The NMR Structure of a DNA Dodecamer in an Aqueous Dilute Liquid Crystalline Phase

Nico Tjandra,*,† Shin-ichi Tate,*,‡,§ Akira Ono,§ Masatsune Kainosho,*,§ and Ad Bax*,||

Laboratory of Biophysical Chemistry, NHLBI, Building 3-418, National Institutes of Health, Bethesda, Maryland 20892-0380, Center for New Materials, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Tatsunokuchi, Ishikawa 923-1292, Japan, Tokyo Metropolitan University, 1-1 Minami-ohsawa, Hachioji, Tokyo 192-0397, Japan, and Laboratory of Chemical Physics, NIDDK, Building 5-126, National Institutes of Health, Bethesda, Maryland 20892-0520

Received January 28, 2000. Revised Manuscript Received April 10, 2000

Abstract: The solution structure of the DNA dodecamer d(CGCGAATTCGCG)$_2$ has been studied in an aqueous liquid crystalline medium containing 5% w/v bicelles. These phospholipid particles impose a small degree of orientation on the DNA duplex molecules with respect to the magnetic field and permit the measurement of dipolar interactions. Experiments were carried out on several samples with different isotopic labeling patterns, including two complementary samples, in which half of the nucleotides were uniformly enriched with ^{13}C and deuterated at the H2″and H5′ positions. From this, 198 ^{13}C−^{1}H and 10 ^{15}N−^{1}H one-bond dipolar coupling restraints were derived, in addition to 200 approximate ^{1}H−^{1}H dipolar coupling and 162 structurally meaningful NOE restraints. Although loose empirical restraints for the phosphodiester backbone torsion angles were essential for obtaining structures that satisfy all experimental data, they do not contribute to the energetic penalty function of the final minimized structures. Except for additional regular Watson−Crick hydrogen bond restraints and standard van der Waals and electrostatic terms used in the molecular dynamics-based structure calculation, the structure is determined primarily by the dipolar couplings. The final structure is highly regular, without any significant bending or kinks, and with C2′-endo/C1′-exo sugar puckers corresponding to regular B-form DNA. Most local parameters, including sugar puckers, glycosyl torsion angles, and propeller twists, are also tightly determined by the NMR data. The precision of the determined structures is limited primarily by the uncertainty in the exact magnitude and rhombicity of the alignment tensor. This causes considerable spread in parameters such as the degree of base-pair opening and the width of the minor groove, which are relatively sensitive to the alignment tensor values used.

Over the past 15 years, multidimensional NMR[1] has become a well-established procedure for determination of the three-dimensional structure of proteins. Initial studies were based almost exclusively on two-dimensional homonuclear NOE and *J* coupling experiments,[2] but the introduction of ^{13}C and ^{15}N labeling has greatly expanded the power of this approach.[3,4] Isotopic enrichment not only provides dramatic simplification of the spectral assignment and NOE analyses steps but also offers access to new structural parameters such as heteronuclear *J* couplings and ^{13}C chemical shifts.[5] The high density of protons in proteins, and the ability to measure a vast number of NOEs between them, make protein structure determination by NMR into a relatively robust technique.

Structure determination of nucleic acids by NMR tends to be more difficult than for proteins. The low proton density in this type of molecule, together with the absence of a globular fold and the resulting paucity of NOE restraints between protons on residues far apart in the nucleic acid sequence, makes the NMR study of the overall shape of these molecules more challenging.[6−9] Even accurate determination of local geometry can be difficult because the low redundancy in the measured NOE information makes it necessary to quantify these NOEs more precisely. This in turn may be difficult in the presence of indirect NOE effects, internal motion, and anisotropic rotational diffusion. Although structure determination of nucleic acids also has benefited from isotope enrichment procedures, which permit the measurement of numerous torsion angles,[5,10,11] these serve primarily to improve the local geometry and not so much the overall shape of the molecule.

The need for accurate NMR structure determination is arguably even greater for nucleic acids than for proteins as nucleic acids are known to be relatively flexible structures whose conformation can be significantly affected by crystallization.[12,13]

† Laboratory of Biophysical Chemistry, National Institutes of Health.
‡ Japan Advanced Institute of Science and Technology.
§ Tokyo Metropolitan University.
|| Laboratory of Chemical Physics, National Institutes of Health.

(1) Ernst, R. R.; Bodenhausen, G.; Wokaun, A. *Principles of nuclear magnetic resonance in one and two dimensions*; Clarendon Press: Oxford, U.K., 1987.
(2) Wüthrich, K. *NMR of proteins and nucleic acids*; Wiley: New York, 1986.
(3) Clore, G. M.; Gronenborn, A. M. *Nat. Struct. Biol.* **1997**, *4*, 849−853.
(4) Ikura, M.; Kay, L. E.; Bax, A. *Biochemistry* **1990**, *29*, 4659−4667.
(5) Ippel, J. H.; Wijmenga, S. S.; deJong, R.; Heus, H. A.; Hilbers, C. W.; deVroom, E.; vanderMarel, G. A.; vanBoom, J. H. *Magn. Reson. Chem.* **1996**, *34*, S156−S176.
(6) Schmitz, U.; James, T. L. *Methods Enzymol.* **1995**, *261*, 3−44.
(7) Allain, F. H. T.; Varani, G. *J. Mol. Biol.* **1997**, *267*, 338−351.
(8) Tonelli, M.; James, T. L. *Biochemistry* **1998**, *37*, 11478−11487.
(9) Leijon, M.; Zdunek, J.; Fritzsche, H.; Šklenar, H.; Graslund, A. *Eur. J. Biochem.* **1995**, *234*, 832−842.
(10) Vanwijk, J.; Huckriede, B. D.; Ippel, J. H.; Altona, C. *Method Enzymol.* **1992**, *211*, 286−306.
(11) Zimmer, D. P.; Marino, J. P.; Griesinger, C. *Magn. Reson. Chem.* **1996**, *34*, S177−S186.
(12) Dickerson, R. E.; Goodsell, D. S.; Kopka, M. L.; Pjura, P. E. *J. Biomol. Struct. Dyn.* **1987**, *5*, 557−579.

10.1021/ja000324n CCC: $19.00
Published on Web 06/13/2000

The DNA dodecamer, d(CGCGAATTCGCG)$_2$, which has been crystallized in the B-form, and which has been studied extensively by X-ray diffraction, first by Dickerson et al.,[14,15] and more recently by Williams and co-workers[16,17] and Tereshko et al.,[18,19] shows localized unusual features in its backbone angles and sugar puckers. Remarkably, these studies report asymmetric structures for this palindromic sequence, and some unusual backbone torsion angles and sugar puckers occur in the same regions in these separate X-ray studies. Recently, the role of counterions and hydration waters on the structure of this dodecamer has been the subject of much debate.[16,20,21]

Some of the early NMR studies also have focused on this so-called Dickerson dodecamer.[22−27] Although no evidence for the unusual features in the X-ray structure was observed in solution, the NMR structures were of insufficient resolution to extract detailed information on local geometry and overall shape of the molecule. The detailed structure of small oligomers remains a topic of high interest because it is critical for understanding the basis of DNA bending, which constitutes an essential factor in chromatin and virus assembly.[28−30] Local DNA structure also plays a role in modulating affinity for proteins and other molecules. Next to crystallography and NMR, DNA bending has been studied extensively by gel shift assays, cyclization kinetics, and computational approaches, including molecular dynamics.[31−33]

Recently, a novel approach has been introduced for obtaining structural information by NMR, which relies on obtaining a very small degree of molecular alignment with the magnetic field.[34−38] Provided molecular alignment is sufficiently weak, the NMR spectrum retains the simplicity and resolution of the regular isotropic solution spectrum, but also makes it possible to measure residual dipolar couplings that report on the average orientation of their internuclear vector relative to the magnetic field. As dipolar couplings report on the orientation of internuclear bond vectors relative to a single molecular axis system, that of its alignment tensor, they intrinsically are global restraints. In contrast to NOEs or *J* couplings, these dipolar couplings therefore can define the orientation of one end of the molecule relative to the opposite end, for example. Molecular alignment can be obtained from the molecule's intrinsic anisotropy of its magnetic susceptibility,[34] although the alignment effect tends to be relatively small for routine application to macromolecules.[35,36,39] Larger alignments can be obtained by increasing this anisotropy through chelation of paramagnetic lanthanide ions[40,41] or by using a very dilute liquid crystalline phase consisting either of nearly neutral disk-shaped phospholipid particles, known as bicelles,[42,43] or rod-shaped viruses.[44−46] The liquid crystal alignment method has rapidly made its entry in the protein structure determination arena and has already been demonstrated to considerably improve the accuracy of such structures.[39,47,48]

Measurement of nucleic acid dipolar couplings has been reported both in bicelle-[43] and phage-based liquid crystals,[46] but so far only a single preliminary report has appeared where these dipolar couplings were actually used to improve definition of the nucleic acid's structure.[49] The present study discusses the use of dipolar couplings measured in the Dickerson dodecamer for defining its solution structure. A nearly complete set of one-bond ^{1}H−^{13}C dipolar couplings, supplemented by ^{1}H−^{15}N and ^{1}H−^{1}H dipolar couplings, was obtained in the bicelle-based liquid crystal medium. Several samples, in which complementary halves of the oligomer were labeled, were used in order to minimize spectral overlap.

