o for an actual wormlike chain) with an amplitude that depends on the flexibility of the chain. Thus, the model can be used to represent from rigid to flexible structures just by playing with the bond angle parameters. In the same way, the model can be used to represent structures with local bents just by setting the bond angle located at the bent position to its characteristic value.

A limiting case of the wormlike chain model would be a string of nearly touching beads each embracing a piece of the double helix (see Figure 5). This macromolecular representation, which could be termed as a "fine-grained" model, was initially suggested by Schellman,[90] implemented in Monte Carlo simulations by Hagerman and Zimm[22] and in Brownian dynamics simulation by Allison and McCammon,[30,91] and is useful in representing short fragments of DNA of a few hundreds of base pairs. The bead diameter, which is the same as the connector length, is set to $b = 24.5$ Å. In that way, the diffusion coefficients of a straight string of beads are practically identical to those of a cylinder with a diameter of 20 Å, which is the hydrated diameter of DNA deduced from cylindrical models.[92] Then, the contour length of a model chain with N beads would be $L = Nb = 24.5N$ Å. On the other hand, the length of a double-helical B-DNA related to the number of base pairs is $L = 3.4N_{bp}$ Å so that the relationship between N and N_{bp} is $N = 0.14N_{bp}$. In the Schellman-Hagerman-Zimm model, the flexibility of the chain is represented by a bending potential quadratic in the bending angle, α, subtended by two successive links between neighbor beads. The bending force constant is related to the persistence length, P, by $Q = k_BTP/b$. Therefore,

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

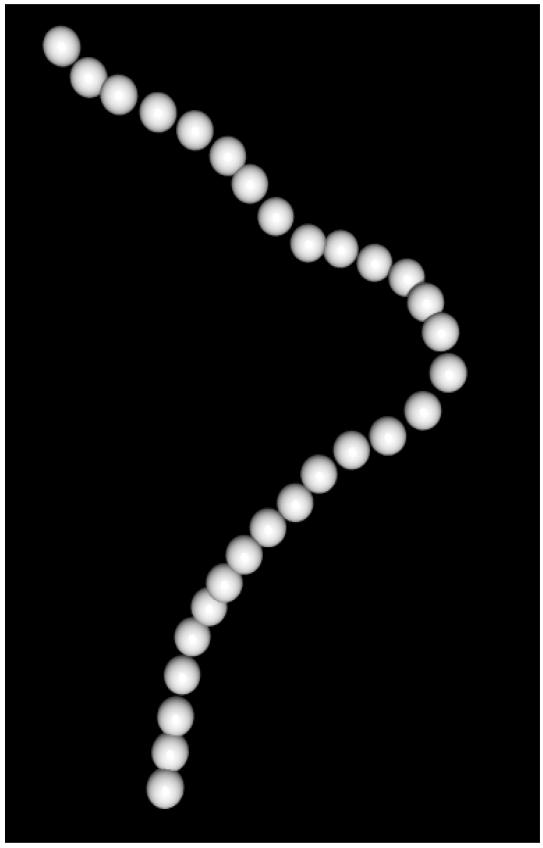

Figure 6. Nearly touching beads model for a DNA of 207 base pairs.

for a typical $P = 500$ Å and a $T = 293$ K, $Q = 8.24 \times 10^{-13}$ erg. The connector length is kept more or less constant (~10% in length fluctuation) by using stiff Fraenkel springs with a spring constant $H = 100k_BT/b^2 = 67.4$ erg/cm^2.

According to the above specifications and values of the model parameters, a DNA of 207 bp was modeled by 29 touching beads as illustrated in Figure 6. We simulated two variants of such a DNA molecule: a) an unbent DNA, with the equilibrium value of all the bond angles equal to zero, and b) a bent DNA, with the equilibrium value of the bond angles zero except for the central one that was set to 40°.

This comparison is intended to analyze the effect of such sharp bents, induced by some special base sequences, in the diffusivity of short DNA molecules.[73,93] For both model chain, BD simulations with HI were carried out by using BROWFLEX. From the generated trajectories, the translational diffusion coefficient and the orientational longest relaxation time of the end-to-end vector coming from both the correlation function based on depolarized dynamic light scattering (applicable to electric birefringence decay) and the correlation function based on the second Legendre polynomial P_2 were computed by using ANAFLEX. As noted above, Hagerman and Zimm[22] anticipated, other works[67,94] confirmed that the RBMC treatment (implemented in MONTEHYDRO)[19] works well for quite stiff semiflexible macromolecules in the prediction of not only translational diffusion but also overall rotational diffusion. Table 3 demonstrates that the RBMC results are in very good agreement with those of the BD simulations.

