Determining DNA Global Structure and DNA Bending by Application of NMR Residual Dipolar Couplings

Annaleen Vermeulen, Hongjun Zhou, and Arthur Pardi*

Contribution from the Department of Chemistry and Biochemistry, University of Colorado at Boulder, Boulder, Colorado 80309-0215

Received May 31, 2000. Revised Manuscript Received August 8, 2000

Abstract: The local structure of nucleic acids can be determined from traditional solution NMR techniques, but it is usually not possible to uniquely define the global conformation of DNA or RNA double helices. This results from the short-range nature of the NOE-distance and torsion angle constraints used in generating the solution structures. However, new alignment techniques make it possible to readily measure residual dipolar couplings, which provide information on the relative orientation of individual bond vectors in the molecule. To determine the effects of incorporating dipolar couplings in the structure determinations of nucleic acids, molecular dynamics calculations were performed with simulated constraints derived from two DNA duplex target molecules. Refinements that included NOE, torsion angle, and dipolar coupling constraints were compared to refinements without dipolar couplings. These results show that dipolar couplings significantly improved the local structure while also dramatically improving the global structure of DNA duplexes. The model simulations also illustrate that molecular dynamics calculations induce changes in the local structure before the global structure, which can have important implications for refinements with dipolar coupling constraints. Results are presented that show that the inclusion of dipolar coupling constraints makes it possible to accurately and precisely reproduce the overall helical bend in a DNA duplex. The implications of including dipolar coupling constraints in defining DNA global structure and DNA bending in solution will be discussed.

Introduction

NMR is the only tool presently available for determining high-resolution three-dimensional structures of nucleic acids in solution. Traditionally, NMR structures are generated from proton−proton distance constraints (<5 Å) obtained from nuclear Overhauser effects (NOEs) and torsion angle restraints obtained from through-bond scalar coupling constants.[1] The quality of structures determined by solution NMR techniques depends critically upon the number and type of experimental constraints, and there is still some debate concerning how accurately and precisely nucleic acid structures can be determined by NMR.[2−6] In most cases, however, nucleic acids are less well-defined than proteins by standard NMR techniques. This arises from the lower density of protons in nucleic acids, and because most of the interresidue NOEs are between neighboring base pairs. Computer simulations that mimic NMR-type data have shown that some local conformational parameters, such as sugar pucker and glycosidic torsion angle, are well-determined by NOEs, whereas other parameters such as helical twist, helical rise, and backbone torsion angles are not very well-defined.[4,5] For a larger RNA such as the hammerhead ribozyme, simulations have shown that the relative positions of the three helical stems surrounding the catalytic core are not well-defined from proton−proton NOE data.[6] Thus, although standard NOE and *J*-coupling constraints are able to define local structure, they are poor at defining global conformation, especially for longer helices or larger nonglobular structures.

Through-space ^{1}H−^{1}H, ^{1}H−^{13}C, and ^{1}H−^{15}N residual dipolar couplings, which are a function of both the distance and the orientation between two nuclei, are a potentially rich source of additional structural information.[7] These dipolar couplings average to zero in isotropic solutions and, therefore, are not normally observed. Recently, solution NMR techniques have been developed in which a liquid crystalline cosolute, such as bicelles, filamentous phage, or purple membranes, are used to induce anisotropic rotational diffusion of the macromolecule, which then allows for the measurement of residual dipolar couplings.[7−11] The size of these residual dipolar couplings can be easily tuned by varying the concentration of the cosolute and only a very small degree of alignment (<0.3%) is needed to achieve an optimal range of residual dipolar couplings (10−50 Hz). The most easily interpreted dipolar couplings are for nuclei at known distances, because the residual dipolar couplings are then primarily a function of the orientation of the internuclear

* To whom correspondence should be addressed. E-mail: arthur.pardi@colorado.edu.