The structure determined by use of these dipolar couplings, together with torsion angles derived from *J* couplings and interproton distances obtained from NOEs, results in a well-defined B-form helix. Overall agreement with the corresponding X-ray structures is reasonable, although several of the irregular features seen in the crystalline state are less pronounced or absent in solution. Most parameters, such as sugar pucker, propeller twist, roll, and the near absence of helical bending, are well determined by the NMR data. In contrast, other parameters such as the base-pair opening and the ^{31}P−^{31}P distance across the minor groove, are less accurately defined

(13) Digabriele, A. D.; Sanderson, M. R.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 1816−1820.
(14) Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature* **1980**, *287*, 755−758.
(15) Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, *149*, 761−786.
(16) Shui, X. Q.; Sines, C. C.; McFail-Isom, L.; VanDerveer, D.; Williams, L. D. *Biochemistry* **1998**, *37*, 16877−16887.
(17) Shui, X. Q.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341−8355.
(18) Tereshko, V.; Minasov, G.; Egli, M. *J. Am. Chem. Soc.* **1999**, *121*, 470−471.
(19) Minasov, G.; Tereshko, V.; Egli, M. *J. Mol. Biol.* **1999**, *291*, 83−99.
(20) Hud, N. V.; Sklenar, V.; Feigon, J. *J. Mol. Biol.* **1999**, *286*, 651−660.
(21) Chiu, T. K.; Kaczor-Grzeskowiak, M.; Dickerson, R. E. *J. Mol. Biol.* **1999**, *292*, 589−608.
(22) Hare, D. R.; Wemmer, D. E.; Chou, S. H.; Drobny, G.; Reid, B. R. *J. Mol. Biol.* **1983**, *171*, 319−336.
(23) Broido, M. S.; James, T. L.; Zon, G.; Keepers, J. W. *Eur. J. Biochem.* **1985**, *150*, 117−128.
(24) Patel, D. J.; Shapiro, L.; Hare, D. *J. Biol. Chem.* **1986**, *261*, 1223−1229.
(25) Nerdal, W.; Hare, D. R.; Reid, B. R. *Biochemistry* **1989**, *28*, 10008−10021.
(26) Denisov, A. Y.; Zamaratski, E. V.; Maltseva, T. V.; Sandstrom, A.; Bekiroglu, S.; Altmann, K. H.; Egli, M.; Chattopadhyaya, J. *J. Biomol. Struct. Dyn.* **1998**, *16*, 547−568.
(27) Lane, A. N.; Jenkins, T. C.; Brown, T.; Neidle, S. *Biochemistry* **1991**, *30*, 1372−1385.
(28) Olson, W. K.; Zhurkin, V. B. In *Biological Structure and Dynamics*; Sarma, R. H., Sarma, M. H., Eds.; Adenine Press: Albany, NY, 1995; pp 341−370.
(29) Ulyanov, N. B.; James, T. L. *Method Enzymol.* **1995**, *261*, 90−120.
(30) Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 15163−15165.
(31) Haran, T. E.; Kahn, J. D.; Crothers, D. M. *J. Mol. Biol.* **1994**, *244*, 135−143.
(32) Gorin, A. A.; Zhurkin, V. B.; Olson, W. K. *J. Mol. Biol.* **1995**, *247*, 34−48.
(33) Young, M. A.; Ravishanker, G.; Beveridge, D. L. *Biophys. J.* **1997**, *73*, 2313−2336.
(34) Gayathra, C.; Bothner-By, A. A.; van Zijl, P. C. M.; MacLean, C. *Chem. Phys. Lett.* **1982**, *87*, 192−196.
(35) Kung, H. C.; Wang, K. Y.; Goljer, I.; Bolton, P. H. *J. Magn. Reson. Ser. B* **1995**, *109*, 323−325.
(36) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9279−9283.
(37) Tjandra, N.; Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6264−6272.
(38) Bax, A.; Tjandra, N. *J. Biomol. NMR* **1997**, *10*, 289−292.
(39) Tjandra, N.; Omichinski, J. G.; Gronenborn, A. M.; Clore, G. M.; Bax, A. *Nat. Struct. Biol.* **1997**, *4*, 732−738.
(40) Beger, R. D.; Marathias, V. M.; Volkman, B. F.; Bolton, P. H. *J. Magn. Reson.* **1998**, *135*, 256−259.
(41) Gochin, M. *J. Am. Chem. Soc.* **1997**, *119*, 3377−3378.
(42) Sanders, C. R.; Schwonek, J. P. *Biochemistry* **1992**, *31*, 8898−8905.
(43) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.
(44) Clore, G. M.; Starich, M. R.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 10571−10572.
(45) Hansen, M. R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065−1074.
(46) Hansen, M. R.; Rance, M.; Pardi, A. *J. Am. Chem. Soc.* **1998**, *120*, 11210−11211.
(47) Ottiger, M.; Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 9825−9830.
(48) Clore, G. M.; Starich, M. R.; Bewley, C. A.; Cai, M. L.; Kuszewski, J. *J. Am. Chem. Soc.* **1999**, *121*, 6513−6514.
(49) Bayer, P.; Varani, L.; Varani, G. *J. Biomol. NMR* **1999**, *14*, 149−155.

by the NMR data because they are quite sensitive to the assumed magnitude and rhombicity of the alignment tensor.

Materials and Methods

NMR Measurements. Six different NMR samples were used for collecting the dipolar coupling and NOE data. All samples contained 40 mM sodium phosphate buffer, pH 7.0. The samples used for NOE data collection were dissolved in 99.9% D_2O, or in 93% H_2O, 7% D_2O, at concentrations of ~0.5 mM duplex. Most measurements were carried out at 35 °C, above the temperature threshold where the bicelle solution changes from an isotropic to a liquid crystalline phase. Bicelle media were prepared as described previously[50] and consisted of 50 mg/mL phospholipids in a 3.15:1 molar ratio of dimyristoyl phosphatidylcholine (DMPC) and dihexanoyl phosphatidylcholine (DHPC). All measurements were carried out in 260-μL Shigemi microcells. Experiments were carried out on Bruker DMX-600 and DMX-750 spectrometers, operating at 600 and 750 MHz ^{1}H frequency, and equipped with triple-resonance, three-axis pulsed field gradient probeheads.

A 2D NOESY spectrum was collected at 750-MHz ^{1}H frequency, in D_2O solution, using unlabeled DNA and a mixing time of 100 ms. Semiquantitative distance restraints were derived from the integrated relative cross-peak intensities, using the intraresidue H1′−H2″distance as a 2.3 Å internal reference, and using an empirical r^{-4} distance dependence which qualitatively accounts for the effect of spin diffusion. For NOEs involving geminal protons, only the most intense of the two NOEs was used, unless the intensity difference for the two cross-peaks was less than 35%. A ±15% tolerance was used on all NOE-derived distances. Distance restraints determined in this manner and used in the calculation have been deposited together with the derived structures in the PDB database (accession codes RCSB010377 and 1DUF).

One-bond $^1J_{CH}$ couplings were measured in isotropic D_2O solution at 600 and 750 MHz, using $^1J_{CH}$ modulated 2D HSQC spectra.[51] These initial experiments were recorded at very high precision as they were aimed at measuring the very small magnetic field dependence of the dipolar contribution to the $^1J_{CH}$ splitting. After subsequent development of liquid crystalline media for inducing molecular alignment, several of these splittings were remeasured from the splitting observed in a (less precise) constant-time ^{1}H−^{13}C HSQC correlation, recorded without ^{1}H decoupling in the F_1 dimension, and found to be in good agreement with a root-mean-square difference (rmsd) of ~1 Hz. All subsequent measurements in the aligned phase were made from F_1-coupled (or F_2-coupled, in case of overlap) HSQC spectra, recorded in H_2O. Isotropic values for the imino $^1J_{NH}$ couplings were measured at lower temperature (23 °C), below the temperature threshold where the bicelle solution switches to a nematic liquid crystalline phase. Dipolar contributions derived from the magnetic field dependence correlate reasonably well with those observed in the liquid crystalline medium ($R = 0.7$) but have the opposite sign. Because of the much larger relative error in the very small dipolar contributions induced by the DNA magnetic susceptibility anisotropy compared to that in the liquid crystal-derived dipolar couplings, only the latter were used in structure calculations.

To minimize resonance overlap in the 2D ^{1}H−^{13}C HSQC spectra, two different sets of samples were used for these heteronuclear J splitting measurements: First, d(**CG**CG**A**A**TT**C**G**CG)$_2$ and d(CG**C**-G**A**A**TT**CG**CG**)$_2$ samples were used, where bold-faced nucleotides are uniformly labeled with ^{13}C and ^{15}N; second, two similarly ^{13}C labeled samples were used in which these bold-faced nucleotides are also deuterated in the H2″ and H5′ positions, using a procedure described previously.[52,53] Measurement of the sum of $^1J_{CH'}$ and $^1J_{CH''}$ splittings[54] was carried out in a sample without this selective deuteration, allowing values for $^1J_{C2'H''}$ and $^1J_{C5'H5'}$ to be derived from the difference between this sum and the separately measured $^1J_{C2'H2'}$ and $^1J_{C5'H5''}$ values.