Table 3. Diffusion Coefficients and Orientational Relaxation Times of Bent and Unbent DNA Obtained via Brownian Dynamics and Monte Carlo Simulations

	bent	unbent
$D_t \times 10^7$ (cm^2 s^{-1}) (RBMC)	2.35	2.33
$D_t \times 10^7$ (cm^2 s^{-1}) (BD)	2.34	2.32
τ_{ee} (μs) (BD-P$_2$)	2.8	3.2
τ_{ee} (μs) (BD-DDLS)	2.5	3.1

6.3. Coarse-Grained Model of the Wormlike Chain. A cruder coarse-grained representation of the wormlike chain mentioned above consists of a string of nontouching beads connected by stiff springs, including a bending potential between successive connectors that determines the persistence length of the chain (see Figure 7). BD simulations on this kind of model were presented years ago by Allison and co-workers.[31,76,91,95]

The only free parameter of this model is the number of beads, N. As long as N is large enough the results are independent of its exact value. Thus, we can assign values to the parameters of our bead-and-spring chain by previously choosing a value for N (it is remarkable that a DNA molecule with 2311 base pairs can be modeled as a chain of only 10 beads).[95] Then, the connector length is fixed by the relationship $b = L/(N - 1)$. The constant of the stiff springs is set to $H = 100k_BT/b^2$, and the equilibrium bending angle is set to $\alpha_0 = 0$. Finally, the constant of the bending potential, Q, is chosen to adjust the persistence length, P, or the radius of gyration, R_g, of the real macromolecular chain, and the bead radius, σ, is chosen to adjust its translational diffusion coefficient, D_t. Using the just described parametrization procedure, we modeled an unbent DNA of 471 bp with $N = 21$ beads and a bent DNA of 471 bp with both $N = 11$ and $N = 21$ beads. The central bent was set to 45° in order to reproduce the kind DNA worked out by Stellwagen and coworkers.[93] Figure 8 shows the evolution of the depolarized dynamic light scattering correlation function for the three simulated DNA model chains. Those correlation functions were obtained after analyzing with ANAFLEX the corresponding Brownian trajectories with a duration $t = 2000$ μs (much longer than the longest relaxation time) generated with BROWFLEX. As observed, the value of N does not influence the results. On the other hand, the difference in dynamics of bent and unbent DNA is easily characterized. That figure is in agreement with Figure 2(b) in ref 93.

6.4. Single-Molecule Stretching of Long DNA in a Flow Field. It is well-known that flexible polymer chains subjected to extensional flows with a rate of strain greater than a certain critical value experience the so-called "coil-

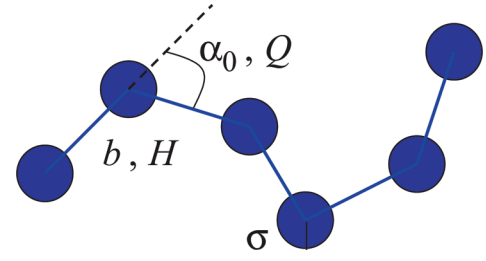

Figure 7. Bead-and-connector model for a wormlike chain.

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

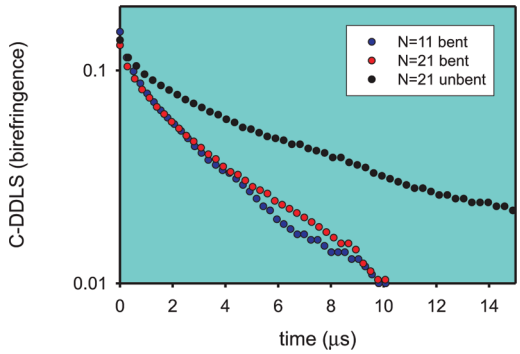

Figure 8. Evolution of the DDLS correlation function for the three DNA molecules modeled as wormlike chains.