(1) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986.
(2) Pardi, A.; Hare, D. R.; Wang, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8785−8789.
(3) Gronenborn, A. M.; Clore, G. M. *Biochemistry* **1989**, *28*, 5978−5984.
(4) Metzler, W. J.; Wang, C.; Kitchen, D. B.; Levy, R. M.; Pardi, A. *J. Mol. Biol.* **1990**, *214*, 711−736.
(5) Ulyanov, N. B.; Gorin, A. A.; Zhurkin, V. B.; Chen, B. C.; Sarma, M. H.; Sarma, R. H. *Biochemistry* **1992**, *31*, 3918−3930.
(6) Allain, F. H.; Varani, G. *J. Mol. Biol.* **1997**, *267*, 338−351.
(7) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.
(8) Hansen, M. R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065−1074.
(9) Clore, G. M.; Starich, M. R.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 10571−10572.
(10) Sass, J.; Cordier, F.; Hoffmann, A.; Cousin, A.; Omichinski, J. G.; Lowen, H.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 2047−2055.
(11) Koenig, B. W.; Hu, J. S.; Ottiger, M.; Bose, S.; Hendler, R. W.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 1385−1386.

10.1021/ja001919l CCC: $19.00
Published on Web 09/26/2000

vector. Therefore, the most commonly measured dipolar couplings are one-bond $^1H-^{13}C$ or $^1H-^{15}N$, although $^1H-^1H$ dipolar couplings have been used to provide additional structural information.[12,13]

Residual dipolar couplings are becoming routinely used in the refinement of protein structures.[14−21] Studies have shown improved precision and accuracy for protein structure determinations when residual dipolar couplings are included in the structure refinement.[15,22] Many proteins have quite globular compact structures in which both the local and global structures are reasonably well-defined by NOE-based NMR techniques. In contrast, nucleic acids often have extended structures that are not well-defined by NOEs. Thus, the residual dipolar couplings may have an even greater impact in nucleic acids than in proteins. The globular structure of proteins often leads to a relatively random distribution of the backbone N−H or C−H bond vector orientations obtained from dipolar coupling data,[23] whereas the regular structure for RNA or DNA helices can lead to a limited distribution of bond vector orientations. Therefore, it is not clear if the improvements observed by including residual dipolar couplings in protein structure determinations will translate into analogous improvements in nucleic acids. There is currently only one example in which dipolar couplings were used in the refinement of a nucleic acid, in which $^{13}C-^1H$ and $^{15}N-^1H$ dipolar couplings in the RNA were included in the structure determination of a U1A protein-RNA complex.[24] In this case, inclusion of the residual dipolar couplings did not lead to a significant overall improvement in the precision of the structure. Thus, it is still an open question to what extent inclusion of residual dipolar couplings will improve the accuracy and precision of the local and global structure of nucleic acids.

Here we examine the benefits of including residual dipolar couplings as additional constraints in the structure determination of nucleic acids. In proteins, the effects of dipolar couplings on solution structure determinations can be studied by comparing the predicted couplings for a molecule with a high-resolution X-ray structure to the measured dipolar couplings of the protein in solution, as was done for ubiquitin.[7] This approach cannot be used for DNA helices, because it is well-known that these duplexes are sensitive to crystal packing forces that can affect their local and global structure.[25,26] Therefore, computer simulations need to be performed to evaluate how accurately and precisely the local and global structure of a target DNA duplex can be determined by inclusion of dipolar coupling data. Various sets of simulations were performed here with different length DNAs, with changes in the number and type of backbone torsion angle constraints, or with changes in the orientation and symmetry of the alignment tensor in order to evaluate how these variables affect the refined structures.

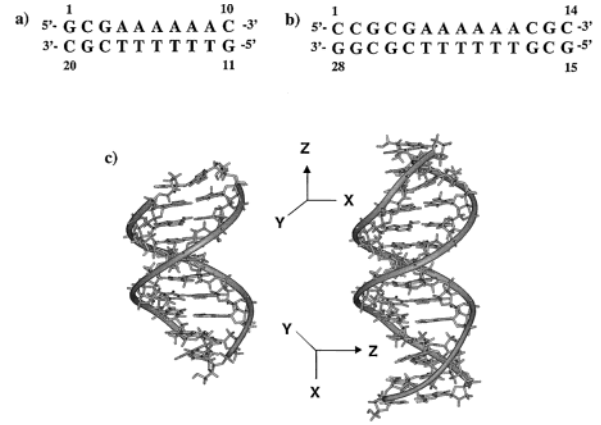

Figure 1. Sequences of the (a) 10mer and (b) 14mer target structures. (c) Target structures of the 10mer (left) and 14mer (right). Orientations of the two alignment tensor axes used in these simulations are also shown, with the principal axis of the alignment tensor (*z*) oriented parallel to the helix axis (top) and perpendicular to the helix axis (bottom).