Values measured for isotropic ^{1}H−^{1}H and ^{1}H−^{31}P J couplings have been reported previously[55,56] and were converted into angular restraints for the backbone torsion angles γ, δ, and ϵ, with relatively loose margins of ±20°. Approximate values for ^{1}H−^{1}H dipolar couplings were derived from the phase-sensitive COSY spectra measured for a sample of d(CGCGAATT**C**GCG)$_2$, where C9 was uniformly labeled with ^{13}C. Splittings measured for C9 were compared with those in d(**CG**C-G**A**A**TT**C**G**CG)$_2$ in order to ensure that nucleic acid alignment was the same as in the other measurements. No correction factor was found to be necessary. COSY cross-peak intensities in the aligned state were converted into approximate values for $|J_{HH} + D_{HH}|$, with the absence of a cross-peak indicating $|J_{HH} + D_{HH}| \leq 6$ Hz; a weak cross-peak, $4 \leq |^1J_{HH} + D_{HH}| \leq 8$ Hz; medium cross-peak, $6 \leq |J_{HH} + D_{HH}| \leq 12$ Hz, and strong cross-peak, $|J_{HH} + D_{HH}| \geq 8$ Hz. In a number of cases where the sign of D_{HH} could not be established a priori, signed values could usually be determined on the basis of preliminary structure calculations which did not include these couplings. For example, all $D_{H1'H2''}$ couplings with a large absolute value were found to be clearly negative. For dipolar couplings whose signs were not determined, the structure calculation only used the magnitude of these couplings.[57] The full set of dipolar couplings, NOE distances, and J coupling-derived torsional restraints is available from the PDB (accession number RCSB010377).

Structure Calculations. A total of 198 D_{CH} values, 10 D_{NH} values, 200 approximate D_{HH} values, 162 NOEs, and 44 J-derived torsional angles were used as input restraints to restrained molecular dynamics calculations, using XPLOR version 3.84,[58] supplemented with home-written routines for inclusion of dipolar coupling restraints.[39] In addition, two (for A−T) or three (for G−C) hydrogen bond distance restraints were used for Watson−Crick base pairing. Although no attempt was made to confirm the actual presence of these hydrogen bonds through J coupling measurement,[59−61] the decreased exchange rates for the hydrogens involved in these bonds[62] and their downfield shift, together with experimental confirmation of their presence in very similar fragments[60,61] indicates that such H-bonds indeed are present in solution. Considering that the ^{31}P chemical shifts for this dodecamer span only the very narrow region that is typically found for regular, undistorted B-DNA, artificial, loose restraints were imposed on the phosphodiester backbone torsion angles α, β, and ζ with values of $-60 \pm 30°$, $180 \pm 30°$, and $-90 \pm 30°$.[63] These artificial restraints do not exert any force on the final structures, but are needed to efficiently converge to a minimum where all other restraints can be satisfied. In the absence of these artificial restraints, low energies are obtained for only a very small fraction of the structures, i.e., for only very few structures can the experimental restraints be approximately satisfied with the protocol used, and even the lowest energy structure is of higher energy than obtained in the presence of these restraints. Clearly, as none of the artificial restraints are consistently violated by more than 0.1°, they only serve to drive the structure to a region of conformational space where the experimental restraints can be satisfied. Similar but less severe convergence problems have been noted when including dipolar restraints in the calculation of protein structures. Presumably, these are related to the fact that a given dipolar coupling corresponds to two allowed cones of bond vector orientations, pointing up and down with respect to the z axis of the alignment tensor, thereby creating an energy landscape with a very large number of local minima.

(50) Ottiger, M.; Bax, A. *J. Biomol. NMR* **1998**, *12*, 361−372.
(51) Tjandra, N.; Bax, A. *J. Magn. Reson.* **1997**, *124*, 512−515.
(52) Ono, A. M.; Oogo, Y.; Tate, S.; Ono, A.; Kainosho, M. *Nucleic Acids Symp. Ser.* **1997**, *37*, 73−74.
(53) Ono, A. M.; Shiina, T.; Ono, A.; Kainosho, M. *Tetrahedron Lett.* **1998**, *39*, 2793−2796.
(54) Ottiger, M.; Delaglio, F.; Marquardt, J. L.; Tjandra, N.; Bax, A. *J. Magn. Reson.* **1998**, *134*, 365−369.
(55) Bax, A.; Lerner, L. *J. Magn. Reson.* **1988**, *79*, 429−438.
(56) Sklenar, V.; Bax, A. *J. Am. Chem. Soc.* **1987**, *109*, 7525−7526.
(57) Tjandra, N.; Marquardt, J.; Clore, G. M. *J. Magn. Reson.*, **2000**, *142*, 393−396.
(58) Brunger, A. T. *XPLOR Manual Version 3.1*; Yale University: New Haven, CT, 1993.
(59) Dingley, A. J.; Grzesiek, S. *J. Am. Chem. Soc.* **1998**, *120*, 8293−8297.
(60) Pervushin, K.; Ono, A.; Fernandez, C.; Szyperski, T.; Kainosho, M.; Wuthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14147−14151.
(61) Dingley, A. J.; Masse, J. E.; Peterson, R. D.; Barfield, M.; Feigon, J.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 6019−6027.
(62) Rajagopal, P.; Gilbert, D. E.; Vandermarel, G. A.; Vanboom, J. H.; Feigon, J. *J. Magn. Reson.* **1988**, *78*, 526−537.
(63) Roongta, V. A.; Jones, C. R.; Gorenstein, D. G. *Biochemistry* **1990**, *29*, 5245−5258.

Table 1. Structural Statistics[a,b]

	NMR	NMR-nodipo	1BNA	355D	Arnott
Rmsd from Distance Restraints (Å)					
NOE (162/142)[c]	0.025	0.022	0.19	0.20	0.33
H-bond (32/26)	0.038	0.037	0.029	0.008	0.0
Rmsd from Dipolar Restraints (Hz)					
D_{CH}^{ribose} (94/74)[d]	1.9	6.1	4.4	5.5	6.7
D_{CH}^{ribose} (64/56)[e]	4.0	11.8	7.1	8.8	12.0
D_{CH}^{base} (24/20)[d]	2.8	4.9	4.0	3.8	4.0
D_{CH}^{base} (12/10)[e]	1.9	3.9	2.4	2.6	3.3
D_{CH}^{methyl} (4/4)	1.0	1.6	1.0	1.2	1.9
D_{NH}^{imino} (10/10)	1.4	2.2	1.7	2.0	2.0
D_{HH} (192/160)[f]	0.9	1.6	1.4	1.7	2.7
Rmsd from Dihedral Restraints (deg)					
all (156/134)	0.2	0.04	13.3	13.3	9.5
Deviations from Idealized Covalent Geometry					
bonds (Å) (819/685)	0.005	0.004	0.02	0.01	0.01
angles (deg) (1479/1237)	1.2	1.1	3.0	2.1	2.1
ampropers (deg) (412/342)	0.26	0.21	1.1	0.8	0.4
Nonbonded Energies					
Lennard-Jones (kcal/mol)	−291	−304	−280	−315	−130
electrostatic (kcal/mol)	−368	−376	−308	−301	−320

[a] Excluding C1, G12, C13, and G24. [b] The final values of the force constants for the various energy terms employed are as follows: 50 kcal mol^{-1} Å^{-2} for interproton distance restraints; 200 kcal mol^{-1} rad^{-2} for dihedral angle restraints; 1000 kcal mol^{-1} Å^{-2} for bonds; 500 kcal mol^{-1} rad^{-2} for bond angles and improper torsion angles; 0.2 kcal mol^{-1} Hz^{-2} for the class 1 and proton−proton dipolar couplings; 0.04 kcal mol^{-1} Hz^{-2} for the class 2 dipolar couplings. [c] The first number in parentheses corresponds to the total number of restraints; the second number is the number of restraints for the center 10 base pairs, for which the reported statistics apply. [d] Class 1 dipolar coupling with estimated errors less than 2 Hz. [e] Class 2 couplings with estimated errors less than 4 Hz. [f] These interproton dipolar couplings are all of known sign. All dipolar couplings for which the sign could not be determined are in the terminal base pairs, and therefore excluded from the statistics.[a]

Families of structures that satisfy all of the measured constraints to within experimental error were calculated by restrained molecular dynamics. Lennard-Jones van der Waals as well as electrostatic energy terms defined by the CHARMM PARNAH1ER1 DNA parameters[64−66] were included in the structure calculation. Switched van der Waals and electrostatic functions, with switching distances of 9.5 and 10.5 Å, were used to truncate the number of possible interactions. A 1/*r*-dependent dielectric constant was used throughout the calculation. The target function that is minimized consists of a quadratic harmonic potential for covalent geometry (bonds, angles, planes, and chirality) and a square well quadratic potential for experimental distance and torsion angle restraints. The empirical potentials for dipolar couplings comprise quadratic harmonic and square well quadratic terms for the one-bond and proton−proton dipolar couplings, respectively. The one-bond dipolar couplings are categorized in two different classes. Class 1 consists of dipolar couplings with estimated errors of less than ±2 Hz, while class 2 is composed of those with estimated errors between ±2 and ±4 Hz, consisting primarily of sites with partial overlap in the HSQC spectrum. The final values of the force constants for the various energy terms are listed in the footnote to Table 1.

Structures were calculated in a two-step procedure. In the first step, approximate starting structures were calculated from completely random initial structures. These starting structures were created by first heating the structure to 1000 K, with only energy terms for bonds, angles, and improper torsions turned on. Next, globally folded structures were calculated using the protocol described by Varani et al. In this stage, only NOE and H-bond restraints were used. Subsequently, NOE, H-bond and dihedral restraints were slowly ramped up during a short refinement stage.[67] A total of 40 independent structures were created in this manner, which were subsequently refined in the manner described below.

