Minor Groove Hydration of DNA in Solution: Dependence on Base Composition and Sequence

Haukur Jóhannesson and Bertil Halle*

Contribution from the Condensed Matter Magnetic Resonance Group, Department of Chemistry, Lund University, P.O. Box 124, S-22100 Lund, Sweden

Received December 22, 1997

Abstract: The hydration of six B-DNA dodecamers with A-tracts of variable length and sequence has been investigated via the nuclear magnetic relaxation dispersion (NMRD) of the water ^{2}H and ^{17}O resonances. By subdividing the aqueous DNA solution into microscopic emulsion droplets, NMRD measurements could be performed at −20 °C, thereby greatly enhancing the sensitivity of the method. The NMRD profiles show that all six dodecamers contain long-lived water molecules. These water molecules are displaced by netropsin and must therefore be located in the minor groove. The number of long-lived water molecules is correlated with the width of the minor groove as seen in crystal structures. The NMRD data are consistent with a single file of 3−9 long-lived water molecules located at the base pair steps and extending 1−2 steps on either side of the A-tract. Dodecamers with central A-tracts of sequence A_4T_4, T_4A_4, and $(AT)_4$ are found to contain seven or nine long-lived water molecules, challenging the common view that T−A steps widen the minor groove and disrupt the hydration structure. The long-lived water molecules observed here are highly ordered with an entropy comparable to that of water molecules in ice, but most of them undergo a symmetric flip motion while residing in the groove. The mean water residence time is essentially the same, 10−15 ns at −20 °C, for all investigated dodecamers, suggesting that water exchange occurs from an open state with a uniformly wide minor groove. From the temperature dependence of the water residence time, an activation enthalpy of 53 kJ mol^{-1} is obtained for this process.

Introduction

Since water−DNA interactions are comparable in strength to the noncovalent interactions within the double helix, hydration and helical fine structure are inextricably linked. In fact, water is often regarded as an integral part of DNA.[1−4] This applies in particular to the ordered water molecules found in narrowed A·T rich regions (A-tracts) of the minor groove in crystal structures of B-DNA.[5,6] Despite the confounding influence of crystal forces,[7] it seems clear that A-tracts have an intrinsic preference for a narrow minor groove and, hence, for a single-file hydration motif ("spine of hydration").

The current knowledge about DNA fine structure and hydration in solution is incomplete. The minor groove width cannot be determined with useful accuracy by solution NMR techniques.[8] NMR studies of the rate of imino proton exchange have demonstrated anomalously long base pair lifetimes in A-tracts. This has been taken as evidence for a distinct (B′) structure of such segments also in solution,[9,10] but the relation between base pair kinetics and DNA morphology is not well understood. Studies of sequence-dependent DNA cleavage by hydroxyl radicals and nucleases[11,12] tend to support the general notion of minor groove narrowing in A-tracts, but the sequence-dependent effects are not in line with the crystal structures. Much recent work on the sequence-dependent fine structure of B-DNA has been motivated by the finding that long restriction fragments of DNA containing short (4−6 base pairs) homopolymeric A-tracts repeated in phase with the helix period acquire a large-scale curvature manifested by anomalously low gel mobility.[13,14] While there is agreement that such spontaneous curvature is the cumulative effect of the peculiar fine structure of A-tracts (and their junctions with canonical B-form DNA), the precise mechanism of DNA bending remains controversial.[15] Minor groove hydration may well be important here. DNA hydration in solution has also been studied via measurements of the apparent molar volume and adiabatic compressibility,[16−18] but

* Correspondence to: Bertil Halle, Condensed Matter Magnetic Resonance Group, Department of Chemistry, Lund University, P.O. Box 124, S-22100 Lund, Sweden. Tel: +46 - 46 222 9516. Fax: +46 - 46 222 4543. E-mail: bertil.halle@fkem2.1th.se.

(1) Saenger, W. *Annu. Rev. Biophys. Biophys. Chem.* **1987**, *16*, 93−114.
(2) Westhof, E. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 125−144.
(3) Beveridge, D. L.; Swaminathan, S.; Ravishanker, G.; Withka, J. M.; Srinivasan, J.; Prevost, C.; Louise-May, S.; Langley, D. R.; DiCapua, F. M.; Bolton, P. H. In *Water and Biological Macromolecules*; Westhof, E., Ed.; CRC Press: Boca Raton, 1993; pp 165−225.
(4) Berman, H. M. *Curr. Opin. Struct. Biol.* **1994**, *4*, 345−350.
(5) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535−556.
(6) Chandrasekaran, R.; Radha, A.; Park, H.-S. *Acta Crystallogr. D* **1995**, *51*, 1025−1035.
(7) Dickerson, R. E. *Methods Enzymol.* **1992**, *211*, 67−111.
(8) Leijon, M.; Zdunek, J.; Fritzsche, H.; Sklenar, H.; Gräslund, A. *Eur. J. Biochem.* **1995**, *234*, 832−842.
(9) Leroy, J.-L.; Charretier, E.; Kochoyan, M.; Guéron, M. *Biochemistry* **1988**, *27*, 8894−8898.
(10) Moe, J. G.; Russu, I. M. *Nucleic Acid Res.* **1990**, *18*, 821−827.
(11) Burkhoff, A. M.; Tullius, T. D. *Nature* **1988**, *331*, 455−457.
(12) Fox, K. R. *Nucleic Acid Res.* **1992**, *20*, 6487−6493.
(13) Hagerman, P. J. *Annu. Rev. Biochem.* **1990**, *59*, 755−781.
(14) Crothers, D. M.; Haran, T. E.; Nadeau, J. G. *J. Biol. Chem.* **1990**, *265*, 7093−7096.
(15) Dickerson, R. E.; Goodsell, D.; Kopka, M. L. *J. Mol. Biol.* **1996**, *256*, 108−125.
(16) Buckin, V. A.; Kankiya, B. I.; Bulichov, N. V.; Lebedev, A. V.; Gukovsky, I. Y.; Chuprina, V. P.; Sarvazyan, A. P.; Williams, A. R. *Nature* **1989**, *340*, 321−322.

S0002-7863(97)04316-3 CCC: $15.00
Published on Web 07/03/1998

the microscopic interpretation of these macroscopic quantities is highly model-dependent.

Among the techniques that can probe biomolecular hydration in aqueous solution, NMR spectroscopy is arguably the most informative. NMR studies of DNA hydration have a long history,[19] but it is only in the last few years that the methodology has matured to a stage where the orientational order and dynamics of water molecules at specific hydration sites can be quantitatively characterized. Two principal NMR techniques are currently available for studying DNA hydration in solution: one of them measures intermolecular nuclear Overhauser effects (NOEs) between water and specific DNA protons[20] and the other records the nuclear magnetic relaxation dispersion (NMRD) of the water ^{2}H and ^{17}O resonances.[21] Since they probe different aspects of the water−DNA interaction and have different ranges of applicability, the NOE and NMRD techniques are highly complementary.[22]

The first DNA duplex to be investigated by the NOE technique[23,24] and the only one so far investigated by NMRD[22,25] is the benchmark dodecamer CGCGAATTCGCG. The observation of positive NOESY cross-peaks between water and the adenine base protons A5 H2 and A6 H2 was taken as "evidence for a spine of hydration" also in solution.[23,24] Since the number of water molecules responsible for these cross-peaks could not be determined, however, it was not really possible to distinguish the single file and double file hydration motifs (and even less to establish the presence of both layers of the originally proposed spine). A subsequent joint NOE and multinuclear NMRD study[22] of the same dodecamer showed that the residence time of the five minor-groove waters in the A-tract of this duplex is 0.9 ± 0.1 ns at 4 °C (from ^{2}H and ^{17}O NMRD) and 0.6 ns at 10 °C (from NOE data). These residence times agree quantitatively with the result of a recent temperature-dependent 1D NOESY study of the same dodecamer, yielding a zero-crossing for the A6 H2 cross-peak at 15 °C which corresponds to a residence time of 0.5 ns at this temperature.[26]

The convergent results on the hydration kinetics of the CGCGAATTCGCG dodecamer, obtained by several groups using different NMR techniques and different DNA preparations, demonstrate that the two NMR methodologies are sound and that quantitatively reliable results can be expected also for other DNA duplexes. The logical next step is to identify the determinants of residence times of minor groove water by investigating duplexes with different base sequences. Until now, this approach has been pursued only with the NOE technique.[24,26−31] The emerging picture of sequence-dependent minor groove hydration in solution is rather complex and leaves several important aspects unresolved. For example, are long-lived water molecules confined to A-tracts or can they also be located near G·C base pairs (where they would not be detected by the NOE technique)? In our view, water residence times are interesting primarily because they reflect the structure and dynamics of the DNA molecule. If the mechanism of water exchange between minor groove and bulk solvent can be established, it might be possible to use water residence times to gauge groove width in solution and to sample rare structural fluctuations involving conformational substates of functional significance. The present joint ^{2}H and ^{17}O NMRD study of six self-complementary dodecamers and one drug−DNA complex was designed with these questions in mind.

A problem encountered in the previous NMRD study of DNA hydration at 4 °C is that a residence time of ca. 1 ns puts the dispersion at the high-frequency edge of the experimentally accessible frequency window and strongly attenuates the magnitude of the dispersion step.[22] The measurable effect of minor groove hydration is then too small to accurately characterize sequence variations. In the present work, we have overcome this limitation by performing the relaxation measurements at −20 °C, where the dispersion step is an order of magnitude larger and the dispersion occurs in the middle of the experimental frequency window. Under these conditions, a single water molecule with a residence time comparable to (or longer than) the tumbling time of the duplex can be detected.

Materials and Methods

DNA Solutions. Oligonucleotides were synthesized on a 10 μmol scale using the phosphoramidite method. The dodecamer CGCGAATTCGCG, obtained from Biosource International (Camarillo, CA, U.S.A.), was purified by reverse-phase HPLC (before detritylation) and then subjected to two ethanol precipitations in NaCl to produce the sodium salt of DNA. The purity of this material was checked by denaturing reverse-phase and anion-exchange HPLC and by denaturing gel electrophoresis, indicating > 95% full-length oligomer. The NMRD profiles from this purified material and from a crude preparation of the same dodecamer were found to be identical within the experimental error.[22] The other dodecamers, synthesized on an Applied Biosystems ABI 394 instrument at the Biomolecular Unit of Lund University, were therefore used without HPLC purification. The crude product was subjected to an ether extraction and two ethanol precipitations in NaCl. The fraction full-length oligomer was estimated to 80−90% by denaturing anion-exchange HPLC.