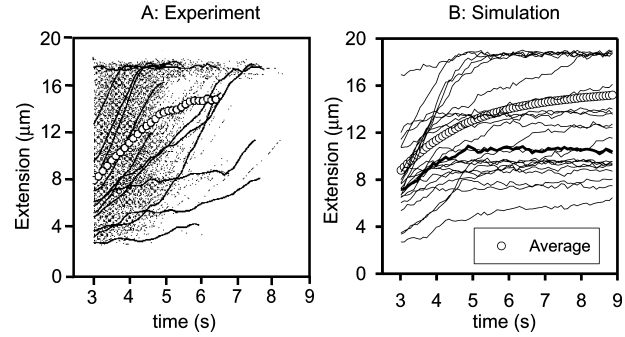

Figure 9. Time evolution of the extension of individual DNA molecules subjected to elongational flow. Comparison of experimental (left graph adapted from ref 41) and simulation (right graph) results.

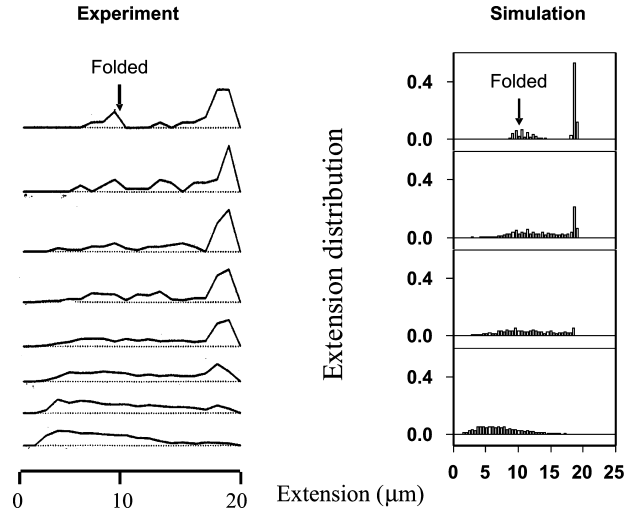

Figure 10. Time evolution of the distribution of the extension of a sample of DNA molecules. Comparison of experimental (left graph adapted from ref 41) and simulation (right graph) results.

stretch" transition.[96] This phenomenon consists of the abrupt unraveling of the random coil to a stretched conformation. In an already classical series of single-molecule experiments carried out with DNA, Chu and co-workers showed that the coil−stretch transition occurs in a particular manner for each chain in a given sample.[41,97,98] This is called molecular individualism.[68] Brownian dynamics simulation has been revealed to be an adequate technique to reproduce such a behavior.[17,39,40]

We show that BROWFLEX is able to reproduce the experimental results of a sample of DNA molecules as those employed in one of the pioneering experiments by Chu et al.[41] In that experiment a circular λ-DNA with 48.5×10^3 base pairs was converted into a linear chain by thermal treatment. Then, each macromolecular chain was stained fluorescently, subjected to elongational flow, and its stretching behavior was monitored by means of fluorescence microscopy. According to the authors of that work the stained DNA employed had a contour length $L = 22$ μm and a radius of gyration $\langle s^2 \rangle^{1/2} = 0.73$ μm (as estimated from the measured translational diffusion coefficient, $D_t = 0.47$ μm^2/s). Thus, assuming a characteristic persistence length for DNA of $P = 0.05$ μm, the ratio $L/P \simeq 440$ ensures that DNA is a flexible chain with random coil equilibrium conformation. Finally authors inform that the solvent viscosity is $\eta_s = 41$ cP and the working temperature is $T = 22.7$ °C. Since we are interested in reproducing the global conformation and dynamics of a quite large and flexible DNA chain, we can use a rough coarse-grained model. The model consists of a linear chain of $N = 20$ beads connected by $N - 1 = 19$ FENE springs that are able to capture both the Gaussian statistic that appears at low strain rate and the finite extensibility that plays a role at high strain rate. Thus, each bead represents a large DNA segment. Taking into account that $L = l_{max}(N - 1)$, we get a maximum spring length $l_{max} = 1.16$ μm. Then, by using the experimental radius of gyration, we get a value for the equilibrium spring length that, after some fitting refinements, turns out to be $b = 0.448$ μm.