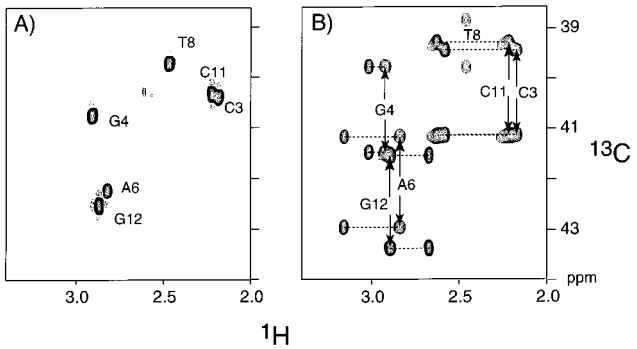

Figure 1. H2′/H2″−C2′ regions of the HSQC spectra recorded for the d(CG**CG**AA**TT**CG**CG**)$_2$ in the aligned state, where boldfaced residues are enriched in ^{13}C. (A) Regular CT-HSQC spectrum for the sample in which the H″ position was deuterated. (B) The same spectral region for the sample without stereospecific deuteration of the methylene sites (except for nucleotide T8) and recorded in the absence of ^{1}H decoupling in the F_1 dimension. The center component of the ^{13}C−{^{1}H$_2$} triplets have near zero intensity, so the F_1 splitting corresponds to the sum of the $^1J_{C2'H'}$ and $^1J_{C2'H''}$ splittings.

Final sets of structures were calculated by starting from the above generated structures and also by starting from idealized A-DNA and B-DNA,[68] the original Dickerson X-ray structure[15] (1BNA), and the more recent X-ray structure by Shui et al.[17] Although starting from these relatively regular helical structures is less general than starting from the fully randomized initial structures, it allows us to evaluate whether in the complete absence of NOE information (required for going from randomized to roughly folded starting structures) dipolar couplings are sufficient to define the proper conformation of these helical DNA molecules. The second step of the protocol used in calculating the final structures consists of an initial equilibration stage where the dipolar force constants were slowly increased from zero to their final values over 50 cycles of molecular dynamic runs, corresponding to a total of 15 ps. This is followed by a 25-ps restrained molecular dynamics calculation at 300 K. The representative final structure was calculated from the minimized average of the last 10-ps trajectory. A total of 40 structures were calculated starting from the 40 initial folds, and 20 structures for each of the four A- and B-DNA starting structures, using randomized initial velocities for each of these. When, for each group of different starting structures, restricting the evaluation to the 25% of structures with lowest total energy, these NMR structures are essentially indistinguishable from one another, with rmsd of less than 0.1 Å for the center 10 base pairs.

Results and Discussion

All dipolar couplings were measured from the difference in splitting measured for DNA samples dissolved in water and in the bicelle-containing liquid crystalline phase. Liquid crystalline samples were stable for at least several months before hydrolysis of the phospholipid resulted in phase separation of the solution.[50] Although the heteronuclear dipolar couplings were collected on several different samples, the liquid crystal concentration and composition were carefully prepared to be as similar as possible. Indeed, the observed solvent ^{2}H quadrupole splitting was very similar (8.3 ± 0.4 Hz) for all samples, suggesting that errors introduced by small variations in the alignment parameters, and thereby in the dipolar couplings, are less than ±5%.

Measurement of Dipolar Couplings. Figure 1 shows the H2′−C2′ regions of the HSQC spectra recorded for the

(64) Brooks, B. R.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(65) Reiher, W. E. Ph.D. Thesis; Harvard University: Cambridge, MA, 1985.
(66) Nilsson, L.; Karplus, M. *J. Comput. Chem.* **1986**, *7*, 591−616.
(67) Varani, G.; Aboulela, F.; Allain, F. H. T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1996**, *29*, 51−127.
(68) Arnott, S.; Hukins, D. W. L. *Biochem. Biophys. Res. Comm.* **1972**, *47*, 1504−1510.

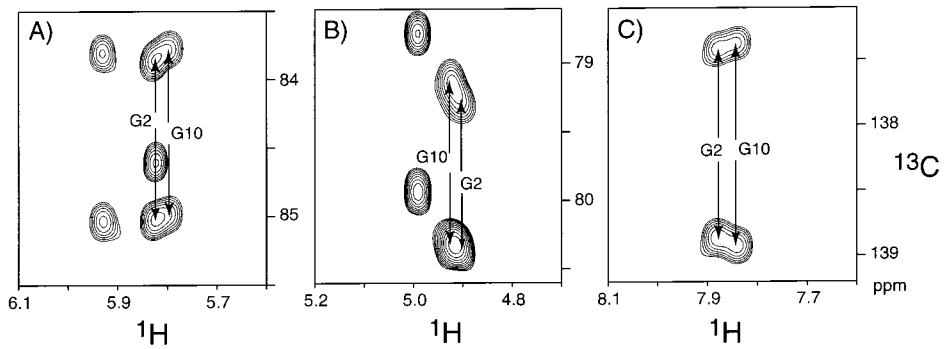

Figure 2. Small regions displaying the (A) H1′/C1′, (B) H3′/C3′, and (C) H8/C8 correlations for G2 and G10 in the aligned phase, recorded at 600 MHz. Although isotropic shifts and *J* couplings are virtually the same for G2 and G10 (not shown), in the aligned phase, the splittings differ substantially.

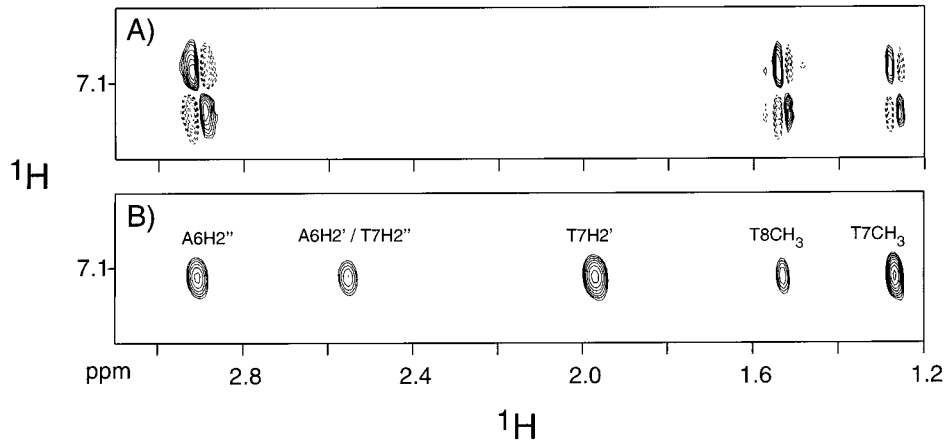

Figure 3. Comparison of a small region of the 750-MHz phase-sensitive ^{1}H−^{1}H P-COSY spectrum[90] (A) recorded in the liquid crystalline phase, with the corresponding region in the NOESY spectrum (B). The region shown includes the cross-peaks from T7−H6 to the H2′/H2" and thymidine methyl protons. Dashed contours correspond to negative intensity.

d(CG**CG**A**A**T**T**CG**CG**)$_2$ in the aligned state, where boldfaced residues are enriched in ^{13}C. Figure 1A was recorded for the sample in which the H2″ position was deuterated, and where the $^1J_{CH}$ splitting is measured from a modulation experiment.[54] The first spectrum recorded in this series of *J*-modulated HSQC spectra corresponds to a regular, decoupled ^{1}H−^{13}C correlation spectrum (Figure 1A). Figure 1B shows the same spectral region, recorded for the sample without stereospecific deuteration of the methylene sites (except for nucleotide T8), and recorded in the absence of ^{1}H decoupling in the F_1 dimension. So, the F_1 splitting corresponds to the sum of the $^1J_{C2'H'}$ and $^1J_{C2'H''}$ splittings. These experiments were also conducted in the isotropic phase, and $D_{C2'H''}$ was extracted from the difference. The same set of spectra yielded $D_{C5'H'}$ and $D_{C5'H''}$ and all other D_{CH} couplings for sugar residues and thymidine methyl groups.

The first four and last four nucleotides of the Dickerson dodecamer have the same sequence, CGCG. As expected, therefore, resonances of G2 and G10 have very similar chemical shifts and so do those of C3 and C11. Similarly, the isotropic $^1J_{CH}$ and $^3J_{HH}$ couplings measured for these nucleotides are nearly indistinguishable. However, in the aligned state, substantial differences in the splittings are observed. For example, Figure 2 shows three small regions displaying the H1′/C1′, H3′/C3′, and H8/C8 correlations for G2 and G10 in the aligned state. Although there is partial overlap between the doublets, in all three cases it is clear that the splittings are substantially different, i.e., that the dipolar contributions to these splittings are different for G2 and G10. Similarly large differences are observed between corresponding dipolar couplings measured for C3 and C11 (spectra not shown). Although in Figure 2 all of the G2 splittings are smaller than for G10, and similarly most of the corresponding C3 splittings are smaller than those for C11, this cannot be attributed to a dynamic effect. First, the line widths and relaxation properties for these two pairs of nucleotides are indistinguishable from one another or from the remainder of the oligomer, except for the first and last base pairs. These terminal C1−G24 and G12−C13 base pairs exhibit increased mobility as judged by longer ^{13}C T_2 values, absence of slowly exchanging imino resonances and sugar ^{1}H−^{1}H *J* couplings, which are indicative of significant averaging of the sugar pucker.[10,55] Even for those terminal base pairs, the degree of motional averaging of the dipolar couplings is not extreme, however. For example D_{C5H5} is 19 Hz for nucleotide C1, which is less than 30% smaller than the largest D_{CH} coupling observed for any of the bases, suggesting that the order parameter ***S*** for this base is at most 30% smaller than for the remainder of the oligonucleotide.