The lyophilized sodium salt of the dodecamers was dissolved in millipore (0.2 μm) filtered water enriched in the ^{2}H and ^{17}O isotopes. Two water batches were used. For CGCGAATTCGCG and CGCAAATTTGCG, we used D_2O (99.95 atom % ^{2}H, 21.9 atom % ^{17}O, 61.9 atom % ^{18}O) obtained from Ventron (Karlsruhe, Germany). For the other dodecamers, we used a HDO mixture (51.8 atom % ^{2}H, 17 atom % ^{17}O and ca. 30 atom % ^{18}O) made from equal weights of ^{17}O-enriched H_2O (35 atom % ^{17}O) and D_2O (99.9 atom % ^{2}H, low paramagnetic content), both from Cambridge Isotope Laboratories (Andover, MA, U.S.A.). The difference in isotope composition between the two water batches is of no consequence for the NMRD measurements, except for a trivial scaling of the rotational correlation time of the DNA duplex by the solvent viscosity[32,33] (7.6 cP for the D_2O mixture

(17) Rentzeperis, D.; Kupke, D. W.; Marky, L. A. *Biopolymers* **1992**, *32*, 1065−1075.
(18) Chalikian, T. V.; Sarvazyan, A. P.; Plum, G. E.; Breslauer; K. J. *Biochemistry* **1994**, *33*, 2394−2401.
(19) Jacobson, B.; Anderson, W. A.; Arnold, J. T. *Nature* **1954**, *173*, 772−773.
(20) Otting, G. *Progr. NMR Spectrosc.* **1997**, *31*, 259−285.
(21) Halle, B.; Denisov, V. P.; Venu, K. In *Modern Techniques in Protein NMR*; Berliner, L. J., Krishna, N. R., Eds.; Plenum: New York, 1998; Vol. 16B.
(22) Denisov, V. P.; Carlström, G.; Venu, K.; Halle, B. *J. Mol. Biol.* **1997**, *268*, 118−136.
(23) Kubinec, M. G.; Wemmer, D. E. J. *Am. Chem. Soc.* **1992**, *114*, 8739−8740.
(24) Liepinsh, E.; Otting, G.; Wüthrich, K. *Nucleic Acids Res.* **1992**, *20*, 6549−6553.
(25) Zhou, D.; Bryant, R. G. *J. Biomol. NMR* **1996**, *8*, 77−86.
(26) Phan, A. T.; Leroy, J.-L.; Guéron, M. *Eur. J. Biophys.* **1997**, *26*, 53.
(27) Fawthrop, S. A.; Yang, J.-C.; Fisher, J. *Nucleic Acids Res.* **1993**, *21*, 4860−4866.
(28) Liepinsh, E.; Leupin, W.; Otting, G. *Nucleic Acids Res.* **1994**, *22*, 2249−2254.
(29) Jacobson, A.; Leupin, W.; Liepinsh, E.; Otting, G. *Nucleic Acids Res.* **1996**, *24*, 2911−2918.
(30) Jenkins, T. C.; Lane, A. N. *Biochim. Biophys. Acta* **1997**, *1350*, 189−204.
(31) Lane, A. N.; Jenkins, T. C.; Frenkiel, T. A. *Biochim. Biophys. Acta* **1997**, *1350*, 205−220.
(32) Osipov, Y. A.; Zheleznyi, B. V.; Bondarenko, N. F. *Zh. Fiz. Khim.* **1977**, *51*, 1264−1265.
(33) Kestin, J.; Imaishi, N.; Nott, S. H.; Nieuwoudt, J. C.; Sengers, J. V. *Physica A* **1985**, *134*, 38−58.

Table 1. Results Derived from ^{2}H and ^{17}O NMRD Data on DNA Dodecamers at −20 °C[a]

sequence	short form	C_{DNA}[b] (mM)	$N_\beta S_\beta^2(^{17}O)$	$S_\beta^2(^2H)/S_\beta^2(^{17}O)$	$10^{-3}\, N_\alpha \rho_\alpha(^{17}O)$	$\rho_\alpha(^2H)/\rho_\alpha(^{17}O)$	τ_β (ns)	τ_W[c] (ns)
CGCTC**TA**GAGCG	TA	7.84	1.2 ± 0.3	0.60 ± 0.21	1.6 ± 0.1	0.77 ± 0.04	9.2 ± 1.9	12 ± 3
CGCG**AATT**CGCG	A2T2	7.63	5.9 ± 0.6	0.70 ± 0.10	1.5 ± 0.1	0.91 ± 0.06	8.8 ± 0.7	11 ± 1
A2T2 + netropsin		7.63	1.6 ± 0.4	0.91 ± 0.35	1.5 ± 0.1	0.76 ± 0.05	9.0 ± 2.2	11 ± 3
CGC**AAATTT**GCG	A3T3	8.21	5.1 ± 0.6	0.55 ± 0.10	1.4 ± 0.1	0.90 ± 0.07	10.8 ± 1.1	14 ± 2
CG**AAAATTTT**CG	A4T4	8.64	8.6 ± 0.4	0.58 ± 0.04	2.3 ± 0.1	0.94 ± 0.04	10.4 ± 0.4	14 ± 1
GC**TTTTAAAA**GC	T4A4	7.98	5.8 ± 0.7	0.46 ± 0.08	2.4 ± 0.1	0.95 ± 0.07	11.0 ± 1.2	15 ± 2
CG**ATATATAT**CG	(AT)4	8.36	7.2 ± 0.5	0.52 ± 0.06	1.8 ± 0.1	0.97 ± 0.10	8.1 ± 0.5	10 ± 1

[a] All NMRD parameters were obtained from simultaneous fits to ^{2}H and ^{17}O data, with τ_β constrained to be the same for both nuclei. [b] Duplex concentration. [c] Mean water residence time obtained from eq 6 with $\tau_R = 50$ ns (D_2O) or 39.6 ns (HDO).

and 6.0 cP for the HDO mixture at the experimental temperature of −20 °C). Solution pH (not corrected for isotope effects) was adjusted to 7.00 ± 0.02 by addition of small amounts (ca. 5 μL) of 0.1−1 M HCl or NaOH. No buffers were added.

DNA concentrations were determined from the optical density at 260 nm, measured on a GBC UV/vis 920 instrument after 1500-fold dilution with phosphate buffer (pH 7.0, 0.1 M KCl) of a small fraction of each DNA solution used for NMRD measurements. Single-strand extinction coefficients were calculated from the nucleotide sequence[34] and corrected for the effect of hairpin formation.[22,35] Duplex concentrations for all NMRD samples are given in Table 1. These concentrations correspond to $N_T = 6000-7100$ water molecules per DNA duplex. (For comparison, the first hydration layer comprises about 300 water molecules.) Prior to emulsification, all duplexes were annealed by slow (ca. 30 min) cooling from 75 °C to room temperature. For the difference-NMRD experiment with the drug−DNA complex, the DNA solution was divided in two parts, to one of which was added 1.01 ± 0.01 equiv of crystalline netropsin (Boehringer Mannheim, Germany).

Preparation of Emulsion Samples. Water and aqueous solutions can be deeply supercooled if the aqueous phase is dispersed into microscopic droplets, as in a water-in-oil emulsion.[36] Since the probability that a micrometer sized droplet contains a heterogeneous nucleator particle is vanishingly small, essentially all droplets remain unfrozen down to the homogeneous nucleation temperature (−38 °C for H_2O at 1 atm).[37] The emulsion technique has been used for low-temperature NMR studies of bulk water and electrolyte solutions[38−40] but apparently not within the field of biomolecular NMR.

The emulsion samples used for the NMRD experiments were prepared by mixing ca. 1 mL of aqueous DNA solution with an equal volume of *n*-heptane (>99%, HPLC grade, from Sigma) supersaturated with 5 wt % of the nonionic surfactant sorbitan tristearate (trade name Span 65, from Sigma). After vortex mixing, yielding droplet diameters of order 100 μm, the emulsion was repeatedly (ca. 100 times) slashed through packed 635 mesh stainless steel gauze mounted on a filter holder with an attached syringe.[36,39] The resulting emulsion was examined under a microscope, revealing a fairly monodisperse droplet size distribution with a mean diameter of ca. 15 μm.

For studies of supercooled biomolecular solutions, emulsions are clearly superior to cryofluids. Perturbation of the aqueous phase is minimal since emulgator and oil have vanishing solubilities in water and since only a minute fraction of the biomolecules are near the interface. (A droplet of 15 μm diameter has only 0.4% of its volume within 100 Å of the interface.) Furthermore, the nonpolar carrier fluid contains a negligible fraction of the water molecules in the sample. The solubility of water in *n*-heptane is 2.3 mM at 25 °C,[41] meaning that only one in 24 000 water molecules is located in the organic phase. Although not a thermodynamically stable phase, an emulsion sample prepared as described above remains stable for months.

To verify that an aqueous biomolecular solution is indeed unperturbed by emulsification, we recorded the ^{2}H NMRD profiles from three samples containing the same solution of the protein BPTI in D_2O (12.1 wt %, pH 5.3): (*i*) bulk solution, (*ii*) a vortexed emulsion (100 μm droplets), and (*iii*) a finely dispersed emulsion (15 μm droplets), all at 27 °C. (For this initial test, a silicon oil was used as carrier fluid.) Within the experimental error of <1%, the three NMRD profiles were identical.

Relaxation Dispersion Measurements. The longitudinal relaxation rate, $R_1 = 1/T_1$, of the water ^{2}H and ^{17}O magnetizations was measured at 6−8 magnetic field strengths in the range 0.38−4.7 T as described elsewhere.[42] Because of the low temperature, the ^{17}O relaxation time was in some cases as short as 450 μs. In a control experiment where the 90°-pulse length was varied from 9 to 33 μs, the T_1 value showed only random scatter within the experimental uncertainty. Any effects of multiple-quantum coherences induced by relaxation during the pulses were thus negligible. Although ^{17}O relaxation is in principle triexponential under non-extreme-narrowing conditions,[43] no deviation from single-exponential inversion recovery was seen. A full relaxation matrix analysis with relaxation parameters representative for the DNA samples demonstrated that the analytical single-exponential approximation for ^{17}O relaxation[44] remains highly accurate under the present experimental conditions.

The ^{17}O transverse relaxation rate, $R_2 = 1/T_2$, was measured at some fields using a spin-echo pulse sequence. Due to their lower accuracy, R_2 data were not included in the dispersion fit. Nevertheless, they were important for establishing that the zero-frequency limit had been essentially reached at the lowest frequency of the R_1 dispersion: the finding that $R_1 = R_2$ rigorously rules out the possibility of a further dispersion step at lower frequencies.[43] In bulk water at pH 7, a relaxation dispersion step in the kHz range is expected due to hydrogen-exchange modulation of the scalar *J* coupling between ^{17}O and ^{2}H or ^{1}H.[45−47] This would make $R_2 > R_1$ in the investigated MHz frequency range. As previously[22] found at 4 °C, however, DNA-promoted exchange catalysis virtually eliminates the ^{17}O scalar relaxation contribution to R_2.