We performed BD simulations of that FENE chain under elongational flow without EV interactions, which implies theta conditions, and including fluctuating HI, with a value of the hydrodynamic parameter $h^* = 0.25$, which corresponds to a hydrodynamic bead radius $\sigma = 0.257b$.

One of the optional input files, intended to describe steady or time-dependent flows (see BROWFLEX user guide), is used to specify a steady elongational rate $\dot{\varepsilon} = 0.86$ s^{-1}. That is the maximum elongational rate employed in ref 41, a value for which the molecular individualism is more clearly appreciated. The property monitored in the experiments by Chu and co-workers was the chain extension along the flow direction; therefore, we studied also the time evolution of that chain property. In order to get a good ensemble, we simulated 1500 molecules. In this study, ANAFLEX works in the multimolecule mode, supplying the time evolution of the sample-average and single-molecule properties (the extension as measured by Chu and co-workers is among the numerous properties that the program can analyze).

Figure 9 is a comparison of the evolution of the molecular extension computed from our simulations to that obtained experimentally.[41] The molecular individualism is appreciated as the particular paths followed by the time evolution of the chain extension: different chains experience coil−stretch transition at different times and reach different amount of extension. As observed, simulation results are in agreement with experiments as well as with other computer simulations.[99] Furthermore, in Figure 10 we compare the time

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

evolution of the histogram representing the probability distribution of the molecular extension obtained from simulations with that reported by Perkins et al. (Figure 1(B) in ref 41). Both histograms series are in excellent agreement. Initially, when the Hencky strain is small $\dot{\epsilon}t = 2.5$, most of the chains are close to their coil conformation although some of them can be stretched, and therefore both histograms are slightly broad and present a maximum at low extension ($\sim$5 μm). As time goes by and strain increases, histograms start to get broader and eventually a peak at a extension corresponding to the fully stretched conformation starts to develop owing to the increase number of chains that become stretched. That peak occurs at the same strain value in both experiment and simulation. Interestingly, at a higher strain a second peak at half the maximum extension arises. That second peak corresponds to folded "hairpin" conformations, which have an extension approximately half of the contour length. Again, the simulation was able to reproduce the experimental evidence.

7. Concluding Remarks

We provide a useful tool for the Brownian dynamics simulation and analysis of flexible and semiflexible bead-and-spring macromolecular chain models, SIMUFLEX, which consists of two programs, BROWFLEX and ANAFLEX. The programs are easy to use and were designed to be quite general. Thus, BROWFLEX can handle macromolecular models with any topology and include a number of common interaction potentials that can be easily extended in future versions. A key feature of this BD simulation tool is the inclusion of fluctuating HI, that allows to perform more realistic simulations. On the other hand, ANAFLEX is quite simple to employ and contemplates the analysis of a number of solution properties (in both steady-state and time-dependent conditions) and time correlation functions, which can also be extended in future versions. The examples employed in this paper have tried to show how the programs work and have in common to correspond to simulations of DNA dynamics in different scales. Thus, it was shown how SIMUFLEX is able to produce results comparable to experiments for different DNA problems that require to model the DNA chain with a different coarse-grain level.

8. Computer Methods

The SIMUFLEX suite that includes the BROWFLEX and ANAFLEX programs, as well as MONTEHYDRO and other related public-domain software, can be freely downloaded from our Web site at http://leonardo.inf.um.es/macromol.

Acknowledgment. This work was performed within a *Grupo de Excelencia de la Región de Murcia* (grant 04531/GERM/06). Support was also provided by grant CTQ-2006-06831 from the *Ministerio de Educación y Ciencia* (MEC), including FEDER funds. J.G.H.C. was the recipient of a *Ramón y Cajal* postdoctoral contract. R.P. acknowledges a postdoctoral fellowship from Fundación Séneca-CARM.

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

References

(1) Teraoka, I. *Polymer Solutions*; John Wiley and Sons: New York, 2002.

(2) Serdyuk, I. N.; Zaccai, N. R.; Zaccai, J. *Methods in Molecular Biophysics. Structure, Dynamics, Function*; Cambridge University Press: New York, 2007.

(3) Schlick, T. *Molecular Modeling and Simulation*; Springer: New York, 2002.

(4) García de la Torre, J.; Bloomfield, V. A. *Q. Rev. Biophys.* **1981**, *14*, 81–39.