Another question that will be addressed in these simulations is whether the orientational information obtained from the dipolar couplings can be used to define DNA bending in solution. Gel electrophoresis experiments unambiguously show that A-tract sequences give rise to DNA bending in solution; however, the molecular mechanism leading to the bending is still controversial.[25,27−30] To help address this DNA bending question, we chose an A-tract sequence as the target structure for the calculations (Figure 1). Thus, these simulations will not only provide insight into the role of residual dipolar couplings in defining local and global structures of nucleic acids, but will also address the feasibility of using dipolar couplings to determine DNA bending in solution.

Materials and Methods

Generating Target Structures. A 10-base-pair DNA duplex target structure, shown in Figure 1a and referred to here as the 10mer, was generated from the coordinates of the central 10 base pairs in the X-ray structures of the A-tract DNA dodecamer solved by Steitz and coworkers (PDB #1D89).[25] Protons were added using InsightII (Molecular Simulations, Inc.), followed by a 1000 step Powell minimization using X-PLOR[31] with hydrogen bond constraints as well as bond, angle, and van der Waals energy terms to ensure satisfactory geometry. Proton−proton distance constraints were then generated from this minimized structure and included all distances less than 5 Å with an error of ±0.2 Å. This structure was further minimized with the above distance constraints and additional distance constraints in the major groove, to help straighten out the helix.

(12) Hansen, M. R.; Rance, M.; Pardi, A. *J. Am. Chem. Soc.* **1998**, *120*, 11210−11211.
(13) Tjandra, N.; Marquardt, J.; Clore, G. M. *J. Magn. Reson.* **2000**, *142*, 393−396.
(14) Bax, A.; Tjandra, N. *J. Biomol. NMR* **1997**, *10*, 289−292.
(15) Tjandra, N.; Omichinski, J. G.; Gronenborn, A. M.; Clore, G. M.; Bax, A. *Nat. Struct. Biol.* **1997**, *4*, 732−738.
(16) Baber, J.; Libutti, D.; Levens, D.; Tjandra, N. *J. Mol. Biol.* **1999**, *289*, 949−962.
(17) Fischer, M. W.; Losonczi, J. A.; Weaver, J. L.; Prestegard, J. H. *Biochemistry* **1999**, *38*, 9013−9022.
(18) Losonczi, J. A.; Prestegard, J. H. *Biochemistry* **1998**, *37*, 706−716.
(19) Markus, M. A.; Gerstner, R. B.; Draper, D. E.; Torchia, D. A. *J. Mol. Biol.* **1999**, *292*, 375−387.
(20) Olejniczak, E. T.; Meadows, R. P.; Wang, H.; Cai, M. L.; Nettesheim, D. G.; Fesik, S. W. *J. Am. Chem. Soc.* **1999**, *121*, 9249−9250.
(21) Mueller, G. A.; Choy, W. Y.; Yang, D. W.; Forman-Kay, J. D.; Venters, R. A.; Kay, L. E. *J. Mol. Biol.* **2000**, *300*, 197−212.
(22) Drohat, A. C.; Tjandra, N.; Baldisseri, D. M.; Weber, D. J. *Protein Sci.* **1999**, *8*, 800−809.
(23) Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Magn. Reson.* **1998**, *133*, 216−221.
(24) Bayer, P.; Varani, L.; Varani, G. *J. Biomol. NMR* **1999**, *14*, 149−155.
(25) DiGabriele, A. D.; Steitz, T. A. *J. Mol. Biol.* **1993**, *231*, 1024−1039.
(26) Dickerson, R. E.; Goodsell, D. S.; Neidle, S. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3579−3583.
(27) Crothers, D. M.; Haran, T. E.; Nadeau, J. G. *J. Biol. Chem.* **1990**, *265*, 7093−7096.
(28) Hagerman, P. J. *Annu. Rev. Biochem.* **1990**, *59*, 755−781.
(29) Dickerson, R. E.; Goodsell, D.; Kopka, M. L. *J. Mol. Biol.* **1996**, *256*, 108−125.
(30) Hud, N. V.; Sklenar, V.; Feigon, J. *J. Mol. Biol.* **1999**, *286*, 651−660.
(31) Brünger, A. T. *X-PLOR 3.1: A System for X-ray Crystallography and NMR*; Yale University Press: New Haven, 1992.