During all relaxation measurements, the sample temperature was maintained at −20.0 ± 0.2 °C by a thermostated flow of nitrogen gas. Considerable care was devoted to sample temperature control, which is a major issue in accurate NMRD work,[21] especially for measurements far from ambient temperature. The sample temperature was monitored frequently, and the same calibrated copper-constantan thermocouple was used with all spectrometers. As a further check on temperature, a pure water reference sample (with frequency-independent T_1) was usually measured together with the DNA sample. Any spurious T_1 variation caused by a temperature deviation would then have been

(34) Gray, D. M.; Hung, S.-H.; Johnson, K. H. *Methods Enzymol.* **1995**, *246*, 19−34.
(35) Marky, L. A.; Blumenfeld, K. S.; Kozlowski, S.; Breslauer, K. J. *Biopolymers* **1983**, *22*, 1247−1257.
(36) Rasmussen, D. H.; MacKenzie, A. P. In *Water Structure at the Water-Polymer Interface*; Jellinek, H. H. G., Ed.; Plenum: New York, 1972; pp 126−145.
(37) Angell, C. A. *Annu. Rev. Phys. Chem.* **1983**, *34*, 593−630.
(38) Hindman, J. C.; Svirmickas, A. *J. Phys. Chem.* **1973**, *77*, 2487−2489.
(39) Lang, E.; Lüdemann, H.-D. *J. Chem. Phys.* **1977**, *67*, 718−723.
(40) Lang, E. W.; Lüdemann, H.-D. *NMR Basic Princ. Progr.* **1990**, *24*, 129−187.
(41) Wolfenden, R.; Radzicka, A. *Science* **1994**, *265*, 936−937.
(42) Denisov, V. P.; Halle, B. *J. Mol. Biol.* **1995**, *245*, 682−697, 698−709.
(43) Abragam, A. *The Principles of Nuclear Magnetism*; Clarendon Press: Oxford, 1961.
(44) Halle, B.; Wennerström, H. *J. Magn. Reson.* **1981**, *44*, 89−100.
(45) Meiboom, S. *J. Chem. Phys.* **1961**, *34*, 375−388.
(46) Halle, B.; Karlström, G. *J. Chem. Soc., Faraday Trans. 2* **1983**, *79*, 1031−1046.
(47) Noack, F. *Prog. NMR Spectrosc.* **1986**, *18*, 171−276.

evident in the temperature-dependent reference T_1 value. The overall accuracy of the reported longitudinal relaxation rates is estimated to 0.5−1.5%. From a large number of measurements on the two bulk water reference samples at −20 °C, we obtained for the D_2O mixture $R_1(^2H) = 17.4 \pm 0.1\ s^{-1}$ and $R_1(^{17}O) = 1391 \pm 5\ s^{-1}$ and for the HDO mixture $R_1(^2H) = 13.3 \pm 0.1\ s^{-1}$ and $R_1(^{17}O) = 1010 \pm 5\ s^{-1}$.

Analysis of NMRD Data. The water ^2H and ^{17}O longitudinal relaxation data were analyzed with the usual Lorentzian dispersion function, containing the three parameters α, β, and τ_β.[21,22] By combining the dispersion amplitude parameter β with the known rigid-lattice nuclear quadrupole frequency and the DNA concentration, we obtain the product $N_\beta S_\beta^2$ of the number N_β of long-lived water molecules and their mean-square orientational order parameter S_β^2.[21] From the high-frequency plateau parameter α and the bulk water relaxation rate (directly measured on a reference sample), we obtain the product $N_\alpha \rho_\alpha$ of the number N_α of dynamically perturbed short-lived water molecules (at the surface of the DNA duplex) and their mean dynamic retardation factor $\rho_\alpha = \langle\tau_\alpha\rangle/\tau_{bulk} - 1$.[21,22] The correlation time τ_β defines the time scale on which the orientation of long-lived water molecules is randomized. This can occur by either (or both) of two independent processes: rotational diffusion of the entire DNA duplex with rotational correlation time τ_R or exchange of long-lived DNA-associated water molecules with bulk water at a rate $1/\tau_W$. These competing processes determine τ_β according to[21]

$$1/\tau_\beta = 1/\tau_R + 1/\tau_W \qquad (1)$$

The rotational diffusion of a B-DNA dodecamer duplex is well described by hydrodynamic theory,[48] modeling the duplex as a rigid cylinder of radius 10 Å and length 12 × 3.4 Å.[49] Using the viscosity at −20 °C for the D_2O and HDO water mixtures employed here (see above), we obtain (mean) rotational correlation times τ_R of 42 and 33 ns, respectively. These values apply to infinite dilution. The limited data available on the concentration dependence of τ_R for oligonucleotide duplexes indicate that τ_R is constant up to about 1.5 mM duplex but increases sharply at higher concentrations.[50] At a duplex concentration of 4 mM, a 40% increase of τ_R was reported for CGCGAATTCGCG.[49] At 8 mM duplex, as in our solutions, τ_R may well have doubled as compared to infinite dilution. We thus estimate that τ_R lies in the range 50−100 ns (probably closer to the higher value) in our samples. For such long τ_R values, the NMRD method becomes an exquisitely sensitive probe of long-lived water molecules; even a single moderately ordered water molecule with $\tau_W > \tau_R$ would be easily detected.

While labile DNA protons have been shown to dominate the ^1H dispersion,[22] the corresponding contribution to the ^2H dispersion for the dodecamer CGCGAATTCGCG at 4 °C was deemed insignificant: the shortest hydrogen exchange times were estimated to a few milliseconds, an order of magnitude longer than the (zero-frequency) intrinsic ^2H relaxation rate of these labile hydrogens.[22] At the present temperature of −20 °C, the exchange times will be considerably longer while the relaxation times will be shorter (due to the viscosity effect on τ_R). All labile DNA hydrogens should then be well into the slow-exchange regime with respect to ^2H relaxation. Both the ^2H and ^{17}O R_1 dispersions can therefore be attributed entirely to water molecules.

For each measured relaxation dispersion, the parameters α (yielding $N_\alpha\rho_\alpha$), β (yielding $N_\beta S_\beta^2$), and τ_β were determined by a nonlinear least-squares fit. Parameter error estimates were derived by assigning an error of 0.5−1.5% to the experimental R_1 data, to yield a reduced χ^2 value close to 0.5. These error estimates correspond to one standard deviation and were confirmed by Monte Carlo simulation.[51] In most of the dispersions, the R_1 values measured at two fields (^2H frequencies 4.58 and 6.86 MHz) displayed a systematic deviation in the fits. These deviations are about 3%, significantly larger than the random error as judged by reproducibility tests, and cannot be accounted for by invoking a multi-Lorentzian spectral density function. These points were therefore assumed to be affected by a systematic error of unknown nature (presumably associated with the probe head used for these two fields) and were not given any weight in the fits. In the fast-exchange limit and in the absence of effects of anisotropic rotational diffusion (since $\tau_W \ll \tau_R$), the ^2H and ^{17}O dispersions should yield the same τ_β value. This was indeed the case (within the error limits) for all investigated dodecamers. The ^2H and ^{17}O data could therefore be fitted together with a common τ_β parameter, i.e., five parameters were fitted to 12−16 data points. Since β and τ_β are correlated, the τ_β constraint reduces the parameter errors.

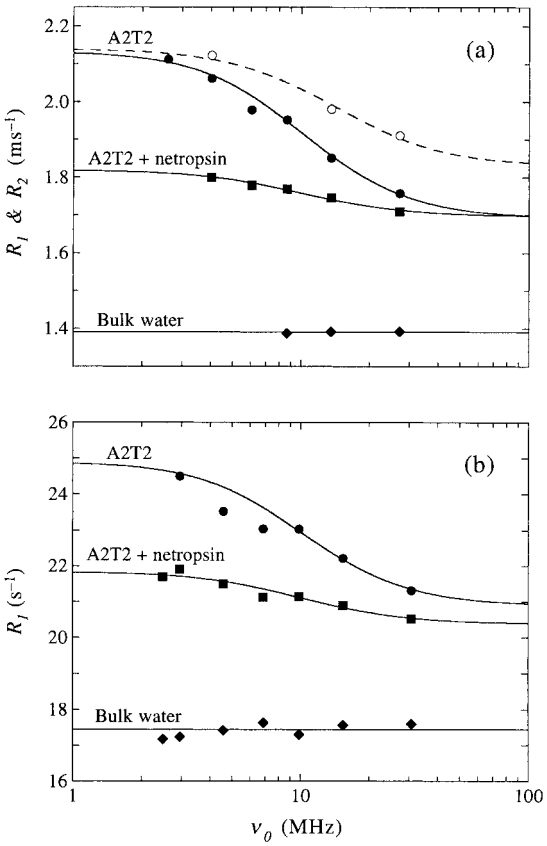

Figure 1. Dispersion of the water ^{17}O (a) and ^2H (b) longitudinal relaxation rate R_1 for the dodecamer duplex A2T2 (●) and for its 1:1 complex with netropsin (■). The curves resulted from joint fits to the combined ^2H and ^{17}O R_1 data (see Materials and Methods). For the ^{17}O dispersion, the transverse relaxation rate R_2 (○) was also measured at a few fields; the dashed curve is the R_2 dispersion calculated with the parameters obtained from the R_1 fit. Data from a pure water (D_2O) reference sample (◆) are also shown, with the line representing the average value.

Results and Discussion

Overview of NMRD Results. We have measured the water ^2H and ^{17}O R_1 relaxation dispersions for six self-complementary dodecamer duplexes (henceforth referred to by the abbreviated names in Table 1) in aqueous solution at −20 °C. Before presenting a detailed analysis of the NMRD data and comparing the results with other data on DNA hydration, we shall briefly summarize the NMRD results in a qualitative way.

Figure 1 shows the ^2H and ^{17}O dispersions for the benchmark dodecamer A2T2 and for its 1:1 complex with netropsin. The parameter values derived from the fits to these and all other dispersions are collected in Table 1. Since the drug is known

(48) Tirado, M. M.; Garcia de la Torre, J. *J. Chem. Phys.* **1980**, *73*, 1986−1993.
(49) Nuutero, S.; Fujimoto, B. S.; Flynn, P. F.; Reid, B. R.; Ribeiro, N. S.; Schurr, J. M. *Biopolymers* **1994**, *34*, 463−480.
(50) Pecora, R. *Science* **1991**, *251*, 893−898.
(51) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes. The Art of Scientific Computing*; Cambridge University Press: Cambridge, 1986.

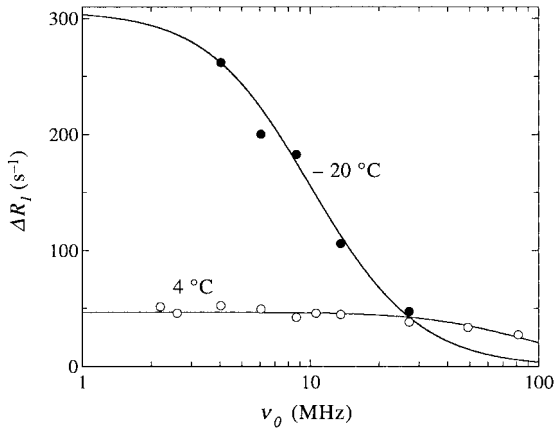

Figure 2. Reduction of the ^{17}O R_1 relaxation dispersion induced by netropsin binding to A2T2 at −20 °C (●) and at 4 °C (○). The curve associated with the −20 °C data resulted from a one-parameter fit to the ^{17}O R_1 data, with $\alpha = 0$ and τ_β fixed at 8.9 ns (the average of the τ_β values deduced from the separate A2T2 and A2T2-netropsin fits). The 4 °C data, which have been reported previously,[22] refer to a 5% higher DNA concentration.

to bind to the central part of the minor groove,[22,52,53] the observed reduction of the dispersion step on addition of netropsin shows that most of the long-lived water molecules responsible for the A2T2 dispersion are located in this part of the minor groove. The previously reported dispersions at 4 °C for A2T2 and its netropsin complex, although less completely characterized due to the one order of magnitude higher dispersion frequency, lead to the same conclusion.[22] Now, however, a small but well-defined dispersion is seen also for the drug−DNA complex.

