Diffusion and Electrophoretic Mobility of Single-Stranded RNA from Molecular Dynamics Simulations

In-Chul Yeh and Gerhard Hummer
Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland

ABSTRACT Hydrodynamic properties of small single-stranded RNA homopolymers with three and six nucleotides in free solution are determined from molecular dynamics simulations in explicit solvent. We find that the electrophoretic mobility increases with increasing RNA length, consistent with experiment. Diffusion coefficients of RNA, corrected for finite-size effects and solvent viscosity, agree well with those estimated from experiments and hydrodynamic calculations. The diffusion coefficients and electrophoretic mobilities satisfy a Nernst-Einstein relation in which the effective charge of RNA is reduced by the charge of transiently bound counterions. Fluctuations in the counterion atmosphere are shown to enhance the diffusive spread of RNA molecules drifting along the direction of the external electric field. As a consequence, apparent diffusion coefficients measured by capillary zone electrophoresis can be significantly larger than the actual values at certain experimental conditions.

INTRODUCTION

Gel electrophoresis is widely used to separate large biological macromolecules (Cantor and Schimmel, 1980). To separate small nucleic acid molecules, capillary zone electrophoresis (CZE), with charged molecules moving in free solution, has proven to be useful (Righetti et al., 2002). Separation of double-stranded deoxyribonucleic acid (DNA) by CZE can be achieved for lengths up to ~170 basepairs, the regime in which the electrophoretic mobility increases with length before reaching a plateau (Stellwagen and Stellwagen, 2002). With the high electric fields generated in the constriction of a capillary, CZE can thus be applied in ultrafast bioanalytical techniques (Jacobson et al., 1998; Plenert and Shear, 2003; Stuart and Sweedler, 2003). CZE can not only be used to *separate* charged (bio)polymers, but also to *characterize* their hydrodynamic properties—in particular, the electrophoretic mobility, translational diffusion coefficient, and hydrodynamic radius (Nkodo et al., 2001; Stellwagen and Stellwagen, 2002; Stellwagen et al., 2001).

However, there has been some controversy about the equivalence of diffusion coefficients obtained by CZE and by nonelectrophoretic measurements (Muthukumar, 1997; Nkodo et al., 2001; Righetti et al., 2002; Slater et al., 2002; Stellwagen and Stellwagen, 2002; Stellwagen et al., 2001). Computer simulations can be useful in testing the basic principles of CZE and possibly resolving the controversy. In molecular dynamics (MD) simulations (Allen and Tildesley, 1987), the dynamical properties of single nucleic acids moving under applied fields can be studied in molecular detail, which is only beginning to become possible experimentally (Smith et al., 1999; Volkmuth and Austin, 1992). With recent advances in MD simulation techniques such as particle-mesh Ewald summation (Darden et al., 1993; Essmann et al., 1995) for the calculation of electrostatic interactions (Simonson, 2003), and the continued development of nucleic acids force fields (Cornell et al., 1995; Foloppe and MacKerell, 2000), it is now possible to perform stable MD simulations of highly charged nucleic acids (Cheatham and Kollman, 2000; Cheatham, et al., 1995; Norberg and Nilsson, 2002). Here we present the results of all atom MD simulations of small single-stranded ribonucleic acid (ssRNA) molecules at applied electric fields with explicit solvent and counterions. We calculate hydrodynamic properties such as diffusion coefficients and electrophoretic mobilities for both RNAs and counterions. Calculated diffusion coefficients are corrected for finite-size effects caused by long-ranged hydrodynamic interactions. The role of counterions on the electrophoretic mobility is carefully examined. We show in particular that fluctuations in the number of counterions bound to individual RNA molecules lead to an increase in the apparent diffusion coefficient in an electrophoretic measurement.

METHODS AND THEORY

Computer simulations

We performed MD simulations of tri- and hexanucleotide ssRNA homopolymers of polyadenylic acid (poly A) and polyuridylic acid (poly U) in explicit water. For the simulations, we used the program NAMD (Kale et al., 1999) with the AMBER 94 force field (Cornell et al., 1995). For water we used the TIP3P model (Jorgensen et al., 1983). Initial configurations of ssRNA were taken from the A-form duplex of poly A and poly U, as generated by the *nucgen* command of AMBER 6.0 (Pearlman et al., 1995), with the C3′-*endo* ribose conformation corresponding to that of stacked ssRNA (Nowakowski and Tinoco, 1999). Both 3′ and 5′ ends were capped by hydroxyl groups, resulting in net charges of $-2e$ and $-5e$ for RNAs with tri- and hexanucleotides, respectively. The RNA molecules were then solvated with 1738 and 1739 water molecules for hexanucleotide poly A

Submitted July 8, 2003, and accepted for publication September 26, 2003.
Address reprint requests to Gerhard Hummer, National Institutes of Health, Bldg. 5, Rm. 132, Bethesda, MD 20892-0520. Tel.: 301-402-6290; Fax: 301-496-0824; E-mail: gerhard.hummer@nih.gov.

0006-3495/04/02/681/09 $2.00

(A_6) and poly U (U_6), respectively, and 695 water molecules for trinucleotide homopolymers (A_3 and U_3). The tri- and hexanucleotide systems were neutralized by two and five potassium ions (K^+), respectively, for which the potential parameters of Åqvist (1990) were used.

Cubic simulation cells were used in all simulations. A constant temperature of 300 K and a pressure of 1 bar were maintained with the Berendsen thermostat (Berendsen et al., 1984) and Langevin piston barostat (Feller et al., 1995), respectively. The velocity-Verlet algorithm (Allen and Tildesley, 1987) with a single time step of 2 fs was used in the time integration. Structures obtained after 1 ns of simulation without electric field were used as starting structures for runs with electric field. Each simulation lasted for 20 ns or longer after an equilibration period of at least 100 ps, as summarized in Table 1, with a combined total production time of 860 ns. The size of the simulation cell and the orientational polarization induced by the electric field reached equilibrium values within ~10 ps. Bonds involving hydrogen atoms were constrained with the SHAKE algorithm (Ryckaert et al., 1977). Long-range electrostatic interactions were treated with the particle-mesh Ewald method (Darden et al., 1993; Essmann et al., 1995) under conducting boundary conditions (Allen and Tildesley, 1987). A grid width of 1 Å or less, an Ewald real-space screening coefficient of 0.31 $Å^{-1}$, and a real-space cutoff radius of 10 Å were used for particle-mesh Ewald calculations. The same cutoff was used for Lennard-Jones interactions. Simulations were performed with and without external electric fields applied along the z direction. For Ewald summation with conducting boundary conditions, the total electric field (E) is equal to the applied external electric field (E_0) (Neumann, 1983; Yeh and Berkowitz, 1999a). Therefore, we used E_0 as the field strength E in our analysis. In Appendix A, we test this explicitly for the hydrated RNA system.

