

# Long-range DNA-water interactions

Abhishek K. Singh,<sup>1</sup> Chengyuan Wen,<sup>1</sup> Shengfeng Cheng,<sup>1,2,3</sup> and Nguyen Q. Vinh<sup>1,2,3,\*</sup>

<sup>1</sup>Department of Physics and Center for Soft Matter and Biological Physics, <sup>2</sup>Macromolecules Innovation Institute, and <sup>3</sup>Department of Mechanical Engineering, Virginia Tech, Blacksburg, Virginia

**ABSTRACT** DNA functions only in aqueous environments and adopts different conformations depending on the hydration level. The dynamics of hydration water and hydrated DNA leads to rotating and oscillating dipoles that, in turn, give rise to a strong megahertz to terahertz absorption. Investigating the impact of hydration on DNA dynamics and the spectral features of water molecules influenced by DNA, however, is extremely challenging because of the strong absorption of water in the megahertz to terahertz frequency range. In response, we have employed a high-precision megahertz to terahertz dielectric spectrometer, assisted by molecular dynamics simulations, to investigate the dynamics of water molecules within the hydration shells of DNA as well as the collective vibrational motions of hydrated DNA, which are vital to DNA conformation and functionality. Our results reveal that the dynamics of water molecules in a DNA solution is heterogeneous, exhibiting a hierarchy of four distinct relaxation times ranging from  $\sim 8$  ps to 1 ns, and the hydration structure of a DNA chain can extend to as far as  $\sim 18$  Å from its surface. The low-frequency collective vibrational modes of hydrated DNA have been identified and found to be sensitive to environmental conditions including temperature and hydration level. The results reveal critical information on hydrated DNA dynamics and DNA-water interfaces, which impact the biochemical functions and reactivity of DNA.

**SIGNIFICANCE** Water plays an important role in the dynamics, flexibility, and stability of DNA molecules. It also strongly affects their recognition processes and thus largely determines their biochemical functions. Strong water absorption and a broad continuous vibrational density of states in the megahertz to terahertz frequency range have prevented optical identification of water-DNA interaction. In this report, we provide new, to our knowledge, insights into the hydration dynamics as well as terahertz-collective motions of hydrated DNA probed over a wide frequency region with an unprecedented signal/noise ratio. Combining megahertz to terahertz spectroscopy and molecular dynamics simulations, we have determined the impact of water on DNA dynamics and vice versa, which influence biochemical functions of DNA.

## INTRODUCTION

Hydration plays an important role in the dynamics, flexibility, and stability of DNA molecules. It also strongly affects their recognition processes and thus largely determines their biochemical functions. Specifically, hydration significantly affects the binding of DNA with proteins and/or ligands, as well as the intercalation of foreign molecules, including anticancer drugs, into DNA molecules (1–6). Without hydration, the observed structures of DNA molecules would be unstable because of electrical field effects from the negatively charged phosphate backbone and the positively charged counterions. Water molecules form hydration layers around DNA and screen these charges,

making an entangled structure stable. Under usual physiological conditions, the B-DNA form is the most abundant, but the A-DNA and Z-DNA forms can emerge when the hydration level is below a certain threshold (2,7–9). For all these forms, the hydration shells are necessary to stabilize the double-helical structure of DNA and allow the formation of grooves (1–4). In these grooves, water molecules serve as bridges between charged groups (10–12), and as a result, ordered spine-like hydration structures are observed on a large scale. Such spines of hydration are regarded as an integral part of DNA because they support DNA stability. All these observations suggest that hydration water is one of the fundamental factors determining the structures and functions of DNA. At the same time, the dynamics of water is also influenced by the presence and conformational transitions of DNA molecules. The dynamics of water molecules in the hydration layers of DNA is especially complicated because of the confined environment provided by the major

Submitted May 17, 2021, and accepted for publication October 18, 2021.

\*Correspondence: [vinh@vt.edu](mailto:vinh@vt.edu)

Editor: Elsa Yan.

<https://doi.org/10.1016/j.bpj.2021.10.016>

© 2021 Biophysical Society.

and minor grooves. Understanding the behavior of these water molecules and their effects on DNA dynamics therefore provides key insights into the biological functions of DNA.

The impact of hydration layers on DNA dynamics can be revealed by probing the long-range structural vibrations of DNA. In the past decade, the collective motions of DNA and the water dynamics in hydration layers have been characterized experimentally and investigated theoretically and computationally using molecular dynamics (MD) simulations (13–22). These simulations suggest that the hydration-bond dynamics and reorientation time of water molecules in DNA hydration shells span a wide range of timescales from picoseconds to a few nanoseconds (13–17,19–21). Based on entropy calculations from atomistic MD simulations, Jana et al. (19) have shown that the water molecules are more restricted in the minor and major grooves as compared to free water. A diverse range of experimental techniques have also been employed to investigate the dynamics of hydrated DNA molecules from different angles. X-ray crystallography experiments performed at cryogenic temperatures have provided a detailed view of DNA and the hydration structure of DNA, revealing the presence of a spine of hydration in the minor grooves (23–25). NMR spectroscopy detects the motion of probe nuclei at different timescales related to the water dynamics in hydration shells. NMR studies revealed that the residence time of water molecules in the minor grooves is around 1 ns or longer and indicated that a spine of hydration is an integral part of the DNA minor grooves (26–30). Neutron scattering (31,32) and Raman scattering techniques (8,33,34) have revealed information on the structure of water molecules in DNA hydration shells, as well as the collective vibrational modes of DNA. Time-resolved Stokes shift experiments (16,35–37), time-resolved polarization measurements (38), dielectric spectroscopy (9,39–45), time-resolved infrared spectroscopy (46), pump-probe experiments (47), vibrational sum-frequency generation measurements (12,48), optical Kerr effect spectroscopy (49), and NMR spectroscopy (28–30) have been used to characterize the collective solvation dynamics around DNA and reveal multiexponential decay kinetics of the orientation of water molecules, which are connected to the hydrogen-bond network dynamics of water. For examples, Berg and co-workers carried out extensive research on power law and logarithmic dynamics of water molecules over picosecond to nanosecond timescales (16,35–37). Among these techniques, a combination of dielectric relaxation measurements (9,39,40) and terahertz spectroscopy (41–45) stands out as a unique approach, as they can simultaneously address the dynamical and structural aspects of nucleic acids and hydration water around them over a wide range of timescales. Such measurements can thus reveal the connection between hydration and DNA dynamics.

The goal of our study is to comprehensively analyze the collective motions of hydrated DNA molecules as well as

the hydration dynamics of water by combining the dielectric relaxation and terahertz spectroscopy. The dielectric spectra of aqueous DNA solutions, which span a wide range of frequencies in the electromagnetic spectrum, are encoded with signatures related to a number of biological processes involving inter- and intramolecular dynamics of DNA and water. The dielectric relaxation at kilohertz to megahertz frequencies was generally viewed as a result of the polarization of condensed counterions along a DNA polyanion (i.e.,  $\alpha$ -relaxation), as well as the relaxation of permanent dipoles and the motion of polar groups on the DNA molecule (i.e.,  $\beta$ -relaxation) (50). At higher frequencies, the dielectric spectroscopy of aqueous DNA solutions was presented separately either at megahertz to gigahertz or terahertz frequencies (9,39–41). Thus, a complete spectrum with high precision covering the wide range of frequencies from megahertz to terahertz will enable a more reliable analysis of DNA hydration and its impact on DNA dynamics. We have recently achieved this goal with an improved dielectric spectrometer (51–56). The new, to our knowledge, spectrometer has a signal/noise ratio enhanced by several orders of magnitude and a large spectral bandwidth extending from megahertz to terahertz frequencies, enabling high-precision dielectric spectroscopic measurements for studying the hydration dynamics in aqueous biomolecular solutions and the collective motions of biomolecules. Specifically, the megahertz to gigahertz dielectric response allows us to directly determine the dynamics of water molecules, as well as to estimate the total number of water molecules whose kinetics are altered by the presence of biomacromolecules in the solution. The terahertz spectroscopic data reveal the collective vibrational motions of hydrated biomolecules in a solution. Using effective-medium approximation, we are able to map out the dynamics of hydrated DNA molecules as well as the number of water molecules that constitute the first hydration shell of a DNA chain. The experimental results are further corroborated and supplied with molecular-level insight revealed by all-atom MD simulations.

## MATERIALS AND METHODS

### Dielectric megahertz to terahertz spectroscopy

The measurements were performed on a frequency-domain dielectric spectrometer that covers a broad range of the electromagnetic spectrum from 100 MHz to 1.12 THz ( $0.000334\text{--}37.36\text{ cm}^{-1}$ ). The setup allows simultaneous measurements of the refractive index and absorbance (i.e., the dielectric dispersion and loss) of an aqueous DNA solution at megahertz to terahertz frequencies. The spectrometer consists of a commercial Vector Network Analyzer from Agilent (PNA N5225A; Santa Clara, CA) that covers the frequency range from 10 MHz to 50 GHz. Frequency extenders from Virginia Diodes (Charlottesville, VA), were used to generate terahertz waves. Seven different frequency extender modules have been used to cover a frequency range of 60 GHz–1.12 THz as reported in our earlier publications (51–56). A variable pathlength sample cell made from aluminum for dielectric experiments was used. Two parallel

windows were mounted inside the aluminum sample cell, with one fixed and the other in a mobile position, to vary the thickness of a liquid sample. An ultraprecision linear motor stage (Newport XMS50; Irvine, CA) was employed to perform nanometer-precision motion with a travel range of 50 mm. The sample cell was equipped with Peltier coolers (Custom Thermoelectric, 27115L31-03CK; Bishopville, MD). High-power resistors were embedded in the aluminum sample cell to provide a controlled heating. The temperature of the sample was monitored and controlled with an accuracy of  $\pm 0.02^\circ\text{C}$ , using the LakeShore 336 temperature controller (LakeShore, Westerville, OH).

## Molecular dynamics simulations

Atomistic simulations were performed with the large-scale atomic/molecular massively parallel simulator (57). A DNA chain with sequence 5'-CGCGAATTCGCG-3' was built with a DNA generator, and 6720 water molecules were generated by NAMD tools (a computer software for MD simulations). The AMBER force field (a family of force fields for MD of biomolecules) was adopted for the DNA, and the SPC/E (extended simple point charge) model was employed for water. The Lennard-Jones 12-6 potential was truncated at 12 Å for DNA and 7.9 Å for water. Coulombic interactions were fully accounted for, with the long-range part computed with the particle-particle particle-mesh method. The geometric mixing rule was adopted for the DNA-water interactions. The equations of motions were integrated with a velocity Verlet algorithm with a time step of 1 fs. The system was equilibrated at 278 K for 1 ns in an NPT (constant-temperature and constant-pressure) ensemble with the temperature and pressure controlled through the Berendsen algorithm. The temperature of the system was increased to 35°C in increments of 10°C. In this way, systems at four different temperatures were obtained. At each temperature, the system was further equilibrated for 1 ns in an NPT ensemble based on the Berendsen algorithm. After equilibration, an NVT (constant-temperature and constant-pressure) ensemble was used for production runs with the system temperature controlled by a Nosé-Hoover thermostat. For each system, the production run was at least 10 ns, and a snapshot of the system was dumped every 1 ps for further analyses. Additional details regarding sample preparations, experimental methods, and MD simulations are provided in the [Supporting materials and methods](#).

## RESULTS

Molecules exhibit various frequency-dependent behaviors in an alternating electrical field, including reorientation, collective oscillation, and vibration. In this work, a high-precision megahertz to terahertz frequency-domain dielectric spectrometer is employed to probe the DNA dynamics and DNA-water interfaces in aqueous solutions (51–56). Details of sample preparation and experimental setup are provided in the [Materials and methods](#) and [Supporting materials and methods](#). The absorption coefficients,  $\alpha(\nu)$ , and refractive indices,  $n(\nu)$ , obtained simultaneously for pure water and aqueous DNA solutions at 25°C, are shown in [Fig. S1](#) for the frequency range of 100 MHz to 1.12 THz. Both quantities show a strong frequency dependence, where  $\alpha(\nu)$  monotonically increases while  $n(\nu)$  decreases with rising frequency. These quantities are used to determine the complex refractive index,  $n^*(\nu) = n(\nu) + ic\alpha(\nu)/(4\pi\nu)$ , and the square root of the dielectric constant,  $\sqrt{\epsilon_{sol}^*(\nu)}$ , with  $c$  being the speed of light. The results are

more conveniently presented in terms of the dielectric response as

$$\epsilon_{sol}^*(\nu) = \epsilon'_{sol}(\nu) + i(\epsilon''_{sol}(\nu) + \sigma/2\pi\nu\epsilon_0), \quad (1)$$

where  $\epsilon'_{sol}(\nu)$  and  $\epsilon''_{sol}(\nu)$  are the dielectric dispersion (real part) and dielectric loss (imaginary part) of the relative permittivity, respectively;  $\epsilon_0$  is the permittivity of the vacuum; and  $\sigma$  is the electrical conductivity of the solution. By measuring the absorption coefficient, refractive index, and electrical conductivity, we have obtained the complex dielectric spectra of aqueous DNA solutions in the megahertz to terahertz frequency range at various temperatures. [Fig. 1](#) shows the dielectric response of pure water and aqueous solutions at various DNA concentrations at 25°C.

## Megahertz to gigahertz spectroscopy

The complex dielectric spectroscopy at the megahertz to gigahertz frequencies reveals relaxation processes related to the cooperative orientations of water and biomolecules in aqueous solutions. At the molecular level, a molecule with a permanent dipole rotates to follow an alternating electrical field. [Fig. 1](#) shows the complex dielectric spectra of aqueous DNA solutions at 25°C, providing information on the cooperative reorientation dynamics of both hydration and bulk water. The position of the dielectric loss peak at 19.25 GHz (i.e., 8.27 ps) remains unchanged (*inset* of [Fig. 1 B](#)), which is well established as one of the relaxation modes of bulk water (53,58,59). The dielectric response induced by the motion of condensed counterions, the motion of polar groups, and the reorientation relaxation of permanent dipoles from DNA molecules generally occur in the range of 1 kHz–10 MHz, which is outside the frequency range probed here (50,55,60). The contribution of orientational polarizations of water molecules in the interfacial region around DNA is complex and less understood. The dielectric response shown in [Fig. 1 B](#) clearly reveals an emerging contribution from hydration water molecules around DNA. The contribution of such water increases significantly as the concentration of DNA is increased. It is understood that water molecules can form hydration layers around a DNA molecule but may not uniformly coat its surface. Some of them, including those in the inner hydration shells in the minor and major grooves of DNA, are tightly bound to the DNA surface, whereas those lying in the outer shells or interacting with  $\text{Na}^+$  counterions are only loosely bound. The megahertz to gigahertz dielectric response reveals the hydration properties and dynamics corresponding to the cooperative orientation relaxation of bulk water, as well as water molecules in the entire hydration sheath.

