Conformational Dynamics of Hyaluronan Oligomers in Solution. 3. Molecular Dynamics from Monte Carlo Replica-Exchange Simulations and Mode-Coupling Diffusion Theory

ARTICLE *in* MACROMOLECULES · AUGUST 2004

Impact Factor: 5.8 · DOI: 10.1021/ma049641v

4 AUTHORS, INCLUDING:

Sara Furlan
Bracco Imaging, Ivrea, Italy
33 PUBLICATIONS **397** CITATIONS
SEE PROFILE

G. La Penna
Italian National Research Council
71 PUBLICATIONS **655** CITATIONS
SEE PROFILE

Attilio Cesàro
Università degli Studi di Trieste
165 PUBLICATIONS **2,354** CITATIONS
SEE PROFILE

Conformational Dynamics of Hyaluronan Oligomers in Solution. 3. Molecular Dynamics from Monte Carlo Replica-Exchange Simulations and Mode-Coupling Diffusion Theory

Sara Furlan,[†,‡] Giovanni La Penna,[‡,§] Angelo Perico,[‡] and Attilio Cesàro*[,†]

Department of Biochemistry, Biophysics and Macromolecular Chemistry, UdR INSTM, University of Trieste, I-34127 Trieste, Italy, and Institute for Macromolecular Studies, Section of Genova, National Research Council, via De Marini 6, I-16149 Genova, Italy

Received February 23, 2004; Revised Manuscript Received June 10, 2004

ABSTRACT: We present the calculation of dynamic properties of the hyaluronan oligomer $(UA)_4$ based on the second-order solution of the diffusion equation in the Smoluchowski limit. The equilibrium averages necessary for the solution are calculated from replica-exchange Monte Carlo simulations in implicit water models. This simulation technique was shown to avoid the trapping of the configurations in local minima typical of simulations in the canonical statistical ensemble. A simple implicit solvent model was applied to $(UA)_4$ taking into account the hydrophobic effect in water solution and hydrophilic local effects due to efficient hydrogen bonds with the solvent. The correction of the statistics due to a finite difference solution of the Poisson equation, not directly included in the simulation, was included through a perturbative method. Several configurational distributions and dynamical parameters related to nuclear magnetic relaxation, sensitive both to the molecular structure and to mobility, were calculated from the replica-exchange Monte Carlo statistics at different temperatures. These data compare fairly well with experiments and with the results obtained by standard molecular dynamics simulation in the explicit water solvent and in the canonical ensemble at $T = 300$ K. Implications in the macromolecular dynamics of hyaluronan were discussed.

I. Introduction

Many studies have been devoted to phenomenological interpretation of hyaluronan (HA, hereafter) unusual behavior and peculiar biological properties. Despite that the conservative chemical structure of HA (see Figure 1) contains elements of intriguing simplicity, yet many of its solution polymeric properties are still eluding a quantitative description at molecular level. On a macromolecular scale, relevant topological features of polysaccharides are revealed in the different degree of disorder going from fibers (and crystal) to concentrated and dilute solutions,[1,2] the overall statistics being described in terms of wormlike chains with a local stiffness.[3,4]

Statistical modeling of polysaccharide chain conformation has been introduced with considerable success on the basis of conformational maps of dimers,[4−6] where all nonbonding interactions are taken into account. In-registry modulation of medium-range and long-range interactions usually provides the peculiar viscoelastic properties. However, the versatility and the probability of the formation of interresidue hydrogen bonds and of other interactions in the HA chain can greatly modify local structure and mobility and, indirectly, change the solvation of the polymer segment.[7−10] As a consequence, it is difficult to provide a realistic representation of HA local structure that contains the essentials of the chain conformational properties.

To describe in atomistic details HA properties, oligomers, from $(UA)_2$ to $(UA)_{40}$, were investigated through nuclear magnetic resonance (NMR) experiments in dilute water solution.[11,12] NMR is widely used because it allows to monitor on an atomic scale the structural and dynamical effects of complicated interactions. In the case of HA oligomers, due to the capabilities of high-field NMR instruments, almost every carbon chemical shift can be assigned up to the $(UA)_4$ oligomer and ^{13}C NMR relaxation can be monitored through measurements of the $R(C_z)$ and $R(C_{xy})$ magnetic relaxation rates (relaxivities, hereafter). Relaxivities are linear combinations of spectral densities[13] that, in turn, are Fourier transformation of second-rank time autocorrelation functions (TCFs, hereafter) of internuclear distance vectors. Therefore, relaxivities contain information both on structure and dynamics of the molecule.

To take into account the effects of hydrogen bond networks, electrostatic interactions, solvation effects combined with short-range torsional potential of wide use in modeling polysaccharides, and numerical techniques were used to obtain the configurational statistics.[6,14] Molecular dynamics simulation algorithm (MD),[15,16] where a deterministic trajectory of every configurational variable is acquired in the canonical ensemble (constant temperature T and particle density ρ), can in principle provide both the dynamics of the variables and their statistical averages and distributions. Nevertheless, most of the MD applications to HA oligomers[11,17−21] did not involve a quantitative derivation of NMR relaxivities from MD trajectories because of the large amount of errors in simulated time correlation functions. Even in the case of peptides, which have been modeled for many years, to date only a few MD and stochastic simulations are long enough to give a good statistical description of molecular rotational tumbling.[22,23] Usually, to compare MD results with experiments, the measured NMR relaxation rates are analyzed within a model-free approach,[24,25] which describes the involved time correlation functions by sepa-

* Author for correspondence. E-mail: cesaro@units.it.
† University of Trieste.
‡ National Research Council of Genova.
§ Magnetic Resonance Center, University of Firenze, Italy.

10.1021/ma049641v CCC: $27.50
Published on Web 07/16/2004

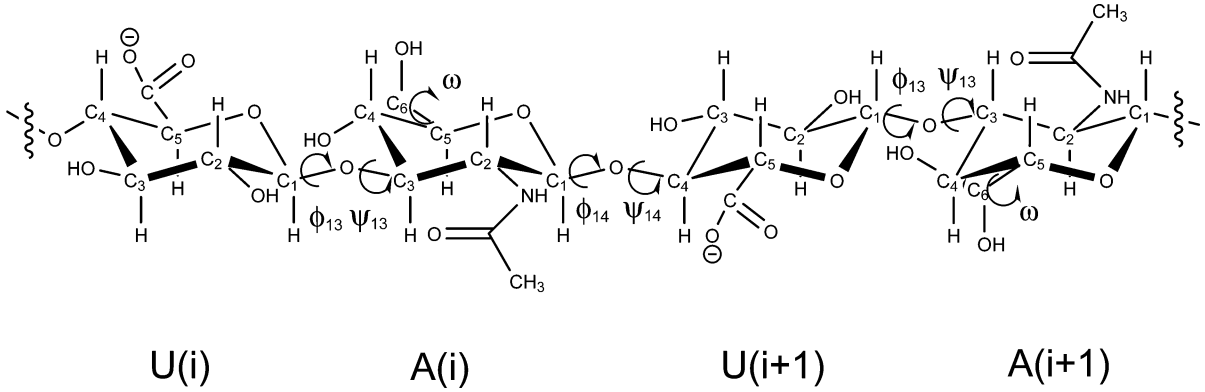

Figure 1. Schematic drawing of the tetrasaccharide GlcA-β-1,3-GlcNAc-β-1,4-GlcA-β-1,3-GlcNAc (U(i)−A(i)−U(i+1)−A(i+1)). Main atomic position and torsion angles at the glycosidic linkages and C6 atom are labeled.

rating overall rotation and internal motions. The parameters involved in fitting experimental results, with procedures that are restricted or extended according to the sophistication of the fitting equations, are essentially overall order parameters and rotational and internal correlation times. The same parameters can be extracted by the MD simulations and compared with those derived by the experiments. More advanced approaches[26] allow a detailed analysis of MD simulations, but again the lower relaxation times computed by MD must be increased in order to compare with experiments.

Calculation of NMR relaxation parameters based on the diffusion theory and mode-coupling approach (MCD theory, hereafter) were performed in different systems, ranging from synthetic polymers[27] to polysaccharides,[14,28] proteins,[29,30] and nucleic acids.[31,32] The application on HA oligomers[14] showed that the method allows the calculation on a microscopic basis of every second-rank TCF in the molecule once the configurational statistics were modeled, without the need for computer simulation of the intrinsic configurational time evolution. The combination of computer simulations with the diffusion theory gives a microscopic description of the NMR relaxivities for biopolymers, without any approximations concerning the separation[26,33] or coupling between time scales.[34]

Many approximations are usually introduced in order to simplify this kind of calculation in a macromolecule, but the most critical assumption is always the necessary numerical model for the configurational statistics. In particular, it has by now been well established that conventional MD or Monte Carlo (MC) simulations in the canonical ensemble construct trajectories that can be trapped in local energy minima.[35−37] This problem limits the application of standard MD to the description of the molecular wobbling of initially biased structures.

On the other hand, the MCD theory needs only statistical averages of configurational variables. To obtain configurational statistics efficiently, avoiding the "trapping" problem, simulation methods based on umbrella sampling[38,39] and generalized statistical ensembles[35,40] were developed and the statistics at a proper temperature can be obtained by reweighting techniques. The first application where replica-exchange MC (REMC) simulations and the MCD theory were combined together to compute ^{1}H-NOESY NMR cross-peaks, showed that, with a single simulation, dynamic and static properties of peptides can be described at several temperatures and in the presence of biologically relevant slow conformational transitions.[30] The application of such advanced statistical algorithm is limited to implicit solvent models. Even if this limitation can be removed, the design and setup of water solvation mean-field is promising for applications of statistical mechanics to large macromolecules.[41,42] Due to this interest, the models are rapidly improving, in the fields of finite difference numerical solution of the Poisson−Boltzmann equation,[43] in the fast computation of generalized Born radii,[44−49] in the calibration of hydrophobic interactions through surface-dependent functions,[48,50] in the applications of the RISM theory,[51] and in functional theory.[52−54]

As a further exploitation of the dynamical behavior of HA in this paper we have sought an application of REMC statistics to a model of the HA $(UA)_4$ oligomer, that is small enough to obtain reliable statistics, but large enough to represent the specific interactions and stiffness of the helical pitch of the polymer. The explicit aim is to reach a rigorous examination of several methods that thereafter could efficiently provide the description of HA large polymer dynamics and conformation, still subject to many controversial interpretations. This aim would therefore justify the some-how extensive use of methodological concepts. Essentially, comparison involves MD simulations in explicit water model (model MD-ES), REMC and MC simulations in vacuo (model S0), and in two implicit water models (models S1 and S2). The ultimate check has been the comparison of the experimental $R(C_z)$ ^{13}C NMR relaxivities with those calculated by using the MCD theory for the above simulated models.