To induce at least nanometer-scale electrophoretic motion on the ns-simulation timescales, high electric field strengths of $E_0 = 3$, 30, 40, and 50 mV/Å were used, exceeding those typically used in capillary electrophoresis experiments by several orders of magnitude (Nkodo et al., 2001; Stellwagen and Stellwagen, 2002). However, in recent microsecond-electrophoresis experiments with μm-separation paths, electric fields in the range of 100 kV/cm (1 mV/Å) were reported (Jacobson et al., 1998; Plenert and Shear, 2003), comparable to the smallest nonzero field used in our simulations, 3 mV/Å.

Diffusion coefficients, hydrodynamic radii, and electrophoretic mobilities

In simulations without electric field, we calculated self-diffusion coefficients (D) of RNA molecules and K^+ ions from the derivative of the mean-square displacement with respect to time,

$$D = \lim_{t\to\infty} \frac{\partial}{\partial t} \frac{\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle}{6}, \tag{1}$$

where $\mathbf{r}(t)$ is the position of the geometric center of a molecule at time t. D was estimated from the slope of a straight-line fit in a time window of 1–10 ps.

For the diffusion coefficients of the RNA molecules calculated by MD simulations, we expect substantial finite-size effects (Dünweg and Kremer, 1993). Hydrodynamic interactions are of long range (~$1/r$), leading to an effective coupling among RNA molecules, the solvent, and their periodic images. The finite-size effects on the diffusivity can be estimated by Ewald summation of the Oseen or Rotne-Prager mobility tensors (Beenakker, 1986). With the Oseen tensor summed over all periodic images (Dünweg and Kremer, 1993; Hummer and Yeh, 2003, unpublished), the system-size dependent *apparent* diffusion coefficient $D_{app}(L)$ is given by

$$D_{app}(L) = D_0 - \frac{k_B T \xi_{EW}}{6\pi\eta L}, \tag{2}$$

with L the box length, D_0 the diffusion coefficient, k_B Boltzmann's constant, T the absolute temperature, η the solvent viscosity, and $\xi_{EW} \approx 2.837297$ the self-term for a cubic lattice (Placzek et al., 1951). Because this expression was derived for the hydrodynamic self-interaction of a point perturbation, deviations from Eq. 2 may be expected for charged polymeric solutes. To account for deviations from the Oseen point-particle limit, we introduce an empirical parameter α into Eq. 2,

$$D_{app}(L) = D_0 - \frac{k_B T \xi_{EW}}{6\pi\eta L}\alpha. \tag{3}$$

We further assume that there are no finite-size effects for the solvent viscosity, in accordance with the results of MD simulations of TIP3P water (Hummer and Yeh, 2003, unpublished). Then, we can estimate D_0 from Eq. 3, where the coefficient α is obtained from simulations of one of the RNA molecules in systems of different size by fitting $D_{app}(L)$ to $1/L$ using Eq. 3.

To correct for the low viscosity of TIP3P water ($\eta_{TIP3P} = 3.1 \times 10^{-4}$ kg m^{-1} s^{-1} at 298 K, Hummer and Yeh, unpublished; Yeh and Hummer, 2002) compared to the measured water viscosity ($\eta_{H_2O} = 8.91 \times 10^{-4}$ kg m^{-1} s^{-1}), we also report scaled diffusion coefficients,

$$D_\eta = \frac{\eta_{TIP3P}}{\eta_{H_2O}} D_0. \tag{4}$$

In addition, we calculate hydrodynamic radii R_H from the Stokes-Einstein relation,

$$R_H = \frac{k_B T}{6\pi\eta_{TIP3P} D_0}. \tag{5}$$

From the simulations under electric field, we estimate electrophoretic mobilities of RNA and potassium ions. In the presence of an electric field (oriented along the z axis), the RNA and potassium ions drift in opposite directions parallel to the field. From the slope of the positions $z(t)$ with respect to time t, we can obtain drift velocities of RNA at different E-values. In the linear regime, the drift velocity, ν, is proportional to the electric field,

$$\nu = \mu E, \tag{6}$$

with the proportionality constant μ defining the electrophoretic mobility. The diffusion coefficient D and the electrophoretic mobility μ are related through the Nernst-Einstein relation,

$$D = \frac{k_B T}{Q_{eff}} \mu, \tag{7}$$

TABLE 1 List of simulation runs

Run	RNA	E (mV/Å)	Water molecules	K^+ ions	Duration (ns)
1–5	A_6	0, 3, 30, 40, 50	1738	5	56.1, 24.2, 20.9, 24,4, 20.9
6–10	U_6	0, 3, 30, 40, 50	1739	5	22.5, 23.7, 22.8, 24.4, 21.7
11–15	A_3	0, 3, 30, 40, 50	695	2	64.7, 20.9, 20.9, 22.9, 20.9
16–20	U_3	0, 3, 30, 40, 50	695	2	61.9, 20.9, 20.9, 22.9, 20.9
21	A_6	0	1738	0	21.0
22–26	A_3	0	1007, 1367, 1766, 2017, 2364	2	37.5, 26.6, 32.0, 25.6, 28.6
27–29	–	0	512, 1367, 2364	1	61.0, 38.0, 30.0

where Q_{eff} defines an effective charge of the drifting particle. This relation can be used to determine Q_{eff} from known μ and D.

Electrophoretic diffusion with fluctuating counterion atmosphere

The ionic cloud surrounding the charged RNA molecules is constantly fluctuating, with bound counterions reducing the effective charge of the RNA. Such charge fluctuations cause transient changes in RNA mobility based on the Nernst-Einstein relation, Eq. 7. In the presence of an external field, the trajectories of RNA molecules will thus spread along the field direction not only because of normal diffusion, but also because of fluctuations in the ion cloud that affect the electrophoretic mobility of individual molecules. As a consequence, the variance in the position $z(t)$ of molecules that started at $z(0) = 0$ at time $t = 0$ will be larger than what is expected from regular diffusion, i.e., $\text{var}[z(t)] \geq 2\,Dt$.