At megahertz to gigahertz frequencies, vibrations, collective oscillations, and the librational motions of molecules in general do not contribute considerably to the

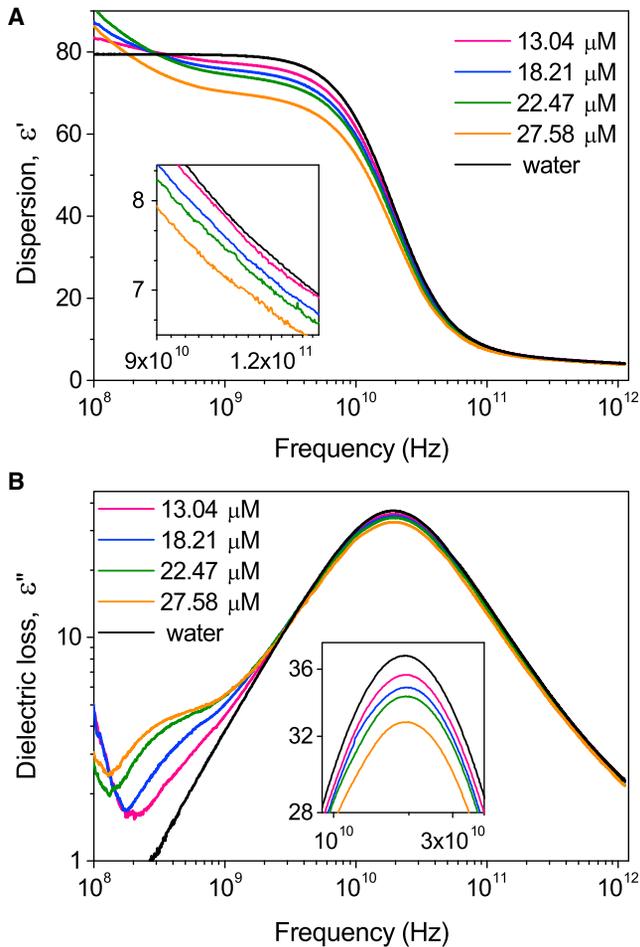


FIGURE 1 The complex dielectric response of aqueous DNA solutions. (A) Dielectric dispersion and (B) dielectric loss spectra of aqueous DNA solutions revealing the interaction between DNA solutions and electromagnetic radiation at megahertz to terahertz frequencies. A magnified view of the spectra of aqueous solutions at various DNA concentrations and pure water is shown in the inset of each figure. The maximum of the dielectric loss spectra occurs at around 19 GHz, similar to the result for pure water, but the magnitude of the maximal dielectric loss decreases as the DNA concentration is increased. A clearly emerging contribution to the dielectric loss from hydration water in the solution can be distinctively seen at low frequencies (between 0.1 and 1 GHz), with the dielectric loss becoming more pronounced at higher DNA concentrations. To see this figure in color, go online.

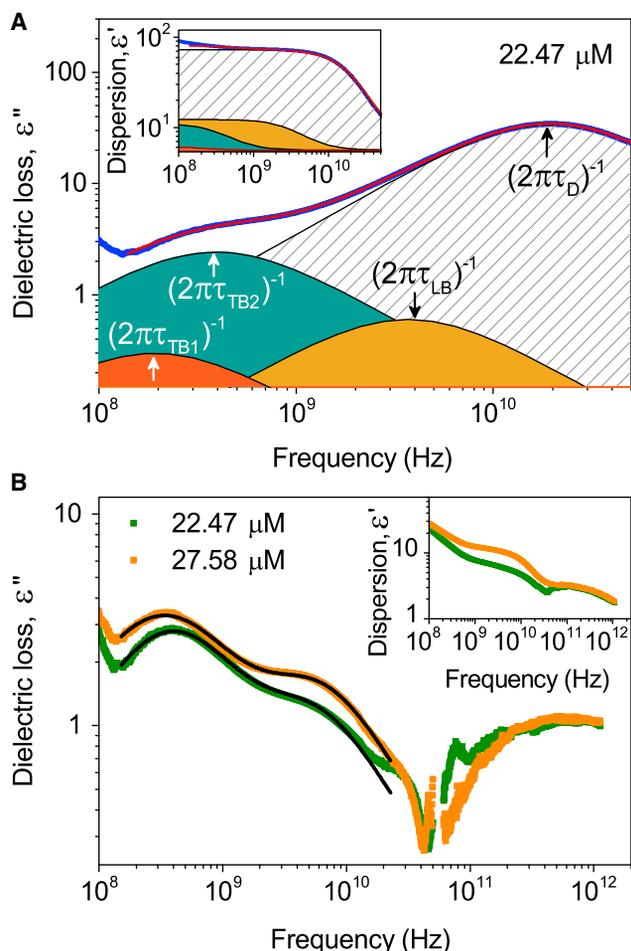
dielectric relaxation of the solution. In this case, the Debye relaxation model can be applied to understand the dielectric property of DNA solutions within error limits (61,62). A superposition of four Debye components is adopted here, which captures the contribution from the tightly bound, loosely bound, and bulk water, respectively. The complex dielectric function has the following form,

$$\begin{aligned} \epsilon_{\text{sol}}^*(\nu) &= \epsilon'_{\text{sol}}(\nu) + i\epsilon''_{\text{sol}}(\nu) \\ &= \epsilon_{\infty} + \frac{\epsilon_S - \epsilon_1}{1 + i\omega\tau_{\text{TB1}}} + \frac{\epsilon_1 - \epsilon_2}{1 + i\omega\tau_{\text{TB2}}} + \frac{\epsilon_2 - \epsilon_3}{1 + i\omega\tau_{\text{LB}}} \\ &\quad + \frac{\epsilon_3 - \epsilon_{\infty}}{1 + i\omega\tau_{\text{D}}}, \end{aligned} \quad (2)$$

where  $\Delta\epsilon_{\text{TB1}} = \epsilon_S - \epsilon_1$ ,  $\Delta\epsilon_{\text{TB2}} = \epsilon_1 - \epsilon_2$ ,  $\Delta\epsilon_{\text{LB}} = \epsilon_2 - \epsilon_3$ , and  $\Delta\epsilon_{\text{D}} = \epsilon_3 - \epsilon_{\infty}$  represent the dielectric contributions of Debye relaxation processes to the total relaxation spectrum from the tightly bound water (including those in the minor grooves and those in the major grooves and around the phosphate groups), loosely bound water, and bulk water, respectively. The corresponding relaxation times are designated as  $\tau_{\text{TB1}}$ ,  $\tau_{\text{TB2}}$ ,  $\tau_{\text{LB}}$ , and  $\tau_{\text{D}}$ . In the equation above,  $\omega = 2\pi\nu$  is the angular frequency of the applied electric field,  $\epsilon_S$  is the static permittivity

given by  $\epsilon_S = \epsilon_{\infty} + \sum_{i=1}^4 \Delta\epsilon_i$ , and  $\epsilon_{\infty}$  includes contributions to the dielectric response from all polarization modes at frequencies much higher than the probed range.

Employing the method, the dielectric response spectra including the dielectric dispersion,  $\epsilon'_{\text{sol}}(\nu)$ , and the dielectric loss,  $\epsilon''_{\text{sol}}(\nu)$ , are fitted simultaneously to Eq. 2. To obtain the best fit, all eight parameters are varied simultaneously, except  $\tau_{\text{D}}$ , which is released after a few iterations. An example of such fits is presented in Fig. 2 A for the 22.47  $\mu\text{M}$  DNA solution, which confirms that Eq. 2 can adequately describe the dielectric response spectrum. If a model similar to Eq. 2 but with only three or fewer Debye relaxation processes is adopted, a residual is clearly discernible between the fit and the measured response (see Fig. S2). Therefore, we stick to Eq. 2 for fitting the dielectric response data to obtain the relaxation time and strength of each mode. For the 22.47  $\mu\text{M}$  DNA solution, the relaxation time for bulk water in the solution,  $\tau_{\text{D}}$ , is found to be  $8.12 \pm 0.20$  ps, similar to the value obtained in pure water, which is 8.27 ps (i.e., 19.25 GHz) (53,58,59). The three longer relaxation times are  $\tau_{\text{TB1}} = 835 \pm 120$  ps (190 MHz),  $\tau_{\text{TB2}} = 400 \pm 45$  ps (398 MHz), and  $\tau_{\text{LB}} = 52 \pm 8$  ps (3.1 GHz), which are separated for the reorientational relaxation processes of the tightly bound water in the minor grooves, the tightly bound water in the major grooves and around the phosphate groups, and the loosely bound ones, respectively. It is noteworthy that  $\text{Na}^+$  counterions have weak interaction with water molecules with a relaxation time of  $\sim 30$  ps (63,64), and therefore, the hydration water molecules interacting with  $\text{Na}^+$  counterions contribute to the loosely bound water. The corresponding dynamics of the three types of bound water show retardation factors of  $\sim 100$ , 50, and 5, respectively. The tightly bound water molecules are in the innermost hydration shells and have direct interactions with the DNA surface. Such interactions are typically strong, but the coverage of the tightly bound water around a DNA molecule is not uniform because the DNA surface is irregular. Some water molecules, such as those in the major grooves, may have direct but relatively weak interactions with the DNA surface (52,55,60). As a result, the tightly bound water exhibits two distinct relaxation times, though both are significantly longer than the bulk relaxation time. The loosely bound water contains all water molecules having weak interactions with the



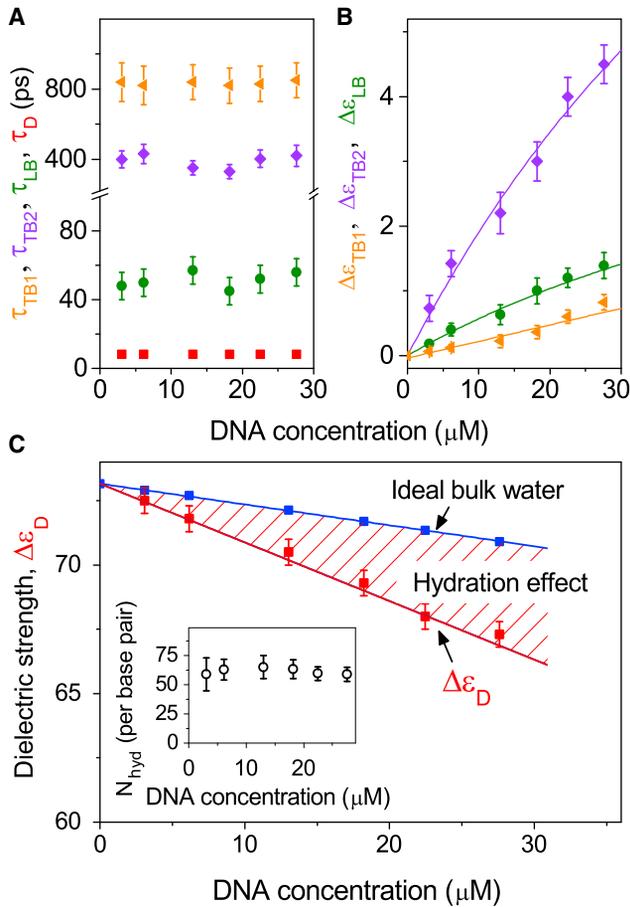
**FIGURE 2** Heterogeneous nature of hydration water. (A) The dielectric loss and dispersion (*inset*) spectra of the 22.47  $\mu\text{M}$  DNA solution reveal the cooperative reorientation dynamics of hydration water. The spectra are deconvoluted into four Debye components, capturing the contributions from the tightly bound water in the minor grooves, the tightly bound water in the major grooves and around the phosphate groups, the loosely bound water, and the bulk water in the solution, respectively. The red curves are the fits to the dielectric response based on these four Debye components from 100 MHz to 50 GHz. (B) The dielectric loss and dispersion (*inset*) spectra, with the contributions from the bulk water subtracted, for the 22.47  $\mu\text{M}$  (green) and 27.58  $\mu\text{M}$  (orange) DNA solutions. The solid lines represent the three-Debye model fit to the contribution of hydration water. To see this figure in color, go online.

charged entities in DNA and  $\text{Na}^+$  counterions, and they are typically located either in the outer hydration shells with respect to DNA surface or interacting with the  $\text{Na}^+$  counterions. Fitting a dielectric spectrum to Eq. 2 also provides the dielectric strength of each relaxation mode, which quantifies the contribution to the overall dielectric response from each group of water molecules. For the 22.47  $\mu\text{M}$  DNA solution,  $\Delta\epsilon_{\text{TB1}} = 0.60 \pm 0.05$ ,  $\Delta\epsilon_{\text{TB2}} = 4.05 \pm 0.30$ ,  $\Delta\epsilon_{\text{LB}} = 1.20 \pm 0.15$ , and  $\Delta\epsilon_{\text{D}} = 68.0 \pm 0.4$ . The fitted values of the relaxation time and dielectric strength of all four types of water molecules in aqueous DNA solutions at different DNA concentrations are included in Table S1. The concen-

trations of  $\text{Na}^+$  counterions at the DNA concentrations probed here typically lie below  $\sim 110$  mM, which is a weak contribution to the dielectric strength of loosely bound water (53).

Dielectric spectroscopy measures the overall response of the entire ensemble of molecules in a solution. As a result, the fitted values in Table S1 represent the macroscopic properties of the tightly bound, loosely bound, and bulk water. To single out the dielectric response solely from the bound water molecules, the dielectric response due to bulk water identified by the dielectric strength,  $\Delta\epsilon_{\text{D}}$ , and the relaxation time,  $\tau_{\text{D}}$ , is subtracted from the full spectrum. The results for two solutions after such subtraction are shown in Fig. 2 B, which represent the contribution to the dielectric response from all bound water molecules in each solution. As expected, each spectrum can be well characterized by a superposition of three Debye relaxation processes, attributed to the reorientation relaxation of three types of bound water. Two of them are tightly bound, with one in the minor grooves of a DNA molecule having the longest relaxation time and the other in the major grooves and around the phosphate groups. The remaining type consists of water molecules loosely bound to the DNA surface and  $\text{Na}^+$  counterions.