The difference in dipolar couplings observed for G2 and C3 relative to G10 and C11 points to a different orientation of these dinucleotides relative to the molecular alignment tensor. As will be discussed later, this is confirmed by the calculated structures.

Figure 3 compares a small region of the ^{1}H−^{1}H COSY spectrum, recorded in the liquid crystalline phase, with the corresponding region in the NOESY spectrum. The region shown includes the cross-peaks from T7−H6 to the H2′/H2″ and thymidine methyl protons. Although the overlapping cross-peaks between T7−H6 and T7−H2″/A6−H2′ are quite intense in the NOESY spectrum, the corresponding cross-peak in the COSY spectrum falls below the detection threshold, indicating both couplings are relatively small. The same applies for the

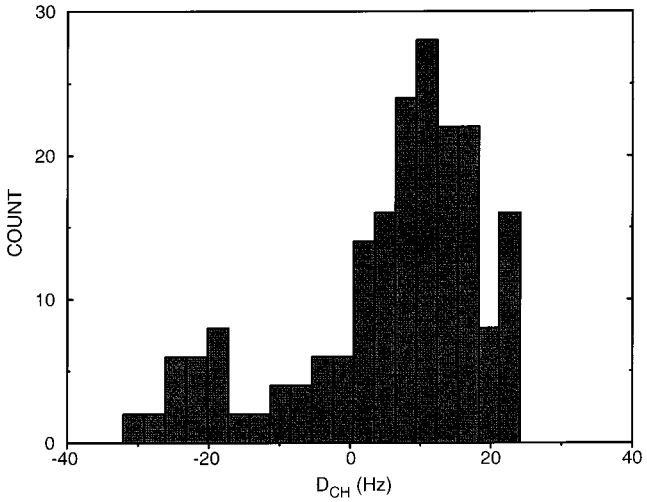

Figure 4. Histogram of the measured one-bond D_{CH} couplings. Extreme dipolar couplings correspond to G4 C8−H8 (+24.2 Hz) and C9 C5′−H5″ (−32.1 Hz).

T7−H6 to T7−H2′ cross-peak. In contrast, an intense COSY cross-peak is observed between T7−H6 and A6−H2″ and also to both the T7 and T8 methyl groups. Preliminary structures, calculated without these $^1H-^1H$ dipolar couplings clearly indicated the signs of these dipolar couplings to be negative for the T7−H6 to A6−H2″ and T7−H6 to T8−CH_3 groups and positive for the T7−H6 to T7−CH_3 interaction.

Determination of Alignment Tensor. In proteins, good estimates of the magnitude and rhombicity (R) of the alignment tensor usually can be obtained from the distribution of observed dipolar couplings, assuming that these are more or less randomly distributed in space.[69] For nucleic acids, such an assumption is clearly not valid, as the B-DNA-type structure is expected to be quite regular, with very few bond vectors pointing parallel to its helix axis. Nevertheless, the distribution of observed dipolar couplings can be used to put stringent lower limits on the possible magnitude of the alignment tensor.

In the principal axis frame of the alignment tensor, **A**, the dipolar coupling between two nuclei, P and Q, as a function of polar coordinates θ and ϕ is given by[69]

$$D_{PQ}(\theta,\phi) = D_a^{PQ}\{(3\cos^2\theta - 1) + {}^3/_2 R \sin^2\theta \cos 2\phi\} \quad (1a)$$

where

$$D_a^{PQ} = -(\mu_o h/16\pi^3)\boldsymbol{S}\gamma_P\gamma_Q\langle r_{pq}^{-3}\rangle A_a \quad (1b)$$

Here, A_a is the dimensionless axial component of the alignment tensor, i.e., $A_{zz} - (A_{xx} + A_{yy})/2$, and $R = (A_{xx} - A_{yy})/A_{zz}$. $\boldsymbol{S}$ is the generalized order parameter,[70] which accounts for the effect of rapid angular vibrations, and the $\langle\ \rangle$ brackets indicate time averaging of the inverse cube of the P−Q distance. To a first approximation, $\boldsymbol{S}$ is assumed to be uniform and its value is subsumed in the magnitude of **A**.

A histogram of the measured one-bond D_{CH} couplings is shown in Figure 4. As expected, there is considerable clustering of the observed couplings, with most values being positive. The extreme values correspond to G4 C8−H8 (+24.2 Hz) and C9 C5′-H5″ (−32.1 Hz). The optimal value of D_a^{CH} and R were derived from a grid search where structures are calculated for each grid point, and the residual dipolar energy term in the structure calculation is evaluated as a function of D_a^{CH} and R.[71] Figure 5 shows a plot of this residual energy. This figure clearly indicates that no low-energy structures can be obtained in the absence of rhombicity in the alignment tensor. Unless stated otherwise, all structures discussed below were calculated using the values $D_a^{CH} = -16$ Hz, $D_a^{NH} = -7.7$ Hz, and $R = 0.26$, which resulted in the lowest total energy and simultaneously the lowest dipolar energy term. The effect of small changes in D_a and R on the structure will be discussed later.

Description of the NMR Structure. Figure 6 shows two views of the lowest energy NMR structure (NMR-dipo), calculated with inclusion of all restraints listed in Table 1. Inspection of Figure 6b immediately reveals the reason for the rhombicity of the alignment tensor: As the length of the dodecamer is only slightly longer than one helical turn, a side view of the molecule reveals a pronounced V-shape, which is caused by the difference in depth of the minor and major grooves and not by bending of the helical axis. In fact, the overall helix axis curvature, as defined by the program Curves,[72] is only 7° (Supporting Information). Clearly, at least several turns of helix are required for a straight B-DNA fragment to approximate cylindrical symmetry, and the rhombicity of this short fragment is therefore not surprising. The R value of 0.26 is in close agreement with the value predicted on the basis of the dodecamer's shape, using a simple steric obstruction model (*M. Zweckstetter, unpublished*).[91]

As can be seen in Table 1, all experimental restraints for the center 10 base pairs are well satisfied. Excluding the first and last base pair, the root-mean-square (rms) displacement of all atoms relative to the average structure of the entire ensemble of calculated structures (<0.1 Å for all atoms) is very low and does not represent the accuracy of the structure (see below). Residual restraint violations are largest in the two terminal base pairs and account for a total penalty that is 38% of the total energy function. It is known that these terminal base pairs are subject to considerable mobility and it is therefore reassuring to see that no single conformation of these nucleotides can satisfy the experimental restraints. However, the fact that the 10 remaining base pairs satisfy experimental restraints within experimental error should not be interpreted as an indication for the absence of motion. It merely means that motions in these base pairs are either of considerably smaller amplitudes than for the terminal base pairs or that alternate conformations are occupied for only a relatively small fraction of time.

Thus, the calculated structures provide a reasonable model for the time-averaged conformation but do not exclude the presence of a C2′-endo/C3′-endo equilibrium, which is strongly shifted to the C2′-endo conformation. Indeed, as can be seen in Table 2, all of the pyrimidine sugar puckers observed fall in the C1′-exo region, which has a sugar pucker pseudorotation angle, P, of 126 ± 18°, intermediate between the C2′-endo (P = 162 ± 18°) and C3′-endo (P = 18 ± 18°). Interestingly, all but one of the purines in the center 10 base pairs have sugar puckers in the C2′-endo region. G10 is the only exception (P = 143°) and falls right on the edge between C2′-endo and C1′-exo. Previous analyses of homo- and heteronuclear J couplings have provided strong indications that such a C2′-endo/C3′-endo equilibrium is present[73,74] with an increase in the population of

(69) Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Magn. Reson.* **1998**, *133*, 216−221.
(70) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4558.
(71) Clore, G. M.; Gronenborn, A. M.; Tjandra, N. *J. Magn. Reson.* **1998**, *131*, 159−162.
(72) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63−91.
(73) Sanderson, M. R.; Mellema, J. R.; van der Marel, G. A.; Wille, G.; van Boom, J. H.; Altona, C. *Nucleic Acids Res.* **1983**, *11*, 3333−3346.
(74) Rinkel, L. J.; Altona, C. *J. Biomol. Struct. Dyn.* **1987**, *4*, 621−649.