The paper is organized as follows. In the "Methods" section the main aspects of different theoretical and computational methods are summarized, and the details regarding the choice of the many parameters involved in each method are reported. In the "Results" section the configuration statistics and the MCD molecular dynamics are analyzed for different models and compared to available experimental data; the conformational and dynamical properties fo HA oligomers are described. In the "Conclusions" section implications and perspectives of this approach in the frame of the HA polymer are summarized.

II. Methods

A. Replica-Exchange Simulation. We will, to begin with, briefly review the replica-exchange Monte Carlo method (REMC), also known as "parallel tempering" MC[55] (see, for instance, refs 35 and 56 for details).

In the REMC method, M noninteracting copies (or, replicas) of the model system in the canonical ensemble at M different temperatures T_m (m = 1, ..., M) were considered. Let X =

d{..., $x_m^{[i]}$, ...} stand for a "state" in this generalized ensemble. Here, $x_m^{[i]}$ stands for the state of the ith replica (at temperature T_m); the superscript i and the subscript m in $x_m^{[i]}$ label the replica and the temperature, respectively. Each state $x_m^{[i]}$ is specified by the coordinates $\boldsymbol{q}^{[i]}$ (and momenta $\boldsymbol{p}^{[i]}$) of all the atoms in replica i.

A simulation of REMC is then realized by alternately performing the following two steps. Step 1: Each replica in canonical ensemble of the fixed temperature is simulated simultaneously and independently for a certain number of MC steps. Step 2: A pair of replicas, e.g., i and j, which are at neighboring temperatures T_m and T_n, respectively, are exchanged:

$$X = \{..., x_m^{[i]}, ..., x_n^{[j]}, ...\} \rightarrow X' = \{..., x_m^{[j]}, ..., x_n^{[i]}, ...\} \quad (1)$$

The transition probability of this replica exchange is given by the Metropolis criterion:

$$w(X \rightarrow X') = \min(1, \exp(-\Delta)) \quad (2)$$

where

$$\Delta = (\beta_m - \beta_n)(U(\boldsymbol{q}^{[j]}) - U(\boldsymbol{q}^{[i]})) \quad (3)$$

with $\beta = 1/RT$ and where $U(\boldsymbol{q}^{[i]})$ and $U(\boldsymbol{q}^{[j]})$ are the potential energies of the ith replica and the jth replica, respectively. In the present work, we employ the standard Metropolis Monte Carlo algorithm[16] for step 1 and we ignore the temperature dependence of the potential energy U.

By repeating step 1 and step 2, a random walk in "temperature space" is realized for each replica, which in turn induces a random walk in potential energy space. This alleviates the problem of getting trapped in states of energy local minima.

The canonical expectation value of a physical quantity A at temperature T_m ($m = 1, ..., M$) was calculated by the usual arithmetic mean as follows:

$$\langle A \rangle_{T_m} = \frac{1}{N_c} \sum_{t=1}^{N_c} A(x_m^{[i]}(t)) \quad (4)$$

where N_c is the total number of measurements made at temperature T_m. Note that the above summation was taken over different replicas i ($i = 1, ..., M$) that happens to correspond to the fixed temperature T_m at the moment of measurement.

Once the statistics of the molecule was obtained with a given potential energy U, averages at a different potential energy $U' = U + \Delta U$ were obtained by

$$\langle A \rangle_{U'} = \langle A \exp(-\beta \Delta U) \rangle_U / \langle \exp(-\beta \Delta U) \rangle_U \quad (5)$$

where the two averages on the right-handside are computed along the REMC trajectory with the potential energy U and temperature T (eq 4). The limits of this equation are related to the appearance of large statistical errors with an increase in the difference between U and U'. The difference is expected to be small when the solvation free energy only is being modified in the model potential energy U.

In this work we have used the SMMP program[57] to perform REMC statistics. The CHARMM force-field developed for HA and previously applied in MD simulations[14,58] were inserted in the SMMP program. It is important to keep in mind that the carboxylate groups are all neutral in the adopted force-field.[14] This is a possible source of differences in comparing these results with other simulations where carboxylate groups are negatively charged and explicit counterions were included.

The $(UA)_4$ model is composed of 187 atoms (all the atoms in the molecule) and 52 dihedral angles were used as degrees of freedom. The dihedral angles in the pyranose rings were kept constant to their initially set values.

The electrostatic interactions were damped and shifted by using the charge neutralization approach[59] so as to allow us the use of a distance cutoff to speed-up simulations and to avoid the usual dependence on electrostatic potential energy from the cutoff. Both Lennard-Jones and damped electrostatic interactions were smoothly switched off between 0.9 and 1.0 nm, and the damping constant for electrostatic interactions was 2 nm^{-1}.

The MC algorithm extracts the statistics from suggested configurational moves. These moves were obtained, as usual, in the torsional space by randomly modifying dihedral angles and keeping bond distances and angles constant, including the ring dihedral angles. Therefore, the bond stretching and angle bending part of the CHARMM force field was never used in this kind of simulation, except in the determination of average geometrical parameters which were to be kept constant. Molecular mechanics studies have shown a great sensitivity of resulting statistics of polysaccharides upon the equilibrium values of distances and angles,[60] particularly for the geometry of the glycosidic linkages. These parameters enter directly into the main chain geometry, affect the molecular size and, therefore, the comparison between calculated and experimental NMR relaxation parameters. In this work, the starting configuration for the $(UA)_4$ oligomer was constructed using bond distances and bending angles calculated by averaging a MD simulation in the vacuum at $T = 300$ K. The usage of MD averaged geometrical parameters was suggested by the observation that glycosidic bending angles changes from 117° in MD to 111.5° in the CHARMM equilibrium value. We decided to use the bond distances and angles averaged along the MD simulation and in equivalent linkages. This correction partially includes the thermal fluctuation in these geometrical quantities due to a temperature close to experiments and the long-range interactions affecting the geometry of the glycosidic linkage. The pyranose rings were always in the $^4C_1$ configuration.

The initial configurations in the REMC simulation, were generated according to the geometrical parameters described above and the ϕ and ψ dihedral angles of the glycosidic linkages (see below for definition) were chosen randomly in the $0-2\pi$ range.

The dihedral angles to be modified were chosen sequentially in the variable set and the dihedral angle random displacement was in the range of $\pm\pi/9$ rad. The number of replicas was 10 in order to match the number of nodes of the cluster of personal computers we had at our disposal and the 10 temperatures of the replicas were 100, 200, 300, 400, 500, 600, 700, 830, 960, and 1100 K. This choice was indicated by the need of an efficient exchange of neighboring temperatures (with an acceptance ratio of about 0.15) and of having the experimentally monitored temperature (300 K) in the temperature set. Temperature exchange was attempted every 200 MC sweeps, where each MC sweep means attempting a change in all the variables. Configurations were saved every 100 MC sweeps and a total amount of 1 000 000 MC sweeps were performed. The averages, both neglecting and including the perturbation (eq 5), were, therefore, computed on 10 000 configurations.

B. Diffusion Theory. In this subsection we will summarize the advances of the diffusion theory[14,29,31,32] which were applied to the calculation of the $R(C_z)$ NMR relaxivities. The approximations in the mode-coupling expansion solution are the same as in the previous application to HA oligomers in ref 14.

Assuming that the ^{13}C nuclear spin relaxation is governed by the dipolar and chemical shift anisotropy mechanisms induced by the hydrogen atom bonded to ^{13}C, the relationships between the $R(C_z)$ relaxivities and spectral densities are[13,61]

$$R(C_z) = dJ(\omega_H - \omega_C) + (3d + c)J(\omega_C) + 6dJ(\omega_H + \omega_C) \quad (6)$$

with

$$d = \frac{\gamma_C^2 \gamma_H^2 \hbar^2}{20}, \quad c = \frac{(\omega_C \delta)^2}{15} \quad (7)$$

γ_i is the gyromagnetic ratio of nucleus i, and δ is the (dimensionless) chemical shift anisotropy (CSA) of each C nucleus. In this work we have assumed that the CSA is −40 ppm for all the monitored C−H bonds.

The spectral densities J are the Fourier transform:

$$J(\omega) = 2\int_0^\infty \cos(\omega t)\ \mathrm{TCF}(t)\ \mathrm{d}t \tag{8}$$

of the time-correlation function (TCF) of second-rank tensor components of the vectors $\boldsymbol{r}$ joining the monitored ^{13}C nucleus with the bonded protons.

These TCF s have the following form[61]

$$\mathrm{TCF}(t) = \left\langle \frac{1}{r^6} \right\rangle \sum_{M=-2}^{2} \langle [D_{M,0}{}^{(2)*}(\Omega(t))][D_{M,0}{}^{(2)}(\Omega(0))] \rangle = \left\langle \frac{1}{r^6} \right\rangle P_2(\cos[\theta(t)]) \tag{9}$$

where $D_{M,0}{}^{(2)}$ are irreducible spherical tensors, P_2 is the second-order Legendre polynomial,[62] and Ω and r are the direction and the modulus, respectively, of the given C−H vector involved in the observed relaxivity. The averages are for thermal equilibrium at a given temperature T. The above equation holds when the effects of the modulus r and orientation of the unit vector (direction) in the TCF are dynamically well separated, as is the case when the two nuclei are covalently bonded.