The hydration structure of DNA is heterogeneous with the bound water exhibiting various distinct dynamics. The percentage of each type of bound water depends on the concentration of DNA in the solution. To reveal this dependence, dielectric spectroscopic measurements are performed at various DNA concentrations. The relaxation time and the corresponding dielectric strength extracted from experimental data are presented in Fig. 3 against the DNA concentration. The four relaxation times for the tightly bound (either in the minor grooves or in the major grooves and around the phosphate groups), loosely bound, and bulk water remain almost unchanged as the DNA concentration is varied from 3 to 28  $\mu\text{M}$  (Fig. 3 A). The quantitative results are  $\tau_{\text{TB1}} = 830 \pm 120$  ps,  $\tau_{\text{TB2}} = 390 \pm 45$  ps,  $\tau_{\text{LB}} = 52 \pm 8$  ps, and  $\tau_{\text{D}} = 8.20 \pm 0.20$  ps. The corresponding dielectric strengths of the bound water molecules,  $\Delta\epsilon_{\text{TB1}}$ ,  $\Delta\epsilon_{\text{TB2}}$ , and  $\Delta\epsilon_{\text{LB}}$ , increase as the DNA concentration is increased and exhibit a tendency to saturate at high DNA concentrations (Fig. 3 B) (52,56). Accordingly, the dielectric strength of bulk water,  $\Delta\epsilon_{\text{D}}$ , decreases with increasing DNA concentration (Fig. 3 C). As expected, before the hydration shells of different DNA molecules start to overlap, the number of bound water molecules is proportional to the number of DNA molecules in a solution. The increase of DNA concentration is at the cost of the bulk water in the solution. Thus, the dielectric strength of bound water increases while that of bulk water decreases as a DNA solution becomes more concentrated. Furthermore, the replacement of water molecules, which strongly absorb megahertz to terahertz radiations, with low-absorbing biomacromolecules also contributes to the decrease of the dielectric strength of the bulk water.



**FIGURE 3** Mechanistic origins of hydration water. (A) The relaxation times from top to bottom are  $\tau_{\text{TB1}}$  for the tightly bound water in the minor grooves (orange triangles),  $\tau_{\text{TB2}}$  for the tightly bound water in the major grooves and around the phosphates groups (purple diamonds),  $\tau_{\text{LB}}$  for the loosely bound water (green circles), and  $\tau_{\text{D}}$  for the bulk water (red squares), respectively. The error bars for  $\tau_{\text{D}}$  are smaller than the symbol size. (B) The dielectric strengths are  $\Delta\epsilon_{\text{TB1}}$  for the tightly bound water in the minor grooves (orange triangles),  $\Delta\epsilon_{\text{TB2}}$  for the tightly bound water in the major grooves and around the phosphate groups (purple diamonds), and  $\Delta\epsilon_{\text{LB}}$  for the loosely bound water (green circles). The solid lines are a guide to the eye. (C) The dielectric strength of the bulk water,  $\Delta\epsilon_{\text{D}}$ , is either extracted from the experimentally measured dielectric response (red squares) or computed on the basis of the ideal bulk water assumption (blue squares). The solid lines are the best linear fits to the data. The difference between the two indicates the hydration effect and can be used to extract the hydration number per DNA basepair,  $N_{\text{hyd}}$ , which is shown in the inset as a function of DNA concentration. To see this figure in color, go online.

The number of water molecules affected by the presence of DNA in an aqueous solution can be estimated from the dielectric response of bulk water in the solution. As discussed previously, water molecules in a DNA solution that relax with a time constant  $\tau_{\text{D}}$  are considered as bulk water. If all water molecules in the solution are assumed to be bulk water, their dielectric strength can be computed based on the amount of water in the solution, which is the curve labeled as “ideal bulk water” in Fig. 3 C. However, the measured dielectric strength,  $\Delta\epsilon_{\text{D}}$ , of the actual bulk water in the

solution, which is extracted by fitting the dielectric spectrum to Eq. 2, is lower than that of the ideal bulk water. As the DNA concentration is increased, the difference between the two becomes larger. This trend is another manifestation of the fact that not all water molecules in a solution relax as bulk water. A significant fraction of water is bound to DNA directly or indirectly via hydrogen bonding. These bound water molecules have longer relaxation times than  $\tau_{\text{D}}$ , as shown in Fig. 3 A, and their number as well as contribution to the dielectric response of the solution grows as the DNA solution becomes more concentrated. The difference between the dielectric strength computed for the ideal bulk water and that measured for the real bulk water in a solution provides a way to quantify the “hydration effect” (Fig. 3 C). By calculating the number of water molecules not participating in the bulk water reorientation dynamics, the number of bound or “slow” water molecules per DNA basepair, which is called the “hydration number”,  $N_{\text{hyd}}$ , can be estimated as

$$N_{\text{hyd}}(c_{\text{DNA}}) = \frac{c_{\text{w}} - \frac{\Delta\epsilon_{\text{w}}}{\Delta\epsilon_{\text{pure}}} c_{\text{pure}}}{n c_{\text{DNA}}}, \quad (3)$$

where  $c_{\text{w}}$  is the molar concentration of water in the aqueous solution,  $c_{\text{pure}} = 55.35 \text{ M}$  is the molarity,  $\Delta\epsilon_{\text{pure}} = 73.25$  is the dielectric strength of neat water at  $25^\circ\text{C}$ ,  $c_{\text{DNA}}$  is the molar concentration of DNA in the solution, and  $n$  is the number of basepairs per DNA. For the  $22.47 \mu\text{M}$  DNA solution,  $N_{\text{hyd}} = 60 \pm 6$ . As the DNA concentration is increased, the value of  $N_{\text{hyd}}$  slightly decreases (inset of Fig. 3 C). As discussed afterward, these hydration water molecules are heterogeneously dispersed over a length scale of  $18 \text{ \AA}$  from the DNA surface. This result is consistent with our estimate of  $N_{\text{hyd}}$ , which will be discussed further along with our MD simulations of DNA hydration. Our results of the hydration number are close to those reported by Ermilova et al. (39), who estimated  $\sim 40$  water molecules per DNA basepair for solutions with a wide range of DNA concentrations using gigahertz dielectric spectroscopy.

The dielectric strengths of the bound water in aqueous DNA solutions show a tendency to saturation at high DNA concentrations (Fig. 3 B). A slight decrease in the hydration number can also be observed in this limit (inset of Fig. 3 C). At low concentrations, DNA macromolecules are well separated from each other, and the hydration structure and dynamics are governed by DNA-water interactions. When the DNA concentration is increased, the average separation between DNA molecules decreases, and the partial overlapping of the hydration layers of different DNA molecules comes into the picture. The slight bending of the dielectric strengths and the weak decrease in the hydration number at high DNA concentrations suggest that the overlapping of hydration layers is not significant in the range of DNA concentration studied here.

Temperature is expected to strongly influence the activity of water and DNA in their mixture. The absorption and refractive index of the 18.21  $\mu\text{M}$  DNA solution at 5, 15, 25, and 35°C are presented in Fig. S3, and the corresponding dielectric response of the DNA solution is shown in Fig. 4. In general, the dielectric spectra shift toward higher frequencies as temperature is increased. Eq. 2 still adequately describes the dielectric response at all temperatures, and the relaxation time constants and dielectric strengths as a function of temperature are presented in Fig. 4, B, C, and D. All the orientation relaxation times and the dielectric strengths of water molecules in the solution decrease with increasing temperature. These trends are expected as molecular orientations relax faster (that is, with higher frequencies) and the hydration number of DNA decreases at higher temperatures

at which thermal fluctuations of molecular motions are enhanced (65). Using the ideal bulk water assumption, we can also compute the anticipated dielectric strength of all water molecules in the solution at a specific temperature if all of them are assumed to relax in the same manner as pure water at that temperature (the *blue solid line* in Fig. 4 D). The difference between this hypothetical response and the measured response of the real bulk water,  $\Delta\epsilon_D$ , reflects the hydration effect. The results in Fig. 4 D show that the difference of the two responses becomes smaller as temperature is raised, which implies that the fraction of water in the solution relaxing as pure water gets larger at higher temperatures. The hydration number per DNA basepair,  $N_{\text{hyd}}$ , computed on the basis of Eq. 3 as discussed previously, is therefore expected to decrease with increasing temperature.

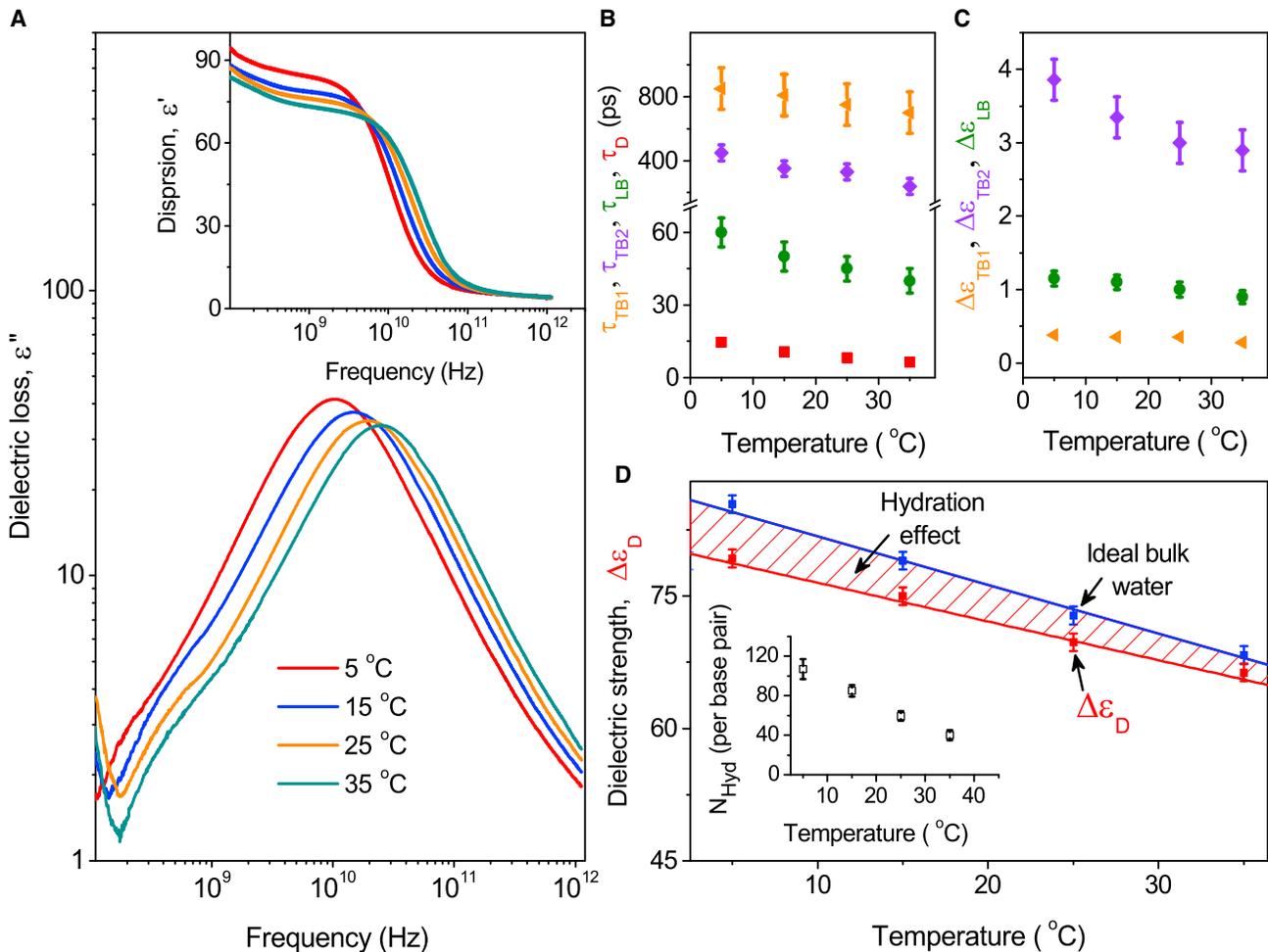


FIGURE 4 Effect of temperature on hydration and DNA dynamics. (A) The dielectric loss and dispersion (*inset*) spectra were collected at 5, 15, 25, and 35°C. (B) The relaxation times from top to bottom are  $\tau_{\text{TB1}}$  for the tightly bound water in the minor grooves (*orange triangles*),  $\tau_{\text{TB2}}$  for the tightly bound water in the major grooves and around the phosphate groups (*purple diamonds*),  $\tau_{\text{LB}}$  for the loosely bound water (*green circles*), and  $\tau_{\text{D}}$  for the bulk water (*red squares*), respectively. (C) The dielectric strengths are  $\Delta\epsilon_{\text{TB1}}$  for the tightly bound water in the minor grooves (*orange triangles*),  $\Delta\epsilon_{\text{TB2}}$  for the tightly bound water in the major grooves and around the phosphate groups (*purple diamonds*), and  $\Delta\epsilon_{\text{LB}}$  for the loosely bound water (*green circles*). (D) The dielectric strength of the bulk water,  $\Delta\epsilon_D$ , is either extracted from the experimentally measured dielectric response (*red squares*) or computed on the basis of the ideal bulk water assumption (*blue squares*). The solid lines are the best linear fits to the data. The inset shows the hydration number,  $N_{\text{hyd}}$ , as a function of temperature. To see this figure in color, go online.

This trend is confirmed by the data shown in the inset of Fig. 4 D.

The observed temperature dependence of the dielectric response can be explained using the concept of Debye screening (66) and counterion condensation (67). The Debye length,  $\lambda_D$ , is a length scale up to which the electrostatic effect of a charged entity persists in a solution. In an ionic solution,  $\lambda_D \sim \sqrt{T/n_0}$ , where  $T$  is temperature and  $n_0$  is the concentration of mobile ions. In a DNA solution, the extent of counterion condensation is controlled by the Manning parameter,  $\xi = \lambda_B/b$ , where  $\lambda_B$  is the Bjerrum length and  $b$  is the average separation between the fixed ions on a DNA chain. Specifically, the fraction of counterions condensed on a DNA chain is given by  $f_c = 1 - \xi_c/\xi$ , where  $\xi_c$  is the critical Manning parameter beyond which counterion condensation starts to occur (67). The Bjerrum length is the separation at which the electrostatic interaction energy between two elementary charges in a solution is equal to the thermal energy scale. Its temperature dependence is  $\lambda_B \sim 1/T$ . When the temperature is raised,  $\lambda_B$  and  $\xi$  get smaller. As a result,  $f_c$  decreases with increasing temperature, implying that more counterions are released to the solution and become mobile at higher temperatures. Therefore,  $n_0$  increases as  $T$  is raised, and its effect on  $\lambda_D$  more than compensates for the effect from the increase of  $T$ . The net result can be that  $\lambda_D$  decreases with an increasing  $T$ . Such a trend has been reported previously for ionic liquids (68,69). The variation of dielectric response with temperature measured in our experiments can therefore be attributed to the enhancement of thermal fluctuations of the mobile ions in a solution as the temperature is raised, which leads to a stronger screening effect on electrostatic interactions. At higher temperatures, the Debye length for DNA molecules thus decreases, and the electrostatic field of a DNA chain decays more rapidly and extends to a smaller distance from its surface. The number of water molecules kinetically influenced by the DNA molecules present in the solution therefore becomes smaller. As a result, the hydration number decreases, and more water molecules behave like bulk water at higher temperatures.