(5) García de la Torre, J.; Navarro, S.; López Martínez, M. C.; Díaz, F. G.; López Cascales, J. J. *Biophys. J.* **1994**, *67*, 530–531.

(6) Carrasco, B.; García de la Torre, J. *Biophys. J.* **1999**, *76*, 3044–3057.

(7) García de la Torre, J.; Huertas, M. L.; Carrasco, B. *Biophys. J.* **2000**, *78*, 719–730.

(8) García de la Torre, J.; Del Río Echenique, G.; Ortega, A. *J. Phys. Chem. B* **2007**, *111*, 955–961.

(9) Allison, S. A. *Biophys. Chem.* **2001**, *93*, 197–213.

(10) Hahn, D. K.; Aragon, S. R. *J. Chem. Theory Comput.* **2006**, *2*, 1416–1428.

(11) Mansfield, M. L.; Douglas, J. F.; Kang, E. H. *Macromolecules* **2007**, *40*, 2575–2589.

(12) Allison, S. A.; Pei, H.; Haynes, M.; Xin, Y.; Law, L.; Labrum, J.; Augustin, D. *J. Phys. Chem. B* **2009**, *112*, 5858–5866.

(13) Rouse, P. E. *J. Chem. Phys.* **1953**, *21*, 1272–1280.

(14) Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*; Oxford University Press: Oxford, 1986.

(15) Ayton, G. S.; Noid, W. G.; Voth, G. A. *Curr. Opin. Struct. Biol.* **2007**, *17*, 192–198.

(16) Prabhakar, R.; Prakash, J. R.; Sridhar, T. *J. Rheol.* **2004**, *48*, 1251–1278.

(17) Sunthar, P.; Prakash, J. R. *Macromolecules* **2005**, *38*, 617–640.

(18) García de la Torre, J.; Pérez Sánchez, H. E.; Ortega, A.; Hernández Cifre, J. G.; Fernandes, M. X.; Díaz Baños, F. G.; López Martínez, M. C. *Eur. Biophys. J* **2003**, *32*, 477–486.

(19) García de la Torre, J.; Ortega, A.; Pérez Sánchez, H. E.; Hernández Cifre, J. G. *Biophys. Chem* **2005**, *116*, 121–128.

(20) Zimm, B. H. *Macromolecules* **1980**, *13*, 592–602.

(21) García de la Torre, J.; Jiménez, A.; Freire, J. J. *Macromolecules* **1982**, *15*, 148–154.

(22) Hagerman, P.; Zimm, B. H. *Biopolymers* **1981**, *20*, 1481–1502.

(23) Harvey, S. H.; García de la Torre, J. *Macromolecules* **1980**, *13*, 960–964.

(24) Allen, M. P.; Tildsley, D. J. *Computer Simulation of Liquids*, 1st ed.; Clarendon: Oxford, 1987.

(25) Díaz, F. G.; Iniesta, A.; García de la Torre, J. *J. Chem. Phys.* **1987**, *87*, 6021–6027.

(26) Díaz, F. G.; García de la Torre, J. *J. Chem. Phys.* **1988**, *88*, 7698–7705.

(27) Díaz, F. G.; García de la Torre, J.; Freire, J. J. *Polymer* **1989**, *30*, 259–264.

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n

(28) Rey, A.; Freire, J. J.; García de la Torre, J. *J. Chem. Phys.* **1989**, *90*, 2035–2041.

(29) Iniesta, A.; García de la Torre, J. *J. Chem. Phys.* **1990**, *92*, 2015–2019.

(30) Allison, S. A.; McCammon, J. A. *Biopolymers* **1984**, *23*, 167–187.

(31) Allison, S. A.; McCammon, J. A. *Biopolymers* **1984**, *23*, 363–375.

(32) Fixman, M. *Macromolecules* **1986**, *19*, 1195–1204.

(33) Fixman, M. *Macromolecules* **1986**, *19*, 1204–1207.

(34) Zylka, W.; Öttinger, H. C. *J. Chem. Phys.* **1989**, *90*, 474–480.

(35) Yamakawa, H. *Modern Theory of Polymer Solutions*; Harper & Row: New York, 1971.

(36) García de la Torre, J.; López Martínez, M. C.; Tirado, M. M.; Freire, J. J. *Macromolecules* **1983**, *16*, 1121–1127.