The difference between the ^{17}O dispersions measured for A2T2 and for the A2T2−netropsin complex is displayed directly in Figure 2. A direct one-parameter fit to the difference-NMRD data (see legend) yields $N_\beta S_\beta^2 = 4.1 \pm 0.2$, in agreement with the difference (4.3 ± 0.7) of the $N_\beta S_\beta^2$ values from the individual fits in Figure 1a (see Table 1). The inclusion in Figure 2 of the partial dispersion observed[22] at 4 °C clearly illustrates the effect of reduced temperature on the frequency and amplitude of the dispersion.

Figures 3 and 4 show the ^{2}H and ^{17}O dispersions for all six investigated dodecamers. The dispersion step is seen to depend strongly on both base composition and base sequence. The smallest and largest dispersions correspond to the shortest and longest A-tracts (Figure 4), as anticipated. The nonmonotonic variation of the dispersion step in the A_NT_N set (Figure 3) and the insignificant effect of the three T−A steps in the (AT)4 dodecamer (Figure 4) are more surprising. A quantitative comparison of the ^{2}H and ^{17}O dispersion amplitudes shows that $N_\beta S_\beta^2$ is substantially smaller for ^{2}H. As discussed below, this indicates that most of the long-lived water molecules are flipping around their dipole axis while residing in the minor groove. Despite the large sequence dependent variation in the number of long-lived water molecules, all dispersions yield correlation times close to 10 ns. Since this is nearly an order of magnitude shorter than the tumbling time of the duplex, it can be identified with the residence time of these water molecules. The consequent near-invariance of the residence time is perhaps the most remarkable finding of the present study.

Location of Long-Lived Water Molecules. It might be expected that the charged phosphate backbone is the most

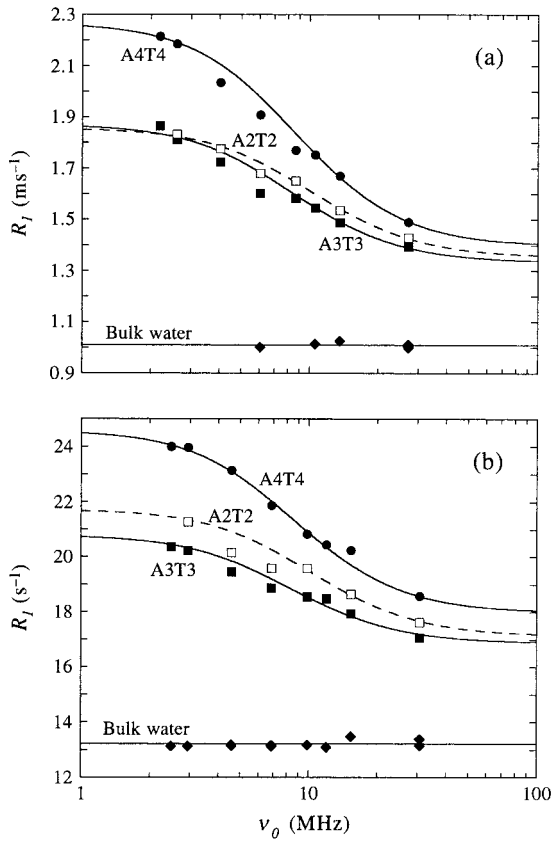

Figure 3. Dispersion of the water ^{17}O (a) and ^{2}H (b) longitudinal relaxation rate R_1 for the dodecamer duplexes A2T2 (□), A3T3 (■), and A4T4 (●). The curves resulted from joint fits to the combined ^{2}H and ^{17}O R_1 data (see Materials and Methods). To allow a direct visual comparison, the A2T2 and A3T3 data (measured in D_2O) and fits have been scaled to the HDO solvent of the A4T4 sample and to its DNA concentration (N_T value), assuming that R_1 scales as R_{bulk} + constant/N_T. Data from a pure water (HDO) reference sample (◆) are also shown, with the line representing the average value.

strongly hydrated part of a B-DNA duplex.[1,54] While this is undoubtedly true in a thermodynamic sense, previous NMRD studies of globular proteins have demonstrated that water molecules in contact with exposed macromolecular surfaces invariably have subnanosecond residence times.[55] With the exception of water molecules coordinated to site-bound multivalent metal ions,[56] long residence times have only been found for water molecules buried in internal cavities or penetrating deeply into narrow crevices or pockets. Water molecules at such shielded sites are often, but not always, extensively hydrogen bonded, but so are many of the more exposed water molecules and, indeed, bulk water. Hydrogen bonds are thus neither sufficient nor necessary for a long residence time. Instead, the essential prerequisite for a long water residence time appears to be steric: the geometry of the site must be such that the trapped water molecule cannot escape into the bulk solvent without passing through an intermediate state where several hydrogen bonds (which may or may not involve the exchanging water molecule) are disrupted. At an exposed site, and in the bulk liquid, water exchange can occur by a concerted rearrangement of the local hydrogen-bond network, whereby new hydrogen bonds are formed as old ones are broken. In a

(52) Patel, D. J. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 6424−6428.
(53) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376−1380.
(54) Westhof, E. In *Water and Biological Macromolecules*; Westhof, E., Ed.; CRC Press: Boca Raton, 1993; pp 226−243.
(55) Denisov, V. P.; Halle, B. *Faraday Discuss.* **1996**, *103*, 227−244.
(56) Denisov, V. P.; Halle, B. *J. Am. Chem. Soc.* **1995**, *117*, 8456−8465.

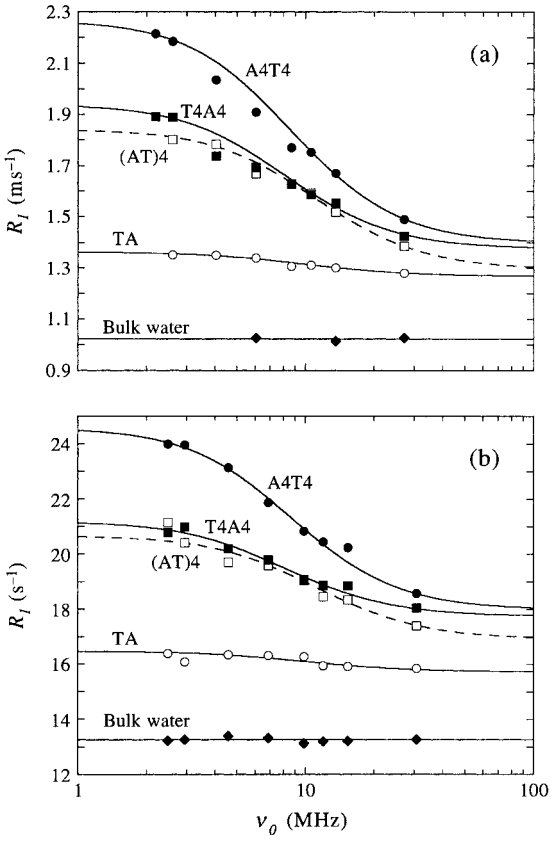

Figure 4. Dispersion of the water ^{17}O (a) and ^{2}H (b) longitudinal relaxation rate R_1 for the dodecamer duplexes A4T4 (●), T4A4 (■), (AT)4 (□), and TA (○). The curves resulted from joint fits to the combined ^{2}H and ^{17}O R_1 data (see Materials and Methods). Data from a pure water (HDO) reference sample (◆) are also shown, with the line representing the average value.

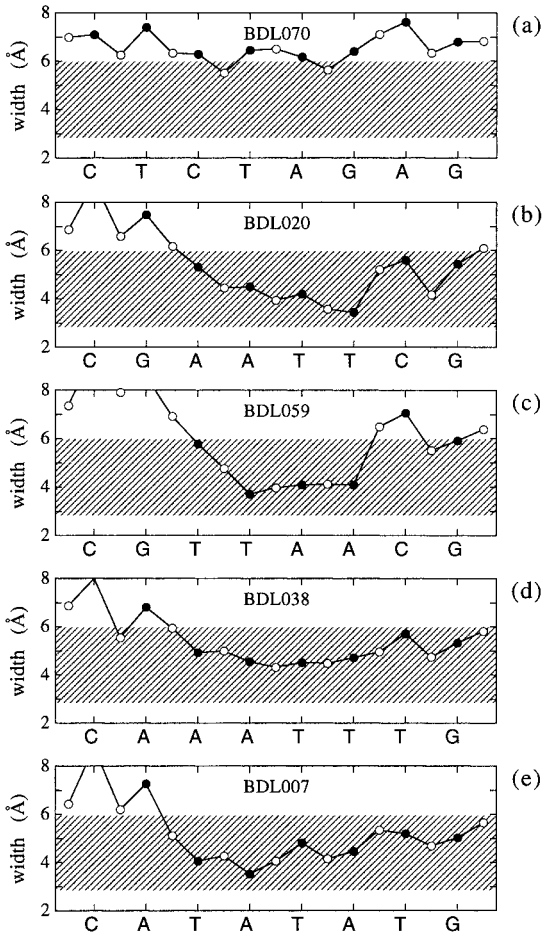

Figure 5. Minor groove width in the crystal structures of five B-DNA dodecamers: (a) CGCTCTAGAGCG,[65] (b) CGCGAATTCGCG,[5,59] (c) CGCGTTAACGCG,[90] (d) CGCAAATTTGCG,[61] and (e) CGCATATATGCG.[71] The Nucleic Acid Database code of each crystal structure is given. The minor groove width is measured as the shortest phosphorous separation across the groove at base pair n less 5.8 Å (two phosphate group van der Waals radii) (●), i.e., between P(n+2) in strand A and P(15−n) in strand B, and by the shortest furanose ring oxygen separation across the groove at base step n − (n+1) less 2.8 Å (two oxygen van der Waals radii) (○), i.e., between O4′(n+2) in strand A and O4′(14−n) in strand B. Here, the A and B strands are both indexed from 1 to 12 in the 5′ to 3′ direction. Points that fall within the shaded band are intermediate between the canonical B (6.0 Å) and B′ (2.8 Å) forms, as found in fibers of mixed-sequence B-DNA and poly(dA)·poly(dT), respectively.[6]

confined geometry, allowing passage of only one water molecule at a time, such a concerted mechanism is not possible. According to this view, the obvious and only candidate for long-lived water sites in B-DNA is the minor groove. Furthermore, the residence time of water molecules in the minor groove should depend on its width.