We can quantify this effect in a simple model that incorporates fluctuations of the effective charge of the polyelectrolyte drifting in the external electric field. In this model, a fluctuating charge of bound (counter)ions, $\delta q(t)$, is added to the charge q_0 of an RNA molecule, resulting in an effective charge of $q(t) = q_0 + \delta q(t)$ at time t. If we assume that the instantaneous drift velocity $\nu(t)$ is given by Eqs. 6 and 7, with $Q_{\text{eff}} = q(t)$,

$$\nu(t) = \frac{Dq(t)E}{k_{\text{B}}T}, \tag{8}$$

then the position $z_{\text{d}}(t)$ at time t because of electrophoretic drift is given by

$$z_{\text{d}}(t) = \int_0^t \nu(t')\,dt' = \frac{DE}{k_{\text{B}}T}\int_0^t q(t')\,dt'. \tag{9}$$

If we assume further that the charge $q(t)$ relaxes exponentially about an average charge $\langle q\rangle$ with a variance $\sigma_{\text{q}}^2 = \langle q^2(t)\rangle - \langle q(t)\rangle^2$ and a relaxation time τ_{q}, then the variance of $z_{\text{d}}(t)$ is given by

$$\text{var}[z_{\text{d}}(t)] = \frac{2D^2E^2\sigma_{\text{q}}^2\tau_{\text{q}}^2}{(k_{\text{B}}T)^2}\left(\frac{t}{\tau_{\text{q}}} + e^{-t/\tau_{\text{q}}} - 1\right), \tag{10}$$

as shown in Appendix B. To obtain the variance in the position $z(t)$ at time t, we thus add the variance of $z_{\text{d}}(t)$ to the normal diffusive spread, $2Dt$:

$$\text{var}[z(t)] = 2Dt + \frac{2D^2E^2\sigma_{\text{q}}^2\tau_{\text{q}}^2}{(k_{\text{B}}T)^2}\left(\frac{t}{\tau_{\text{q}}} + e^{-t/\tau_{\text{q}}} - 1\right). \tag{11}$$

At long times, $t\rightarrow\infty$, the apparent diffusion coefficient deduced from the spread of RNA along the direction of the electric field is thus

$$D_{\text{app}} = \lim_{t\rightarrow\infty}\frac{\partial}{\partial t}\frac{\text{var}[z(t)]}{2} = D + \frac{D^2E^2\sigma_{\text{q}}^2\tau_{\text{q}}}{(k_{\text{B}}T)^2}. \tag{12}$$

Note that both σ_{q}^2 and τ_{q} may depend on the strength of the electric field, as well as the type, concentration, and size of the polyelectrolyte, counterions, and excess salt. A related model with two discrete states has been studied in the context of single-molecule fluorescence experiments (Berezhkovskii et al., 1999; Geva and Skinner, 1998). Such a discrete model should prove useful to describe the electrophoretic motion of the counterions, with fluctuations between polyelectrolyte-bound and free states.

Other fluctuations in the mobility of the charged molecule moving under the influence of the external electric field may further enhance the diffusive spread along the field direction. We expect an effect similar to that of the fluctuating ion atmosphere from dynamic changes in the structure of the drifting polyelectrolyte. Such conformational fluctuations change the friction coefficient (Pastor and Karplus, 1988) and thus the mobility.

RESULTS AND DISCUSSION

Structural properties of RNA

Base stacking is one of the driving forces for nucleic acid folding and stability and a key factor in nucleic acid secondary structure formation (Norberg and Nilsson, 2002). Our simulations started with fully stacked ssRNA configurations. To probe structural properties of RNA during simulation runs, we calculated the number of base stackings between pairs of bases adjacent in the sequence as well as the radius of gyration for each RNA (Table 2). Two bases are considered stacked if they have >25 (adenine) or 15 (uracil) distinct contacts (<5 Å distance) between nonhydrogen atoms. Both hexa- and trinucleotides of poly A with larger purine bases show enhanced base stacking compared to poly U of corresponding length with smaller pyrimidine bases. This is consistent with the results of previous experimental and simulation studies of base stacking in single-stranded nucleic acids (Norberg and Nilsson, 2002; Nowakowski and Tinoco, 1999). The average numbers of stackings between adjacent bases are lower than the values of 5 and 2 expected for fully-stacked hexa- and trinucleotide RNAs, respectively. Note, however, that simulations of 20 to 50 ns are not long enough to sample the configurational space of the single-stranded RNA molecules completely. This is evident in considerable variations of the stacking numbers in runs of the same molecule, without a systematic dependence on the electric field (Table 2).

Diffusion coefficients of RNA and potassium ions

Self-diffusion coefficients D of RNA molecules and K^+ ions calculated from the simulations using Eq. 1 are compiled in Table 3. To correct for finite-size effects, we calculated diffusion coefficients D of RNA molecules from MD simulations of A_3 RNA with 695, 1007, 1367, 1766, 2017, and 2364 water molecules (simulation runs 11 and 22–26 in Table 1). Fig. 1 shows the resulting diffusion coefficients of A_3 RNA as a function of the reciprocal of the box length, $1/L$. Equation 3 provides an excellent fit to the observed system-size dependence with a correction factor $\alpha \approx 0.76$, close to the ideal value of 1. We used Eq. 3 with this α-value

TABLE 2 Structural properties of RNA

	A_6		U_6		A_3		U_3	
RNA	N_{B}	R_{g}	N_{B}	R_{g}	N_{B}	R_{g}	N_{B}	R_{g}
$E = 0$ mV/Å	3.0	7.14	1.8	8.53	1.2	5.20	0.7	5.13
3	3.1	7.08	0.6	6.91	1.6	5.26	0.5	5.44
30	3.2	6.80	3.2	7.82	1.8	5.19	0.5	4.97
40	3.3	6.88	0.7	6.55	1.8	5.13	0.9	5.44
50	4.0	7.23	1.9	6.75	1.9	5.22	0.0	5.16

N_{B}, average number of base stackings between pairs of bases adjacent in the sequence during the initial 20-ns time period. R_{g}, radius of gyration in units of Å.