Integrating the orientational part of the TCF the correlation time allows τ to be obtained:

$$\tau = \int_0^\infty P_2(\cos\beta(t))\mathrm{d}t \tag{10}$$

The mode-coupling diffusion (MCD) theory of the dynamics of a biological macromolecule in solution was adopted for the computation of the above TCF of eq 9. The diffusion theory treats the solvent hydrodynamically and uses a detailed molecular model for the polymer in terms of beads (atoms or groups of atoms) connected by real or effective bonds diffusing in an atomistic potential (the same potential used in the simulation). The beads are represented as points of coordinates $\boldsymbol{r}_i$ with friction coefficients $\zeta_i = 6\pi\eta a_i$, where η is the solvent viscosity. The Stokes' radii a_i were calculated here by using the accessible surface area (ASA) method[63] with a zero probe radius and their values and locations are summarized below.

According to the mode-coupling solution to the diffusion time evolution equation, any TCF of a function f of the variables, bond vectors, is obtained as a multiexponential expansion:

$$\langle f(t)\, f(0)\rangle = \sum_m a_m \exp(-\lambda_m t) \tag{11}$$

with the preexponential coefficients a_m given as

$$a_m = \left(\sum_i c_{i,m}\langle f\phi_i\rangle\right)^2 \tag{12}$$

These equations are obtained after expansion of the solution of the generalized diffusion equation in terms of eigenvalues and eigenvectors of the adjoint of the diffusion operator (see Appendix and eq A1) and expansion of the eigenfunctions in a mode-coupling basis set $\{\phi_i,\ i = 1, ..., M\}$ (eq A3). The constant vector $c_{i,m}$, $i = 1, ..., M$ contains the coefficients of the expansion of eigenfuntion m in the basis set and $-\lambda_m$ are the eigenvalues of the adjoint diffusion operator, with λ_m the relaxation rates of the diffusing polymer system. If we include in the mode-coupling basis set all the infinite powers of the bond variables with the correct rotational symmetry, we obtain the exact solution to the generalized diffusion equation (eq A1). In practice we construct approximations to this exact mode-coupling solution simply by truncating the number of powers included in the basis set for computational convenience.

As a result the time evolution of the TCF is solved in terms of equilibrium averages necessary to calculate the matrix elements of the force and metric matrices **F** and **S** (eqs A6 and A5, respectively) and the ensemble averages of the projections of the function f on the elements of the basis set ($\langle f\phi_i\rangle$). In practice, these equilibrium averages are calculated by time averaging over a MD trajectory or by Monte Carlo averaging. In the present work, we have used the replica-exchange Monte Carlo method for the calculation of these averages (see eq 4) We note that the mode-coupling solution to the generalized diffusion equation completely avoids the classical preaveraging of the hydrodynamic interactions, preserving the full nonlinearity and anisotropy of the diffusion equation. The basis set should be chosen also having the correct rotational symmetry or tensorial rank: first rank tensors are simply vectors, and second rank tensors are spherical functions. Note also that the natural and simplest basis set of first-rank first-order is given by the bonds or equivalently by linear combinations of bonds as the first-rank first-order eigenmodes. This basis set simply provides the exact dynamics of the Gaussian polymer model with preaveraged hydrodynamic interaction.[64,65]

For time-correlation functions of second-rank tensor components, such as those necessary to calculate NMR relaxation (see eq 6) a reduced second-order basis set RM2-II[66] can be generated coupling the first-order/first-rank modes in the proper irreducible tensorial form up to the fourth power. Let e be the number of the lowest-rate first-order/first-rank modes that we choose. The RM2-II basis set contains $e(e+1)/2$ second powers and $[e(e+1)/2]^2$ fourth powers of these e first-rank/first-order modes. Once the second-rank TCF s have been generated, they can be Fourier transformed to the spectral densities (sums of Lorentzian functions) which are required to compute the $R(C_z)$ relaxivities of eq 6. In this work, the choice of hydrodynamic parameters (Stokes' radii, friction location, and hydrodynamic screening constant α) was identical to the previous application of MCD theory to the $(UA)_4$ HA oligomer.[14] Summarizing, the friction points are located on most of the heavy atoms in each monomer, for a total amount of 89 friction points. Each bead represents isotropically the friction of the heavy atom together with the bonded hydrogen atoms. The Stokes' radii of the beads are derived by the zero-probe ASA of the group of atoms centered on each bead. The radii are within 0.073 nm (the single O1 atom) and 0.185 nm (the methyl group of residue A). The screening constant α in eq A7 is 0.25.

The RM2-II basis-set used for the second-order solution of the diffusion equation was built with e lowest rate first-order/first-rank modes. Value up to $e = 10$ were used, to test the convergence of the RM2-II basis-set. Values of e larger than 5 produce a variation within 10% of NMR $R(C_z)$, i.e., within both the NMR experimental error and the best estimate of the standard error on the calculated relaxivities. A more significant difference was observed in both NMR relaxivities and correlation times by increasing e from 3 to 5. This effect was less pronounced in the previous work based on the MD statistics,[14] thus showing that a careful check of the convergence of the addressed relaxation parameters is mandatory when the configuration statistics is changed. The value of $e = 5$ was used for all the following analysis of molecular dynamics.

The second-rank dynamics of the C−H bond orientation, where C is one of the NMR probed ^{13}C nuclei in the molecule, are given by eq 11. This is a multiexponential expression that, according to the rate distribution, can be eventually separated in two parts: a part roughly associated with an average rotational tumbling and the other one to the coupling of rotational and internal modes. Adopting a simple form like that proposed by Lipari and Szabo,[24] the TCF in eq 9 can be described in terms of two decaying exponentials:

$$P_2(\cos[\theta(t)]) = \sum_{i=1}^{5} a_m \exp(-\lambda_m t) + \sum_{i=6}^{M} a_m \exp(-\lambda_m t) \simeq S^2\exp(-t/T_R) + (1 - S^2)\exp(-t/T_I) \tag{13}$$

In the two-exponential approximation S^2 has the meaning of an order parameter describing the fast fluctuations of the molecule: $S^2 = 1$ means that the molecule is rigid and spherical while $S^2 = 0$ means that the C−H bond fluctuates completely independently from the rest of the molecule. This parameter is given as the sums of the five coefficients of the "average rotational tumbling":

$$S^2 = \sum_{i=1}^{5} a_m \tag{14}$$

Then, the mean rotational time may be written in terms of the first five coefficients and relaxation rates as

$$T_R = (S^2)^{-1} \sum_{i=1}^{5} \frac{a_m}{\lambda_m} \tag{15}$$

Note that for second-rank functions the average tumbling is described approximately by the 5 lower rates in the entire rate spectrum. In general, the five lower rates are different corresponding to an average anisotropic rotational diffusion tensor. Here, we shall ignore this anisotropy to estimate with a single number the order of magnitude of the average rotational time.

Finally, the relaxation times relative to the coupled rotational and internal fluctuations, T_R and T_I, and to the pure internal fluctuations, T_F, are

$$T_I = (1 - S^2)^{-1} \sum_{i=1}^{5} \frac{a_m}{\lambda_m} \tag{16}$$

$$T_F = T_I T_R / (T_R - T_I) \tag{17}$$

In addition, by integrating eq 13 we can separate the contributions to the correlation time of the TCF, eq 13, into two parts, due to the average overall tumbling and to the internal fluctuations, respectively:

$$\tau = \tau_R + \tau_I = R\tau + I\tau \tag{18}$$

with

$$R = \tau_R/\tau = \tau^{-1} \sum_{i=1}^{5} \frac{a_m}{\lambda_m} \tag{19}$$

$$I = \tau_I/\tau = \tau^{-1} \sum_{i=6}^{M} \frac{a_m}{\lambda_m} \tag{20}$$

C. Implicit Solvent Model. The solvation free energy for a solute macromolecule in a given configuration is the potential energy of the configuration averaged over all the configurational variables of the solvent at constant temperature and pressure, i.e. in the isobaric−isothermal statistical ensemble. The solvation free energy, ΔG_{solv}, is a mean field contribution to the potential energy $U(\boldsymbol{q})$ used in the REMC method for the solute in the given configuration $\boldsymbol{q}$ (eq 3). ΔG_{solv} can be decomposed in two terms, ΔG_{nopol} and ΔG_{pol}. The first term is the contribution to the solvation free energy due to the formation of a cavity of zero charge density with the shape of the solute and to the creation of the solute−solvent interface. The introduction of a charge density in the space occupied by the solute macromolecule gives the second contribution. The charge density, in the usual classical force fields for atomistic molecular models, is given in terms of the point charges q_i of atom i in the CHARMM force field, where i runs over the N_a atoms in the molecule. These point charges are located in the positions $\boldsymbol{r}_i$ of the given molecular configuration.

A widely used approximation for ΔG_{nopol} is a linear combination of the solvent accessible surface areas (SASAs) of each group in the solute molecule:[50,67]

$$\Delta G_{nopol} = \sum_{i}^{N_a} \sigma_i SASA_i \tag{21}$$

The coefficients σ_i are mainly positive because nonpolar groups are generally hydrophobic groups. Nevertheless, the same relationship was used for hydrophilic groups, with negative σ coefficients, to model the enthalpic contribution due to hydrogen bonds between the solvent and the solute polar atoms. This means implicitely that the electrostatic contribution ΔG_{pol} is assumed to be approximately constant in the data set adopted in order to derive the σ coefficients.

The electrostatic contribution to the solvation free energy, ΔG_{pol}, can be approximated by using the information that the solute charge density is placed within the volume occupied by solute atoms. This volume is assumed to have a relative dielectric permittivity of one, and to be confined in a continuum solvent with a given relative permittivity (80 in the case of water at room conditions). The electrostatic energy for a solute molecule with an arbitrary shape can be obtained by numerical finite difference solution of the Poisson equation (PE).[43] The solvation energy is obtained in terms of the interactions between the atomic point charges and the polarization surface charge density placed on the solute−water interface.[68]

The evaluation of SASA of each atom was performed using the NSC code[69] with a density of 122 points per unit sphere. The probe radius was 0.14 nm and it was kept fixed for all of the following calculations. The σ coefficients of eq 21 were taken from the literature,[50] and the SASA of each group was calculated with the van der Waals radii of CHARMM.