## Terahertz spectroscopy

At terahertz frequencies, electromagnetic waves are coupled to the collective motions of biomolecules present in a solution, including intra- and intermolecular motions of the biomolecules and their immediate hydration layers. In a DNA solution, water molecules form tightly and loosely bound hydration layers around DNA. The tightly bound water molecules have strong and direct contacts with the DNA surface. They become an integral part of the DNA molecule and cannot move easily (52,54,55). The terahertz dielectric spectra of an aqueous DNA solution thus encode information on the low-frequency internal helical vibrations of the DNA-water complexes in the solution, involving the collec-

tive motions of the DNA backbones and side chains and the interactions between DNA and tightly bound hydration water. With the refractive index and absorption coefficient of an aqueous DNA solution simultaneously measured with high precision, it is possible to map out the collective vibrational modes of the hydrated DNA molecules and the number of water molecules tightly bound to DNA.

In view of the dielectric model for a heterogeneous system, a DNA solution can be considered as a mixture of water and hydrated DNA, each of which possesses its own complex dielectric property characterized by  $\epsilon_{\text{wat}}^*$  and  $\epsilon_{\text{hDNA}}^*$ , respectively. The combination of their dielectric responses results in the complex dielectric response of the solution,  $\epsilon_{\text{sol}}^*$ , which is determined from the experimentally measured refractive index,  $n(\nu)$ , and absorption coefficient,  $\alpha(\nu)$  (55,62,70). Because the wavelengths of the probing electromagnetic waves are orders of magnitude larger than the size of a hydrated DNA molecule, each solution can be approximated as an effectively homogeneous medium. Several models based on this effective-medium theory have been proposed in the literature (62,71–73). The Bruggeman model (62,70,71), which is suitable for the high dielectric contrast between DNA macromolecules and water, is adopted here in the following form,

$$f_{\text{hDNA}} \frac{\epsilon_{\text{hDNA}}^* - \epsilon_{\text{sol}}^*}{\epsilon_{\text{hDNA}}^* + 2\epsilon_{\text{sol}}^*} + (1 - f_{\text{hDNA}}) \frac{\epsilon_{\text{wat}}^* - \epsilon_{\text{sol}}^*}{\epsilon_{\text{wat}}^* + 2\epsilon_{\text{sol}}^*} = 0, \quad (4)$$

where  $f_{\text{hDNA}}$  is the volumetric fraction of hydrated DNA in a solution. The model can be applied in a wide range of solute concentration under certain conditions. Firstly, a fraction of water molecules is embedded in the hydration shells of DNA macromolecules via strong hydrogen bonds. These water molecules are considered as an integral part of the DNA macromolecules in the aqueous solution. This assumption is in line with the experimental results from Raman spectroscopy (8,33), swelling experiments (7), and infrared measurements (47), which all show that a small fraction of water molecules within the first hydration shell of a DNA molecule are strongly bound to the DNA surface and kinetically distinct from the rest of water molecules in the solution. Secondly, a DNA molecule is approximated as a cylinder of radius  $R_D$  and length  $h$ . Each cylinder is surrounded by water molecules that are an integral part of the DNA. Such water molecules are embedded within a cylindrical shell of an average thickness of  $d$  surrounding the cylinder representing the DNA molecule. The volumetric fraction of the hydrated DNA in the solution is, therefore,  $f_{\text{hDNA}} = (N_D/V)(\pi h)(R_D + d)^2$ , where  $N_D$  is the number of DNA molecules and  $N_D/V$  is thus the concentration of DNA in the solution (62). Lastly, the molecular properties of water molecules in the hydration sheath are assumed to be similar to those of bulk water, but their dynamics are kinetically retarded.

Employing the Bruggeman effective-medium approximation, the dielectric dispersion and loss of hydrated DNA as well as the number of water molecules in the tightly bound hydration layer of DNA have been estimated at various temperatures (Fig. 5 A). The average number of water molecules that are tightly bound to DNA and therefore considered as its integral part is estimated to be  $27 \pm 4$  per basepair, corresponding to a tightly bound hydration layer of a thickness of  $4.5 \pm 0.3$  Å from the DNA surface. This value remains unchanged in the temperature range investigated here. These water molecules are not kinetically frozen but are highly enslaved and participated in the DNA collective motions. The tightly bound water molecules have direct and strong contact with the DNA surface including the minor grooves, the major grooves, and the phosphate groups. With the dielectric response measured at megahertz to gigahertz frequencies, we can estimate the contributions of

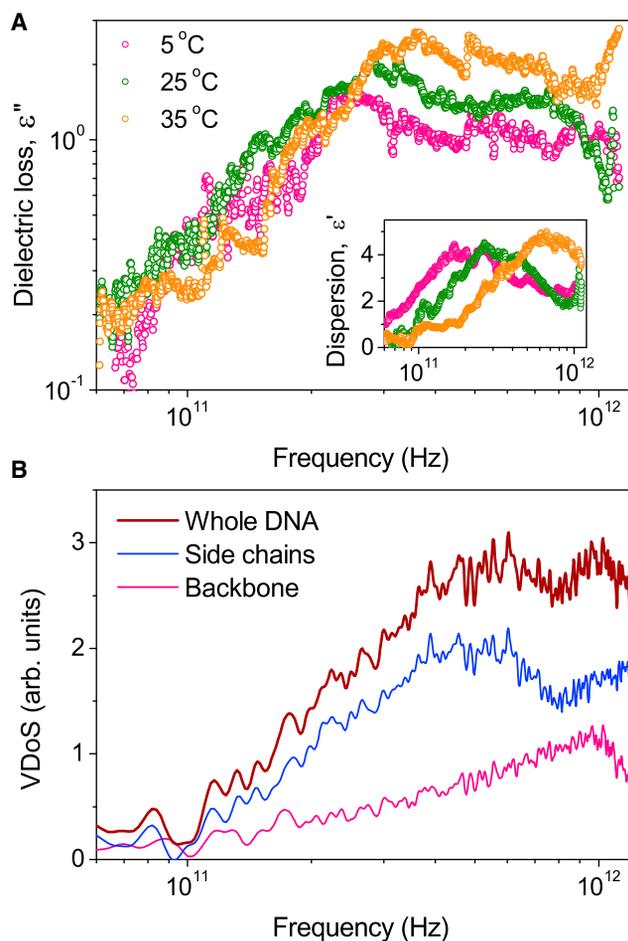


FIGURE 5 Collective vibrational modes of hydrated DNA. (A) The dielectric loss and dispersion (inset) spectra of hydrated DNA in the  $18.21 \mu\text{M}$  DNA solution, estimated using the Bruggeman effective-medium approximation, are shown at 5, 25, and  $35^\circ\text{C}$ . As the temperature is increased, the spectra shift toward higher frequencies. (B) From top to bottom, the VDoS for the whole DNA, the side chains, and the backbone are calculated separately with MD simulations. To see this figure in color, go online.

different types of water to the total dielectric response of a DNA solution. Specifically, the dielectric strength of the tightly bound water molecules in the minor grooves is  $\sim 20\%$  of that of the water molecules in the major grooves and around the phosphate groups. We conclude that  $\sim 4$ – $6$  water molecules per basepair, which have a longer relaxation time of  $\sim 800$  ps, are located in the minor grooves, and  $\sim 20$ – $24$  water molecules per basepair reside in the major grooves and around the phosphate groups. Using different techniques, several earlier studies reported that  $\sim 20$ – $30$  water molecules per basepair are in the innermost hydration shell of a DNA molecule, including 2–6 water molecules that are difficult to remove. Liu et al. (47) estimated that  $\sim 20$ – $30$  water molecules per basepair, including two water molecules permanently bound to DNA, are in the primary hydration shell of DNA using femtosecond pump-probe experiments conducted on moist samples. By employing Raman scattering spectroscopy, Tao et al. (33) predicted that the first hydration shell of DNA contains  $\sim 30$  water molecules per basepair, of which 5–6 water molecules are very difficult to remove from DNA. Lavallo et al. (7) reported that  $\sim 6$  water molecules per basepair are strongly embedded in the B-form of Na-DNA using a swelling technique. All these results suggest that  $\sim 5$ – $6$  water molecules per basepair are integrally attached to DNA in an aqueous solution, forming the “spine of hydration” around the minor grooves and playing a vital role in stabilizing the DNA structure (12). To fully reveal the temporal dynamics and spatial distribution of water molecules around DNA at molecular scales, MD simulations are performed, as discussed in the next section.

### Molecular dynamics simulations

To establish a complete picture of hydrated DNA at the molecular level, we investigate the hydration structure and dynamics of DNA with MD simulations. The goal here is to connect the simulation results to the experimentally measured dielectric responses and provide a microscopic picture of the hydration dynamics and collective motions of DNA. The details of the simulations are included in the [Materials and methods](#) and [Supporting materials and methods](#). We compute the vibrational density of states (VDoS) of a hydrated DNA chain, a quantity characterizing the collective vibrational modes of the chain, and then we separate the contributions from its sugar-phosphate backbone and side chains. The results of a dilute DNA solution at  $25^\circ\text{C}$  are shown in Fig. 5 B. The VDoS spectrum of the backbone has peak intensity at a frequency around 1 THz, whereas that of the side chains has a slightly broader peak at  $\sim 0.5$  THz. As a result, the VDoS spectrum of the whole DNA chain exhibits two peaks at frequencies around 0.5 and 1 THz, respectively. These characteristic frequencies are in line with the frequencies of 0.35 and 0.8 THz at which the dielectric loss spectrum displays peak intensities (Fig. 5 A).

The relaxation time of water is expected to be different in the minor or major grooves of DNA. Usually, a radial distribution function can be used to quantify the water distribution at a distance  $r$  from a macromolecule surface. However, the DNA surface is irregular, and the hydration shells of DNA can be better quantified with a proximal radial distribution function,  $\text{pRDF}(r)$ , which determines the water distribution at a distance  $r$  along a local normal direction from the DNA surface (15). We compute the  $\text{pRDF}(r)$  for water molecules in the minor and major grooves of DNA and around the phosphate groups by employing a grid-bin algorithm introduced by Makarov et al. (74). The technical details on the construction of a DNA surface and the determination of a distance profile,  $r$ , from that surface are discussed in the [Supporting materials and methods](#). The  $\text{pRDF}(r)$  curves for water molecules surrounding DNA at 25°C show that the distribution of water is significantly different in the minor and major grooves as well as around the phosphate groups (Fig. 6 A). The narrower space of the minor grooves provides a stronger confinement on water molecules, resulting in a sharp and distinct peak for the first hydration layer at  $r \cong 3 \text{ \AA}$  and a much weaker peak for the second hydration shell at  $r \cong 5 \text{ \AA}$ . The  $\text{pRDF}(r)$  curve for water molecules in the major grooves shows a broad peak at  $r \cong 3 \text{ \AA}$  and then oscillates slightly before it saturates at large values of  $r$ , indicating a more dispersed distribution of water molecules. For the phosphate groups, a more distinct hydration structure can be observed with the first hydration layer at a larger distance,  $r \cong 3.8 \text{ \AA}$ , followed by a valley of  $\text{pRDF}(r)$  around  $r \cong 4.4 \text{ \AA}$ . Two more weaker but still distinct peaks can be seen at  $r \cong 5.6$  and  $r \cong 7.8 \text{ \AA}$ . However, the more distinct layering of water around the phosphate groups does not mean that water molecules are more strongly bound to them and is just a reflection of the more regular surface of a phosphate group compared to the surface of the minor and major grooves.

Water molecules in the minor and major grooves as well as around the phosphate groups show distinct dynamics in their reorientational relaxation process. The relaxation can be quantified through the orientational autocorrelation function,  $C(t)$ , as explained in detail in the [Supporting materials and methods](#). The long-time correlation function of water molecules in the minor and major grooves and around the phosphates exhibits a monoexponential behavior, with a correlation time of  $1125 \pm 50$ ,  $419 \pm 50$ , and  $330 \pm 45$  ps for each group, respectively (Fig. 6 B). These results indicate that water molecules in the minor grooves are the most tightly bound to DNA and thus relax with the longest reorientation time, whereas those in the major grooves and around the phosphate groups have comparable relaxation times and are relatively less tightly bound. The simulations thus corroborate our early analysis based on the dielectric response measurements, in which two types of tightly bound water are identified with relaxation times around 830 and

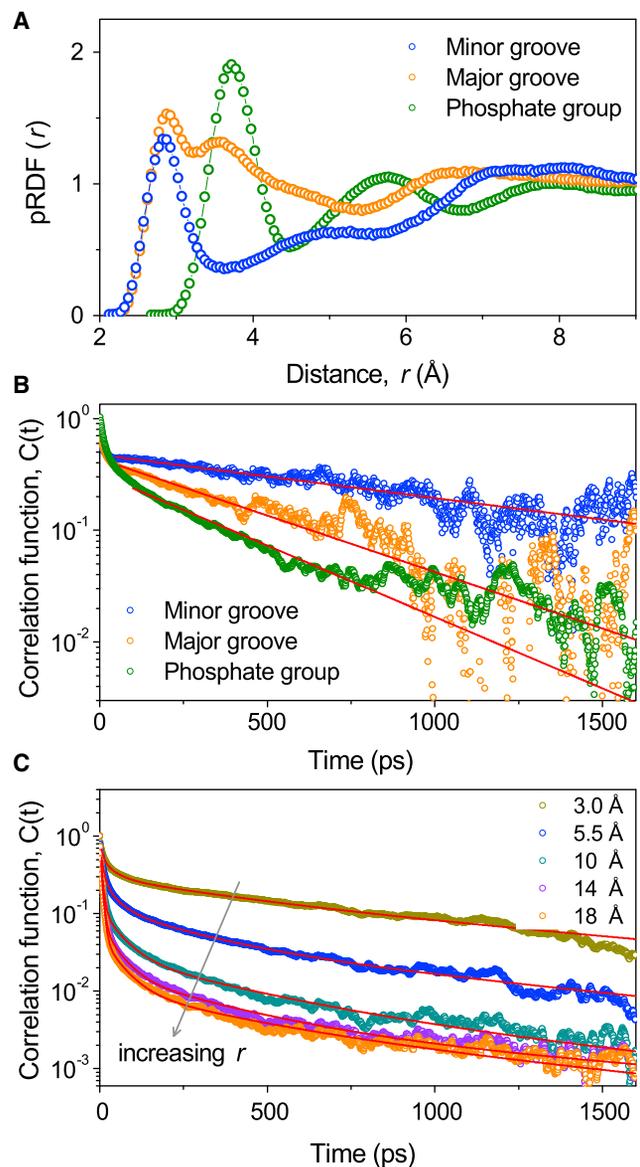


FIGURE 6 Dynamics of bound water molecules at the molecular level. (A) The proximal radial distribution function,  $\text{pRDF}(r)$ , is computed for the tightly bound water molecules in the minor grooves, in the major grooves, and around the phosphate groups, where  $r$  is the distance from the DNA surface. (B) From top to bottom, the orientational autocorrelation function for the tightly bound water molecules in the minor grooves, in the major grooves, and around the phosphate groups. The solid lines are the monoexponential fits to  $C(t)$ , revealing relaxation times of  $1125 \pm 50$ ,  $419 \pm 50$ , and  $330 \pm 45$  ps, for the three types of tightly bound water, respectively. (C) The orientational autocorrelation functions,  $C(t)$ , for water molecules within various distances from the DNA surface. The solid lines represent the fits to  $C(t)$  using Eq. 5. To see this figure in color, go online.