TABLE 3 Calculated diffusion coefficients in units of 10^{-5} cm^2 s^{-1}

	A_6	U_6	A_3	U_3
		RNA		
$D_{app}(L)$	0.291 (0.003)	0.287 (0.006)	0.395 (0.006)	0.423 (0.005)
D_0	0.695 (0.017)	0.691 (0.017)	0.940 (0.018)	0.970 (0.023)
D_η	0.242 (0.006)	0.240 (0.006)	0.327 (0.006)	0.337 (0.008)
		K^+		
$D_{app}(L)$	3.14 (0.03)	3.21 (0.02)	3.04 (0.02)	3.07 (0.03)

Results are listed for uncorrected diffusion coefficients, $D_{app}(L)$; diffusion coefficients corrected for finite-size effects, D_0; and diffusion coefficients corrected in addition for the low viscosity of TIP3P water, D_η. The diffusion coefficients are averages of particle trajectories in the same run (K^+ ions with hexanucleotide RNAs), trajectories divided into blocks of equal length (RNA molecules) or combinations of both (K^+ ions with trinucleotide RNAs). Estimated statistical errors are shown in parentheses (mean ±1 standard deviation).

to obtain the corrected diffusion coefficients D_0 of other RNA molecules shown in Table 3 along with D_η values corrected in addition for the solvent viscosity of TIP3P water. We also performed a MD simulation of A_6 in water without added counterions (simulation run 21 in Table 1). The diffusion coefficient of A_6 estimated from that simulation is $(2.86 \pm 0.02) \times 10^{-6}$ cm^2 s^{-1}, which is close to the corresponding diffusion coefficient for A_6 with counterions shown in Table 3.

The corrected diffusion coefficients D_η of RNA from the MD simulations decrease with the RNA length, without a significant dependence on the RNA sequence. Recently, Stellwagen and Stellwagen (2002) estimated the diffusion coefficient of a single-stranded DNA (ssDNA) with 20 nucleotides to be $D_{20} = 1.52 \times 10^{-6}$ cm^2 s^{-1} at 20°C by capillary electrophoresis. Nkodo et al. (2001) estimated the diffusion coefficient of ssDNA with 18 nucleotides to be $D_{18} = (0.98 \pm 0.05) \times 10^{-6}$ cm^2 s^{-1} at 30°C in 7 M urea solvent. Adapting the empirical relationship $D \propto n^{-0.67}$ established for diffusion of double-stranded DNA with n basepairs

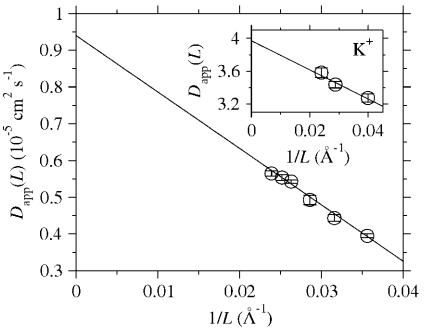

FIGURE 1 Diffusion coefficients $D_{app}(L)$ of A_3 RNA as a function of the inverse box length, $1/L$ (*circles*). The solid line was obtained by fitting Eq. 3 to the calculated diffusion coefficients. The inset shows the corresponding plot for a single K^+ ion in water. Error bars (mean ±1 standard deviation) were estimated from block averages.

(Stellwagen et al., 2001), one may attempt to extrapolate from the measured diffusion coefficients D_{18} and D_{20} (Nkodo et al., 2001; Stellwagen and Stellwagen, 2002) to estimate diffusion coefficients of tri- and hexanucleotide single-stranded RNA. The resulting diffusion coefficients extrapolated from the 18-mer and 20-mer values without correction for the temperature dependence are $D_3 \approx 3.3$ and 5.4×10^{-6} cm^2 s^{-1}, respectively, and $D_6 \approx 2.1$ and 3.4×10^{-6} cm^2 s^{-1}. These values compare well with the corrected diffusion coefficients D_η of tri- and hexanucleotide ssRNA estimated from the simulations. We also calculated diffusion coefficients of RNA using the program HYDROPRO (de la Torre et al., 2000; Fernandes et al., 2002) with the viscosity of TIP3P water and representative structures from our MD simulations as input. They are 0.845(±0.012), 0.869(±0.016), 0.678(±0.011), and 0.644(±0.012) $\times 10^{-5}$ cm^2 s^{-1} for A_3, U_3, A_6, and U_6, respectively. These values are in good agreement with the corresponding diffusion coefficients D_0 corrected for finite-size effects, as listed in Table 3. The hydrodynamic radii R_H calculated from the Stokes-Einstein relation, Eq. 5, are 7.54, 7.31, 10.20, and 10.26 Å for A_3, U_3, A_6, and U_6, respectively.

In the presence of an electric field, the charged RNA molecules drift along the field. However, in the plane orthogonal to the field, the motion should be diffusive. Therefore, we also estimated diffusion coefficients from x and y components of the mean-square displacement of RNA from RNA/water simulations with electric fields along the z direction, adapting Eq. 1 for two-dimensional diffusion. Diffusion coefficients calculated by this method from the data with the electric field are in close agreement with zero field values of Table 3. No systematic correlation between the resulting diffusion coefficients D for motion normal to the field and the field strength is observed. This implies that the RNA diffusion normal to the field is largely independent of E even at electric fields as high as ~50 mV/Å, which is consistent with a recent experimental observation at electric fields E several orders of magnitude smaller (Nkodo et al., 2001).

Table 3 also lists diffusion coefficients of K^+ ions. We observe only small differences between the K^+ diffusion coefficients in simulations with different RNA molecules. A reduction in the diffusion coefficients of K^+ in the systems with smaller RNAs may be caused by the relatively smaller system size (Dünweg and Kremer, 1993). To correct for the system-size dependence, and the binding of K^+ ions to RNA, we performed additional MD simulations of a single K^+ ion in solution with 512, 1367, and 2364 water molecules (simulation runs 27–29 in Table 1). The resulting apparent diffusion coefficients can again be fitted nicely to the $1/L$ dependence of Eq. 3, as shown in the inset of Fig. 1. From this fit, we obtain a finite-size corrected diffusion coefficient $D_0 = 4.0 \times 10^{-5}$ cm^2 s^{-1} for a K^+ ion, a viscosity-corrected diffusion coefficient of $D_\eta = 1.38 \times 10^{-5}$ cm^2 s^{-1}, and a correction factor of $\alpha = 0.88$. The

experimental value of the diffusion coefficient of K^+ at 25°C is 1.96×10^{-5} cm^2 s^{-1} (Atkins, 1990).