The finite difference solution of the Poisson equation for the electrostatic contribution to the solvation free energy was performed by using the approach described in ref 43 in order to compute the electrostatic potential in a grid of at maximum 144 points per side. Each configuration of the $(UA)_4$ molecule was placed in the center of an ortorhombic cell providing at least 1 nm of boundary dielectric in every direction of space. The Coulomb boundary conditions were adopted; i.e., the electrostatic potential was zero at the cell boundaries. The grid size was kept fixed at 0.0425 nm.

The relative dielectric permittivities of the molecule and of the solvent were 1 and 80, respectively. The definition of SASA needed to assign points to the interior and exterior of the molecule was performed with the same NSC code used for eq 21 in the REMC simulation. The correction proposed in ref 68 for computing the electrostatic solvation energy in terms of the Coulomb interaction between polarization charges positioned on the molecular surface and the point charges positioned on the atomic nuclei was adopted.

D. Conformational Definitions. In the next section, we will use several configurational properties to describe the $(UA)_4$ statistics at different temperatures and for the different solvent models. Here is a brief summary of the related definitions.

The end-to-end distance is defined as the distance in space between C4 (U_1) and C1 (A_4) atoms.

The ϕ and ψ dihedral angles in the glycosidic linkages are defined as the dihedral angles between the bonds H1(U_i)−C1(U_i)−O1(U_i)−C3(A_i) and C1(U_i)−O1(U_i)−C3(A_i)−H3(A_i) for UA and H1(A_i)−C1(A_i)−O1(A_i)−C4(U_{i+1}) and C1(A_i)−O1(A_i)−C4(U_{i+1})−H4(U_{i+1}) for AU, respectively, where the dihedral angle is 180° in the trans conformation.

Conformational maps of UA and AU dimers linkages will be used for comparisons. These maps were obtained through simulated annealing trajectories of the UA and AU molecules in the vacuum. A MC trajectory of 10 000 sweeps was performed for each pair of fixed ϕ and ψ dihedral angle modified in steps of 10°. The temperature of the MC trajectory is to be exponentially decreased according to rule $\beta = \beta_0(\beta_1/\beta_0)^{i-1/N-1}$ with i as the current MC sweep, N the total number of MC sweeps, the initial temperature $T_0 = 1000$ K

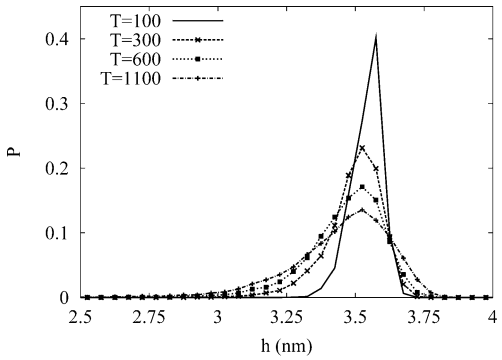

Figure 2. End-to-end distance distribution for the replica-exchange MC (REMC) simulation with the implicit short-range solvent model S1: $T = 100$ K (solid line), $T = 300$ K (long dashes and ×), $T = 600$ K (short dashes and ■) and $T = 1100$ K (dotted line and +).

and the final temperature $T_1 = 100$ K. The potential energy corresponding to the minimum explored during the annealing is being assigned to the corresponding point in the conformational map.

The HB population can be monitored, as usual, by identifying and counting the X−H···Y configurations where the X···Y distance is within 0.35 nm and the X−H···Y angle is in the range 135−180°. Every heavy atom was here considered as candidate X or Y atoms: only populations larger than 10% are to be considered significant.

III. Results

Before presenting the results, let here briefly summarize the models and the acronyms that will be encountered in the following. In this paper, three models will be studied by replica-exchange Monte Carlo (REMC) simulations and mode-coupling diffusion (MCD) theory. The first model (model S0) is in the vacuum. A second model (model S1) includes only the short-range contribution to ΔG_{nopol} modeled by eq 21 with the coefficients tabulated by Ooi et al.[50] The statistics obtained with model S1 are reweighted through eq 5 by using as incremental potential energy the electrostatic contribution to the solvation free energy obtained by finite difference solution of Poisson equation (model S2). The results of the previous molecular dynamics (MD) simulation in the explicit solvent (model MD-ES) are the only ones used for comparisons.

A. Configuration statistics of HA. The REMC simulations allows the calculation of distributions and averages at the temperatures used in the exchange process. In Figure 2, the distributions of the end-to-end distance for the S1 model are compared at four selected temperatures, 100, 300, 600, and 1100 K. The distribution shows a sharp peak at 100 K and becomes broader at 300 K. At temperatures beyond 600 K the width of the distribution increases only slightly, thus suggesting that the configurational landscape in a given force field is characterized by barriers that can be overtaken by the REMC simulation. In Figure 3, the trajectories in the ϕ/ψ glycosidic dihedral angles plane obtained for the U(3)−A(3) (a−c) and A(2)−U(3) (d−f) linkages at the three temperatures $T = 100$, 300, and 1100 K are shown: the higher density of dots indicates a higher population in the given ϕ/ψ domain. Both the UA and AU maps display sets of accessible conformations located more or less in the central region of the figure, although with increasing temperature ($T < 300$ K) UA population only broadens its contour range, while AU seems to elongate toward other regions with different

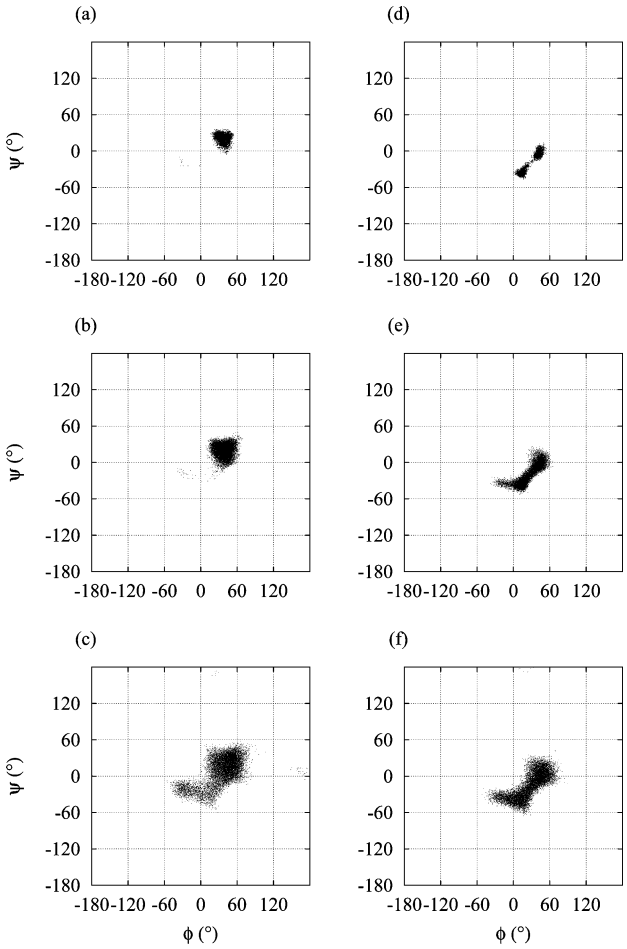

Figure 3. Trajectory of ϕ and ψ glycosydic dihedral angles of the U(3)−A(3) (a−c) and A(2)−U(3) (d−f) glycosydic linkages in the replica-exchange MC (REMC) simulation with the implicit short-range solvent model S1: $T = 100$ K (a, d), $T = 300$ K (b, e), and $T = 1100$ K (c, f).

topologies. This originates from the presence of secondary minima with different energy that displace the populations with the temperature in a different way. Furthermore, in the plots at the highest (1100 K) temperature other sets of (higher) energy minima become slightly populated showing that the exploration occurs over the whole map notwithstanding the presence of high energy barriers between the minima. It will be shown below that these different features of the two maps as a function of temperature have a counterpart in the scheme and stability of hydrogen bonding.

The trajectories in the ϕ/ψ plane obtained through REMC at 300 K can be compared with the MD simulation of $(UA)_5$ in the explicit water solvent.[19] Both simulations display a population concentrated in one and two minima for the UA and AU linkages, respectively. The locations of these minima are only slightly different, and as expected, MD distributions are broader due to the accounted bending flexibility. The comparison of these two different models is good despite the remarkable differences in intramolecular and solvation potential.

In Figure 4, the lower energy part of the potential energy maps obtained through simulated annealing of the UA and AU S0 models are displayed. The comparison between Figures 3 and 4, clearly shows that the REMC trajectories at high temperature move within the given conformational map overtaking the potential

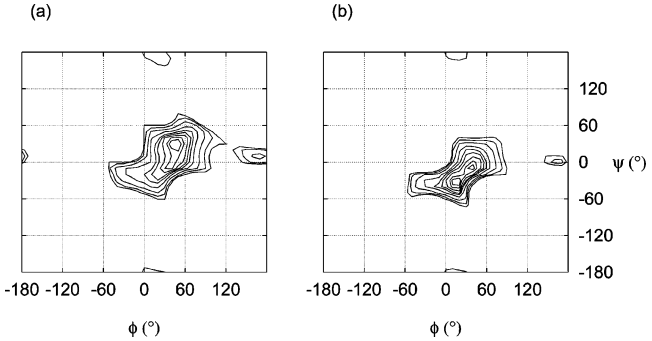

Figure 4. Conformational maps of UA (a) and AU (b) dimers in the vacuum obtained by simulated annealing (see text for details) with fixed ϕ and ψ dihedral angles in the glycosydic linkages. Contour levels are at 0, 0.5, 1, 2, 3, 4, 6, 10, 15, and 20 kcal/mol with respect to the absolute minimum calculated for each map.

energy barriers between different local minima. The different shape of the AU landscape and its temperature variation of accessible conformational domain make this glycosidic linkage less flexible, consistently with the MD-ES results. This overall behavior becomes a characteristic feature of the UA and AU linkages and independent of their location along the HA sequence, with the exception of the first and last terminal linkages (data not shown), as it will be discussed below. It is also worth mentioning that MD-ES simulation in the explicit solvent did not display any significant additional barrier transition and the MD-ES trajectory samples simply shown deeper energy of each glycosydic linkage (data not shown). Summing up, the accessibility of the local minima within the region of the highest population ($\phi \simeq 50°/\psi \simeq 30°$ in UA and $\phi \simeq 20°/\psi \simeq -30°$ in AU) explains the broadening of the end-to-end distribution up to 600 K.