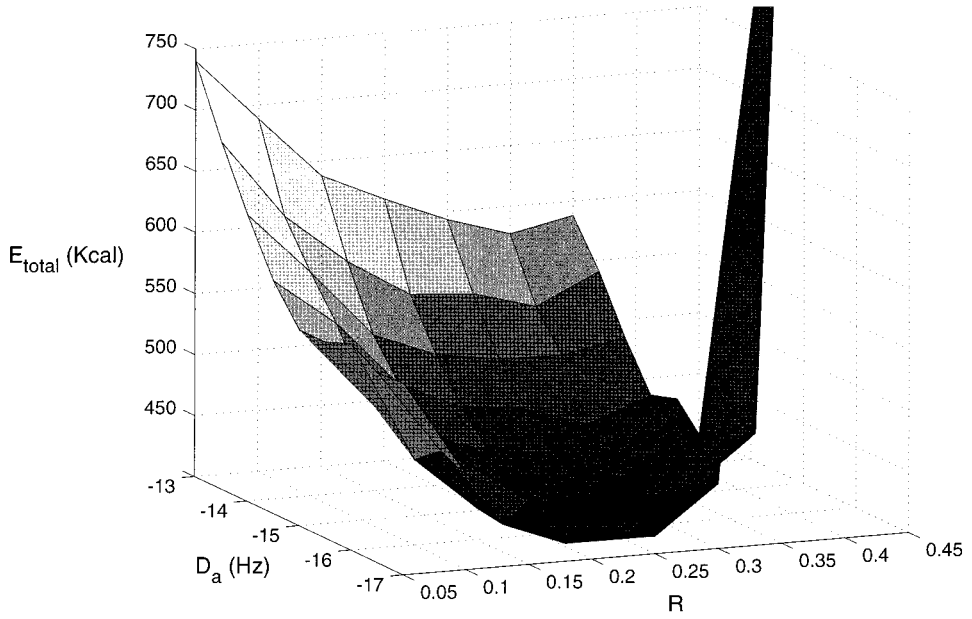

Figure 5. Plot of the total energy for the lowest energy structure of ensembles obtained as a function of D_a^{CH} and *R*. The plot is derived from a grid search where structures are calculated for each grid point.

Table 2. Pseudorotation Phase and Amplitude for the Deoxyribose Sugars in the Lowest Energy NMR Structure of $d(CGCGAATTCGCG)_2$

residue	phase P (deg)	amplitude (deg)	pucker
C(1)	***124***	*27*	C1′-exo
G(2)	148	*30*	C2′-endo
C(3)	***120***	*29*	C1′-exo
G(4)	155	*35*	C2′-endo
A(5)	*152*	*29*	C2′-endo
A(6)	***149***	*29*	C2′-endo
T(7)	**119**	*30*	C1′-exo
T(8)	**136**	*28*	C1′-exo
C(9)	***116***	*32*	C1′-exo
G(10)	143	***29***	C1′-exo
C(11)	***110***	*32*	C1′-exo
G(12)	***113***	*28*	C1′-exo

[a] Italicized or boldfaced entries differ by more than 20° from the corresponding pseudorotation angle, or more than 10° in pucker amplitude, in either of the two strands of the 1BNA (italics) or 355D (bold) X-ray structures.

the C3′-endo fraction in pyrimidines relative to purines[8,10,75] and our data are consistent with such a model. In this respect it is also interesting to note that the amplitude of the sugar puckers is artifactually small, presumably reflecting the average sugar conformation and not the average sugar pucker amplitude.

The relatively high energy computed for a deoxyribose in the C1′-exo form[76] also suggests that the conformations observed in our study reflect time-averaged structures. However, no quantitative interpretation of this dynamic equilibrium has been attempted at this stage because such an analysis also must take into account that the alignment tensor itself will be modulated by changes in the backbone conformation. Even for the analysis of dipolar couplings in small molecules, studied in the liquid crystalline phase, analysis of such motions gives rise to frequently insurmountable problems.[77]

(75) Rinkel, L. J.; Vandermarel, G. A.; Vanboom, J. H.; Altona, C. *Eur. J. Biochem.* **1987**, *166*, 87−101.

(76) Zhurkin, V. B.; Lysov, Y. P.; Florentiev, V. L.; Ivanov, V. I. *Nucleic Acids Res.* **1982**, *10*, 1811−1830.

(77) Emsley, J. W. In *Encyclopedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; Wiley: Chichester, U.K. 1996; Vol. 4, pp 2781−2787.

Table 3. Backbone Torsion Angles α−ζ, and Glycosyl Angles χ for the Lowest Energy NMR Structure[a]

	angles (deg)						
residue	α	*β*	*γ*	*δ*	*ϵ*	*ζ*	*χ*
C(1)			60	118	−170	−81	−126
G(2)	*−73*	***−162***	45	132	**172**	**−99**	−103
C(3)	**−64**	**178**	57	***117***	178	−85	−123
G(4)	−73	**−165**	*50*	138	***173***	***−104***	−102
A(5)	−57	***−174***	47	134	176	−95	−114
A(6)	−67	−172	55	130	177	−97	−105
T(7)	−55	173	53	116	177	**−82**	−121
T(8)	−58	−174	*41*	124	179	−90	−107
C(9)	−69	175	58	**113**	***176***	−83	−122
G(10)	−75	***−163***	48	*129*	***−177***	***−94***	**−105**
C(11)	*−56*	***172***	53	***112***	−174	**−87**	**−124**
G(12)	−77	−177	58	114			−113

[a] Italicized or bold faced entries differ by more than 20° from the corresponding angle in either of the two strands of 1BNA (italics) or 355D (bold) X-ray structures.

Table 3 lists the average backbone torsion angles α−ξ and the glycosyl angle χ, observed for the bundle of structures. Standard deviations from these average numbers are extremely small (≤1°) and do not reflect the accuracy at which they can be determined (see below). Structures were calculated without the pseudononcrystallographic symmetry energy term in the XPLOR protocol, and symmetry of the duplex therefore only results from the similarity in the restraints used for the two strands. As virtually none of the NOE or dihedral restraints exhibit any violation in the final structure, the small standard deviations in the bundle of obtained structures reflect primarily how tight these time-averaged angles are determined by the dipolar coupling data. Again, it is important to note that the small value of the standard deviation does not reflect the rms amplitude of the angular fluctuations.

Overall, the NMR dodecamer structure (NMR-dipo) is highly regular, with only modest variations in the backbone and glycosyl torsion angles and parameters such as propeller twist, and inter-base-pair rise and twist (Tables 2−4). A pronounced narrowing of the minor groove is observed near the center, although the exact width of this groove is relatively sensitive to the magnitude and rhombicity of the alignment tensor used

Table 4. Helical Parameters for the Lowest Energy NMR Structure of d(CGCGAATTCGCG)$_2$, as Derived from the Program Curves[a]

	propeller twist ω (deg)	rise D_z (Å)[b]	tilt τ (deg)[b]	roll ρ (deg)[b]	helical twist Ω (deg)[b]
C(1)−G(24)	−15 (−17, −23)	4.1 (3.6, 3.3)	−6 (−4, −3)	7 (4, 9)	32 (44, 35)
G(2)−C(23)	−10 (−13, −19)	3.2 (3.6, 3.7)	0 (2, 4)	−2 (−6, −13)	35 (36, 43)
C(3)−G(22)	−9 (−4, −9)	3.4 (3.1, 3.0)	0 (3, 1)	7 (13, 17)	32 (28, 26)
G(4)−C(21)	−13 (−17, −11)	3.4 (3.5, 3.4)	−2 (−4, −3)	4 (−1, −1)	36 (40, 37)
A(5)−T(20)	−16 (−27, −20)	3.2 (3.3, 3.3)	−1 (−1, −1)	−3 (0, −1)	37 (35, 37)
A(6)−T(19)	−24 (−27, −19)	3.4 (3.3, 3.3)	0 (3, 0)	−3 (−3, −4)	36 (34, 33)
T(7)−A(18)	−24 (−24, −20)	3.2 (3.3, 3.2)	1 (3, 1)	−3 (−1, 0)	37 (35, 34)
T(8)−A(17)	−16 (−28, −23)	3.4 (3.4, 3.6)	2 (2, 2)	4 (−2, −6)	36 (40, 41)
C(9)−G(16)	−13 (−25, −13)	3.4 (3.2, 3.0)	0 (−3, −3)	7 (9, 9)	32 (32, 29)
G(10)−C(15)	−10 (−9, −10)	3.2 (3.7, 3.4)	−1 (−6, −3)	−2 (−15, −14)	35 (39, 40)
C(11)−G(14)	−10 (−27, −23)	4.1 (3.1, 3.1)	6 (3, −1)	8 (−2, 10)	32 (34, 35)
G(12)−C(13)	−15 (−5, −7)				

[a] D_z, τ, ρ, and Ω are the local inter-base-pair parameters. For reference, values for 1BNA[15] and 355D[17] are included in parentheses. Nearly perfect symmetry of the NMR structure and helical parameters results from the symmetric input restraints. To permit comparison with the asymmetric crystal structures, inter-base-pair parameters are shown for the full sequence. [b] Values relative to the next base-pair.