390 ps at 25°C (Table S1), respectively. These results are in line with a highly nonexponential behavior due to multiple contributions to hydration water dynamics ranging from picosecond to nanosecond timescales reported by other groups (13,16,19–22,35).

To understand the extent of the hydration layer around DNA, the orientational autocorrelation function,  $C(t)$ , is computed for water molecules located in a cylindrical shell extending from the DNA surface to a given distance,  $r$ , from that surface. Results on  $C(t)$  are shown for  $r = 3.0, 5.5, 10, 14,$  and  $18 \text{ \AA}$  with a dilute DNA solution at  $25^\circ\text{C}$  (Fig. 6 C). As the value of  $r$  increases,  $C(t)$  decays faster with  $t$ , indicating that the average relaxation time gets shorter. Note that water molecules further away from the DNA surface have weaker interactions with the DNA molecule and therefore relax faster. With a larger value of  $r$ , more of such water molecules are included in the calculation of  $C(t)$ , and as a result, the average relaxation time decreases with increasing  $r$ .

To make the analysis quantitative, we fit  $C(t)$  to a superposition of four exponential functions as in

$$C(t) = A\exp\left(-\frac{t}{\tau_1}\right) + B\exp\left(-\frac{t}{\tau_2}\right) + C\exp\left(-\frac{t}{\tau_3}\right) + D\exp\left(-\frac{t}{\tau_4}\right), \quad (5)$$

where  $A, B, C,$  and  $D$  are the weight coefficients and  $\tau_1, \tau_2, \tau_3,$  and  $\tau_4$  are the relaxation time constants. This choice of the fitting function is motivated by the fact that the dielectric spectroscopic results reveal four types of water molecules in a DNA solution. The choice is corroborated by the finding that all the results in Fig. 6 C can be fitted to Eq. 5 with  $\tau_1 = 1200 \pm 130, \tau_2 = 335 \pm 35, \tau_3 = 54 \pm 5,$  and  $\tau_4 = 8.5 \pm 1 \text{ ps}$  at  $r = 5.5 \text{ \AA}$ . These relaxation times are in line with the experimental results discussed previously, confirming that it is plausible to classify water molecules in a DNA solution into four groups on the basis of their relaxation characteristics. The contribution of each type of water molecules to  $C(t)$  varies with  $r$ , as shown in Fig. 6 C. For example, at  $r = 3 \text{ \AA}$ , the orientational correlation function is dominated by water molecules relaxing with  $\tau_1$  and  $\tau_2$ . The contribution from the  $\tau_2$  mode appears to be significant until  $r = 5.5 \text{ \AA}$ , beyond which the relaxation modes with shorter relaxation times start to come into play. The shortest relaxation time of water molecules bound to DNA,  $\tau_3$ , is  $54 \pm 5 \text{ ps}$ , and this relaxation mode is effective all the way until  $r = 18 \text{ \AA}$ , indicating that the hydration layer around a DNA extends to  $\sim 18 \text{ \AA}$  from the DNA surface. Water molecules beyond  $18 \text{ \AA}$  from the DNA surface relax as bulk water with  $\tau_4 = 7.5 \pm 1 \text{ ps}$ .

## DISCUSSION

Combining the simulation and experimental results, water molecules within the hydration sheath of DNA can be subdivided into three categories based on their relaxation characteristics, which are controlled by the water-DNA interactions. The slowest component corresponds to those “long-lived”

water molecules that are quite firmly embedded in the minor grooves. The orientational relaxation time of these tightly bound water molecules is on the order of  $\sim 1 \text{ ns}$ , as revealed by both MD simulations and dielectric spectroscopy. The second component is the tightly bound water molecules with a relaxation time around  $400 \text{ ps}$ . It is noteworthy that the macroscopic analysis via the correlation function approach and the experimental dielectric response may not distinguish the temporal disparity between water molecules in the major grooves and those near the phosphate groups. Therefore, the second slowest component is attributed to an aggregate response from the water molecules directly interacting with the major grooves and phosphates. Based on the MD data, the water molecules in the minor and major grooves and around the phosphate groups amount to  $24 \pm 3$  water molecules per DNA basepair, of which  $4 \pm 1$  water molecules are in the minor grooves. They constitute the tightly bound hydration layer around DNA and become an integral part of the DNA. The third group includes water molecules outside the tightly bound hydration layer. They are loosely bound to DNA and the  $\text{Na}^+$  counterions and have a relaxation time of  $\sim 50 \text{ ps}$ . Both the terahertz spectroscopic data and the MD simulation results suggest that such water molecules are distributed in a spatial region as far as  $18 \text{ \AA}$  from the DNA surface. The implication is that the electrostatic field of DNA can extend to a relatively large distance from its surface, despite the screening effect of mobile ions in the solution. Finally, the relaxation time of bulk water is found to be around  $7 \pm 1 \text{ ps}$  from both the dielectric response data and the simulation results. Thus, our results from MD simulations and dielectric spectroscopy simultaneously show that the influence of DNA can extend as far as  $18 \text{ \AA}$  away for the surface. In comparison with DNA, the length scale of electrostatic perturbations induced by proteins has been widely debated lately (55,75–77). The evidence is there for long-range (76,77) and short-range (55,75) protein-water interactions. It is noteworthy that nucleic acids are highly negatively charged as compared to conventional proteins and therefore can influence the dynamics of water molecules more effectively (78). For example, the kinetic retardation factor is  $\sim 100$  for the slowest water in the aqueous DNA solutions studied here, which is about twice the value measured in the aqueous solutions of conventional globular proteins (55).

The orientational correlation function is also computed for water molecules within  $5.5 \text{ \AA}$  from the DNA surface at  $5, 15, 25,$  and  $35^\circ\text{C}$ , as shown in Fig. S4 A. As expected, the relaxation is faster at higher temperatures. By fitting the  $C(t)$  curves to Eq. 5, four well-separated characteristic relaxation times can be extracted at each temperature (open symbols in Fig. S4 B). All four relaxation times decrease with an increasing temperature. The values are generally in line with the relaxation times identified early with the dielectric spectroscopic measurements (solid symbols in Fig. S4 B) for the tightly bound water molecules in the minor grooves, those in the major grooves and around

the phosphate groups, the loosely bound water molecules, and the bulk water in a DNA solution. For example, the data of  $\tau_2$  and  $\tau_{TB2}$  and those of  $\tau_4$  and  $\tau_D$  almost overlap at all temperatures. Although  $\tau_1$  from the simulation is  $\sim 30\%$  larger than  $\tau_{TB1}$  from the experiment and  $\tau_3$  is slightly larger than  $\tau_{LB}$ , both show similar temperature dependences. Considering that the force field parameters adopted in the MD simulations are not optimized for the DNA molecules probed in the experiments reported here, it is certainly satisfying that the simulation results on the relaxation times are close to the experimental values, and more importantly, the two sets of data exhibit the same hierarchy in the range of temperature explored here.

The collective vibrational modes of hydrated DNA depend strongly on the environmental temperature. The dielectric response reflects the collective vibrational motions of hydrated DNA, including inter- and intramolecular contributions, both of which are temperature dependent. Specifically, the locations of the peak intensities, which are correlated to the characteristic frequencies of the collective motions of hydrated DNA, shift toward higher frequencies at higher temperatures. This trend is also clearly reflected by the dispersion spectra shown in the inset of Fig. 5 A. This temperature dependence can be understood on the basis that the thermally activated vibrational fluctuations of hydrated DNA, including the water molecules tightly bound to the DNA chain, are enhanced when temperature is raised, and therefore, the peak intensities shift toward higher frequencies. The findings are in close agreement to those reported by Urabe et al. (34), who detected the collective vibrational motions of hydrated B-DNA using low-frequency Raman spectroscopy. They reported that the slowest collective vibrational modes of a DNA chain together with its tightly bound water molecules occur at frequencies around 0.48 THz ( $\sim 16 \text{ cm}^{-1}$ ) and 2.5 THz ( $\sim 85 \text{ cm}^{-1}$ ). The two collective modes shown in Fig. 5 A at terahertz frequencies from the dielectric spectra can be associated to the vibrational motions of the side chains and the DNA backbone, as revealed by the MD simulations discussed below. Using terahertz time-domain spectroscopy, several groups have probed the collective vibrational motions of DNA in moist or gel samples (45,79,80), in which DNA molecules are not fully covered with water. A broad absorption peak at terahertz frequencies has been observed for moist DNA, but no quantitative information of water molecules coupled to the collective motions of hydrated DNA has been provided in these studies. Here, we have identified the collective vibrational modes of DNA and hydration water at terahertz frequencies, and these motions are sensitive to temperature.

## CONCLUSIONS

In conclusion, the conformations and functions of a DNA chain strongly depend on its hydration level. The hydration

layers provide channels for efficient and rapid energy dissipation resulting from ultraviolet photon absorption and thus protect the delicate double-helix structure of the DNA. Employing a high-precision megahertz to terahertz dielectric spectrometer and MD simulations, we have investigated the hydration structure of DNA, the cooperative reorientation dynamics of water molecules, as well as the collective motions of hydrated DNA in aqueous DNA solutions. We have probed the concentration and temperature dependence of the hydration number of DNA and the orientational relaxation times of hydration water. A hierarchy of four distinct relaxation times has been observed, indicating the heterogeneous nature of the dynamics of hydration water surrounding DNA. The strong electrostatic affinity of water molecules to the DNA surface makes the dynamics of hydration water significantly retarded as compared to bulk water. The dielectric spectroscopic measurements and MD simulations synergistically suggest that the water molecules hydrating a DNA chain can be separated into three groups. These include the tightly bound water in the minor grooves with a relaxation time around 1.0 ns, the tightly bound water molecules in the major grooves and those bound to the phosphate groups with a relaxation time around 390 ps, and the loosely bound water with a relaxation time around 52 ps. The hydration water is distributed over a spatial region extending to as far as 18 Å from the DNA surface. The hydration number, including all water molecules retarded kinetically by the presence of DNA, is estimated to be  $60 \pm 6$  per DNA basepair. The terahertz dielectric response data, analyzed on the basis of the effective-medium approximation, suggest that  $\sim 27 \pm 4$  water molecules per basepair are tightly bound to DNA, of which around five water molecules per basepair are located in the minor grooves, whereas the rest are mainly in the major grooves and around the phosphate groups. The water molecules in the minor grooves can figuratively be identified as the “long-lived” water molecules that form the “spine of hydration” of DNA. The terahertz dielectric spectroscopy also allows us to map out the collective vibrational response of hydrated DNA and serves as a promising tool in general to provide fresh insight into the collective motions of hydrated biomacromolecules.

## SUPPORTING MATERIAL

Supporting material can be found online at <https://doi.org/10.1016/j.bpj.2021.10.016>.

## AUTHOR CONTRIBUTIONS

N.Q.V. and A.K.S. designed the research program. A.K.S. performed the terahertz experiments supervised by N.Q.V. C.W. performed the MD simulations supervised by N.Q.V. and S.C. All the authors discussed the results on the manuscript, and N.Q.V., A.K.S., and S.C. wrote the manuscript.

## ACKNOWLEDGMENTS

Authors gratefully acknowledge financial support by the Air Force Office of Scientific Research, United States under award number FA9550-18-1-0263 and the National Science Foundation, United States (CHE-1665157). We acknowledge Advanced Research Computing at Virginia Tech for providing computational resources and technical support that have contributed to the results reported within this manuscript.