Figure 5 shows the minor groove width profiles in the crystal structures of five dodecamers, including the TA, A2T2, and A3T3 sequences investigated here. For the other three investigated sequences, with eight central A·T base pairs, crystal structures have not been reported. Significant trends can be inferred, however, from analogous A-tracts in a variety of decamers and dodecamers. Substantial minor groove narrowing is evident for all dodecamers in Figure 5 with one exception: the TA duplex has an even wider minor groove than the canonical B-form. The TA dodecamer also stands out in terms of dispersion amplitude (see Table 1 and Figure 4): $N_\beta S_\beta^2$ is a factor 7 smaller than for A4T4. This finding strongly indicates that the long-lived water molecules responsible for the dispersions are located in the minor groove. Conversely, it shows that not all water molecules in the minor groove are long-lived. Further support for the proposition that long-lived water molecules reside exclusively in the minor groove comes from the netropsin binding experiment: the dispersion amplitude is reduced by a factor 3.7 (see Table 1 and Figure 1) when netropsin binds to the central part of the minor groove, displacing its single-file hydration structure.[22,52,57] The numerical values deduced for the quantity $N_\beta S_\beta^2$ are also consistent with the minor groove as the sole habitat of long-lived water molecules. For the single-file hydration motif, one water molecule is positioned at each base pair step. If a long residence time is conditional on a narrowed minor groove, we would thus expect that $N_\beta S_\beta^2 \leq 11$, as indeed observed (see Table 1 and below).

Minor Groove Hydration: Dependence on Base Composition. If, as argued above, long-lived water molecules are located at the base pair steps of a single-file hydration motif in the narrow part of the minor groove, then $N_\beta \leq 11$ for a dodecamer. A lower bound on N_β is provided by the experimental $N_\beta S_\beta^2$ value: since $S_\beta^2 \leq 1$, it follows that $N_\beta \geq N_\beta S_\beta^2$. We consider first the $N_\beta S_\beta^2$ values derived from the ^{17}O dispersions as these invariably exceed the ^{2}H values (see Table 1) and therefore furnish the most restrictive lower bounds on N_β. We can also impose a symmetry restriction on N_β: since a self-complementary duplex must exhibit dyad symmetry in solution, N_β must

(57) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *J. Mol. Biol.* **1985**, *183*, 553−563.

be odd if the long-lived water molecules are located at the base pair steps. By applying these "rules" to the ^{17}O $N_\beta S_\beta^2$ data in Table 1, we arrive at several interesting conclusions.

For the free A2T2 dodecamer, with $N_\beta S_\beta^2 = 5.9 \pm 0.6$, N_β must be 7, 9, or 11. Although an even larger number of water molecules are seen at the floor of the minor groove in the crystal structure, at most five seem to conform to a regular single-file hydration motif with the water O atoms within hydrogen bonding distance of adenine N3 and thymine O2 atoms.[5,58,59] Accordingly, the single-file motif is generally thought to be confined to A-tracts and to be disrupted by the guanine N2 amino group.[5] Since the NMRD data indicate at least seven long-lived water molecules, it appears that, in solution, the single-file hydration motif extends into the C·G regions including, at least, the two C-G steps adjacent to the A-tract. This conclusion is supported by the finding that the netropsin complex contains at least two long-lived water molecules ($N_\beta S_\beta^2 = 1.6 \pm 0.4$). If there were only five long-lived water molecules in free A2T2, the complex would not have produced any dispersion at all since the bound netropsin molecule occupies five base steps.[22,52,57] The shorter single-file hydration seen in the A2T2 crystal structure can probably be ascribed to the intrusion into the minor groove of the terminal guanine of an adjacent duplex, an interaction which also induces a 19° bend of the helix axis at the A·T/G·C junction.[60] This crystal interaction is also responsible for the pronounced asymmetry in the minor groove width profile seen in Figure 5b. It seems plausible, therefore, that the minor groove of A2T2 in solution is narrowed (with respect to the canonical B-form) over at least seven base steps. It should be noted that a more extended single-file hydration of A2T2 is fully consistent with the NOE data available for this duplex:[22,24,26] while positive NOESY cross-peaks were seen only with the adenine H2 protons of the four central base pairs, the lack of suitable reporter nuclei would have rendered any long-lived water molecules in the G·C regions invisible. Furthermore, since the two adenine H2 cross-peaks disappear on netropsin binding,[22] the long-lived water molecules in the complex must be located outside the A-tract.

Turning now to the A3T3 dodecamer, conventional wisdom would predict a larger number of long-lived water molecules since the A-tract is 50% longer than in A2T2. Contrary to this expectation, we see a slightly smaller dispersion for A3T3 ($N_\beta S_\beta^2 = 5.1 \pm 0.6$) than for A2T2 (see Figure 3 and Table 1). This result correlates with the minor groove width profiles in the crystal structures, shown in Figure 5(parts b and d). Although the profiles have the same shape, the groove is uniformly slightly wider in A3T3, the average groove width in the A-tracts being 5.1 versus 4.2 Å. (The two dodecamers crystallize in the same space group with the same interlocking minor groove packing motif.) Even more significantly, the minor groove hydration is more disordered in the A3T3 crystal with only one water molecule (as opposed to four in A2T2) within hydrogen bonding distance of adenine N3 or thymine O2 atoms.[61]

Two dodecamers with a central A_3T_3 region have been investigated by the NOE technique. For CCTAAATTTGCC, positive NOESY cross-peaks were seen for several adenine H2 protons (at the underlined positions) indicating that most of the A-tract is occupied by long-lived water molecules.[27] The absence of a cross-peak from A4 was tentatively ascribed to groove widening at the T−A step. For the self-complementary dodecamer CGCAAATTTGCG (= A3T3 of the present study) positive cross-peaks were seen for all three adenine H2 protons.[31] A quantitative analysis yielded residence times in the range 0.3−1.5 ns at 10 °C, similar to the value 0.6 ns deduced for A2T2 at the same temperature.[22] The NOE results for A3T3 and A2T2 are fully consistent with the present NMRD results for the same dodecamers. The fact that positive cross-peaks (indicative of nearby long-lived water molecules) were observed for six base pairs in A3T3 but only for four base pairs in A2T2 does not imply that there are more long-lived water molecules in A3T3, because any long-lived water molecules in G·C regions would have escaped detection (see above).

Measurements of imino proton exchange rates have shown that the base pair lifetimes are anomalously long in the A3T3 dodecamer[10] (and in a homologous A_3T_3 decamer[9]) as well as in other A-tracts with more than four consecutive A·T base pairs (without a T−A step). This was correlated with the large-scale helix curvature inferred from the electrophoretic mobility and, by extension, with a local B′-DNA structure featuring a strongly narrowed minor groove. In other solution studies, the A-tract of A3T3 was found to be considerably more protected against hydroxyl radical cleavage than the A-tract of A2T2. The present NMRD results clearly challenge the views that slowing down of imino proton exchange[9,10] and protection against radical cleavage[11,12,62] are correlated with minor groove narrowing. Instead we suggest that these processes involve highly excited conformational substates of little relevance for the subset of low-energy conformations that determine the ensemble-averaged groove width in solution.

The A4T4 dodecamer produced the largest dispersion step of all the investigated sequences (see Table 1 and Figures 3 and 4). The ^{17}O dispersion yields $N_\beta S_\beta^2 = 8.6 \pm 0.4$, implying that N_β is 9 or (less likely) 11. The minor groove hydration in A4T4 is thus approaching the maximum length of the single-file hydration motif for a dodecamer. As in the case of A2T2, the single file appears to extend beyond the A-tract. No crystal structure has been reported for this sequence, but the homoadenine A-tracts in $CGCA_6GCG$[63] and $CGCGA_6CG$[64] show minor groove narrowing similar to that in A2T2 throughout their A-tracts. The NOE technique has been applied to a decamer with an even longer A-tract: AAAAATTTTT.[24,26] Positive cross-peaks were observed for the underlined adenine H2 protons, consistent with the presence of long-lived water molecules at all steps except the outermost ones. The present NMRD results for A4T4 are also consistent with long-lived hydration at all steps except the outermost ones, but in this case the single-file motif includes one G-A step and one T-C step at the extremities of the A-tract.

The TA dodecamer produces the smallest dispersion step by far among the six investigated sequences (see Figure 4 and Table 1). The ^{17}O dispersion yields $N_\beta S_\beta^2 = 1.2 \pm 0.3$, suggesting that N_β is 1 or (more likely) 3. This is in line with the wide minor groove seen in the crystal structure of this dodecamer[65] (see Figure 5a). If this geometry persists in solution, we would not expect any long-lived water molecules at all. The small but significant dispersion observed for this duplex therefore suggests that the minor groove exhibits local narrowing in solution. This explanation is consistent with the structural

(58) Kopka, M. L.; Fratini, A. V.; Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1983**, *163*, 129−146.
(59) Westhof, E. *J. Biomol. Struct. Dyn.* **1987**, 5, 581−600.
(60) Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, *149*, 761−786.
(61) Edwards, K. J.; Brown, D. G.; Spink, N.; Skelly, J. V.; Neidle, S. *J. Mol. Biol.* **1992**, *226*, 1161−1173.
(62) Burkhoff, A. M.; Tullius, T. D. *Cell* **1987**, *48*, 935−943.
(63) Nelson, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A. *Nature* **1987**, *330*, 221−226.
(64) DiGabriele, A. D.; Steitz, T. A. *J. Mol. Biol.* **1993**, *231*, 1024−1039.

Table 1. Mean Square Orientational Order Parameter S_β^2 and the Most Probable Number N_β of Long-Lived Water Molecules, Both Derived from ^{17}O NMRD Data

	$S_\beta^2(^{17}O)$						
sequence	$N_\beta = 1$	$N_\beta = 3$	$N_\beta = 5$	$N_\beta = 7$	$N_\beta = 9$	$N_\beta = 11$	N_β
TA	1.18 ± 0.28	$\mathbf{0.39 \pm 0.09}$	0.24 ± 0.06	0.17 ± 0.04	0.13 ± 0.03	0.11 ± 0.03	3
A2T2				$\mathbf{0.84 \pm 0.08}$	0.65 ± 0.06	0.53 ± 0.05	7
A2T2 difference			$\mathbf{0.81 \pm 0.04}$	0.58 ± 0.03	0.45 ± 0.02	0.37 ± 0.02	5
A3T3			1.03 ± 0.12	$\mathbf{0.73 \pm 0.08}$	0.57 ± 0.07	0.47 ± 0.05	7
A4T4					$\mathbf{0.96 \pm 0.04}$	0.79 ± 0.03	9
T4A4				$\mathbf{0.82 \pm 0.10}$	0.64 ± 0.07	0.52 ± 0.06	7
(AT)4				1.03 ± 0.08	$\mathbf{0.80 \pm 0.06}$	0.65 ± 0.05	9

variability of the CTAG segment observed in crystal structures of several free oligomers and of the *met* and *trp* repressor-operator complexes.[65,66]

Minor Groove Hydration: Dependence on Base Sequence. We now compare the dispersions from the dodecamers A4T4, T4A4, and (AT)4, which differ in the order of the adenine and thymine bases within the A-tract (see Figure 4 and Table 1). As already discussed, A4T4 yields a large dispersion step corresponding to highly ordered and long-lived water molecules at nine of the 11 base steps. Reversal of all base steps, as in T4A4, reduces $N_\beta S_\beta^2$ by 33% to 5.8 ± 0.7, similar to A2T2. The alternating A−T sequence (AT)4, with three T−A steps, yields $N_\beta S_\beta^2 = 7.2 \pm 0.5$, intermediate between A4T4 and T4A4.