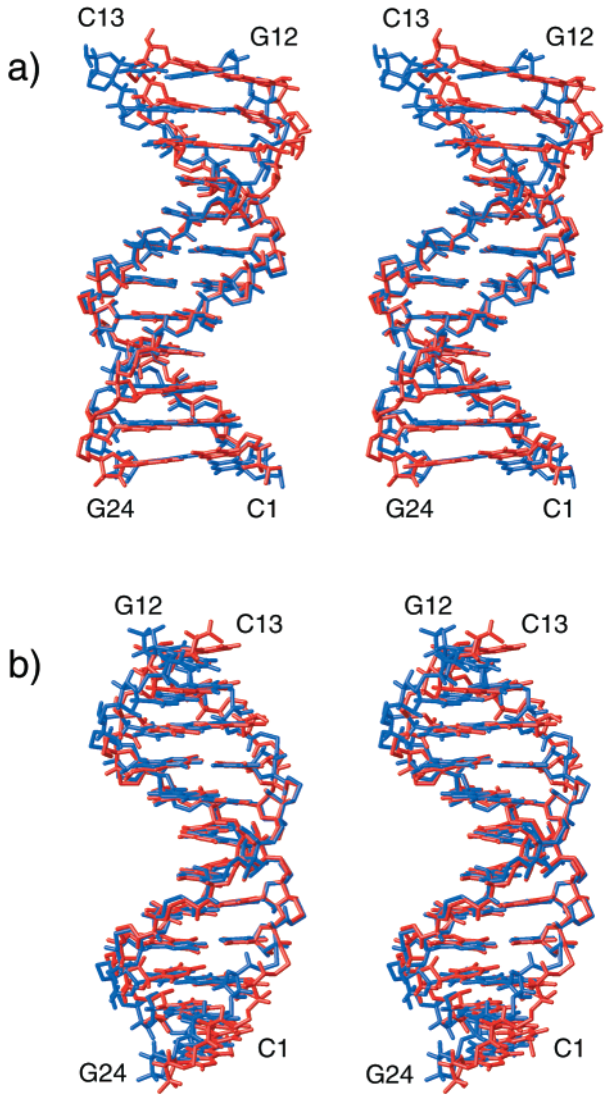

Figure 6. Two stereoviews of the NMR structure of d(CGCGAATTCGCG)$_2$. In blue is shown the NMR-dipo structure and in red the 1BNA[15] X-ray structure. (B) is rotated by 90° around the helix axis relative to (A) and highlights the deviation from cylindrical symmetry, caused by the very different widths of the minor and major grooves.

in the calculations (see below and Supporting Information, Table 1). As discussed below, the structure differs substantially from the one calculated in the absence of dipolar couplings (NMR-nodipo) and from previous NMR structures, and although somewhat closer to the X-ray structures, it lacks several of the unusual features observed in the crystalline state.

Evaluation of Accuracy. Evaluation of the accuracy of NMR structures is a notoriously difficult problem.[29,78−80] For nucleic acids, such evaluations are even harder than for proteins because, for example, no equivalent criterion to the "most-favored region of the Ramachandran map" is available, and no clear correlation between backbone and glycosyl torsion angles applies. In the present case, the bundle of structures calculated in the presence of dipolar restraints is very narrow (not shown). However, substantial changes in the NMR-dipo structure occur when the magnitude and/or rhombicity of the alignment tensor are varied within the range of their uncertainty. Supporting Information, Table 2, lists the magnitude of the changes in the various helical parameters that occur upon increasing or decreasing D_a by 0.5, and upon increasing or decreasing R by 15%. In particular, the base-pair opening is quite sensitive to these parameters and therefore cannot be interpreted quantitatively. The width of the minor groove is also quite sensitive to D_a and R (Supporting Information, Table 1) and narrows by 1 Å when either D_a or R is increased. However, within the range of uncertainty in D_a and R, the narrowing of the minor groove in the center of the dodecamer is quite pronounced and closely mimics the minor groove width observed in the X-ray crystal structures.[15,17] When considering the entire set of NMR structures, calculated with $D_a = -15.5$, -16, and -16.5 Hz, and $R = 0.22$, 0.26, and 0.3, the rmsd relative to their average equals 0.33 Å for the center 10 base pairs and 0.30 Å for the center six. This full set of structures has been deposited in the PDB.

Another qualitative indication of the accuracy of the NMR-dipo structure can be obtained by comparing the agreement between the structures of two two-base-pair fragments, (G2−C23, C3−G22) and (G10−C15, C11−G14), when superimposing G2 on G10 (not on G14). As mentioned before, the ^{1}H and ^{13}C chemical shifts and J couplings for G2 and G10 are nearly identical, and the same applies for C3 and C11. Therefore, the structures of these dinucleotide base-paired fragments should be very similar, despite the fact that they are aligned differently relative to the alignment tensor frame and exhibit quite different dipolar coupling patterns (Figure 2). Indeed, a very close structural resemblance between these two sets of adjacent base pairs is observed for the NMR-dipo structure (rmsd 0.28 Å). In contrast, considerably less similarity is observed in the crystalline state (1.00 Å for 1BNA; 0.75 Å for 355D). This strongly suggests that much of the difference between our solution NMR

(78) Macarthur, M. W.; Thornton, J. M. *Prot. Struct. Func. Genet.* **1993**, *17*, 232−251.
(79) Doreleijers, J. F.; Rullmann, J. A. C.; Kaptein, R. *J. Mol. Biol.* **1998**, *281*, 149−164.
(80) Doreleijers, J. F.; Vriend, G.; Raves, M. L.; Kaptein, R. *Prot. Struct. Func. Genet.* **1999**, *37*, 404−416.

structure and the corresponding crystal structures results from the rather different environments in which the molecules are studied and not from inaccuracy of the NMR coordinates.

Inspection of Table 1 indicates that the measured dipolar couplings show comparable agreement with both X-ray structures and somewhat worse agreement with the NMR structure calculated without dipolar couplings (NMR-nodipo). Dipolar couplings calculated for Arnott B-DNA[68] deviate the most from the measured dipolar couplings. Not surprisingly, the structure calculated with dipolar couplings included (NMR-dipo) agrees much better with experimental values. More significantly, even when omitting a small, random fraction of dipolar couplings, agreement for this omitted fraction with the calculated structure is considerably improved relative to the structure calculated in the absence of dipolar couplings. However, the relatively small number of dipolar couplings and their nonuniform distribution make it more difficult to quantitatively evaluate structural accuracy in this manner than for proteins.

Comparison with Other Structures. The superpostion of the 1BNA X-ray structure[15] on the NMR-dipo structure is shown in Figure 6. The rmsd is relatively large when measured over the full length of the dodecamer (1.5 Å to both 1BNA and 355D[17]). In contrast, the two X-ray structures fit considerably better to one another (rmsd of 0.90 Å over all 12 base pairs). Inspection of Figure 6 indicates that the larger difference between the X-ray and NMR structures is caused primarily by the asymmetric small kinks, observed in the first and last four base pairs of both X-ray structures, which are absent in the NMR-dipo structure.

When only the center six base pairs are compared, the agreement is considerably better (Xray (1BNA) vs Xray (355D), 0.86 Å; NMR vs 1BNA, 0.53 Å; NMR vs 355D, 0.99 Å). When the NMR-dipo structure is compared with the average structure calculated in the absence of dipolar coupling data (NMR-nodipo), but using the same NOE and torsion angle restraints, the rmsd is rather large (0.89 Å over the center 6 base pairs; 1.7 Å over all 12 base pairs). The width of the bundle of calculated NMR-nodipo structures is also much larger (0.46 Å over the center 6 base pairs; 1.2 Å over all 12 base pairs) but significantly less than the change in the structure. Remarkably, as can be seen from Table 1, the NMR-dipo structure fits the nondipolar restraints about as well as the NMR-nodipo structures. Excluding the dipolar energy term, the difference in total restraint violation energy between the NMR-dipo and NMR-nodipo structures is only 0.8 kcal/mol. The large difference between the NMR-dipo and NMR-nodipo structures therefore is not meaningful and simply reflects the broad range of NMR structures that is compatible with the NMR data in the absence of dipolar couplings. Presumably as a result of minute differences in calculated energy, the entire ensemble of conformations compatible with the restraint data is not fully represented in the ensemble of calculated structures. So, the true ensemble of NMR-nodipo structures compatible with experimental restraints is much larger than reflected in the rmsd of the calculated bundle.

The NMR-dipo bundle width is much narrower than for NMR-nodipo, and is also expected to underestimate the true uncertainty in the derived structure. As discussed above, this bundle width increases when accounting for uncertainties in the alignment tensor, but even then may be smaller than the true uncertainty. Therefore, only rather pronounced differences are discussed below. In Tables 2 and 3, angles that differ by more than 20° from those in either of the two strands of one of the X-ray structures are italicized (1BNA) or boldfaced (355D). A more detailed comparison of the sugar pucker and helical parameters is provided in Table 4 and in the Supporting Information. As can be seen from Table 3, for the center four base pairs, the backbone angles are quite similar to the X-ray structures, but differ notably for the adjacent three base pairs (G2 to G4 and C9−C11), which are proximate to two hydrated Mg^{2+} ions observed in the X-ray structures. These ions have been postulated to cause the small bends observed in these regions of the crystal structures,[16,21] No such bends are observed in the NMR-dipo structure, however. Although the NMR data were collected in the absence of Mg^{2+}, no chemical shift changes are observed in the NMR spectrum upon addition of 2 mM Mg^{2+}, indicating that in solution this Mg^{2+}-induced distortion is not significantly populated.

Table 4 shows some of the characteristic parameters of the NMR-dipo structure, as computed using the program Curves.[72] Comparison of these numbers with those of the original X-ray structure (1BNA)[15] and the more recent one (355D)[17] mostly show reasonable agreement, but also some parameters that differ significantly. The overall length of the DNA duplex for the NMR-dipo and the two X-ray structures are within 1 Å from one another. This contrasts with the NMR-nodipo structure which is shorter by 3−4 Å, similar to NMR structures reported previously.[25,27] Interestingly, the NMR structure by Denisov et al.[26] (2DAU) does not show this shortening, but instead appears to be underwound relative to NMR-dipo and the X-ray structures.