Electrophoretic mobility of RNA and potassium ions

Fig. 2 shows the average drift velocity as a function of the applied electric field. The drift velocity grows linearly with the electric field even at fields as high as 50 mV/Å. Values of the mobility μ calculated from the slope of lines in Fig. 2 according to Eq. 6 are 2.91, 2.96, 4.86, and 5.24×10^{-4} cm^2 V^{-1} s^{-1} for A_3, U_3, A_6, and U_6, respectively. Based on the Nernst-Einstein relation, Eq. 7, one can assume that the mobility has the same finite-size and solvent-viscosity dependence as the diffusion coefficient. With that assumption, we obtain corrected mobility values of 2.41, 2.36, 4.04, and 4.38×10^{-4} cm^2 V^{-1} s^{-1} for A_3, U_3, A_6, and U_6, respectively. For short single-stranded polythymine DNA at low excess-salt concentration (0.001 mol l^{-1}), Hoagland et al. (1999) measured mobilities of comparable magnitude, ~4 and 5×10^{-4} cm^2 V^{-1} s^{-1}, for tri- and hexanucleotide DNA. The molecular charge appears to be the dominant factor for the mobility of the small RNA molecules because the larger hexanucleotide RNAs, A_6 and U_6, with five net-negative charges, display larger values of μ than the smaller trinucleotide RNAs, A_3 and U_3, with two net-negative charges. This is consistent with the experimental finding that the mobility of DNA increases with an increasing number of basepairs until it becomes constant at ~170 basepairs (Stellwagen and Stellwagen, 2002). The faster drift of U_6 compared to A_6 at $E = 40$ and 50 mV/Å despite a smaller diffusion coefficient at zero field can be understood from differences in the structures of RNA in these simulation runs. We found that RNA molecules with larger radii of gyration diffuse more slowly, with D determined from the x and y displacements normal to the field. This is expected from the Stokes-Einstein relation, Eq. 5, and provides an explanation for the faster drift of U_6 with a smaller radius of gyration R_g compared to A_6 at $E = 40$ and 50 mV/Å. We also estimated the mobility of K^+ ions in solution with RNA, and obtained values of 1.1×10^{-3} cm^2 V^{-1} s^{-1} independent of the RNA type.

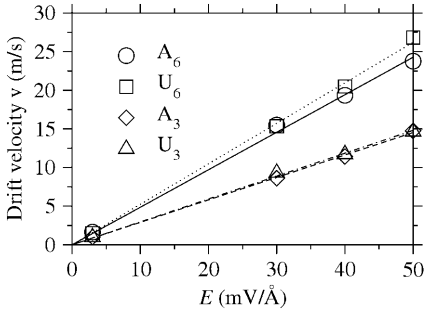

FIGURE 2 Drift velocity of RNA as a function of the electric field E. Solid, dotted, dashed, and dot-dashed lines are the results of linear fits of Eq. 6 for A_6, U_6, A_3, and U_3, respectively.

Coupling of electrophoretic drift of RNA to counterion binding

Although the RNA and K^+ ions on average drift electrophoretically along the field direction z, they also spread diffusively about the average drift position. If fluctuations about the average stem entirely from uncorrelated thermal random motion, the distributions of z displacements at a given time interval should be symmetric and Gaussian. Then we could estimate D by measuring the variance of z displacements at a fixed time interval. This forms the basis of D measurements by capillary electrophoresis (Nkodo et al., 2001; Stellwagen and Stellwagen, 2002). However, as discussed in Methods and Theory, fluctuations in the ionic atmosphere surrounding the drifting RNA molecules are expected to increase the apparent diffusion coefficients of RNA molecules for motion along the direction of the electric field. This is indeed what we find in the simulations. Fig. 3 shows var[$z(t)$] for the A_6 RNA at a field of $E = 50$ mV/Å. Equation 11 is found to provide an excellent description of the data, with a standard deviation of the fluctuating counterion charge bound to the RNA of $\sigma_q = 0.57\ e$ and a corresponding relaxation time of $\tau_q = 72$ ps. These fitted values agree well with the rough estimates obtained directly from the simulations under the assumption that only ions in the first solvation shell affect the RNA mobility, $\sigma_q = 0.51\ e$ and $\tau_q = 53$ ps, where τ_q was estimated from the exponential decay at long times in the autocorrelation function of the number of bound counterions. We conclude from this that diffusion coefficients obtained from the spread of charged polymers along the field direction should be corrected for the effect of fluctuations in the counterion charge. We expect

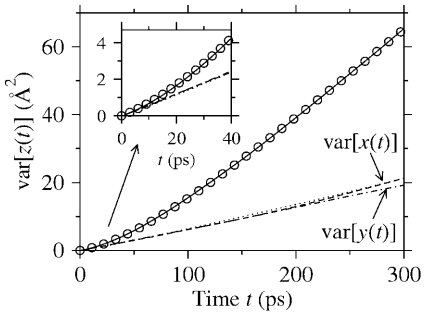

FIGURE 3 Variance of displacements $z(t)$ along the direction of the electric field ($E = 50$ mV/Å) for A_6 RNA as a function of time t (*solid line*). Circles represent the results of fitting Eq. 11. Also shown are the variance of the displacement $z(t)$ of A_6 without electric field (*dotted line*), and variances of A_6 in directions x and y normal to the electric field with dashed and dot-dashed lines, respectively. Differences between the variances in $x(t)$ and $y(t)$ at finite electric field and $z(t)$ at zero field are within the statistical uncertainties. The inset shows a magnified view of the initial nonlinear diffusive spread.

this effect to be small at low electric fields (E^2 term in Eq. 11), such as the ~100 V/cm (10^{-3} mV/Å) fields in the measurements of Stellwagen et al. (2001), and at high ion concentrations. In fact, var[$z(t)$] calculated at $E = 3$ mV/Å is close to that calculated from the data at zero field. However, quantitative predictions would require a detailed analysis of the dependence of σ_q^2 and τ_q on the electric field, ion type and concentration, and polyelectrolyte type and length, which is beyond the scope of the present work.