Several lines of evidence have suggested that a T−A step inserted into an A-tract has the effect of widening the minor groove and disrupting the single-file hydration motif.[9,11,67,68] While we do observe a 33% reduction of $N_\beta S_\beta^2$ on going from A4T4 to T4A4, this is nothing like the all-or-none effect suggested previously. Moreover, if a T−A step widens the minor groove, we would expect $N_\beta S_\beta^2$ to decrease on going from T4A4 with one T−A step to (AT)4 with three consecutive T−A steps, whereas, in fact, we see an increase (see Table 1). The NMRD results thus appear to challenge some of the previous interpretations. The available crystallographic data bearing on this issue are not entirely conclusive but seem to suggest that T−A steps are more flexible, and hence more easily perturbed by crystal forces, than A−A or A−T steps.[6,69−71] While measurements of the apparent molar volume and adiabatic compressibility have been taken to imply an anomalously high level of hydration for poly(dA)·poly(dT) as compared to poly(dAT)·poly(dAT),[16,17,72] no difference was found between A_4T_4 and $(AT)_4$.[16]

The available NOE results on the effect of T−A steps on minor groove hydration are also somewhat ambiguous. Positive cross-peaks (indicating a residence time > 0.3 ns) were reported for the underlined adenine H2 protons in GTGGAATTCCAC, whereas the cross-peaks in GTGGTTAACCAC were negative.[28] On the assumption that a T−A step widens the minor groove, it was therefore proposed that a long residence time requires a narrow minor groove.[28] Structural data on these duplexes are not available. Whereas GTGGTTAACCAC showed negative cross-peaks (short residence time), a subsequent NOE study showed positive cross-peaks (long residence time) for the decamer GCATTAATGC (but a strongly negative cross-peak for A3).[29] Like the present NMRD results on T4A4 and (AT)4, this finding invalidates the conjecture that a T−A step disrupts the single-file hydration structure[68] and leads to short residence times.[28] Unfortunately, no crystal structure has been reported for this decamer. In the decamer GCATTAACGC, the adenine H2 cross-peaks progressed from positive to negative on going from the central two A·T base pairs to the outermost (3/18) A·T base pair,[29] while in the decamer GCCTTAAAGC there was no indication of long-lived hydration at all.[29]

It is not easy to make general statements about the sequence-dependence of minor groove hydration on the basis of the available NOE data. This could, of course, reflect a complicated sequence-structure relationship for DNA in solution, as has been found in the crystalline state (although sometimes attributed to crystal forces). On the other hand, the inability to detect long-lived water molecules near C·G base pairs introduces a bias in the NOE-derived picture of minor groove hydration and tends to exaggerate relative differences. Furthermore, it appears that, in the temperature range 4−10 °C where most NOE studies have been carried out, many water molecules in the minor groove happen to have residence times not far from the NOE zero-crossing of 0.3 ns (for a ^{1}H frequency of 600 MHz). Variations in the sign of cross-peaks could then result from rather subtle variations in water dynamics and DNA structure. In view of the present results, it seems likely that all of the oligomers investigated by the NOE technique at 4 °C would produce a substantial relaxation dispersion at −20 °C. To date, only three oligonucleotides have been investigated by both the NOE and NMRD techniques.[22−24,26,31,73] Further comparative studies are clearly desirable.

Minor Groove Hydration: Orientational Order. The long-lived water molecules in the minor groove that give rise to the relaxation dispersion interact with the DNA duplex primarily via hydrogen bonds to base O2 and N3 and furanose O4′ atoms. Any local reorientation of these water molecules during their residence time (ca. 10 ns, see below) has the effect of partially averaging the anisotropic electric quadrupole interaction that induces spin relaxation.[43] This effect is described by the second-rank order parameter S_β, ranging from 0 (for a water molecule undergoing an isotropic local motion) to 1 (for a water molecule rigidly attached to the DNA duplex).[21,74]

Table 2 shows the range of possible S_β^2 values, obtained by dividing the $N_\beta S_\beta^2$ derived from the ^{17}O dispersions with any of

(65) Urpi, L.; Tereshko, V.; Malinina, L.; Huynh-Dinh, T.; Subirana, J. A. *Nature Struct. Biol.* **1996**, *3*, 325−328.
(66) Hunter, W. N.; D'Estaintot, B. L.; Kennard, O. *Biochemistry* **1989**, *28*, 2444−2451.
(67) Hagerman, P. J. *Nature* **1986**, *321*, 449−450.
(68) Chuprina, V. P. *Nucl. Acid Res.* **1987**, *15*, 293−311.
(69) Leslie, A. G. W.; Arnott, S.; Chandrasekaran, R.; Ratliff, R. L. *J. Mol. Biol.* **1980**, *143*, 49−72.
(70) Yuan, H.; Quintana, J.; Dickerson, R. E. *Biochemistry* **1992**, *31*, 8009−8021.
(71) Yoon, C.; Privé, G. G.; Goodsell, D. S.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6332−6336.
(72) Marky, L. A.; Kupke, D. W. *Biochemistry* **1989**, *28*, 9982−9988.
(73) Sunnerhagen, M.; Denisov, V. P.; Venu, K.; Bonvin, A. M. J. J.; Carey, J.; Halle, B.; Otting, B. *J. Mol. Biol.* **1998**, in press.
(74) Halle, B.; Wennerström, H. J. *Chem. Phys.* **1981**, *75*, 1928−1943.

the allowed N_β values (see above). Entries in boldface correspond to the most probable number of long-lived water molecules, arrived at as follows. Since netropsin occupies five base steps, we take $N_\beta = 5$ for the A2T2 difference dispersion, resulting in $S_\beta^2 = 0.81 \pm 0.04$ as an average for these five water molecules. Unless the additional long-lived water molecules in the free A2T2 duplex are much less ordered, we must have $N_\beta = 7$ here. Since the long-lived water molecules in A3T3 should not be much more ordered than those in A2T2 (considering the wider groove), $N_\beta = 5$ is rejected. We thus set $N_\beta = 7$ also here. For A4T4, we set $N_\beta = 9$ since water molecules at the terminal base steps are exposed to bulk water and, hence, are not likely to be long-lived (even if fraying is minimal at −20 °C). For T4A4 and (AT)4 we take the lowest N_β that yields an order parameter that is not higher than in A4T4. For TA, the wide groove suggests $N_\beta = 3$ as more likely than $N_\beta = 1$. For still larger N_β, S_β^2 becomes too small to be compatible with a long residence time. While these assignments are clearly not rigorous, they should not be far from the truth.

With the exception of the extreme cases TA and A4T4, S_β^2 is close to 0.8. To get an idea of what this means in terms of motional amplitudes, we show in Figure 6a the calculated variation of $S_\beta^2(^{17}O)$ with the rms angular amplitude for an anisotropic harmonic libration model.[75] The three libration modes refer to motion around the dipole axis (twist), around an axis perpendicular to the molecular plane (rock), and around an axis parallel to the H−H vector (wag). The ^{17}O order parameter is seen to be most sensitive to the twist mode, which can account for the deduced S_β^2 values if the libration amplitude is 10−15° (and 6° for A4T4). With the possible exception of the TA sequence, the long-lived water molecules in the minor groove of all the investigated DNA duplexes are thus highly ordered.

Rather than directly analyzing the $N_\beta S_\beta^2$ values derived from the ^2H dispersions, we consider the ratio of the $N_\beta S_\beta^2$ values from the ^2H and ^{17}O dispersions, which yields the order parameter ratio $S_\beta^2(^2H)/S_\beta^2(^{17}O)$ without any assumptions about N_β. This ratio exhibits a remarkable invariance, clustering around 0.6 (see Table 1). Figure 6b shows that, within the anisotropic harmonic libration model, a ratio of 0.6 would correspond to very large rock or wag amplitudes and is entirely inconsistent with a twist libration (which provides a plausible interpretation of the ^{17}O order parameter). The situation changes radically, however, if (some of) the water molecules undergo 180° flips around the water dipole (C_2) axis, while they reside in the minor groove (i.e., on time scales shorter than ca. 10 ns). For symmetry reasons, a C_2 flip has no effect on the ^{17}O order parameter.[56] In contrast, $S_\beta^2(^2H)$ is reduced to 0.366 by this motion alone. As seen from Figure 6b, the $S_\beta^2(^2H)/S_\beta^2(^{17}O)$ ratio is hardly affected by small-amplitude librations superimposed on a C_2 flip. Since the experimental $S_\beta^2(^2H/S_\beta^2(^{17}O)$ ratios are larger than 0.366, not all long-lived water molecules are flipping. Assuming that all N_β water molecules in a given duplex have the same twist libration amplitude, we find that the experimental ^2H and ^{17}O order parameters can be accounted for if all but 1−3 water molecules are flipping rapidly.

In view of the suggestion that the release of ordered hydration water contributes to entropy driven drug binding[76,77] and premelting transitions,[78,79] it is of interest to estimate the rotational entropy corresponding to the order parameters obtained here. Since the C_2 flip interconverts two indistinguishable states, it does not contribute to the configurational entropy. Considering only the twist libration mode, $S_\beta^2(^{17}O)$ values of 0.96, as for A4T4, and 0.80, as for (AT)4, correspond to rotational entropies of 0.3 and 1.7 cal mol^{-1} K^{-1}, respectively.[75] Proceeding as in a previous analysis of water in protein cavities,[75] we find for the entropy of transfer (at 300 K) from the minor groove to bulk water 6.8 cal mol^{-1} K^{-1} for A4T4, the same value as for transfer from ice to water, and 5.3 cal mol^{-1} K^{-1} for (AT)4. These values may be compared with $\Delta S = 33$ cal (mol drug)$^{-1}$ K^{-1} for netropsin binding to poly(dA)·poly(dT) at 25 °C.[77] Since one netropsin molecule displaces five long-lived water molecules, it does indeed appear that release of ordered water could make a major contribution to the binding entropy. It should be remembered, however, that several assumptions are involved in our entropy estimates. If other libration modes (apart from twist) that do not affect the ^2H and ^{17}O order parameters much are significantly excited, then we have overestimated the entropy of transfer. This would also be the case if the libration amplitude were significantly larger at room temperature than at −20 °C.