## REFERENCES

- Sinden, R. R. 2012. DNA Structure and Function, First Edition. Academic Press, San Diego, CA.
- Wood, B. R. 2016. The importance of hydration and DNA conformation in interpreting infrared spectra of cells and tissues. *Chem. Soc. Rev.* 45:1980–1998.
- Saenger, W., W. N. Hunter, and O. Kennard. 1986. DNA conformation is determined by economics in the hydration of phosphate groups. *Nature.* 324:385–388.
- Laage, D., T. Elsaesser, and J. T. Hynes. 2017. Water dynamics in the hydration shells of biomolecules. *Chem. Rev.* 117:10694–10725.
- Reddy, C. K., A. Das, and B. Jayaram. 2001. Do water molecules mediate protein-DNA recognition? *J. Mol. Biol.* 314:619–632.
- Wilhelm, M., A. Mukherjee, ..., R. Lavery. 2012. Multistep drug intercalation: molecular dynamics and free energy studies of the binding of daunomycin to DNA. *J. Am. Chem. Soc.* 134:8588–8596.
- Lavalle, N., S. A. Lee, and A. Rupprecht. 1990. Counterion effects on the physical properties and the A to B transition of calf-thymus DNA films. *Biopolymers.* 30:877–887.
- Lee, S. A., N.-J. Tao, and A. Rupprecht. 2013. A Raman scattering study of the interactions of DNA with its water of hydration. *J. Biomol. Struct. Dyn.* 31:1337–1342.
- Umehara, T., S. Kuwabara, ..., S. Yagihara. 1990. Dielectric study on hydration of B-, A-, and Z-DNA. *Biopolymers.* 30:649–656.
- Kochoyan, M., and J. L. Leroy. 1995. Hydration and solution structure of nucleic acids. *Curr. Opin. Struct. Biol.* 5:329–333.
- Berman, H. M. 1994. Hydration of DNA - take 2. *Curr. Opin. Struct. Biol.* 4:345–350.
- McDermott, M. L., H. Vanselow, ..., P. B. Petersen. 2017. DNA's chiral spine of hydration. *ACS Cent. Sci.* 3:708–714.
- Mukherjee, S., S. Mondal, ..., B. Bagchi. 2018. DNA solvation dynamics. *J. Phys. Chem. B.* 122:11743–11761.
- Duboué-Dijon, E., A. C. Fogarty, ..., D. Laage. 2016. Dynamical disorder in the DNA hydration shell. *J. Am. Chem. Soc.* 138:7610–7620.
- Saha, D., S. Supekar, and A. Mukherjee. 2015. Distribution of residence time of water around DNA base pairs: governing factors and the origin of heterogeneity. *J. Phys. Chem. B.* 119:11371–11381.
- Sardana, D., K. Yadav, ..., S. Sen. 2019. Origin of slow solvation dynamics in DNA: DAPI in minor groove of dickerson-drew DNA. *J. Phys. Chem. B.* 123:10202–10216.
- Auffinger, P., and E. Westhof. 2000. Water and ion binding around RNA and DNA (C,G) oligomers. *J. Mol. Biol.* 300:1113–1131.
- Oosawa, F. 1970. Counterion fluctuation and dielectric dispersion in linear polyelectrolytes. *Biopolymers.* 9:677–688.
- Jana, B., S. Pal, ..., B. Bagchi. 2006. Entropy of water in the hydration layer of major and minor grooves of DNA. *J. Phys. Chem. B.* 110:19611–19618.
- Pal, S., P. K. Maiti, ..., J. T. Hynes. 2006. Multiple time scales in solvation dynamics of DNA in aqueous solution: the role of water, counterions, and cross-correlations. *J. Phys. Chem. B.* 110:26396–26402.
- Bagchi, B. 2014. Anomalous power law decay in solvation dynamics of DNA: a mode coupling theory analysis of ion contribution. *Mol. Phys.* 112:1418–1426.
- Andreatta, D., J. L. Pérez Lustres, ..., M. A. Berg. 2005. Power-law solvation dynamics in DNA over six decades in time. *J. Am. Chem. Soc.* 127:7270–7271.
- Kopka, M. L., A. V. Fratini, ..., R. E. Dickerson. 1983. Ordered water structure around a B-DNA dodecamer. A quantitative study. *J. Mol. Biol.* 163:129–146.
- Schneider, B., D. Cohen, and H. M. Berman. 1992. Hydration of DNA bases: analysis of crystallographic data. *Biopolymers.* 32:725–750.
- Egli, M., V. Tereshko, ..., M. Manoharan. 1998. X-ray crystallographic analysis of the hydration of A- and B-form DNA at atomic resolution. *Biopolymers.* 48:234–252.
- Liepinsh, E., G. Otting, and K. Wüthrich. 1992. NMR observation of individual molecules of hydration water bound to DNA duplexes: direct evidence for a spine of hydration water present in aqueous solution. *Nucleic Acids Res.* 20:6549–6553.
- Zhou, D., and R. G. Bryant. 1996. Water molecule binding and lifetimes on the DNA duplex d(CGCGAATTCGCG)2. *J. Biomol. NMR.* 8:77–86.
- Johannesson, H., and B. Halle. 1998. Minor groove hydration of DNA in solution: dependence on base composition and sequence. *J. Am. Chem. Soc.* 120:6859–6870.
- Halle, B., and V. P. Denisov. 1998. Water and monovalent ions in the minor groove of B-DNA oligonucleotides as seen by NMR. *Biopolymers.* 48:210–233.
- Franck, J. M., Y. Ding, ..., S. Han. 2015. Anomalously rapid hydration water diffusion dynamics near DNA surfaces. *J. Am. Chem. Soc.* 137:12013–12023.
- Bastos, M., V. Castro, ..., J. Teixeira. 2004. Hydration of ds-DNA and ss-DNA by neutron quasielastic scattering. *Biophys. J.* 86:3822–3827.
- Nakagawa, H., Y. Yonetani, ..., H. Kono. 2014. Local dynamics coupled to hydration water determines DNA-sequence-dependent deformability. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 90:022723.
- Tao, N. J., S. M. Lindsay, and A. Rupprecht. 1989. Structure of DNA hydration shells studied by Raman spectroscopy. *Biopolymers.* 28:1019–1030.
- Urabe, H., H. Hayashi, ..., M. Tsuboi. 1985. Collective vibrational-modes in molecular assembly of DNA and its application to biological-systems - low-frequency Raman-spectroscopy. *J. Chem. Phys.* 82:531–535.
- Shweta, H., and S. Sen. 2018. Dynamics of water and ions around DNA: what is so special about them? *J. Biosci.* 43:499–518.
- Sen, S., D. Andreatta, ..., M. A. Berg. 2009. Dynamics of water and ions near DNA: comparison of simulation to time-resolved Stokes-shift experiments. *J. Am. Chem. Soc.* 131:1724–1735.
- Berg, M. A., R. S. Coleman, and C. J. Murphy. 2008. Nanoscale structure and dynamics of DNA. *Phys. Chem. Chem. Phys.* 10:1229–1242.
- Pal, S. K., L. Zhao, and A. H. Zewail. 2003. Water at DNA surfaces: ultrafast dynamics in minor groove recognition. *Proc. Natl. Acad. Sci. USA.* 100:8113–8118.
- Ermilova, E., F. F. Bier, and R. Hölzel. 2014. Dielectric measurements of aqueous DNA solutions up to 110 GHz. *Phys. Chem. Chem. Phys.* 16:11256–11264.
- Saif, B., R. K. Mohr, ..., T. A. Litovitz. 1991. On the mechanism of dielectric relaxation in aqueous DNA solutions. *Biopolymers.* 31:1171–1180.
- Glancy, P. 2015. Concentration-dependent effects on fully hydrated DNA at terahertz frequencies. *J. Biol. Phys.* 41:247–256.
- Fischer, B. M., M. Walther, and P. Uhd Jepsen. 2002. Far-infrared vibrational modes of DNA components studied by terahertz time-domain spectroscopy. *Phys. Med. Biol.* 47:3807–3814.
- Niessen, K. A., M. Xu, ..., A. G. Markelz. 2019. Protein and RNA dynamical fingerprinting. *Nat. Commun.* 10:1026.
- Polley, D., A. Patra, and R. K. Mitra. 2013. Dielectric relaxation of the extended hydration sheath of DNA in the THz frequency region. *Chem. Phys. Lett.* 586:143–147.

45. Markelz, A. G., A. Roitberg, and E. J. Heilweil. 2000. Pulsed terahertz spectroscopy of DNA, bovine serum albumin and collagen between 0.1 and 2.0 THz. *Chem. Phys. Lett.* 320:42–48.
46. Siebert, T., B. Guchhait, ..., T. Elsaesser. 2016. Range, magnitude, and ultrafast dynamics of electric fields at the hydrated DNA surface. *J. Phys. Chem. Lett.* 7:3131–3136.
47. Liu, Y., B. Guchhait, ..., T. Elsaesser. 2017. Molecular couplings and energy exchange between DNA and water mapped by femtosecond infrared spectroscopy of backbone vibrations. *Struct. Dyn.* 4:044015.
48. Campen, R. K., T. T. Ngo, ..., M. Bonn. 2010. Molecular restructuring of water and lipids upon the interaction of DNA with lipid monolayers. *J. Am. Chem. Soc.* 132:8037–8047.
49. González-Jiménez, M., G. Ramakrishnan, ..., K. Wynne. 2016. Observation of coherent delocalized phonon-like modes in DNA under physiological conditions. *Nat. Commun.* 7:11799.
50. Tomić, S., S. D. Babić, ..., R. Podgornik. 2007. Dielectric relaxation of DNA aqueous solutions. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 75:021905.
51. George, D. K., A. Charkhesht, and N. Q. Vinh. 2015. New terahertz dielectric spectroscopy for the study of aqueous solutions. *Rev. Sci. Instrum.* 86:123105.
52. George, D. K., A. Charkhesht, ..., N. Q. Vinh. 2016. New insights into the dynamics of zwitterionic micelles and their hydration waters by gigahertz-to-terahertz dielectric spectroscopy. *J. Phys. Chem. B.* 120:10757–10767.
53. Vinh, N. Q., M. S. Sherwin, ..., K. W. Plaxco. 2015. High-precision gigahertz-to-terahertz spectroscopy of aqueous salt solutions as a probe of the femtosecond-to-picosecond dynamics of liquid water. *J. Chem. Phys.* 142:164502.
54. Vinh, N. Q., S. J. Allen, and K. W. Plaxco. 2011. Dielectric spectroscopy of proteins as a quantitative experimental test of computational models of their low-frequency harmonic motions. *J. Am. Chem. Soc.* 133:8942–8947.
55. Charkhesht, A., C. K. Regmi, ..., N. Q. Vinh. 2018. High-precision megahertz-to-terahertz dielectric spectroscopy of protein collective motions and hydration dynamics. *J. Phys. Chem. B.* 122:6341–6350.
56. Charkhesht, A., D. Lou, ..., N. Q. Vinh. 2019. Insights into hydration dynamics and cooperative interactions in glycerol-water mixtures by terahertz dielectric spectroscopy. *J. Phys. Chem. B.* 123:8791–8799.
57. Plimpton, S. 1995. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* 117:1–19.
58. Ellison, W. J. 2007. Permittivity of pure water, at standard atmospheric pressure, over the frequency range 0–25 THz and the temperature range 0–100°C. *J. Phys. Chem. Ref. Data.* 36:1–18.
59. Buchner, R., and G. Hefter. 2009. Interactions and dynamics in electrolyte solutions by dielectric spectroscopy. *Phys. Chem. Chem. Phys.* 11:8984–8999.
60. Cametti, C., S. Marchetti, ..., G. Onori. 2011. Dielectric relaxation spectroscopy of lysozyme aqueous solutions: analysis of the  $\delta$ -dispersion and the contribution of the hydration water. *J. Phys. Chem. B.* 115:7144–7153.
61. Debye, P. 1929. *Polar Molecules*. The Chemical Catalogue Company, New York.
62. Choy, T. C. 1999. *Effective Medium Theory: Principles and Applications*. Clarendon Press, Oxford, UK.
63. Chowdhuri, S., and A. Chandra. 2001. Molecular dynamics simulations of aqueous NaCl and KCl solutions: effects of ion concentration on the single-particle, pair, and collective dynamical properties of ions and water molecules. *J. Chem. Phys.* 115:3732–3741.
64. Koneshan, S., J. C. Rasaiah, ..., S. H. Lee. 1998. Solvent structure, dynamics, and ion mobility in aqueous solutions at 25°C. *J. Phys. Chem. B.* 102:4193–4204.
65. Harrison, G. 1976. *The Dynamic Properties of Supercooled Liquids*. Academic Press, London, UK.
66. Debye, P., and E. Hückel. 1923. The theory of electrolytes. I. Freezing point depression and related phenomena. *Phys. Z.* 24:185–206.
67. Manning, G. S. 1969. Limiting laws and counterion condensation in polyelectrolyte solutions. I. Colligative properties. *J. Chem. Phys.* 51:924–933.
68. Gebbie, M. A., H. A. Dobbs, ..., J. N. Israelachvili. 2015. Long-range electrostatic screening in ionic liquids. *Proc. Natl. Acad. Sci. USA.* 112:7432–7437.
69. Lee, A. A., C. S. Perez-Martinez, ..., S. Perkin. 2017. Scaling analysis of the screening length in concentrated electrolytes. *Phys. Rev. Lett.* 119:026002.
70. Hernandez-Cardoso, G. G., A. K. Singh, and E. Castro-Camus. 2020. Empirical comparison between effective medium theory models for the dielectric response of biological tissue at terahertz frequencies. *Appl. Opt.* 59:D6–D11.
71. Bruggemann, D. A. G. 1935. Berechnung Verschiedener Physikalischer Konstanten von Heterogenen Substanzen. *Ann. Phys. Leipzig.* 24:636–664.
72. Hanai, T. 1960. Theory of the dielectric dispersion due to the interfacial polarization and its application to emulsions. *Colloid Polym. Sci.* 171:23–31.
73. Garnett, J. C. M. 1904. Colours in metal glasses and in metallic films. *Philos. Trans. R. Soc. Lond. Ser. A.* 203:385–420.
74. Makarov, V. A., B. K. Andrews, and B. M. Pettitt. 1998. Reconstructing the protein-water interface. *Biopolymers.* 45:469–478.
75. Persson, F., P. Söderhjelm, and B. Halle. 2018. The spatial range of protein hydration. *J. Chem. Phys.* 148:215104.
76. Ebbinghaus, S., S. J. Kim, ..., M. Havenith. 2007. An extended dynamical hydration shell around proteins. *Proc. Natl. Acad. Sci. USA.* 104:20749–20752.
77. Ding, T., R. Li, ..., R. J. Falconer. 2010. Terahertz and far infrared spectroscopy of alanine-rich peptides having variable ellipticity. *Opt. Express.* 18:27431–27444.
78. Lipfert, J., S. Doniach, ..., D. Herschlag. 2014. Understanding nucleic acid-ion interactions. *Annu. Rev. Biochem.* 83:813–841.
79. Brucherseifer, M., M. Nagel, ..., R. Buttner. 2000. Label-free probing of the binding state of DNA by time-domain terahertz sensing. *Appl. Phys. Lett.* 77:4049–4051.
80. Laman, N., S. S. Harsha, ..., J. S. Melinger. 2008. High-resolution waveguide THz spectroscopy of biological molecules. *Biophys. J.* 94:1010–1020.