The conformational maps displayed in Figure 4 are comparable with recent results reported in the literature[70] as for the location of the deepest minima in both linkages. The most significant difference, due to the different force field, is related to the inaccessibility of far minima in the present model.

The detailed profile of the potential energy landscape in the conformational ϕ/ψ maps is obviously modulated by intra- and interresidue hydrogen bonds (HBs). On the other hand, the relevance of HBs (in terms of percentual formation) is affected by temperature when other conformational states lacking of HBs are accessible. To show the actual temperature gradient of HBs stability, the percentage of most relevant HBs in each dimer units is reported as a function of temperature for the S1 model in Figure 5. These HBs are of the type $O3(U_i)-HO3\cdots O5(A_{i-1})$ bridging across the AU linkage and $O4(A_i)-HO4\cdots O5(U_i)$ bridging across the UA linkage. Two different trends are immediatly recognized; the population of HBs across AU residues is strongly decreased by increasing temperature, while the percentage of HBs that stiffen the UA junctions seems only slightly affected. However, an exception occurs for the terminal A reducing end which shows the opposite trend, that is the absolute minimum does not coincide with the formation of HB for the terminal U(4)−A(4) junction. Some minor difference also exists for the nonreducing terminal U(1)−A(1) junction, although limited to the lower temperatures. All populations seem to converge at a value of 27 at very high temperature. This value should correspond to the fraction of HB

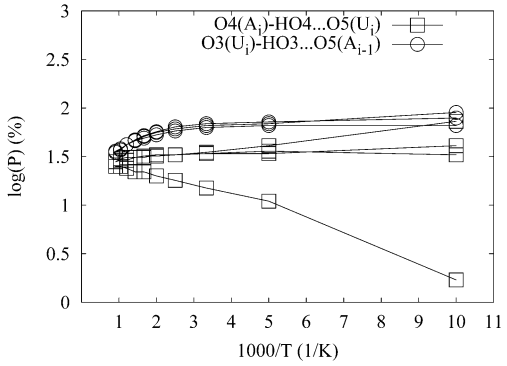

Figure 5. Populations for $O3(U_i)-HO3\cdots O5(A_{i-1})$ (○) and $O4(A_i)-HO4\cdots O5(U_i)$ hydrogen bonds (□) obtained from replica-exchange MC simulation with the implicit short-range model S1 for different temperatures T.

Table 1. Population of HB (%) at 300 K Where Only Populations Larger than 10% Were Considered

HB pair	S0	S1	S2	MD-ES
O2(U1)−HO2(U1)··O22(A1)	15			
O6(A1)−HO6(A1)··O3(U2)	14			
O6(A2)−HO6(A2)··O3(U3)	17			
O6(A3)−HO6(A3)··O3(U4)	16			
O4(A1)−HO4(A1)··O5(U1)	64	35	16	24
O3(U2)−HO3(U2)··O5(A1)	80	66	32	62
O4(A2)−HO4(A2)··O5(U2)	57	34	50	23
O3(U3)−HO3(U3)··O5(A2)	82	69	52	67
O4(A3)−HO4(A3)··O5(U3)	57	34	48	26
O3(U4)−HO3(U4)··O5(A3)	81	63	80	63
O4(A4)−HO4(A4)··O5(U4)	24	15	42	18

coverage in the accessible conformational space, independent of the residue location.

The lower flexibility of the AU linkage with respect to the UA linkage at $T \sim 300$ K is related to the most frequent HBs reported in Figure 5. The stronger HBs across the AU glycosidic linkages determine the extent of thermal fluctuations for exploring the region of the global minimum and the shape of the conformational distribution in the ϕ/ψ plane (Figure 4). The different stability of HBs across UA and AU linkages were already pointed out in other MD simulations in the explicit water,[19] while strong HBs like $O2-HO2(U_i)\cdots O22(A_i)$, observed elsewhere, have here a population smaller than 10% both in MD-ES and in REMC-S1.

The structure of the HB network is sensitive to the solvation environment. In Table 1, the HBs with a population larger than 10% at $T = 300$ K are reported for REMC simulations in the vacuum, in the S1 implicit solvent model and for the MD-ES simulation. The populations are similar in the two solvated models and are significantly different from the simulation in the vacuum. All the HBs involving O6 of A residues were destroyed by the solvent, due to the larger propensity of the group to be solvated by hydrogen bonded water molecules. On the other hand, the O5 atoms are better screened by HO3/HO4 of the next residue in the absence of solvent, both implicit and explicit.

It is now possible to understand the differences between the end-to-end distance distribution obtained by the different models (Figure 6). The main effect of the above-described structural features obtained through the S1 model allows an estimate of the molecular size consistent with the more computationally expensive MD-ES model: the distribution has in fact the maximum located at 3.5 nm instead of 3.6 nm in the MD-ES model and both distributions have similar width, while

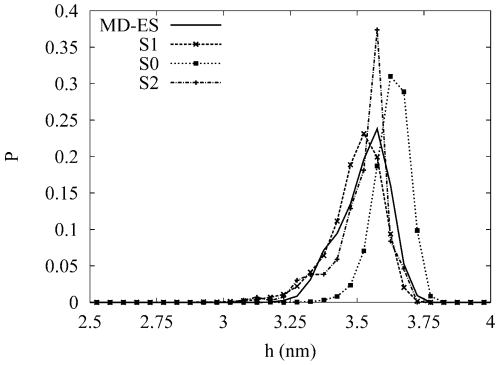

Figure 6. End-to-end distance distribution: obtained from MD in the explicit solvent (MD-ES, solid line); from replica-exchange MC (REMC) with the implicit short-range solvent model (S1, long dashes and ×); from replica-exchange MC simulation with S1 solvent and using the electrostatic solvation contribution as perturbation (S2, dotted line and +); MC simulation in the vacuum (S0, short dashes and ■).

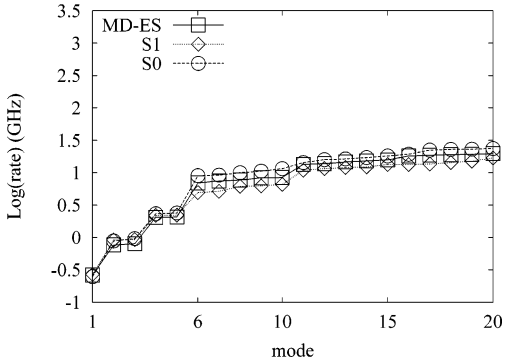

Figure 7. Second-rank/second-order relaxation rates obtained by mode-coupling diffusion theory combined with different statistics: MD simulation in the explicit solvent (MD-ES, □); replica-exchange MC with the implicit short-range solvent model (S1, ◇); MC in the vacuum (S0, ○). Only the first 20 modes are displayed.

the S0 distribution is significantly different both in shape and in the location of the maximum (3.7 nm). Seeing the effect of introducing the electrostatic term as obtained by solving the Poisson equation for the molecule in a continuum description of pure bulk water (model S2), we remark on a slight increase of the location of the maximum of the end-to-end distance population with respect to the S1 model, which, almost exactly, matches the location of the maximum obtained with the MD-ES model. Nevertheless, the reweighting in terms of the Boltzmann factor in eq 5 needs a very significant statistics in order to weight the more relevant configurations sufficiently, and therefore, the displayed wings in the end-to-end distance distribution are due to statistical errors.

The change of shape of the molecule by using the solvation free energy perturbation is accompanied by a change in the HB network, especially in the populations of HBs across the glycosydic linkages (see Table 1). After the application of the perturbation of the S1 statistics, the HBs change significantly, in some cases (A(3)−U(4)) becoming as strong as in the statistics in the vacuum (S0). Although the statistical errors in eq 5 do not allow reliable quantitative estimates, the main effect of the electrostatic contribution is to select configurations similar to those selected in the S0 statistics.

As far as the configuration statistics are concerned, the S1 implicit model provides the same results of the MD-ES model, the major difference being a significantly wider exploration at the 300 K of the conformational landscape within the global minima of the glycosydic linkages. Both solvated models predict a molecular size significantly smaller than in the vacuum: this property is expected to be important in determining the global molecular dynamics and is strongly affected by the presence of intramolecular HBs. Moreover the REMC statistics indicate that the conformational transitions are accessed only at very high temperatures, given the CHARMM force field and the implicit S1 solvent model.

B. Comparison with NMR Relaxation Experiments. In Figure 7, the second-rank second-order rates obtained by MCD from the three S0, MD-ES, and S1 statistics are compared. The five lowest relaxation rates may be associated with an "average" description of the rotational motion of the fluctuating molecule. Note that the MD-ES and S1 models give almost the same rotational rates: this means that the average diffusion tensor eigenvalues are similar for the two models. This is particularly important because, as pointed out in the previous paper,[14] the NMR relaxivities are strongly affected by the anisotropy of the average diffusion tensor. The lowest rate in the S0 model is only slightly lower than that of the MD-ES and S1 models, thus reflecting the fact that in the vacuum the former model produces a slight expansion of the oligomer: the location of the maximum in the end-to-end distance distributions (see Figure 6) is in fact 3.7 nm in vacuo compared to 3.5−3.6 in the S1 and MD-ES models, respectively. The rates from 2 to 5 are significantly lower in the MD-ES model than in the S0 and S1 models, conseguently suggesting that the anisotropy of the average diffusion tensor is slightly lower in the explicit solvent MD model.

The most significant difference in the relaxation rates' spectrum can be observed for the higher "internal" rates, where the MD-ES curve is in the middle between the higher S0 and the lower S1 curves. Even if the higher rates do not directly affect the NMR relaxivities at the experimental NMR frequency $\nu(^{13}C) = 75$ MHz,[14] the different structure of the related modes are important, because these relaxation modes affect the amplitudes of the five lowest-rank relaxation modes.