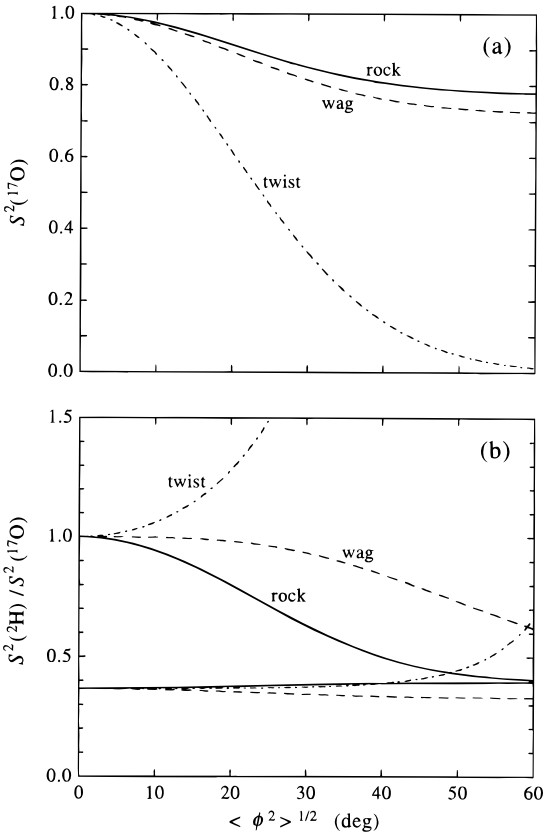

Figure 6. Variation of the ^{17}O order parameter (a) and the ratio of the ^2H and ^{17}O order parameters (b) with the rms libration angle for the rock (solid curves), wag (dashed), and twist (dash-dot) modes in the anisotropic harmonic librabon model.[21,75] In (b), the lower set of curves include the effect of a fast C_2 flip. Electric field gradient asymmetry parameters of 0.11 (^2H) and 0.93 (^{17}O) and a HOH angle of 107° were used in the calculations.

Minor Groove Hydration: Residence Time and Temperature Dependence. In contrast to $N_\beta S_\beta^2$, which varies by a factor 7, the correlation time τ_β is nearly the same, about 10

(75) Denisov, V. P.; Venu, K.; Peters, J.; Hörlein, H. D.; Halle, B. *J. Phys. Chem. B* **1997**, *101*, 9380−9389.
(76) Wilson, W. D.; Wang, Y.−H.; Krishnamoorthy, C. R.; Smith, J. C. *Biochemistry* **1985**, *24*, 3991−3999.
(77) Breslauer, K. J.; Remeta, D. P.; Chou, W.−Y.; Ferrante, R.; Curry, J.; Zaunczkowski, D.; Snyder, J. G.; Marky, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8922−8926.
(78) Herrera, J. E.; Chaires, J. B. *Biochemistry* **1989**, *28*, 1993−2000.
(79) Breslauer, K. J. *Curr. Opin. Struct. Biol.* **1991**, *1*, 416−422.

ns, for the six investigated dodecamers (see Table 1). This invariance would not have been remarkable if all water molecules that contribute to the observed dispersions have residence times τ_W long compared to the rotational correlation time τ_R of the duplex. According to eq 6, we would then have $\tau_\beta = \tau_R$, and τ_R should be virtually the same for all B-DNA dodecamers. This explanation of the τ_β invariance cannot be correct, however, since the high viscosity of water at −20 °C (7.6 cP for our D_2O solvent) implies a τ_R of ca. 40 ns at infinite dilution, increasing to 60−100 ns at our high DNA concentration (see Materials and Methods). It then follows from eq 1 that the observed τ_β is essentially the water residence time. The last column of Table 1 shows the τ_W values obtained from eq 1 by setting $\tau_R = 50$ ns (for the D_2O solvent), a conservative lower bound on τ_R. Since τ_β is always a lower bound on τ_W, the actual residence time should lie in the narrow range defined by the last two columns of Table 1.

The residence time of the long-lived water molecules in the minor groove thus appears to be essentially independent of base composition and sequence, at least for the sequences investigated here. Moreover, for a given dodecamer, the several long-lived water molecules responsible for the dispersion must have similar residence times. As illustrated in Figure 7a, even a single ordered water molecule with a residence time comparable to or longer than the tumbling time of the duplex would have been clearly revealed in the NMRD data, primarily through a large difference between R_2 and R_1 at low frequencies. On the other hand, the NMRD data would not distinguish a modest number of ordered water molecules with τ_W of about 1 ns among the ca. 300 water molecules in contact with the DNA surface (see Figure 7b and next section). The detection limit of the present NMRD data is a few ordered water molecules with a τ_W of a few nanoseconds.

To appreciate the implications of the observed τ_W invariance, it is necessary to consider the mechanism of exchange of water molecules from a single-file hydration motif in the minor groove into the bulk solvent. Two limiting models can be distinguished, which we refer to as closed-state exchange (CSE) and open-state exchange (OSE). In the CSE model, water exchange is essentially decoupled from fluctuations in DNA structure. The residence time would then be controlled by the local minor groove morphology and by the hydrogen bonding status of the water molecule. Since both of these factors are expected (largely on the basis of crystallographic data) to vary with base composition and sequence, the observed τ_W invariance argues against the CSE model.

In the OSE model, the rate-limiting step for water exchange is the opening up (widening) of the minor groove and τ_W is the mean lifetime of the closed state with a narrow minor groove. The open state presumably has a much shorter lifetime and, hence, will not significantly affect measurements that only probe the static aspects of the equilibrium ensemble of DNA structures. Nevertheless, the lifetime of the open state should be long compared to the residence time (of order 50 ps at −20 °C) of exposed water molecules so that water exchange occurs during each opening event. Two variants of the OSE model can be envisaged, depending on whether groove opening is local, allowing only one water molecule to exchange, or global, allowing all long-lived water molecules to exchange during one opening event. The observed τ_W invariance seems to favor a global OSE model, where, moreover, the dynamics of groove width fluctuations are insensitive to base composition and sequence. We note that all six investigated duplexes are of the same length (12 base pairs). In contrast, a local OSE model must be invoked for imino proton exchange to account for the substantial variation of base pair lifetimes within an A-tract.[9,10] Since base pair opening is 7−8 orders of magnitude slower than water exchange from the minor groove, it is clear that the two processes probe different regions of the conformational energy landscape. The disruption of Watson−Crick base pairing, required for imino proton exchange, may well be more sensitive to local interactions than the groove width fluctuations that seem to control water exchange.

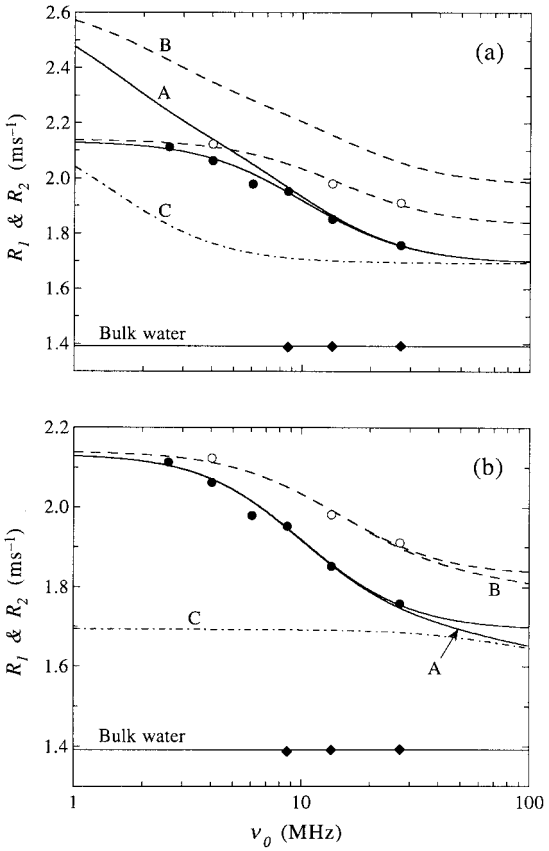

Figure 7. Dispersion of the water ^{17}O longitudinal (●, solid curves) and transverse (○. dashed curves) relaxation rates for the dodecamer duplex A2T2. The data and fits are the same as in Figure la. The three new curves, labeled A, B, and C, simulate the effect of long-lived water molecules in addition to the ones deduced from the experimental data ($N_\beta S_\beta^2 = 5.9$, $\tau_\beta = 8.8$ ns). Curve A is the new R_1 dispersion, curve B the new R_2 dispersion, and curve C the R_1 dispersion that would have been produced by the new class of long-lived water molecules alone. In panel (a), a single ordered water molecule ($N_\beta S_\beta^2 = 1$) with $\tau_\beta = 60$ ns is added. In panel (b), ten ordered water molecule ($N_\beta S_\beta^2 = 10$) with $\tau_\beta = 1$ ns are added with a corresponding reduction of α so that the low-frequency plateau is unaffected.

If water exchange is strongly coupled to groove width fluctuations, as in the OSE model, water can be said to be an integral part of the DNA duplex also in a dynamic sense. This applies, however, only to the relatively few water molecules in the single-file hydration motif in a narrowed minor groove. Recent MD simulations of the A2T2 dodecamer,[80,81] although still subject to force-field uncertainties,[82] suggest very large (up to 5 Å) fluctuations of minor groove width (on a subnanosecond time scale at 25 °C). We note also that normal mode calculations have emphasized the coupling between minor

(80) Duan, Y.; Wilkosz, P.; Crowley, M.; Rosenberg, J. M. *J. Mol. Biol.* **1997**, *272*, 553−572.
(81) Young, M. A.; Ravishanker, G.; Beveridge, D. L. *Biophys. J.* **1997**, *73*, 2313−2336.
(82) Feig, M.; Pettitt, B. M. *J. Phys. Chem. B* **1997**, *101*, 7361−7363.

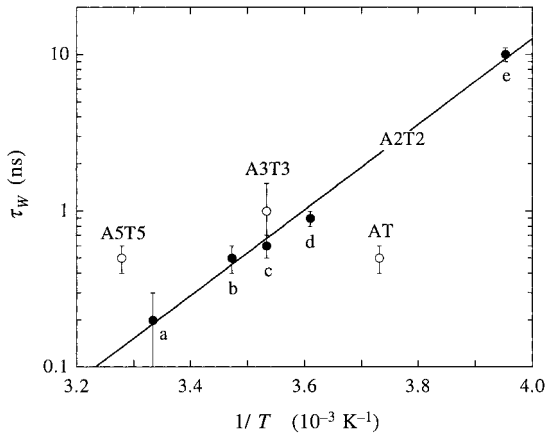

Figure 8. Temperature variation of the residence time of water molecules in the minor groove of B-DNA duplexes, determined by the NMRD and NOE techniques. The line resulted from an Arrhenius fit to the five points (●) that refer to the A2T2 dodecamer. The sources and methods for these data are as follows: (a) NMRD, low-frequency plateau,[22] (b) zero-crossing of NOESY cross-peaks,[26] (c) ratio of NOESY and ROESY cross-peak intensities,[22] (d) NMRD, partial dispersion,[22] and (e) NMRD, full dispersion (this work). Results for three other duplexes (○) are also shown: the A3T3 dodecamer, with τ_W obtained from the ratio of NOESY and ROESY cross-peak intensities,[31] and the decamers A_5T_5 and CGCGATCGCG, with τ_W obtained from the zero-crossing of NOESY cross-peaks.[26] The residence times determined here for five more dodecamers and one DNA−drug complex are close to point (e), but are omitted for clarity.

groove morphology and the presence of ordered single-file hydration in poly(dA)·poly(dT).[83,84] It can be argued that the sensitivity of DNA morphology to helix−helix interactions in crystals and to the details of force fields in solution-state MD simulations both suggest a relatively flat conformational energy landscape and a correspondingly flexible DNA structure.