The 14mer duplex target structure (see Figure 1b) was generated by adding one base pair to each end of the X-ray crystal structure.[25] Protons were added using InsightII, and 1000 steps of Powell minimization were performed as described above, except that no distance constraints were used. These 10mer and 14mer target structures showed low energies using the X-PLOR potentials and, therefore, could be used as target structures for the computer simulations.

Distance Constraints. Distance constraints were generated from the target structures to simulate an experimental NOE data set. Only proton−proton distances less than 4.5 Å were selected, but no distance constraints were included for the H5′ and H5″ protons, the amino protons of guanine and adenine, or for any intraresidue sugar−sugar interactions. Because the two amino protons of cytosine undergo chemical exchange in solution by rotation about the C−N bond, only the shorter distance constraint for each amino proton pair was used in the constraint set. The distances involving methyl groups were calculated from the center of the three protons. All the NOE constraints were defined with an uncertainty of ±0.5 Å except for distances to a methyl group for which the uncertainty was increased to ±1.0 Å. The minimum lower bound for all distance constraints was set to 1.8 Å. To try to mimic a realistic experimental data set in which resonance overlap prevents unambiguous assignment of some of the NOEs, the final set of distance constraints included only a random sample of 70% of these constraints.

Hydrogen bond constraints for base pairs were determined from distances in the target structures. Distance constraints were generated between the proton and the hydrogen bond acceptor, as well as between the two heavy atoms of each hydrogen bond, and these constraints used uncertainties of ±0.2 Å. Two- and three-hydrogen bond constraints were used for each A−T and G−C base pair, respectively.

Torsion Angle Constraints. The program CURVES 5.2[32] was used to determine the pucker, phase, and amplitude of pseudorotation for all sugars in the target structure.[33] If the observed sugar pucker could be unambiguously determined from standard ^{1}H−^{1}H ^3J-coupling constants, ν_1 and ν_3 sugar pucker constraints were included for this residue.[34] Thus, torsion angle constraints were included for 14 and 20 of the sugars in the 10mer and 14mer duplexes, respectively. For one set of calculations, backbone torsion angle constraints were also included. To simulate realistic backbone torsion angle data, a random sample of 70% of the β, γ, and ϵ backbone torsion angles were included as constraints.[34] The torsion angles were determined from the target structures with CURVES 5.2 and were constrained with ranges of ±30°.

Dipolar Coupling Constraints. ^{1}H−^{13}C dipolar coupling constraints were generated for each target structure using the following equations[23]

$$D_{CH} = D_a[(3\cos^2\theta - 1) + {}^3/_2 R(\sin^2\theta \cos 2\phi)] \quad (1)$$

$$D_a = [-(h\mu_o S\gamma_C\gamma_H)/(16\pi^3 \langle r^3 \rangle)]A_a \quad (2)$$

$$D_r = [-(h\mu_o S\gamma_C\gamma_H)/(16\pi^3 \langle r^3 \rangle)]A_r \quad (3)$$

where D_{CH} is the residual dipolar coupling, A_a and A_r are the axial and rhombic components of the molecular alignment tensor, respectively, and θ and ϕ are the angles in spherical coordinates that define the orientation of the internuclear vector. The rhombicity, R, is given by D_r/D_a. Both D_a and D_r contain the averaged internuclear distance (r), Planck's constant (h), the permeability of a vacuum (μ_o), the generalized order parameter (S), and the gyromagnetic ratios for the two nuclei (γ). The dipolar coupling constraints were generated with $D_a = -20$ Hz and $D_r = 0$ Hz unless otherwise indicated. The orientation of the alignment tensor is defined relative to the molecular reference frame by three Euler angles. In all but one of the simulations, the Euler angles were chosen so that the principal axis of the alignment tensor was parallel to the helix axis of the DNA; in the other simulation, the principal axis was chosen to be perpendicular to the helix axis.