The separation between the rotational tumbling of the average molecule and the wobbling of C−H bonds about the average molecule can be quantified observing Figure 7. The five lower rates differ from the first internal rate (associated with the sixth mode) by about 1 order of magnitude, and therefore, there is a good time scale separation between rotational and internal modes. On this basis, the analysis of the second-rank dynamics of the C−H bonds in the S1 model in terms of the two decaying exponential description of eq 13 is useful. For sake of simplicity the average anisotropic rotational decay is estimated by a single time constant.

We report in Table 2 the S^2, R, T_R, and T_F parameters as obtained for all the probed ^{13}C nuclei in the terminal (1 and 4) and central (2 and 3) residues of type U and A, respectively. The parameters of equivalent carbons in U (or A) in the terminal (and in the central) residues differ very little from each other and in the table we report simply the average value. The S^2 parameter is significantly lower for terminal residues (0.46) than for central residues (0.75), with the exception of C6 in A2 and A3 (0.47). The C6−H6 bonds in residues A are more free to rotate with respect to the attached pyranose ring and the lower order parameter is expected. However the

Table 2. Parameters Describing the Second-Rank Time Correlation Function of C−H Bonds Obtained by Mode-Coupling Diffusion Calculations[a]

residue	atom	S^2	R	T_R (ns)	T_F (ps)
U1,4	C4	0.46	0.93	1.17	95
	C3	0.44	0.94	1.47	93
	C2	0.45	0.94	1.51	92
	C1	0.46	0.94	1.29	94
	C5	0.45	0.94	1.34	93
A1,4	C4	0.46	0.94	1.53	98
	C3	0.49	0.94	1.05	92
	C2	0.46	0.94	1.37	97
	C5	0.46	0.94	1.54	98
	C1	0.45	0.95	1.74	99
	C6	0.40	0.98	1.99	52
U2,3	C4	0.75	0.98	1.00	77
	C3	0.74	0.98	1.23	77
	C2	0.74	0.98	1.27	77
	C1	0.75	0.98	1.10	77
	C5	0.74	0.98	1.11	77
A2,3	C4	0.76	0.98	1.34	112
	C3	0.77	0.97	0.97	119
	C2	0.76	0.98	1.24	116
	C1	0.77	0.98	1.44	112
	C5	0.76	0.98	1.34	113
	C6	0.47	0.99	2.08	86

[a] See eqs 14, 19, 15, and 17 in text for definitions. The values are averages over the two indicated equivalent residues.

rotation of C6−H6 around the C5−C6 bond is hindered by the interactions of O6 with other groups and with the solvent. Therefore, a quantification of the degree of mobility of C6−H6 bonds in A is important because the C6 nucleus can be a probe of both the intramolecular potential and the molecular hydration (see the discussion below). The main chain C−H bonds have almost the same order parameter within each residue and the order is sensitive to the position of the residue (terminal or central).

The R values in Table 2 show that 94−98% of the integral of each TCF is given by the rotational tumbling. The integral depends on the long-time behavior of the TCF and, therefore, on the lowest rates and on the structure of the lowest-rate modes. The contribution of structure and rate of high-rate modes (modes with index larger than five) appear in the NMR relaxivities where the percentage of relaxivity given by rotational tumbling may decrease to 90% and 75% in the central and terminal residues, respectively (data not shown).

It must be noticed that in this molecule the variation of T_R and T_F in different residues is significant. In Table 2, T_R for bonds in the terminal residues are, on average, about 1.4 ns, while in the central residues they are about 1.2 ns: T_R increases in the terminal residues and, in particular, in A(1) and A(4). This behavior is due to the average orientation of terminal residues with respect to the axes of the diffusion tensor and it is, therefore, related to the order parameters. The details of the effects of average orientation are in the five different amplitudes of the five lowest-rate relaxation modes and they are beyond the summary given by the single order parameter S^2. The effects of orientation and of molecular diffusion anisotropy are also in the behavior of T_F that is significantly larger in A(2) and A(3) with respect to central U and terminal residues. To have a complete description of NMR relaxivities, all the details provided by the full spectrum of relaxation modes given by the MCD approach will be used in the following discussion.

In Figure 8, the TCF correlation times are plotted for

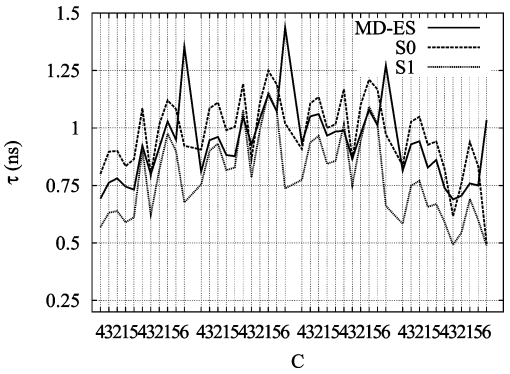

Figure 8. Correlation times obtained by mode-coupling diffusion theory combined with different statistics: MD in the explicit solvent (MD-ES, solid line); replica-exchange MC with the implicit short-range solvent model (S1, dotted line); MC in the vacuum (S0, dashed line).

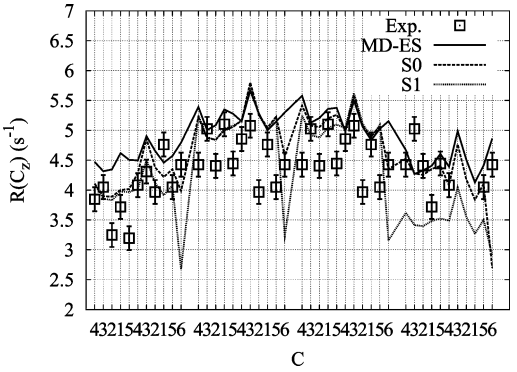

Figure 9. NMR $R(C_z)$ relaxivities at $\nu(^{13}C)$ = 75 MHz, experimental results (squares) and calculated results obtained by combining mode-coupling diffusion theory with different statistics: MD in the explici solvent (MD-ES, solid line); replica-exchange MC with the implicit short-range solvent model (S1, dotted line); MC in the vacuum (S0, dashed line).

each of the NMR monitored C−H bonds in the oligomer. It is clearly shown that the MC simulation in the vacuum gives almost the same pattern of the MD-ES simulation and the plot is only shifted toward higher values. A slower dynamics for the S0 model is expected because of the lower value for the lowest relaxation rate (see Figure 7) and because of the larger molecular size. On the other hand, the almost identical pattern of the correlation times means that the relaxation modes have a similar structure. Concerning the S1 model, the correlation times pattern is very close to the MD-ES model in the U residues, while it is significantly different in the A residues and particularly as for C6. This means that the time scale of the relaxation process is captured by the almost identical five lowest relaxation rates, but the different structure of the modes, as revealed by the relaxation spectrum shown in Figure 7, has significant local contributions in the A residues.

This effect is even more pronounced in the comparison of ^{13}C NMR relaxivities measured at 75 MHz. In Figure 9, the three models are compared with the experimental values for the same oligomer in water solution.[12] The S0 results are here very close to the MD-ES data and to most of the experimental values, while the S1 results are different from the other calculated data sets in the last UA pair in the chain and in the C6 nuclei in all the A residues.

Let us point out that the $R(C_z)$ parameters are very sensitive to the structure of the five lowest relaxation

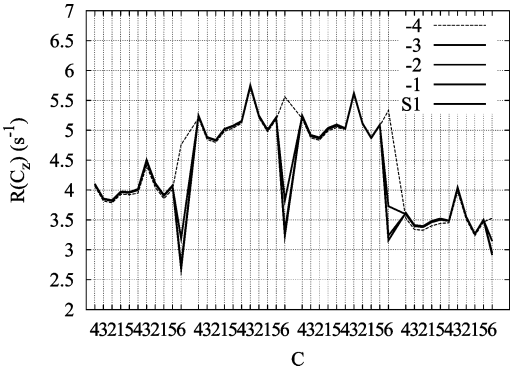

Figure 10. Same as Figure 9 for the replica-exchange MC simulation with the S1 solvent model when the configurations with 1, 2, and 3 (thick lines) and 4 (thin line) ω in the range $240° < \omega < 360°$ are discarded from the statistics.

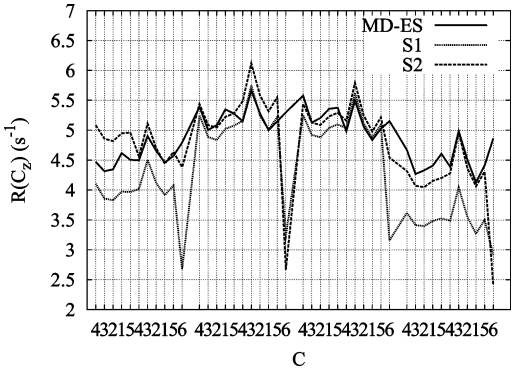

Figure 11. Same as Figure 9: MD in the explicit solvent (MD-ES, solid line); replica-exchange MC with the implcit short-range solvent model (S1, dotted line); replica-exchange MC with S1 solvent model and the electrostatic solvation contribution as perturbation (S2, dashed line).

modes and that these latter ones are sensitive to the molecular configuration statistics. In the case of the C6 nuclei the differences in relaxivities are due to the differences in the distribution of the O5−C5−C6−O6 (ω) dihedral angle in the A residues within the different models. As explained above, this dihedral angle affects the orientation of C6−H6 bonds with respect to the oligomer backbone. The correct description of the population assumed by this angle is particularly relevant for the assessment of the conformational topologies of polysaccharides contening 1 → 6 linkages.[71,72] In the S0 model, at $T = 300$ K, two maxima are displayed in the distribution (data not shown) at 60 and 300°, with the second peak much lower than the first one in population. In the MD-ES model, three of the dihedral angles are at 60° and the A(4) angle is at 300° for the entire trajectory. Therefore, the average ω distribution appears similar to the S0 distribution, but it reflects a dynamically frozen conformational landscape. In the S1 model at 300 K, two maxima are again populated, at 60 and 300°, by all the addressed dihedral angles, but with a higher ratio between the second and first peaks in terms of population. However, this ratio is still lower than one. The effect of population at 300° on the NMR relaxivities can be monitored by deleting progressively from the S1 statistic configurations with 4, 3, 2 and 1, respectively, of O5−C5−C6−O6 dihedral angles in the region $240° < \phi < 360°$. In Figure 10 it is shown that the effect of entirely canceling these torsional states from the statistics is to increase the C6 relaxivity up to the values of the MD-ES statistics, without changing the other values. This example demonstrates that this torsional state affects the relaxation modes only locally, without changing the amplitudes of relaxation for bonds different from C6−H6.