To further test the counterion binding model of RNA diffusion, we estimated instantaneous drift velocities from linear fits to 100-ps time intervals of $z(t)$, as illustrated in the inset of Fig. 4. We correlate the drift velocity of A_6 to the number of counterions bound to A_6 during the 100-ps intervals. Ions are classified as bound to RNA if they are closer than 3.4 Å to any atom of the RNA molecule. That distance corresponds to the first minimum in the distribution of the closest distance between an ion and any atom in the RNA. The second minimum occurs at 5.9 Å. In Fig. 5, we plot the average drift velocity of A_6 with respect to the average number of bound counterions (N_{K^+}) during the 100-ps time interval as well as probability distributions of N_{K^+}. We find that RNA molecules with larger numbers N_{K^+} of neutralizing counterions bound to them drift more slowly than those with smaller numbers. In the limit of zero counterions bound to the RNA ($N_{K^+} = 0$), we recover the drift velocity predicted from the Nernst-Einstein relation in Eq. 8, $\nu = D|Q|E/k_BT$, for a hexanucleotide RNA with full charge, $Q = -5e$. Moreover, if we simply correct for the potassium charge contained in the first shell and use a modified Nernst-Einstein relation,

$$\nu = \frac{D|Q + eN_{K^+}|E}{k_BT}, \tag{13}$$

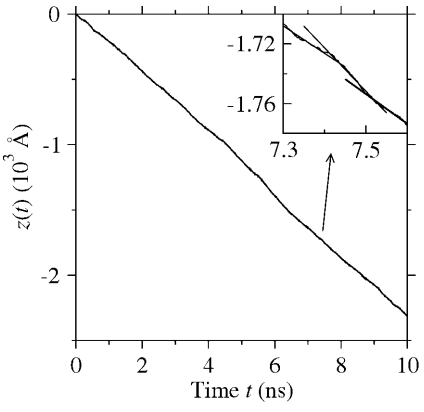

FIGURE 4 Position $z(t)$ of the geometric center of A_6 as a function of time along the direction of the electric field (E = 50 mV/Å). The inset is a magnified view of $z(t)$ during a 300-ps time interval. The three straight lines in the inset are the best linear fits for three consecutive 100-ps blocks, with slopes of 18.37, 29.15, and 19.96 m s^{-1}, respectively. The average number of ions bound to the RNA during those 100-ps time intervals are 0.50, 0.03, and 0.14, respectively.

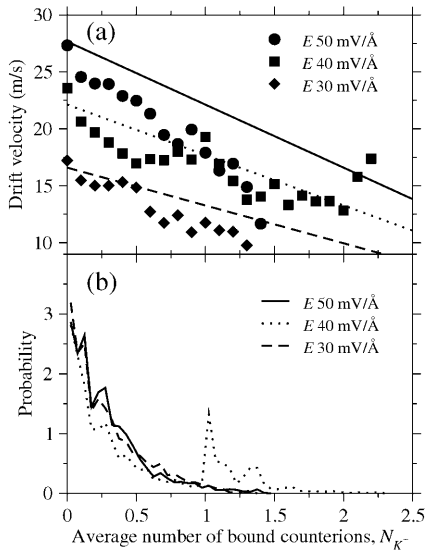

FIGURE 5 (*a*) Average drift velocity of A_6 plotted as a function of the average number of bound counterions N_{K^+}, calculated for 100-ps time intervals. Circles, squares, and diamonds represent the results at $E = 50$, 40, and 30 mV/Å, respectively. Solid, dotted, and dashed lines represent the drift velocities predicted by the modified Nernst-Einstein formula, Eq. 13, for E = 50, 40, and 30 mV/Å, respectively, calculated with the diffusion coefficient of A_6 RNA in the MD simulation without counterions, $D = 2.86 \times 10^{-6}$ cm^2 s^{-1}. (*b*) Probability distributions of the average number of bound counterions during 100-ps time intervals.

where D is the diffusion coefficient of RNA in the absence of electric fields, we obtain good agreement with the observed drift velocities ν as a function of the number of bound ions N_{K^+} (Fig. 5 *a*).

Fluctuations in the counterion binding also affect the diffusion of K^+ ions. We calculated the distribution of K^+ displacements in the field direction (z) at a fixed 100-ps time interval. The peak positions in the probability distributions of z displacements of K^+ ions in the A_6/K^+/water mixture at applied electric fields agree well with the positions expected from free ionic drift, as indicated by arrows in Fig. 6. However, the probability distributions are asymmetric and non-Gaussian and do not follow those of random thermal fluctuations. At $E = 40$ mV/Å, we even see a small peak near $z = -17.9$ Å which corresponds to the approximate average displacement of A_6 with one counterion bound at that field. This indicates that there are K^+ counterions transiently bound to RNA and moving with it against the field, which affects the coupled motion of RNA and K^+ under the influence of electric fields. Such transient trapping of counterions near a charged polymer forms the basis of the theory of ionic self-diffusion in a polyelectrolyte solution developed by Lifson and Jackson (1962).

Table 4 lists the average number of counterions within 3.4 Å and 5.9 Å obtained from the MD simulations. We also estimated effective RNA charges Q_{eff} from the Nernst-Einstein relation, Eq. 7, as −4.31, −4.72, −1.90, and −1.81 e for A_6, U_6, A_3, and U_3, respectively. These Q_{eff} values are

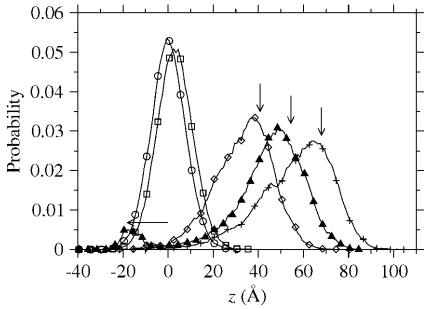

FIGURE 6 Probability distributions of displacements $z(t)$ of K^+ ions during time intervals of $t = 100$ ps in the A_6/K^+/water solution at 0, 3, 30, 40, and 50 mV/Å shown as lines with circles, squares, diamonds, solid triangles, and crosses, respectively. Vertical arrows indicate the positions expected from free ionic drift at fields of 30, 40, and 50 mV/Å with a mobility of $D_K^+ e/k_B T$, where $D_K^+ = 3.515 \times 10^{-5}$ cm^2 s^{-1} is the diffusion coefficient of a free K^+ ion estimated from the inset in Fig. 1 for the system size of the A_6 simulation. The horizontal arrow indicates the expected average drift position of a K^+ ion bound to A_6 RNA, $-4e\ D_{RNA}\ Et/k_B T$, at $E = 40$ mV/Å, where D_{RNA} is the diffusion coefficient $D_{app}(L)$ of A_6 listed in Table 3.

similar to the charges obtained by adding the average charge of bound counterions (Table 4) to the RNA charges.