**Biophysical Journal, Volume 120**

**Supplemental information**

**Long-range DNA-water interactions**

**Abhishek K. Singh, Chengyuan Wen, Shengfeng Cheng, and Nguyen Q. Vinh**

# Supplemental Information

## Long-Range DNA-Water Interactions

Abhishek K. Singh,<sup>1</sup> Chengyuan Wen,<sup>1</sup> Shengfeng Cheng,<sup>1,2,3</sup> and Nguyen Q. Vinh<sup>1,2,3\*</sup>

<sup>1</sup>Department of Physics and Center for Soft Matter and Biological Physics

<sup>2</sup>Macromolecules Innovation Institute

<sup>3</sup>Department of Mechanical Engineering, Virginia Tech, Blacksburg, Virginia 24061, USA

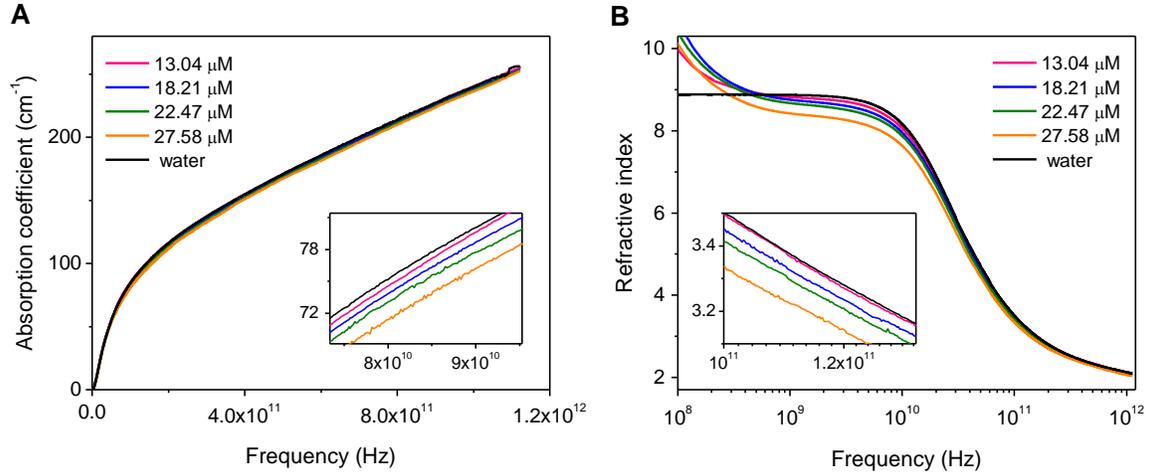
\*Correspondence: vinh@vt.edu (Nguyen Q. Vinh)

### Sample preparation

Deoxyribonucleic acid (DNA) sodium salt from salmon testes (D1626), received from Sigma Aldrich, was used to probe the DNA-water interactions as well as the DNA collective vibrations. The material was used as received without any further purification. The salmon testes DNA consists of ~2,000 base pairs with a molecular weight of  $1.3 \times 10^6$  Da. To prepare aqueous solutions, the DNA was dissolved in ultrapure water. In order to accurately determine the volume-filling factor,  $f_{\text{DNA}}$ , of DNA and the molar concentration, partial specific volumes of individual components were measured while the solutions were prepared.

### Megahertz to terahertz dielectric spectroscopy

The dielectric measurements were performed on a frequency-domain dielectric spectrometer that covers a broad spectral range of electromagnetic waves from 100 MHz to 1.12 THz ( $0.000334 - 37.36 \text{ cm}^{-1}$ ). The system allows us to accurately measure both the refractive index and absorbance of a solution at megahertz to terahertz frequencies. The spectrometer consists of a commercial Vector Network Analyzer (VNA) from Agilent (PNA N5225A) that covers a frequency range of 10 MHz to 50 GHz. Frequency extenders from Virginia Diodes, Inc. (Charlottesville, VA), were used to generate terahertz waves. Seven different frequency extender modules were used to cover a frequency range of 60 GHz to 1.12 THz as reported in the earlier publications (1-9). A variable path-length sample cell made from aluminum was used for dielectric experiments. The sample cell was equipped with Peltier coolers (Custom Thermoelectric, 27115L31-03CK) to control temperature during measurements. High power resistors were embedded in the aluminum sample cell to provide a controlled heating. The temperature was monitored and controlled with an accuracy of  $\pm 0.02$  °C using a Lakeshore 336 temperature controller.



**FIGURE S1.** DNA – water interactions at megahertz-terahertz frequencies. (A) Absorption coefficient and (B) refractive index of pure water and aqueous DNA solutions at 25 °C depend strongly on the frequency of the probing electromagnetic wave.

Using the above-described spectrometer and sample cell, we have measured the change of intensity and phase in aqueous solutions as a function of the thickness (path-length) of the sample (1-9). The intensity and phase of the terahertz radiation passing through a sample are described,

$$I_{\text{trans}}(l, \nu) = I_0(\nu) \cdot e^{-\alpha(\nu) \cdot l}, \quad (\text{S1a})$$

$$\phi_{\text{trans}}(l, \nu) = \phi_0(\nu) + n(\nu) \cdot l \cdot 2\pi\nu/c, \quad (\text{S1b})$$

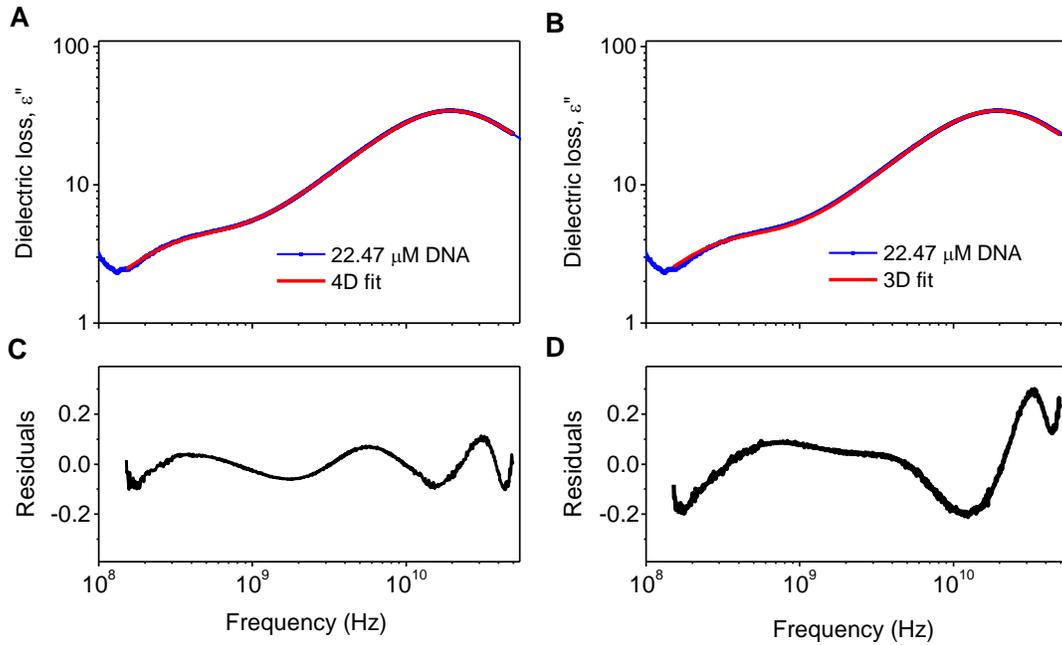
where  $I_0$ ,  $\phi_0$ ,  $I_{\text{trans}}$ , and  $\phi_{\text{trans}}$  are the intensity and phase of the radiation before and after the sample, respectively.  $\nu$ ,  $c$ ,  $l$ ,  $\alpha(\nu)$ ,  $n(\nu)$ , and  $l$  are the frequency, speed of light, thickness of the sample, and absorption coefficient and refractive index as a function of radiation frequency, respectively. When the radiation passes through a material, the intensity and phase are changed. Both characteristics have been used to define the complex refractive index of a material as follows,

$$n^*(\nu) = n(\nu) + i\kappa(\nu), \quad (\text{S2})$$

where  $\kappa(\nu)$  is the extinction coefficient of the material and is related to the absorption coefficient through  $\kappa(\nu) = c\alpha(\nu)/(4\pi\nu)$ . We have measured the intensity and phase shift of water and aqueous solutions as functions of the path-length,  $l$ , at the given temperature. The absorption coefficient and the refractive index are determined by the slope of linear fits of  $\ln(I(l, \nu)/I_0(\nu))$  and  $(\phi_{\text{transmission}}(l, \nu) - \phi_0(\nu))$  values to the path-length,  $l$ , respectively, without the need for precise determination of the absolute path-length and intrinsic optical properties of the sample cell. The absorption coefficient and refractive index are shown in Fig. S1 as a function of frequency for pure water and aqueous solutions at various DNA concentrations at 25 °C. Note that Eq. S1b does not have a form of  $(n(\nu) - 1)$  as in other measurements since we measure the phase shift when we vary the sample thickness. The standard errors of the mean of replicate

measurements are typically smaller than 0.2%. Using the spectrometer, we have measured precisely the absorption coefficient and refractive index of the strong absorption materials like water and aqueous solutions at the terahertz frequencies (Fig. S1). The error bars of our absorption and refractive index measurements are within the thickness of the lines.

The choice of the thickness of liquid water samples depends on the dynamics of frequency bands and the absorption of the material. The thickness varies from 0.5 mm for WR1.0 to 2.0 mm for the WR10 band in our system (1). To mitigate problems associated with multiple reflections of the incident light (standing waves and etalon effect), the thickness of our shortest path-length was selected to be long enough to ensure strong attenuation of the incident radiation (transmission  $<10^{-2}$ ). At each frequency, we examine an average of 100 different path-lengths with increments ranging from 0.1 to 20  $\mu\text{m}$ , depending on the absorption strength of the sample.



**FIGURE S2.** Examples of fitting to (A) four-Debye and (B) three-Debye models to the dielectric spectra of 22.47  $\mu\text{M}$  DNA solution are shown. The corresponding residuals are shown in (C) and (D), respectively. The four-Debye model provides a better fit to the dielectric response.

In parallel, the complex dielectric constant of a solution can be expressed as

$$\epsilon_{\text{sol}}^*(\nu) = \epsilon'_{\text{sol}}(\nu) + i\eta''_{\text{sol}}(\nu), \quad (\text{S3})$$

where  $\epsilon'_{\text{sol}}(\nu)$  and  $\eta''_{\text{sol}}(\nu)$  are the real (dielectric dispersion) and imaginary (total dielectric loss) components of the relative permittivity, respectively. The imaginary component of the dielectric response of a DNA solution contains two contributions,

$$\eta''_{\text{sol}}(\nu) = \epsilon''_{\text{sol}}(\nu) + \eta''_{\sigma}(\nu) = \epsilon''_{\text{sol}}(\nu) + \sigma/(2\pi\nu\epsilon_0), \quad (\text{S4})$$

where  $\epsilon''_{\text{sol}}(\nu)$  represents the dielectric loss,  $\eta''_{\sigma}(\nu)$  is the Ohmic loss due to the drift of ions contained in the DNA solution,  $\epsilon_0$  is the vacuum permittivity, and  $\sigma$  is the electrical conductivity of the solution. The electrical conductivity values of the DNA solutions investigated here are listed in Table S1. The dielectric dispersion and loss can be extracted from the refractive index and absorption coefficient through

$$\begin{aligned}\epsilon'_{\text{sol}}(\nu) &= n^2(\nu) - \kappa^2(\nu) = n^2(\nu) - (c\alpha(\nu)/4\pi\nu)^2 \\ \epsilon''_{\text{sol}}(\nu) &= 2n(\nu) \cdot \kappa(\nu) = 2n(\nu)c\alpha(\nu)/4\pi\nu - \sigma/(2\pi\nu\epsilon_0)\end{aligned}\quad (\text{S5})$$

Most biomolecules, including DNA, show a very low absorbance at terahertz frequencies as compared to bulk water (10), which explains that the absorption coefficient and refractive index of aqueous DNA solutions are only slightly lower than those of pure water and decrease gradually as the DNA concentration is increased. Since the polarization relaxation of condensed counterions along a DNA chain (i.e.,  $\alpha$ -relaxation), the relaxation of permanent dipoles, and the motion of polar groups in the DNA chain (i.e.,  $\beta$ -relaxation) occur at frequencies well below 100 MHz (11), their contributions to the dielectric response of a DNA solution in the megahertz to terahertz frequency range are negligible and not considered in the present work.

**TABLE S1.** Parameters extracted by fitting the dielectric response to a model based on four Debye relaxation processes. The error for fitting parameters is smaller than 8 %.

DNA ( $\mu\text{M}$ )	$\tau_{\text{TB1}}$ (ps)	$\tau_{\text{TB2}}$ (ps)	$\tau_{\text{LB}}$ (ps)	$\tau_{\text{D}}$ (ps)	$\Delta\epsilon_{\text{TB1}}$	$\Delta\epsilon_{\text{TB2}}$	$\Delta\epsilon_{\text{LB}}$	$\Delta\epsilon_{\text{D}}$	$\epsilon_{\infty}$	$\sigma$ ( $\text{S}\cdot\text{m}^{-1}$ )	$N_{\text{hyd}}$
3.08	825	400	48	8.24	0.08	0.73	0.12	72.5	6.2	0.05	$58 \pm 11$
6.15	820	430	55	8.21	0.14	1.35	0.33	71.8	6.0	0.15	$63 \pm 8$
13.04	850	340	59	8.19	0.22	2.10	0.57	70.5	5.9	0.36	$65 \pm 10$
18.21	830	330	45	8.22	0.35	3.00	1.00	69.2	5.2	0.44	$64 \pm 8$
22.47	835	400	52	8.12	0.60	4.05	1.20	68.0	5.0	0.49	$60 \pm 6$
27.58	814	419	56	8.17	0.80	4.50	1.40	67.2	4.9	0.56	$59 \pm 6$

The complex dielectric relaxation spectrum of a liquid, in which molecules relax via  $n$  individual processes, can be fit by the Havriliak-Negami function as (12),