(37) Hernández Cifre, J. G.; García de la Torre, J. *J. Rheol.* **1999**, *43*, 339–358.

(38) Rossky, P. J.; Doll, J. D.; Friedman, H. L. *J. Chem. Phys.* **1978**, *69*, 4628–4633.

(39) Larson, R. G.; Hu, H.; Smith, D. E.; Chu, S. *J. Rheol.* **1999**, *43*, 267–304.

(40) Hur, J. S.; Shaqfeh, E. S. G.; Larson, R. G. *J. Rheol.* **2000**, *44*, 713–742.

(41) Perkins, T. T.; Smith, D. E.; Chu, S. *Science* **1997**, *276*, 2016–2021.

(42) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosiewicz, J.; Gilson, M. K.; Bagheri, B.; Scott, L. R.; McCammon, J. A. *Comput. Phys. Commun.* **1995**, *91*, 57–95.

(43) Klenin, K.; Merlitz, H.; Langowski, J. *Biophys. J.* **1998**, *74*, 780–788.

(44) Del Río Echenique, G.; Rodríguez Schmidt, R.; Freire, J. J.; Hernández Cifre, J. G.; García de la Torre, J. *J. Am. Chem. Soc.* **2009**, *131*, 8548–8556.

(45) Fraenkel, G. K. *J. Chem. Phys.* **1952**, *20*, 642–647.

(46) Ermak, D. L.; McCammon, J. A. *J. Chem. Phys.* **1978**, *69*, 1352–1360.

(47) Warner, H. R. *Ind. Eng. Chem. Fundam.* **1972**, *11*, 379–387.

(48) Pamies, R.; Hernández Cifre, J. G.; García de la Torre, J. *J. Polym. Sci. Pol. Phys.* **2007**, *45*, 1–9.

(49) Carrasco, B.; Pérez Belmonte, A.; López Martínez, M. C.; García de la Torre, J. *J. Chem. Phys.* **1996**, *100*, 9900–9905.

(50) Navarro, S.; Carrasco, B.; López Martínez, M. C.; García de la Torre, J. *J. Polym. Sci. Pol. Phys.* **1997**, *35*, 689–697.

(51) Fernandes, M. X.; Castanho, M. A. R. B.; García de la Torre, J. *BBA-Biomembranes* **2002**, *1565*, 29–35.

(52) Chirico, G.; Langowski, J. *Macromolecules* **1992**, *25*, 769–775.

(53) Öttinger, H. C. *Stochastic Processes in Polymer Fluids*; Springer: Berlin, 1996.

(54) Jian, H. M.; Vologodskii, A. V.; Schlick, T. *J. Comput. Phys.* **1997**, *136*, 168–179.

(55) Benham, C. J.; Mielke, S. P. *Annu. Rev. Biomed. Eng.* **2005**, *7*, 21–53.

(56) Rotne, J.; Prager, S. *J. Chem. Phys.* **1969**, *50*, 4831–4837.

(57) Yamakawa, H. *J. Chem. Phys.* **1970**, *53*, 436–443.

(58) García de la Torre, J.; Bloomfield, V. A. *Biopolymers* **1977**, *16*, 1747–1763.

(59) Carrasco, B.; García de la Torre, J.; Zipper, P. *Eur. Biophys. J.* **1999**, *28*, 510–515.

(60) Kröger, M.; Alba Pérez, A.; Laso, M.; Öttinger, H. C. *J. Chem. Phys.* **2000**, *113*, 4767–4773.

(61) Jendrejack, R. M.; Graham, M. D.; de Pablo, J. J. *J. Chem. Phys.* **2000**, *113*, 2894–2900.

(62) Schlick, T.; Beard, D. A.; Huang, J.; Strahs, D. A.; Qian, X. *Comput. Sci. Eng.* **2000**, *2*, 38–51.

(63) Huertas, M. L.; Navarro, S.; López Martínez, M. C.; García de la Torre, J. *Biophys. J.* **1997**, *73*, 3142–3153.

(64) Pamies, R.; López Martínez, M. C.; Hernández Cifre, J. G.; García de la Torre, J. *Macromolecules* **2005**, *38*, 1371–1377.

(65) Macosko, C. W. *Rheology. Principles, measurements and applications*; VCH Publishers: New York, 1994.