Both X-ray structures show a small degree of curvature for the dodecamer, primarily caused by two small kinks in the first and last four base pairs, and possibly influenced by the nearby situated Mg^{2+} ions and by crystal packing. In solution, no such kinks are observed and as a result the bending of the DNA is even smaller (7°). The net helical bending observed in the NMR-dipo structure is sensitive to both the magnitude and rhombicity of the alignment tensor, however, and changes by as much as 4° within the range of the uncertainty in these parameters ($D_a^{CH} = 16 \pm 0.5$ Hz; $D_r = 0.26 \pm 0.04$).

The base propeller twists in the NMR-dipo structure agree quite well with those for the two X-ray structures (Table 4), but show somewhat less variation than was seen in the X-ray structures. Interestingly, in both the NMR and X-ray structures, a pronounced decrease in propeller twist is seen for the C3−G22 and G10−C15 base pairs. The inter-base-pair rise, roll, and twist parameters in the NMR-dipo structure (Table 4) also show less variation than seen in the two X-ray structures. For example, the NMR structure does not show the opposite signed pronounced roll at C9/G10 and G10/C11, nor the sharply decreased rise between C3 and G4 and between C10 and G11.

Concluding Remarks

The NMR-dipo structure shows reasonably close agreement with the X-ray crystal structures of this molecule, but is more regular in its appearance. Unusual roll and sugar puckers in the first and last four base pairs of the crystal structures have been attributed to intermolecular hydrogen bonds in the crystal lattice[15] and the presence of tightly coordinated hydrated Mg^{2+} ions in the major groove,[81] and consequently, differences relative to the NMR structure are largest in these regions of the structure. General tendencies in twist angles in the NMR-dipo structure are consistent with average values in solution[82] and crystal

(81) McFail-Isom, L.; Shui, X.; Williams, L. D. *Biochemistry* **1998**, *37*, 17105−17111.
(82) Kabsch, W.; Sander, C.; Trifonov, E. N. *Nucleic Acids Res.* **1982**, *10*, 1097−1104.

structures.[32] For example, twist (AA) is larger than twist (AT), and twist (GC) is larger than twist (CG). As predicted, the roll parameter in the center AATT section is slightly negative; positive roll is observed at the CG steps and very small negative roll at GC.[28,83,84]

The solution NMR structure calculated in this study is primarily determined by the dipolar couplings, measured in the dilute liquid crystalline medium. Although the experimental $^1H-^1H$ dipolar couplings agree rather poorly with the X-ray structures and with the NMR-nodipo structure (Table 1), agreement with NMR structures calculated in the presence of all heteronuclear (but no homonuclear) dipolar couplings and NOEs is much better (data not shown). In fact, the difference between the structures with and without the $^1H-^1H$ dipolar couplings is less than 0.3 Å over the center 10 base pairs. Similarly, after the dynamics protocol had been optimized and the sign of most $^1H-^1H$ dipolar couplings had been established, converging structures were obtained when omitting all NOE restraints, starting either from Arnott A-form or B-form DNA (data not shown). These no-NOE structures differ by less than 0.2 Å from the NMR-dipo structures, calculated with all NOEs included (center 10 base pairs), confirming that the NMR-dipo structure is defined primarily by the dipolar couplings and not by the NOEs. When starting from fully randomized structures, we found NOEs to be essential for obtaining a low-resolution model, before dipolar couplings could be introduced (Experimental Section). So, with the structure calculation protocol used, we could only obtain these no-NOE structures when starting from helical model structures.

It is perhaps remarkable that the C−H dipolar couplings agree more poorly with our NMR-nodipo structure than with either of the X-ray structures (Table 1). Also, even though the NMR-dipo structure fits the NOE distance restraints about equally well as the NMR-nodipo structure, they are quite different in both local and global features. This confirms that the simple approach used in our study, adapted from protein NMR, where NOEs are converted into relatively loose distance restraints, is not able to define the DNA structure accurately. More extensive measurement of NOE buildup curves and more sophisticated approaches for deriving distances from these NOEs reportedly can increase the accuracy of the derived interproton distances and, thereby, of the NOE-derived DNA structures.[85,86] It is likely that combined use of this latter approach and the dipolar couplings will further improve the accuracy of the final structure.

It is important to point out, however, that DNA is a rather flexible molecule, and the derived structure aims to reflect only the time average of the atomic positions. The approach used is based on the premise that angular fluctuations are relatively small and of similar magnitude for all C−H bond vectors. Although this results in reasonable structures, which appear to be of considerably higher quality than attainable in solution without the use of dipolar couplings, they tend to yield sugar pucker amplitudes that are smaller than expected in a pure C2′-endo conformation, and the pseudorotation angle *P* is also shifted in the direction of C3′-endo by about 10−15% of the difference between the idealized C2′-endo and C3′-endo *P* values for the purine nucleotides, and by about 15−25% for the pyrimidines. As the intermediate *P* value and low pucker amplitude correspond to relatively high energy conformations,[76] it is unlikely that this apparent intermediate conformation is significantly populated. Instead, it provides additional support for previous studies which found the sugar conformation to be in a rapid dynamic equilibrium between C2′-endo and a less populated C3′-endo pucker.[8,10,73−75]

Rotations around the helix axis would not be constrained by the dipolar couplings if the alignment tensor were axially symmetric. However, for the short dodecamer studied here, there is a considerable deviation from axial symmetry because it is only slightly longer than one helical turn (Figure 6b), whereas a considerably longer segment would be needed to impose axial symmetry. It is important to note, therefore, that the presence of a rhombic component in the alignment tensor makes it more useful in defining the molecular structure. A similar observation was made in a recent application of dipolar couplings to the study of protein structure.[87]

The structure calculated here was derived using a nearly complete set of measurable one-bond C−H and N−H dipolar couplings and utilized specific deuteration at the C2′ and C5′ sites. Even so, residual error in the calculated average structure remains, in part caused by the uncertainty in the exact magnitude and rhombicity of the alignment tensor, but also by the large uncertainties in the $^1H-^1H$ experimental dipolar couplings. For uniformly labeled oligonucleotides, it frequently will be impossible to obtain such complete sets of C−H and N−H dipolar data. Interestingly, when only using those dipolar couplings that are expected to be readily measurable in such uniformly labeled oligonucleotides (one-bond base and C1′−H1′), together with all $^1H-^1H$ dipolar couplings, dihedral angles, and NOEs, the resulting structure is almost the exact average between the NMR-nodipo and NMR-dipo structures (data not shown). However, if methods can be developed for reliable measurement of one-bond $^{13}C-^{13}C$ and two-bond $^1H-^{13}C$ dipolar couplings, and also for more quantitative measurement of $^1H-^1H$ couplings, these parameters may be able to compensate for the fraction of one-bond interactions that is lost as a result of increased spectral overlap in uniformly labeled oligonucleotides.

One-bond dipolar couplings tightly constrain the relative orientation of different parts of a macromolecule but do not provide direct information on their separation. It recently has been shown that the use of paramagnetic shifts provides a very accurate measure for the distance from the paramagnetic metal(s).[40,41,88,89] This strongly suggests that the combination of long-range distance restraints derived from such paramagnetic shifts, combined with dipolar coupling information measured in the liquid crystalline phase, may yet provide another avenue to characterize the solution structure of oligonucleotides at even higher accuracy.

Acknowledgment. The authors thank Andrew Lane for providing the coordinates of his NOE-derived model, and Victor Zhurkin, Marius Clore, and Attila Szabo for useful discussions. This work was supported, in part, by CREST (Core Research for Evolutional Science and Technology) of the Japan Scientific and Technology Cooperation (JST), and by the AIDS Targeted Anti-Viral Program of the Office of the Director of the National

(83) Bolshoy, A.; McNamara, P.; Harrington, R. E.; Trifonov, E. N. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2312−2316.
(84) Crothers, D. M.; Haran, T. E.; Nadeau, J. G. *J. Biol. Chem.* **1990**, *265*, 7093−7096.
(85) Borgias, B. A.; James, T. L. *Methods Enzymol.* **1989**, *176*, 169−183.
(86) Boelens, R.; Koning, T. M. G.; Kaptein, R. *J. Mol. Struct.* **1988**, *173*, 299−311.
(87) Delaglio, F.; Kontaxis, G.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 2142−2143.
(88) Banci, L.; Bertini, I.; Savellini, G. G.; Romagnoli, A.; Turano, P.; Cremonini, M. A.; Luchinat, C.; Gray, H. B. *Prot. Struct. Func. Genet.* **1997**, *29*, 68−76.
(89) Tu, K.; Gochin, M. *J. Am. Chem. Soc.* **1999**, *121*, 9276−9285.
(90) Marion, D.; Bax, A. *J. Magn. Reson.* **1988**, *80*, 528−533.
(91) Zweckstetter, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 3791−3792.

Institutes of Health. **Accession Codes.** The coordinates of the lowest energy structure and structures calculated for $D_a = -16 \pm 0.5$ and $R = 0.26 \pm 0.04$ have been deposited in the PDB (accession codes RCSB010377 and 1DUF), together with all of the restraints used in the structure calculation.

Supporting Information Available: Four tables containing P−P distances across the minor groove, the influence of D_a and R on backbone and helical parameters, comparisons of sugar conformations in the X-ray and NMR structures, and helical bending information. This material is available free of charge via the Internet at http://pubs.acs.org.

JA000324N