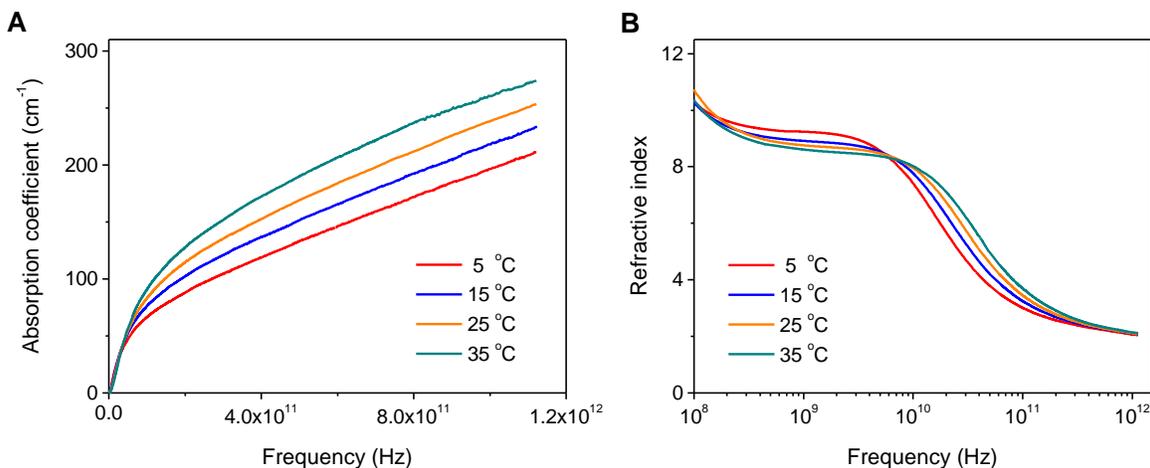
$$\epsilon^*(\nu) = \epsilon_{\infty} + \sum_{j=1}^n \frac{\epsilon_{j-1} - \epsilon_j}{(1 + (i2\pi\nu\tau_j)^{1-\alpha_{\text{HN}j}})^{\beta_{\text{HN}j}}}, \quad (\text{S6})$$

where  $\Delta\epsilon_j = \epsilon_j - \epsilon_{j+1}$  is the dielectric strength of the  $j^{\text{th}}$  process,  $\tau_j$  is the relaxation time of the  $j^{\text{th}}$  process, and  $\epsilon_n = \epsilon_{\infty}$  is the asymptotic dielectric constant at the high-frequency limit capturing contributions of polarization modes with frequencies much higher than the probed range. In the equation,  $\epsilon_0 = \epsilon_s$  is the dielectric constant at the zero-frequency limit or the static permittivity and can be represented as  $\epsilon_s = \epsilon_{\infty} + \sum_{j=1}^n \Delta\epsilon_j$ . The parameters  $\alpha_{\text{HN}j}$  and  $\beta_{\text{HN}j}$  are distribution parameters and must satisfy  $0 \leq \alpha_{\text{HN}j} < 1$ ,

and  $0 < \beta_{\text{HN}_j} \leq 1$ , respectively. The Havriliak-Negami equation reduces to the Cole-Davidson function when  $\alpha_{\text{HN}_j} (\equiv \alpha_{\text{CD}_j}) = 0$  (13), the Cole-Cole distribution when  $\beta_{\text{HN}_j} (\equiv \beta_{\text{CC}_j}) = 1$  (14), and the Debye function (15) when  $\alpha_{\text{HN}_j} (\equiv \alpha_{\text{D}}) = 0$  and  $\beta_{\text{HN}_j} (\equiv \beta_{\text{D}}) = 1$ . A qualitative comparison between four-Debye and three-Debye model fits to dielectric response is demonstrated in Fig. S2.

### Molecular dynamics simulations

The proximal radial distribution function, pRDF( $r$ ), describes the radial distribution of water molecules surrounding a surface and can be computed with molecular dynamics (MD) simulations. In order to distinguish water molecules in the major and minor grooves of a DNA, we first identify atoms on the DNA forming the minor and major grooves. Since the DNA surface is irregular, we use the distance from the oxygen atom of a water molecule to the closest atom (N, O, or P atoms) on the DNA surface as the distance of the water molecule to the DNA surface. To calculate the volume of a layer of a given thickness surrounding an irregular surface, we adopt the algorithm of Makarov *et al.*, (16) The simulation box is split into a three-dimensional rectangular grid. The distances of all atoms on the DNA forming the minor and major grooves to the center of each bin are calculated. The closest distance between the bin and the DNA surface is used to identify the bin belonging to either the major or minor grooves. Water molecules in each bin are also found and the hydration water belonging to the major or minor grooves are thus identified by checking if they are in the bins associated to the major or minor groove. Finally, in the calculation of pRDF( $r$ ), the volume normalization at each distance is done by summing the volumes of all the bins at that distance from the DNA surface.

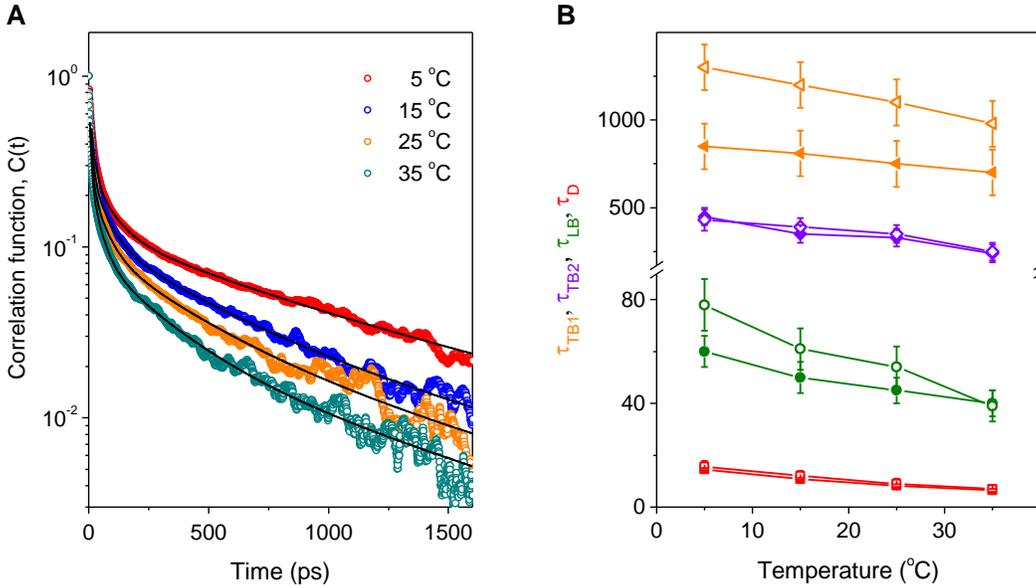


**FIGURE S3.** DNA-water interactions. Effect of temperature variation. (A) Absorption coefficient and (B) refractive index of the 18.21  $\mu\text{M}$  aqueous DNA solution at 5, 15, 25, and 35  $^{\circ}\text{C}$  in the megahertz to terahertz frequency range.

The orientational autocorrelation function,  $C(t)$ , is calculated using

$$C(t) = \left\langle \frac{1}{N} \sum_{i=1}^N \mu_i(t) \mu_i(0) \right\rangle, \quad (\text{S7})$$

where  $\mu_i(t)$  is the unit vector along the direction of the instantaneous electric dipole associated with a water molecule at time  $t$ , and the summation is over all  $N$  water molecules in the solution. An ensemble average is computed using various states of the system as the initial state at  $t = 0$ . Our analyses indicate that orientational autocorrelation functions,  $C(t)$ , can be fit to a superposition of multiple exponential functions,  $\sum_i A_i e^{-(t/\tau_i)}$ , with  $\tau_i$  being the relaxation time of the  $i$ -th relaxation mode, and  $A_i$  being a weighted coefficient for that mode. The identification of multiple relaxation times using  $C(t)$  indicates that the water molecules in a DNA solution relax through multiple relaxation processes, reflecting the various interactions between water and DNA molecules.



**FIGURE S4.** Effect of temperature on orientational dynamics. (A) The orientational correlation function,  $C(t)$ , of water molecules within 5.5 Å from the DNA surface in a dilute DNA solution at 5, 15, 25, and 35 °C, from top to bottom. The fits to Eq. 5 are shown as black lines. (B) The orientational relaxation time for the tightly-bound water in the minor grooves,  $\tau_{\text{TB1}}$ , the tightly-bound water in the major grooves and around the phosphate groups,  $\tau_{\text{TB2}}$ , the loosely-bound water,  $\tau_{\text{LB}}$ , and the bulk water,  $\tau_{\text{D}}$ , is plotted against temperature. The open symbols represent data from MD simulations by computing  $C(t)$  for the water molecules within 5.5 Å from the DNA surface and fitting it to Eq. 5 of the main text, whereas the solid symbols represent the experimental data from Fig. 4 B of the main text.

The orientational correlation functions,  $C(t)$ , of water molecules within 5.5 Å from the DNA surface at 5, 15, 25, and 35 °C, computed with data from MD simulations, are plotted in Fig. S4. The fits using a superposition of four exponential functions as in Eq. 5 of the main text are also included. As expected, when temperature is increased, the orientational relaxation of water molecules becomes faster. The

orientational correlation functions can also be computed for water molecules within various distances from the DNA surfaces. The relaxation times,  $\tau_1$ ,  $\tau_2$ ,  $\tau_3$ , and  $\tau_4$ , as well as the corresponding strength of each relaxation mode extracted by fitting  $C(t)$  to Eq. 5 of the main text are included in Table S2 for a dilute DNA solution at 25 °C.

The vibrational density of states (VDoS) of a DNA molecule is also computed. The DNA molecule is separated into a sugar phosphate backbone and side chains including all the base pairs. The VDoS is computed by calculating the Fourier transform of the velocity autocorrelation functions for the atoms on the backbone and/or the side chains as

$$C_v(t) = \langle \frac{1}{N} \sum_{i=1}^N v_i(t) v_i(0) \rangle, \quad (\text{S8})$$

where the summation is over all the atoms in the group (i.e., the backbone or the side chains of a DNA molecule or both) for which the VDoS is computed.

**TABLE S2.** Parameters extracted by fitting the orientation correlation functions to Eq. 5 of the main text, which is computed for the water molecules within distance  $r$  from the DNA surface in a dilute DNA solution at 25 °C.

$r$ (Å)	$A$	$B$	$C$	$D$	$\tau_1$ (ps)	$\tau_2$ (ps)	$\tau_3$ (ps)	$\tau_4$ (ps)
3	0.100	0.240	0.28	0.38	1300	450	37	7.7
5.5	0.050	0.120	0.21	0.62	1200	335	54	7.5
10	0.010	0.030	0.12	0.84	1250	320	50	6.7
14	0.0050	0.015	0.08	0.90	1100	330	49	6.3
18	0.0050	0.015	0.060	0.92	1050	320	52	6.2

## Supporting References

1. George, D. K., A. Charkhesht, and N. Q. Vinh. 2015. New terahertz dielectric spectroscopy for the study of aqueous solutions. *Rev. Sci. Instrum.* 86:123105.
2. Vinh, N. Q., M. S. Sherwin, S. J. Allen, D. K. George, A. J. Rahmani, and K. W. Plaxco. 2015. High-Precision Gigahertz-to-Terahertz Spectroscopy of Aqueous Salt Solutions as a Probe of the Femtosecond-to-Picosecond Dynamics of Liquid Water. *J. Chem. Phys.* 142:164502.
3. George, D. K., A. Charkhesht, O. A. Hull, A. Mishra, D. G. S. Capelluto, K. R. Mitchell-Koch, and N. Q. Vinh. 2016. New Insights into the Dynamics of Zwitterionic Micelles and Their Hydration Waters by Gigahertz-to-Terahertz Dielectric Spectroscopy. *J. Phys. Chem. B.* 120:10757-10767.
4. Charkhesht, A., C. K. Regmi, K. R. Mitchell-Koch, S. Cheng, and N. Q. Vinh. 2018. High-Precision Megahertz-to-Terahertz Dielectric Spectroscopy of Protein Collective Motions and Hydration Dynamics. *J. Phys. Chem. B.* 122:6341-6350.
5. Charkhesht, A., D. Lou, B. Sindle, C. Y. Wen, S. F. Cheng, and N. Q. Vinh. 2019. Insights into Hydration Dynamics and Cooperative Interactions in Glycerol-Water Mixtures by Terahertz Dielectric Spectroscopy. *J. Phys. Chem. B.* 123:8791-8799.
6. Maurya, D., A. Charkhesht, S. K. Nayak, F. C. Sun, D. George, A. Pramanick, M. G. Kang, H. C. Song, M. M. Alexander, D. Lou, G. A. Khodaparast, S. P. Alpay, N. Q. Vinh, and S. Priya. 2017. Soft phonon mode dynamics in Aurivillius-type structures. *Phys. Rev. B.* 96:134114.
7. Song, M. E., D. Maurya, Y. F. Wang, J. Wang, M. G. Kang, D. Walker, P. A. Thomas, S. T. Huxtable, R. J. Bodnar, N. Q. Vinh, and S. Priya. 2020. Phase Transitions and Phonon Mode Dynamics of  $\text{Ba}(\text{Cu}_{1/3}\text{Nb}_{2/3})\text{O}_3$  and  $\text{Sr}(\text{Cu}_{1/3}\text{Nb}_{2/3})\text{O}_3$  for Understanding Thermoelectric Response. *ACS Appl. Energy Mater.* 3:3939-3945.
8. Vinh, N. Q., S. J. Allen, and K. W. Plaxco. 2011. Dielectric Spectroscopy of Proteins as a Quantitative Experimental Test of Computational Models of Their Low-Frequency Harmonic Motions. *J. Am. Chem. Soc.* 133:8942-8947.
9. Vinh, N. Q. 2016. Probe conformational dynamics of proteins in aqueous solutions by terahertz spectroscopy. *Proc. of SPIE.* 9934:99340R.
10. Markelz, A. G., A. Roitberg, and E. J. Heilweil. 2000. Pulsed terahertz spectroscopy of DNA, bovine serum albumin and collagen between 0.1 and 2.0 THz. *Chem. Phys. Lett.* 320:42-48.
11. Tomic, S., S. D. Babic, T. Vuletic, S. Krca, D. Ivankovic, L. Griparic, and R. Podgornik. 2007. Dielectric relaxation of DNA aqueous solutions. *Phys. Rev. E.* 75:021905.
12. Havriliak, S., and S. Negami. 1967. A Complex Plane Representation of Dielectric and Mechanical Relaxation Processes in Some Polymers. *Polymer.* 8:161-210.
13. Davidson, D. W., and R. H. Cole. 1951. Dielectric Relaxation in Glycerol, Propylene Glycol, and Normal-Propanol. *J. Chem. Phys.* 19:1484-1490.
14. Cole, K. S., and R. H. Cole. 1941. Dispersion and absorption in dielectrics I. Alternating current characteristics. *J. Chem. Phys.* 9:341-351.
15. Debye, P. 1929. Polar molecules. New York.
16. Makarov, V. A., B. K. Andrews, and B. M. Pettitt. 1998. Reconstructing the protein-water interface. *Biopolymers.* 45:469-478.