CONCLUSIONS

We have calculated diffusion coefficients and electrophoretic mobilities of small RNA molecules with three and six nucleotides in free solution using MD simulations. We have found that the electrophoretic mobility increases with increasing RNA length, which is consistent with the experimental results. Differences in transport properties among RNAs with the same number of nucleotides but different base sequences are small. Calculated diffusion coefficients of RNA corrected for finite-size effects and the low viscosity of TIP3P water agree well with those estimated from experiments and hydrodynamic calculations. Diffusion coefficients and electrophoretic mobilities are found to be related through the Nernst-Einstein formula if we use reduced effective charges of RNA that correct for counterion binding. A consequence of this is that variations in the electrophoretic mobility of single molecules caused by fluctuations in their counterion atmosphere can lead to an enhanced diffusive spread along the direction of the electric field. Therefore, apparent diffusion coefficients measured by capillary zone electrophoresis at certain conditions, in particular high electric fields, can be larger than the actual values.

TABLE 4 Average numbers of counterions within a fixed distance from RNA

	A_6		U_6		A_3		U_3	
RNA Distance (Å)	3.4	5.9	3.4	5.9	3.4	5.9	3.4	5.9
$E = 0$ mV/Å	0.20	1.05	0.23	1.06	0.07	0.46	0.08	0.46
3	0.15	0.89	0.19	0.96	0.09	0.49	0.09	0.50
30	0.29	1.11	0.21	0.96	0.08	0.46	0.10	0.48
40	0.49	1.18	0.40	1.12	0.12	0.51	0.12	0.53
50	0.28	1.05	0.23	0.98	0.10	0.48	0.14	0.49

APPENDIX A: ELECTRIC FIELD IN WATER AND RNA SYSTEMS

The average dipolar angle of water with respect to the direction of an external field is a unique function of the strength E of the electric field (Yeh and Berkowitz, 1999a,b). For bulk TIP3P water, we obtained the relationship between the dipolar angle of water and the electric field E by MD simulations. Fig. 7 compares the average dipolar angles of water molecules at least 20 Å away from the geometric center of RNA with those from simulations of pure water. This comparison indicates that the electric field E in the RNA systems is indeed close to the applied external electric field E_0. Slight deviations are expected for an inhomogeneous system under conducting boundary conditions where only the average electric field in the simulation box is E_0.

APPENDIX B: FLUCTUATING CHARGE MODEL OF POLYELECTROLYTE DIFFUSION

To calculate the diffusive spread of charged polymers caused by fluctuations in the ion atmosphere, we assume that the effective charge $q(t)$ of the polymer fluctuates about a mean $\langle q \rangle$ with a variance $\sigma_q^2 = \langle q^2(t) \rangle - \langle q(t) \rangle^2 = \langle \Delta q^2(t) \rangle$ and a correlation function $\langle \Delta q(t)$

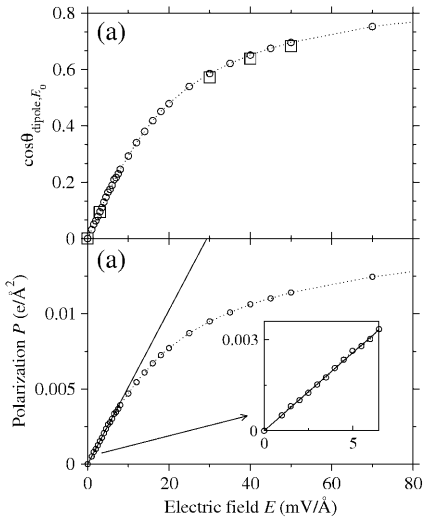

FIGURE 7 (*a*) Cosine of the angle between water dipoles and the field direction as a function of the electric field E. Open circles connected with a dotted line represent results obtained from simulations of bulk TIP3P water. Open squares at 0, 3, 30, 40, and 50 mV/Å show the results calculated from water molecules at least 20 Å away from the RNA center in A_6/water simulations at corresponding electric field strengths E_0. (*b*) Polarization (P) of bulk TIP3P water as a function of the electric field E (*open circles* connected with *dotted line*). The solid line is the linear-response result (Neumann, 1983), $P = (\varepsilon - 1)\ \varepsilon_0\ E$, for a dielectric constant of $\varepsilon = 94$ that was obtained from a linear fit at small electric fields (*inset*), where ε_0 is the vacuum permittivity. The nonlinearity between P and E becomes pronounced at electric fields above $E \sim 7$ mV/Å, resulting in an apparent decrease of ε (Yeh and Berkowitz, 1999a).

$\Delta q(0)\rangle$ with $\Delta q(t) = q(t) - \langle q\rangle$. Then, the mean and variance of the time-integrated charge are given by

$$\left\langle \int_0^t q(t')\,dt' \right\rangle = \langle q\rangle t \tag{B1}$$

$$\mathrm{var}\left[\int_0^t q(t')\,dt'\right] = 2\int_0^t (t-t')\langle \Delta q(t')\Delta q(0)\rangle\,dt'. \tag{B2}$$

For an exponentially relaxing charge, $\langle \Delta q(t)\,\Delta q(0)\rangle = \sigma_q^2 \exp(-t/\tau_q)$, with relaxation time τ_q, the integral of the variance can be calculated analytically. In combination with Eq. 9, we then obtain $z_d(t) = \mu E\, t$ with the mobility corresponding to the Nernst-Einstein relation Eq. 7, $\mu = D\langle q\rangle/k_BT$. The variance of the drift position $z_d(t)$ because of the fluctuating charge of the polymer is given by Eq. 10. For Brownian dynamics of the charge on a harmonic free energy surface, the exponential relaxation is exact, and the distribution of the drift distances $z_d(t)$ is exactly Gaussian—as can be shown, e.g., by using Anderson's formalism for the line shape (Anderson, 1954).

The authors thank Dr. Attila Szabo for many helpful discussions. This study utilized the Biowulf PC/Linux cluster at the National Institutes of Health, Bethesda, MD.

REFERENCES

Allen, M. P., and D. J. Tildesley. 1987. Computer Simulations of Liquids. Oxford University Press, New York.

Anderson, P. W. 1954. A mathematical model for the narrowing of spectral lines by exchange or motion. *J. Phys. Soc. Jpn.* 9:316–339.

Åqvist, J. 1990. Ion-water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem.* 94:8021–8024.

Atkins, P. W. 1990. Physical Chemistry, 4th Ed. Oxford University Press, Oxford, UK.