(66) Larson, R. G. *The Structure and Rheology of Complex Fluids*; Oxford University Press: New York, 1999.

(67) Iniesta, A.; Díaz, F. G.; García de la Torre, J. *Biophys. J.* **1988**, *54*, 269–275.

(68) De Gennes, P. G. *Science* **1997**, *276*, 1999.

(69) Hernández Cifre, J. G.; García de la Torre, J. *J. Non-Cryst. Solids* **1998**, *235−237*, 717–722.

(70) Rey, A.; Freire, J. J.; García de la Torre, J. *Macromolecules* **1990**, *23*, 3948–3953.

(71) Allison, S. A. *J. Chem. Phys.* **1989**, *90*, 3843–3849.

(72) Favro, L. D. *Phys. Rev.* **1960**, *119*, 53–62.

(73) García de la Torre, J. *Colloids Surf. B* **2007**, *56*, 4–15.

(74) Berne, B.; Pecora, R. *Dynamic Light Scattering*; John Wiley and Sons: New York, 1976.

(75) Chu, S. In *Soft Matter Characterization*; Borsali, R., Pecora, R., Eds.; Springer: New York, 2008; Chapter 7.

(76) Lewis, R. J.; Allison, S. A.; Eden, D.; Pecora, R. *J. Chem. Phys.* **1988**, *89*, 2490–2503.

(77) Allison, S. A.; Nambi, P. *Macromolecules* **1992**, *25*, 759–769.

(78) Provencher, S. *J. Chem. Phys.* **1976**, *64*, 2772–2777.

(79) Provencher, S. *Biophys. J.* **1976**, *16*, 151–170.

(80) García de la Torre, J.; Horta, A. *J. Phys. Chem.* **1976**, *80*, 2028–2035.

(81) García de la Torre, J.; Horta, A. *Polym. J.* **1977**, *9*, 33–39.

(82) Drukker, K.; Wu, G.; Schatz, G. C. *J. Chem. Phys.* **2001**, *114*, 579–588.

(83) Tepper, H.; Voth, G. A. *J. Chem. Phys.* **2005**, *122*, 124906.

(84) Malhorta, A.; Tan, R. K.-Z.; Harvey, S. C. *Biophys. J.* **1994**, *66*, 1777–1795.

(85) Devkota, B.; Petrov, A. S.; Lemieux, S.; Boz, M. B.; Tang, L.; Schneemann, A.; Johnson, J. E.; Harvey, S. C. *Biopolymers* **2009**, *91*, 530–538.

(86) García de la Torre, J.; Navarro, S.; López Martínez, M. C. *Biophys. J.* **1994**, *66*, 1573–1579.

(87) Eimer, W.; Pecora, R. *J. Chem. Phys.* **1991**, *94*, 2324–2329.

(88) Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144–150.

(89) Frembgen-Kesner, T.; Elcock, A. H. *J. Chem. Theory Comput.* **2009**, *5*, 242–256.

(90) Schellman, J. A. *Biopolymers* **1974**, *13*, 217–226.

(91) Allison, S. A. *Macromolecules* **1986**, *19*, 118–124.

(92) Tirado, M. M.; López Martínez, M. C.; García de la Torre, J. *J. Chem. Phys.* **1984**, *81*, 2047–2052.

(93) Lu, Y.; Weers, B. D.; Stellwagen, N. *Biophys. J.* **2003**, *85*, 409–415.

(94) García de la Torre, J. *Eur. Biophys. J.* **1994**, *23*, 307–322.

(95) Allison, S. A.; Sorlie, S. S.; Pecora, R. *Macromolecules* **1990**, *23*, 1110–1118.

(96) De Gennes, P. G. *J. Chem. Phys.* **1974**, *60*, 5030–5042.

(97) Smith, D. E.; Chu, S. *Science* **1998**, *281*, 1335–1340.

(98) Smith, D. E.; Babcock, H. P.; Chu, S. *Science* **1999**, *283*, 1724–1727.

(99) Schroeder, C. M.; Shaqfeh, E. S. G.; Chu, S. *Macromolecules* **2004**, *37*, 9242–9256.

CT900269N

Downloaded by UNIV DE MURCIA on October 13, 2009 | http://pubs.acs.org
Publication Date (Web): August 24, 2009 | doi: 10.1021/ct900269n