The presence of significant population corresponding to 300° in this dihedral angle is due to the high hydrophilic nature of O6 in the adopted S1 model: a trial calculation with $\sigma(O6) = 0$ in eq 21 (data not shown), does not contribute to population in this region for any of the O5−C5−C6−O6 dihedral angles in the molecule, thus suggesting that the O6 hydrophilic nature was overestimated in the REMC-S1 model and that the O6 hydroxyl group is not as hydrophilic as the other hydroxyl groups in the molecule are. The definition of correct parameters for the torsion deforming the ω dihedral angle are object of intensive theoretical studies and the calculation of $R(C_z)$ relaxivities to be compared with NMR experiments seems to be a good tool for assessing this crucial component of any force field of general use in the field of polysaccharides.

A second remarkable difference between the S1/S0 and the MD-ES models is in the relaxivities of end residues (U(1)−A(1) and U(4)−A(4)), both being smaller in the S1 than in MD-ES model, and smaller in the last residues in the S1 compared to experiments (see Figure 9). This difference is expected to be related to a change in shape of the molecule in the terminal regions and, therefore, to a change in the electrostatic contribution to the solvation free energy that was not taken into account in the S1 model. This contribution can be partially included in the statistics by the reweighting scheme of eq 5. In Figure 11, the NMR relaxivities obtained by using averages obtained by eq 5 are compared with the S1 and MD-ES model. The effect of electrostatic reweighting is the correction of the S1 relaxivities toward the MD-ES, mostly in the end residues, without significant modifications of the pattern within each residue and without modifying the C6 relaxivity that, as observed above, is mainly caused by a local hydrophilic interaction with the effective solvent.

As shown in Figure 6, the effect of the electrostatic contribution to the solvation free energy is to slightly increase the molecular size toward the value obtained in the explicit solvent. This effect is mainly due to changes in the structure of HBs in the first and last linkages affecting the orientation of the end residues with respect to the four central residues (see Table 1). This statistical modification is also observed in the pattern of NMR relaxivities.

IV. Conclusions

A few important issues were explored in this paper devoted to optimize statistical methods for hyaluronan and to improve the molecular description of conformational and dynamical properties of HA. The outcomes are summarized in the following points which can define new strategies in HA further studies.

1. An insight of the statistics of the $(UA)_4$ oligomer of HA has been calculated through REMC simulations once the geometrical parameters were chosen according to thermal averages obtained by MD simulations in the vacuum at $T = 300$ K. The oligomer was at first simulated in the vacuum (S0 model). An implicit model for the nonpolar contribution to the free energy of water solvation was then included directly in the REMC simulation (S1 model) and the polar (electrostatic) contribution was taken into account through a reweighting approach (S2 model). The configuration sta-

tistics were compared with a previous MD simulation in the explicit solvent (MD-ES model). Interestingly, the statistics of the S1 model well compares with MD-ES, providing the additional information on a wider exploration of the energy landscape in terms of the glycosydic dihedral angles ϕ and ψ at different temperatures with a single highly efficient parallel calculation. Other important dihedral angles, like the ω in the A residues, efficiently explore the ranges of variation.

2. Still more important, the calculation shows that essential interactions can be extracted by the overwhelming information contained even in a single MD trajectory in the explicit solvent by using statistical tools that avoid possible errors due to dynamical trapping in energy minima and by adopting a few assumptions on the solvent model. Althought the statistical tool was widely recognized in the REMC method and in related generalized-ensemble statistics, the use of simplified solvent models based on a continuum description is still matter of debate. Nevertheless, the present calculations allow the identification of important hints for polysaccharides. The solvation nonelectrostatic free energy $\Delta G_{\mathrm{nopol}}$ approximation of eq 21 gives a good description of configurational statistics compared to the computationally expensive use of an explicit solvent and it can be easily improved being the REMC algorithm easily scalable on parallel computers. The inclusion of the ΔG_{pol} electrostatic contribution to the solvation free energy provides an unbiased correction that needs, however, large statistics to obtain quantitative results. A more rapid convergence can be achieved by including a first estimate of ΔG_{pol} by using generalized Born radii in the simulation and using as reweigthing the energy correction provided by the solution of the Poisson equation.

3. The mode-coupling diffusion theory was used to calculate the molecular dynamics from statistical averages obtained by the S1 and S2 statistics. The S1 model was found only slightly weaker, in predicting the molecular dynamics, than MD-ES statistical averages. However, the reasons of model weakness were easily identified in the high hydrophilic nature assigned to the O6 hydroxyl groups in the A residues and in the behavior of end residues that can be partially corrected by the inclusion of ΔG_{pol} in the model. The MCD derived molecular dynamics was then interpreted in terms of a few essential features: the average NMR relaxivity is affected by the molecular size that was well predicted by the S1 model and in the previous MD-ES simulation; the pattern of NMR relaxivities is strongly influenced by the dihedral angle distributions, the dihedral O5−C5−C6−O6 distribution being a significative example; the behavior of end residues is affected by the electrostatic contribution to the solvation free energy. The orientational dynamics of the C−H bonds in the four central residues (U(2)−A(2) and U(3)−A(3)), except for C6−H6 in A, are not very sensitive to the details of interactions included in the MD-ES model and are almost identical in the S1 and MD-ES models.

4. The mode-coupling diffusion theory is presently the more efficient tool for describing at the molecular level the dynamics of short polymers in solution by combining advances in computer simulations of molecular statistics with molecular dynamics. This combination allows the interpretation of a large amount of experiments in the field of NMR. Far from a quantitative prediction of NMR relaxivities, the REMC or other equivalent generalized-ensemble simulation approaches that are able to use simplified molecular models for HA oligomers, are valuable tools in order to understand the essential interactions in modeling HA polymers.

5. The results concerning the oligomer $(UA)_4$ local features are expected to appear by-en-large also in the polymer HA within the proper macromolecular perturbation of long-range interactions. The $(UA)_4$ oligomer has been as a suitable model to test methodologies, but it would be only speculative predict HA chain conformation at this stage. In particular, aiming the polyelectrolite nature due to the negative charge carried by the U carboxyl group (here not yet included in any model) can be tackled within this approach. In the traditional MD-ES simulations the inclusion of ionic force must be done explicitly by adding several counterions, while in this REMC perspective this effect can be included in the Poisson−Boltzmann equation for oligomers and in its generalized Born approximations for models of the HA polymer.

Acknowledgment. This work has been sponsored by INSTM (Florence, Italy). S.F. is recipient of a scholarship of the University of Trieste.

Appendix: The Mode-Coupling Diffusion Theory

For the reader utility, the mode-coupling diffusion approach to describe the dynamics in polymer solutions,[66,73] is here briefly summarized, sticking to the formulation of paper 2 in this series.[14]

Given a polymer of N beads of friction coefficients ζ_i and coordinates $\boldsymbol{r}_i$, connected by $N-1$ bonds ($\boldsymbol{l}_i$, $i = 1, ..., N-1$), the dynamics of the macromolecule described by the variables $\mathbf{l}_i$, is governed by the operator L, adjoint to the diffusion Smoluchowski operator D:

$$\frac{\partial \boldsymbol{l}}{\partial t} = L\boldsymbol{l}$$

$$L = \sum_{i,j=1}^{N} [\nabla_i \cdot \boldsymbol{D}_{i,j} \cdot \nabla_j - (\nabla_i \mathrm{U}/k_B T) \cdot \boldsymbol{D}_{i,j} \cdot \nabla_j] \qquad \text{(A1)}$$

where $\mathbf{l}$ is the $3 \times N - 1$ dimension vector containing all the bonds, $\mathbf{D}$ is the diffusion tensor, U the bead potential energy, k_B the Boltzmann constant and T the temperature. By expanding the conditional probability (the solution to the Smoluchowski equation) in a complete set of eigenfunctions of L, the time autocorrelation function (TCF) of any coordinate-dependent dynamic variable with zero average $f(t)$, may be expressed in the standard form

$$\langle f(t)\, f(0)\rangle = \sum_i \langle f\psi_i\rangle\langle\psi_i f\rangle \exp(-\lambda_i t) \qquad \text{(A2)}$$

where $-\lambda_i$ and ψ_i are the eigenvalues and the normalized eigenfunctions of the operator L. By representing ψ_i in a set of basis functions $\Phi = \{\phi_m, m = 1, ..., M\}$,

$$\psi_i = \sum_{m=1}^{M} C_{m,i}\phi_m \qquad \text{(A3)}$$

the diffusion eigenvalue equation may be written in matrix form

$$\mathbf{FC} = \mathbf{SC\Lambda} \qquad \text{(A4)}$$

with $\mathbf{\Lambda}$ the diagonal matrix with the eigenvalues λ_m as the diagonal elements, $\boldsymbol{C}$ the eigenvector matrix of coefficients $C_{i,m}$, $\boldsymbol{S}$ the metric matrix

$$S_{ij} = \langle \phi_i \phi_j \rangle \tag{A5}$$

and $\boldsymbol{F}$ the equilibrium force matrix

$$F_{ij} = -\langle \phi_i L \phi_j \rangle = \sum_{m,n=1}^{N} \langle (\nabla_m \phi_i) \mathbf{D}_{mn} (\nabla_n \phi_j) \rangle \tag{A6}$$

where

$$\mathbf{D}_{ij} = D_i \mathbf{H}_{ij}$$

$$\mathbf{H}_{ij} = \mathbf{1}\delta_{ij} + \alpha \zeta_i \boldsymbol{T}_{ij}(1 - \delta_{ij})$$

$$\boldsymbol{T}_{ij} = (8\pi\eta r_{ij})^{-1}[\mathbf{1} + \boldsymbol{r}_{ij}\boldsymbol{r}_{ij}/{r_{ij}}^2] \tag{A7}$$

with $\mathbf{H}$ and $\mathbf{T}$ representing the hydrodynamic interaction matrix and Oseen tensor, respectively.

$$D_i = k_B T/\zeta_i \tag{A8}$$

is the diffusion coefficient of each bead and η the solvent viscosity. The factor α in the hydrodynamic interaction of eq A7 is a parameter whose values are often restricted to the range $\alpha < 1$ so as to maintain a positive definite $\mathbf{H}$, while $\alpha = 1$ corresponds to the full interaction strength related to the specific choice of friction coefficients. The ensemble equilibrium averages, indicated within brackets in the above equations, are calculated as

$$\langle ab \rangle = \int_V P_{\mathrm{eq}}(\boldsymbol{r})\, a(\boldsymbol{r})\, b(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} \tag{A9}$$

with $P_{\mathrm{eq}}(\boldsymbol{r})$ the equilibrium distribution function depending on the atomistic potential energy U of eq 3 used in the simulation. As the eigenfunctions ψ_i are orthonormalized, the coefficient matrix $\mathbf{C}$ should satisfy the normalization equation:

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{C} = \mathbf{1} \tag{A10}$$

Accurate approximations to the TCFs evaluation are obtained by taking basis functions in the basis set built with increasing power of the variables, bond vectors or their linear combinations, in the spirit of the mode-coupling theory.[66,73] First and second-order basis sets, which must be given in the proper irreducible tensorial form, are therefore, for second-rank functions, bilinear and tetralinear, respectively, in the bond variables.