Beenakker, C. W. J. 1986. Ewald sum of the Rotne-Prager tensor. *J. Chem. Phys.* 85:1581–1582.

Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. Di Nola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

Berezhkovskii, A. M., A. Szabo, and G. H. Weiss. 1999. Theory of single-molecule fluorescence spectroscopy of two-state systems. *J. Chem. Phys.* 110:9145–9150.

Cantor, C. R., and P. R. Schimmel. 1980. Biophysical Chemistry, Part II. Freeman, San Francisco, CA.

Cheatham, III, T. E., and P. A. Kollman. 2000. Molecular dynamics simulation of nucleic acids. *Annu. Rev. Phys. Chem.* 51:435–471.

Cheatham, III, T. E., J. L. Miller, T. Fox, T. A. Darden, and P. A. Kollman. 1995. Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J. Am. Chem. Soc.* 117:4193–4194.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second-generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.

Darden, T. A., D. M. York, and L. G. Pedersen. 1993. Particle mesh Ewald: an $N\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

de la Torre, J. G., M. L. Huertas, and B. Carrasco. 2000. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* 78:719–730.

Dünweg, B., and K. Kremer. 1993. Molecular dynamics simulation of a polymer chain in solution. *J. Chem. Phys.* 99:6983–6997.

Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.

Feller, S. E., Y. H. Zhang, R. W. Pastor, and B. R. Brooks. 1995. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* 103:4613–4621.

Fernandes, M. X., A. Ortega, M. C. L. Martínez, and J. G. de la Torre. 2002. Calculation of hydrodynamic properties of small nucleic acids from their atomic structure. *Nucleic Acids Res.* 30:1782– 1788.

Foloppe, N., and A. D. MacKerell. 2000. All-atom empirical force field for nucleic acids. I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comp. Chem.* 21:86–104.

Geva, E., and J. L. Skinner. 1998. Two-state dynamics of single biomolecules in solution. *Chem. Phys. Lett.* 288:225–229.

Hoagland, D. A., E. Arvanitidou, and C. Welch. 1999. Capillary electrophoresis measurements of the free solution mobility for several model polyelectrolyte systems. *Macromolecules.* 32:6180–6190.

Jacobson, S. C., C. T. Culbertson, J. E. Daler, and J. M. Ramsey. 1998. Microchip structures for submillisecond electrophoresis. *Anal. Chem.* 70:3476–3480.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Kale, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999. NAMD2: greater scalability for parallel molecular dynamics. *J. Comp. Phys.* 151:283–312.

Lifson, S., and J. L. Jackson. 1962. On the self-diffusion of ions in a polyelectrolyte solution. *J. Chem. Phys.* 36:2410–2414.

Muthukumar, M. 1997. Dynamics of polyelectrolyte solutions. *J. Chem. Phys.* 107:2619–2635.

Neumann, M. 1983. Dipole moment fluctuation formulas in computer simulations of polar systems. *Mol. Phys.* 50:841–858.

Nkodo, A. E., J. M. Garnier, B. Tinland, H. J. Ren, C. Desruisseaux, L. C. McCormick, G. Drouin, and G. W. Slater. 2001. Diffusion coefficient of DNA molecules during free solution electrophoresis. *Electrophoresis.* 22:2424–2432.

Norberg, J., and L. Nilsson. 2002. Molecular dynamics applied to nucleic acids. *Acc. Chem. Res.* 35:465–472.

Nowakowski, J., and I. Tinoco, Jr. 1999. RNA structure in solution. *In* Oxford Handbook of Nucleic Acid Structure. S. Neidle, editor. Oxford Press, New York. 567–602.

Pastor, R. W., and M. Karplus. 1988. Parametrization of the friction constant for stochastic simulations of polymers. *J. Phys. Chem.* 92:2636–2641.

Pearlman, D. A., D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, III, S. DeBolt, D. Ferguson, G. Seibel, and P. Kollman. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Comm.* 91:1–41.

Placzek, G., B. R. A. Nijboer, and L. van Hove. 1951. Effect of short wavelength interference on neutron scattering by dense systems of heavy nuclei. *Phys. Rev.* 82:392–403.

Plenert, M. L., and J. B. Shear. 2003. Microsecond electrophoresis. *Proc. Natl. Acad. Sci. USA.* 100:3853–3857.

Righetti, P. G., C. Gelfi, and M. R. D'Acunto. 2002. Recent progress in DNA analysis by capillary electrophoresis. *Electrophoresis.* 23:1361–1374.

Ryckaert, J. P., G. Ciccotti, and H. J. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comp. Phys.* 23:327–341.

Simonson, T. 2003. Electrostatics and dynamics of proteins. *Rep. Prog. Phys.* 66:737–787.

Slater, G. W., S. Guillouzic, M. G. Gauthier, J. F. Mercier, M. Kenward, L. C. McCormick, and F. Tessier. 2002. Theory of DNA electrophoresis. *Electrophoresis.* 23:3791–3816.

Smith, D. E., H. P. Babcock, and S. Chu. 1999. Single-polymer dynamics in steady shear flow. *Science*. 283:1724–1727.

Stellwagen, E., and N. C. Stellwagen. 2002. Determining the electrophoretic mobility and translational diffusion coefficients of DNA molecules in free solution. *Electrophoresis*. 23:2794–2803.

Stellwagen, N. C., S. Magnusdottir, C. Gelfi, and P. G. Righetti. 2001. Measuring the translational diffusion coefficients of small DNA molecules by capillary electrophoresis. *Biopolymers*. 58:390–397.

Stuart, J. N., and J. V. Sweedler. 2003. Ultrafast capillary electrophoresis and bioanalytical applications. *Proc. Natl. Acad. Sci. USA*. 100:3545–3546.

Volkmuth, W. D., and R. H. Austin. 1992. DNA electrophoresis in microlithographic arrays. *Nature*. 358:600–602.

Yeh, I.-C., and M. L. Berkowitz. 1999a. Dielectric constant of water at high electric fields: molecular dynamics study. *J. Chem. Phys*. 110:7935–7942.

Yeh, I.-C., and M. L. Berkowitz. 1999b. Ewald summation for systems with slab geometry. *J. Chem. Phys*. 111:3155–3162.

Yeh, I.-C., and G. Hummer. 2002. Peptide loop-closure kinetics from microsecond molecular dynamics simulations in explicit solvent. *J. Am. Chem. Soc*. 124:6563–6568.