The available data on residence times of water molecules in the minor groove of B-DNA duplexes are collected in Figure 8. These residence times were derived from NMRD and NOE studies as detailed in the legend. (In most other NOE studies, only bounds on τ_W were reported.) An Arrhenius fit according to $\tau_W(T) = \tau_W^0 \exp[\Delta H^{\#}/(RT)]$ yields an activation enthalpy $\Delta H^{\#} = 52 \pm 3$ kJ mol^{-1} for the A2T2 dodecamer, more than twice the 20−25 kJ mol^{-1} activation enthalpy of solvent viscosity in the same temperature range (the higher figure refers to D_2O). This difference indicates that water exchange from the minor groove of DNA occurs by a mechanism other than the cooperative restructuring of the hydrogen-bond network that controls transport processes in bulk water. If the reaction coordinate for water exchange from the minor groove involves many coupled degrees of freedom, it may be more appropriate to model the temperature dependence of τ_W in terms of a Gaussian energy landscape,[85,86] with $\tau_w(T) = \tau_W^0 \exp[\langle H^2\rangle/(RT)^2]$, rather than in terms of a single smooth barrier (implicit in the Arrhenius picture). The fit of this super-Arrhenius function to the τ_W data is of comparable quality to that shown in Figure 8 and yields for the rms ruggedness of the energy landscape $\langle H^2\rangle^{1/2} = 7.7 \pm 0.2$ kJ mol^{-1}, comparable to the relevant noncovalent interactions. The physical significance of these activation parameters would be affected if the DNA structure, in particular the groove width, varies with temperature. The very similar $N_\beta S^2$ values found at −20 °C (5.9 ± 0.6) and at +4 °C (6.4 ± 0.8) from the ^{17}O A2T2/netropsin difference dispersion, however, seem to rule out significant structural changes in this temperature range.

What, then, is the nature of the open state involved in water exchange from the minor groove? Unless the energy landscape is strongly temperature-dependent, the open state should be more populated at higher temperatures. Moreover, in the global OSE model favored by the NMRD data, the closed → open conformational transition should exhibit some degree of cooperativity. On the basis of these considerations, it is tempting to identify the open state accessed by 10 ns fluctuations at −20 °C with the high-temperature state that prevails above the premelting transition observed in the 30−40 °C range by CD and UV spectroscopy.[78,87,88] This premelting transition has been linked to minor groove widening and consequent disruption of the single-file hydration motif.[15,78,79]

While the water residence times derived from NMRD and NOE data agree in the case of A2T2 (although NMRD detects more long-lived water molecules), there is less agreement regarding the dependence of τ_W on base sequence and location within an A-tract. Whereas we find τ_W to be nearly the same for all investigated sequences, the zero-crossing temperature for NOESY cross-peaks was reported as 32 °C for A_5T_5 and −5 °C for CGCGATCGCG.[26] With an activation enthalpy of 50 kJ mol^{-1}, this corresponds to a 15-fold variation of τ_W at a fixed temperature. (These results are also seen to deviate substantially from the line in Figure 8.) The origin of this discrepancy is not clear to us. It should be stressed, however, that whereas we obtain τ_W directly from the dispersion frequency (under the present conditions, the dispersion profile is essentially the Fourier transform of the survival probability for the exchanging water molecules), the NOE cross-peak intensity is less directly related to the residence time.[22] For example, it is generally assumed that the order parameter for the intermolecular H−H vector is unity as for a water molecule rigidly bound to a rigid DNA duplex. The effect of a fast modulation of the orientation and/or length of this vector is to shift the zero-crossing to longer τ_W and to reduce the intensity of positive NOESY cross-peaks.[22,89] The intermolecular order parameter for a fast C_2 flip is $S^2 = [1 + 2x^3P_2(\cos\theta) + x^6]/[2(1 + x^6)]$, with x the ratio of the H−H separations in the two symmetry-related configurations and θ the angle between the corresponding H−H vectors. Using crystallographic data,[5,59] we find that S^2 is about 0.8 for the water molecules that produce positive NOESY cross-peaks with the A2T2 dodecamer. To this should be added the effect of librations and intermolecular vibrations. When NOESY cross-peaks are observed with several adenine H2 protons in an A-tract, it is usually found that their intensity decreases (sometimes passing through zero) toward the periphery of the A-tract.[24,27−29,31] While this effect is usually interpreted as a reduction of τ_W, it could also be produced by a reduction of the intermolecular order parameter.

Surface Hydration. The high-frequency plateau of the relaxation dispersion provides a global measure of the dynamic perturbation of all short-lived (τ_W < 1 ns) water molecules in contact with the DNA duplex. The number N_α of water

(83) Chen, Y. Z.; Prohofsky, E. W. *Nucl. Acid Res.* **1992**, *20*, 415−419.
(84) Chen, Y. Z.; Prohofsky, E. W. *Biophys. J.* **1993**, *64*, 1385−1393.
(85) Zwanzig, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2029−2030.
(86) Denisov, V. P.; Peters, J.; Hörlein, H. D.; Halle, B. *Nature Struct. Biol.* **1996**, *3*, 505−509.
(87) Chan, S. S.; Breslauer, K. J.; Hogan, M. E.; Kessler, D. J.; Austin, R. H.; Ojemann, J.; Passner, J. M.; Wiles, N. C. *Biochemistry* **1990**, *29*, 6161−6171.
(88) Park, Y.−W.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 1551−1555.
(89) Otting, G.; Liepinsh, E.; Wüthrich, K. *Science* **1991**, *254*, 974−980.
(90) Balendiran, K.; Rao, S. T.; Sekharudu, C. Y.; Zon, G.; Sundaralingam, M. *Acta Crystallogr. D* **1995**, *51*, 190−198.

molecules in contact with the DNA surface can be estimated from the solvent accessible surface area or from computer simulations.[3,81] Both approaches yield slightly more than 20 water molecules per base pair. Taking end effects into account, we thus estimate a coordination number $N_\alpha = 300$ for a B-DNA dodecamer. With $N_\alpha\rho_\alpha = 1500$, as found from the ^{17}O dispersions in the first four entries of Table 1, we then obtain a dynamic retardation factor $\rho_\alpha = 5$ (see Materials and Methods). The somewhat higher values of $N_\alpha\rho_\alpha$ obtained for the last three sequences in Table 1 may reflect a contribution from water molecules with residence times of order 1 ns (see above) or may simply be due to the slightly higher impurity content of these samples as indicated by HPLC analysis. For compact globular proteins, ρ_α is in the range 4−6 (at 27 °C).[55] The average dynamic perturbation in the primary hydration layer of B-DNA is therefore not significantly different from that of typical proteins, despite the highly charged DNA phosphate backbone. This conclusion is in line with the insignificant pH dependence of ρ_α in proteins, showing that ionic residues are not particularly potent in perturbing water dynamics.[55] With two exceptions, the ratio of ^{2}H and ^{17}O retardation factors is in the range 0.9−1.0 (see Table 1), indicating nearly isotropic reorientation of surface water. For globular proteins, this ratio is also close to unity (0.9−1.1), except in a few cases (at low or high pH) where labile hydrogens contribute to $\alpha(^2\text{H})$.[55]

Conclusions

The present ^{2}H and ^{17}O NMRD data, measured on emulsified samples at −20 °C, display an order-of-magnitude enhancement of the dispersion step as compared to previous NMRD studies of DNA. This improved sensitivity made it possible to study the dependence of minor groove hydration on nucleotide base sequence for the first time with the NMRD technique. The principal conclusions of this study are as follows.

(1) Highly ordered and long-lived (ca. 10 ns) water molecules are found in all six investigated dodecamers. The NMRD data are consistent with localization of these water molecules at the base steps in the minor groove, forming a single-file hydration motif. This is supported by three observations: (*i*) the reduction of the dispersion amplitude induced on netropsin binding corresponds to five water molecules, as expected from independent structural data, (*ii*) the estimated number of long-lived water molecules ranges from 3 to 9, consistent with the presence in a dodecamer of nine base step sites with low accessibility to external solvent, and (*iii*) the variation of the number of long-lived water molecules is correlated with the width of the minor groove in the available crystal structures.

(2) The single-file hydration motif extends beyond the A-tract, including one or two base steps on either side. Ordered water molecules at C-G steps have not been detected by the NOE technique because of the lack of suitable reporter nuclei and are often perturbed in crystal structures by neighboring helices penetrating the outer regions of the minor groove.

(3) The NMRD data do not support the common view that a T−A step within an A-tract widens the minor groove and disrupts the single-file hydration motif. For the A-tracts AAAATTTT, TTTTAAAA, and ATATATAT, we find only marginal variations of the number of long-lived water molecules (and their residence times). This suggests that the minor groove width is similar in these A-tracts and challenges the usual interpretation of anomalous gel migration, hydroxyl radical cleavage, and imino proton exchange in terms of minor groove narrowing in A-tracts devoid of T−A steps.

(4) Most, but not all, long-lived water molecules undergo a symmetric flip motion during their residence time but are otherwise highly ordered with only 10−15° libration amplitude. The entropy associated with transfer of these water molecules into bulk solvent is comparable to the entropy difference between ice and bulk water. This finding supports the earlier suggestion that the elimination of ordered water provides an important entropic driving force for drug binding and premelting transition.

(5) For all investigated dodecamers, the residence time of the long-lived water molecules is nearly the same, 10−15 ns at −20 °C. The presence of even a single water molecule with a significantly longer residence time in any of the dodecamers can be ruled out. There also appears to be little variation of the residence time for the individual water molecules in the single-file hydration motif of a given dodecamer sequence. The available NMRD and NOE results on the temperature-dependent residerice time of long-lived water molecules in the A2T2 dodecamer are mutually consistent and yield an Arrhenius activation enthalpy of 52 ± 3 kJ mol^{-1} or a super-Arrhenius rms ruggedness of the "energy landscape" of 7.7 ± 0.2 kJ mol^{-1}.

(6) The invariance of the residence time with respect to base sequence and location within the minor groove as well as the large activation enthalpy suggest that water exchange occurs from an open state with a uniformly wide minor groove. Within a two-state model, the water residence time of 10−15 ns can then be identified with the lifetime (at −20 °C) of the closed state with a narrowed B′-form minor groove morphology. This is consistent with recent MD simulations, showing large fluctuations in the minor groove width of the A2T2 dodecamer.

(7) Apart from the single-file hydration motif in the minor groove, all water molecules coordinating the B-DNA duplex, including its highly charged phosphate backbone, have short residence times (≪10 ns at −20 °C), and their reorientational dynamics do not differ significantly from that of water molecules at the surface of globular proteins.

Acknowledgment. We are grateful to Dr. Elmar Lang and Dr. Vladimir Denisov for helpful advice. This work was supported by grants from the Crafoord Foundation and the Swedish Natural Science Research Council (NFR).

JA974316R