References and Notes

(1) Rao, V. S. R.; Qasba, P. K.; Balaji, P. V.; Chandrasekaran, R. *Conformation of Carbohydrates*; Harwood Academic Publishers: Amsterdam, The Netherlands, 1998.
(2) Rees, D. *Polysaccharides Shapes;* Chapman & Hall: London, U.K., 1977.
(3) Brant, D. A. *Curr. Opin. Struct. Biol.* **1999**, *9*, 556−562.
(4) Burton, B. A.; Brant, D. A. *Biopolymers* **1983**, *22*, 1769−1792.
(5) Flory, P. J. *Statistical Mechanics of Chain Moleculs*; Hanser: New York, 1989.
(6) Imberty, A.; Pérez, S. *Chem. Rev.* **2000**, *100*, 4567−4588.
(7) Kennedy, J. F., Phillips, G. O., Williams, P. A., Hascall, V. C., Eds.; *Hyaluronan: Chemical, Biochemical and Biological Aspects*; Woodhead Pub. Ltd.: Cambridge, U.K., 2002.
(8) Almond, A.; Brass, A.; Sheehan, J. K. *Glycobiology* **1998**, *8*, 973−980.
(9) Almond, A.; Brass, A.; Sheehan, J. K. *J. Mol. Biol.* **1998**, *284*, 1425−1437.
(10) Scott, J. E.; Heatley, F. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4850−4855.
(11) Cavalieri, F.; Chiessi, E.; Paci, M.; Paradossi, G.; Flaibani, A.; Cesàro, A. *Macromolecules* **2001**, *34*, 99−109.
(12) Cowman, M. K.; Feder-Davis, J.; Hittner, D. M. *Macromolecules* **2001**, *34*, 110−115.
(13) Peng, J.; Wagner, G. *J. Magn. Reson.* **1992**, *98*, 308−332.
(14) Letardi, S.; La Penna, G.; Chiessi, E.; Perico, A.; Cesàro, A. *Macromolecules* **2002**, *35*, 286−300.
(15) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids;* Clarendon Press: Oxford, U.K., 1989.
(16) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: San Diego, CA, 1996.
(17) Holmbeck, S. M. A.; Petillo, P. A.; Lerner, L. E. *Biochemistry* **1994**, *33*, *3*, 14246−14255.
(18) Kaufmann, J.; Möhle, K.; Hofmann, H.-J.; Arnlod, K. *J. Mol. Struct.* **1998**, *422*, 109−121.
(19) Almond, A.; Brass, A.; Sheehan, J. K. *J. Phys. Chem. B* **2000**, *104*, 5634−5640.
(20) Donati, A.; Magnani, A.; Bonechi, C.; Barbucci, R.; Rossi, C. *Biopolymers* **2001**, *59*, 434−445.
(21) Tafi, A.; Manetti, F.; Corelli, F.; Alcaro, S.; Botta, M. *Pure Appl. Chem.* **2003**, *75*, 359−366.
(22) Stocker, U.; van Gunsteren, W. F. *Proteins* **2000**, *40*, 145−153.
(23) Peter, C.; Daura, X.; van Gunsteren, W. F. *J. Biomol. NMR* **2001**, *20*, 297−310.
(24) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4570.
(25) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Groenenborn, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 4989−4991.
(26) Prompers, J. J.; Brüschweiler, R. *J. Am. Chem. Soc.* **2002**, *124*, 4522−4534.
(27) La Penna, G.; Carbone, P.; Carpentiero, R.; Rapallo, A.; Perico, A. *J. Chem. Phys.* **2001**, *114*, 1876−1886.
(28) Perico, A.; Mormino, M.; Urbani, R.; Cesàro, A.; Tylianakis, E.; Dais, P.; Brant, D. A. *J. Phys. Chem. B* **1999**, *103*, 8162−8171.
(29) La Penna, G.; Fausti, S.; Perico, A.; Ferretti, J. A. *Biopolymers* **2000**, *54*, 89−103.
(30) La Penna, G.; Mitsutake, A.; Masuya, M.; Okamoto, Y. *Chem. Phys. Lett.* **2003**, *380*, 609−619.
(31) Fausti, S.; La Penna, G.; Paoletti, J.; Genest, D.; Lancelot, G.; Perico, A. *J. Biomol. NMR* **2001**, *27*, 333−349.
(32) La Penna, G.; Genest, D.; Perico, A. *Biopolymers* **2003**, *69*, 1−14.
(33) Prompers, J. J.; Scheurer, C.; Brüschweiler, R. *J. Mol. Biol.* **2001**, *305*, 1085−1097.
(34) Tugarinov, V.; Liang, Z.; Shapiro, Y. E.; Freed, J. H.; Meirovitch, E. *J. Am. Chem. Soc.* **2001**, *123*, 3055−3063.
(35) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers (Peptide Sci.)* **2001**, *60*, 96−123.
(36) Hansmann, U. H. E. *Comput. Phys. Comm.* **2002**, *147*, 604−607.
(37) Berg, B. A. *Comput. Phys. Comm.* **2003**, *153*, 397−406.
(38) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(39) Bartels, C.; Karplus, M. *J. Phys. Chem. B* **1999**, *110*, 8254−8282.
(40) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *J. Chem. Phys.* **2003**, *118*, 6676−6688.
(41) Honig, B. *J. Mol. Biol.* **1999**, *293*, 283−293.
(42) Onufriev, A.; Bashford, D.; Case, D. *J. Phys. Chem. B* **2000**, *104*, 3712−3720.
(43) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435−445.
(44) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(45) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *246*, 122−129.
(46) Qiu, D.; Shenkin, P.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005−3014.
(47) Ghosh, A.; Rapp, C. S.; Friesner, R. A. *J. Phys. Chem. B* **1998**, *102*, 10983−10990.
(48) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. *J. Comput. Chem.* **2002**, *23*, 517−529.
(49) Lee, M. S.; Salsbury, F. R.; Brooks, C. L., III. *J. Chem. Phys.* **2002**, *116*, 10606−10614.
(50) Ooi, T.; Oobatake, M.; Némethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086−3090.
(51) Mitsutake, A.; Kinoshita, M.; Okamoto, Y.; Hirata, F. *Chem. Phys. Lett.* **2000**, *329*, 295−303.
(52) Gavryushov, S.; Zielenkiewicz, P. *J. Phys. Chem. B* **1997**, *101*, 10903−10909.

(53) Borgis, D.; Lévy, N.; Marchi, M. *J. Chem. Phys.* **2003**, *119*, 3516−3528.
(54) Urbani, R.; Cesàro, A. *Polymer* **1991**, *32*, 3013−3020.
(55) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604−1608.
(56) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.
(57) Eisenmenger, F.; Hansmann, U. H. E.; Hayryan, S.; Hu, C.-K. *Comput. Phys. Commun.* **2001**, *138*, 192−212.
(58) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832−3846.
(59) Wolf, D.; Keblinski, P.; Phillpot, S. R.; Eggebrecht, J. *J. Chem. Phys.* **1999**, *110*, 8254−8282.
(60) Gobel, C. V.; Dimpfi, W. L.; Brant, D. A. *Macromolecules* **1970**, *3*, 644−654.
(61) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Skelton, N. J. *Protein NMR Spectroscopy;* Academic Press: San Diego, CA, 1996.
(62) Rose, M. E. *Elementary Theory of Angular Momentum*; John Wiley & Sons: New York, 1957.
(63) Pastor, R. W.; Karplus, M. *J. Phys. Chem.* **1988**, *92*, 2636−2641.
(64) Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*; Clarendon: Oxford, U.K., 1986.
(65) Zwanzig, R. *J. Chem. Phys.* **1974**, *60*, 2717−2720.
(66) La Penna, G.; Pratolongo, R.; Perico, A. *Macromolecules* **1999**, *32*, 506−513.
(67) Eisenberg, D.; McLachlan, A. D. *Nature (London)* **1986**, *319*, 199−203.
(68) Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. *J. Comput. Chem.* **2002**, *23*, 128−137.
(69) Eisenhaber, F.; Lijnzaad, P.; Argos, P.; Sander, C.; Scharf, M. *J. Comput. Chem.* **1995**, *16*, 273−284.
(70) Haxaire, K.; Braccini, I.; Milas, M.; Rinaudo, M.; Pérez, S. *Glycobiology* **2000**, *10*, 587−594.
(71) Kadkhodaei, M.; Wu, H.; Brant, D. A. *Biopolymers* **1991**, *31*, 1581−1592.
(72) Angioletti, C. Ph.D. Thesis; University of Trieste, Trieste, Italy, 2003.
(73) Perico, A.; Pratolongo, R. *Macromolecules* **1997**, *30*, 5958−5969.

MA049641V