Table 1. Simulated NMR Constraints Used in the Structure Calculations

	10mer	14mer
total NOE constraints	194	287
intraresidue	78	109
interresidue	116	178
sequential	91	141
nonsequential	25	37
hydrogen bond constraints	48	72
total torsion angle constraints	66	95
sugar pucker (ν_1 and ν_3)	2 × 14	2 × 20
phosphate backbone (β, γ, and ϵ)	38	55
^{1}H−^{13}C dipolar coupling constraints	91	123

To simulate a realistic set of dipolar coupling constraints for DNA, only a certain percentage of all possible one-bond ^{1}H−^{13}C dipolar couplings was selected. The type and percentage of dipolar couplings used were similar to what were experimentally observed in measurements of ^{1}H−^{13}C dipolar couplings at natural abundance in the Dickerson dodecamer (A.V., H.Z., A.P., unpublished results). Thus, we chose 45% of the C2′H2′ and C2′H2″, 70% of the C1′H1′, C3′H3′, and C4′H4′, and 90% of the aromatic CH dipolar couplings as constraints. Table 1 gives a summary of the constraints for the 10mer and 14mer, and Tables S1 and S2 in Supporting Information give the detailed lists of dipolar coupling constraints for the 10mer and 14mer, respectively.

Two other sets of dipolar coupling constraints were generated for the 10mer. In one case, a rhombic component was introduced into the alignment tensor with values of $D_a = -20$ Hz and $D_r = -4$ Hz ($R = 0.2$), and in the second set, the alignment tensor had no rhombic component ($R = 0$), but the principal axis of the alignment tensor was perpendicular to the helix axis (see Figure 1c).

Refinement with Distance and Torsion Angle Constraints. All structure calculations were performed using a modified version of X-PLOR 3.8 (kindly provided by G. M. Clore, NIH) that included refinement with dipolar couplings.[15,31] The refinement with distance and torsion angle constraints employed a molecular dynamics (MD)/simulated annealing protocol similar to what was previously described.[35,36] The starting structures were generated by randomizing the backbone torsion angles (α, β, γ, ϵ, ζ), the glycosidic torsion angle (χ), and the phase and amplitude of pseudorotation for the sugar pucker. Several rounds of simulated annealing (for a total of 37 ps) at high temperature were performed with low force constants for the van der Waals (vdw) interactions and with zero lower bounds for hydrogen bonding interactions, which allows the atoms to move through each other in order to satisfy constraints. The additional rounds of simulated annealing (for a total of 75 ps) used normal lower bounds for the hydrogen bonds. The force constants for the various potentials used in the last rounds of MD are given in Table 2. The 15 lowest-energy structures were subjected to a final 500-step Powell minimization with the nonbonded repel term in X-PLOR replaced by a Lennard−Jones potential.

Sugar pucker torsion angle constraints were next added to the MD calculations (for 7.5 ps), followed by a 500-step Powell minimization with a Lennard−Jones potential. The phosphate backbone torsion angle (β, γ, and ϵ) constraints were then added in another round of MD refinement (for 15 ps), followed by Powell minimization.

Determining the Axial and Rhombic Components of the Alignment Tensor. The version of X-PLOR that was used here employs fixed magnitudes for the axial and rhombic components of the alignment tensor during the refinement. Thus, a grid search was used to determine the initial values of the magnitudes for the alignment tensor components, D_a and R.[37] The input structures were from the refinement with NOE

(32) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dynam.* **1989**, *6*, 655−667.
(33) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(34) Wijmenga, S. S.; Mooren, M. M. W.; Hilbers, C. W. In *NMR of Macromolecules*; Roberts, G. C. K., Ed.; Oxford University Press: Oxford, 1993; pp 217−288.
(35) Allain, F. H. T.; Varani, G. *J. Mol. Biol.* **1995**, *250*, 333−353.
(36) Hoogstraten, C. G.; Legault, P.; Pardi, A. *J. Mol. Biol.* **1998**, *284*, 337−350.
(37) Clore, G. M.; Gronenborn, A. M.; Tjandra, N. *J. Magn. Reson.* **1998**, *131*, 159−162.

Table 2. Structural Statistics for the 10mer and 14mer Structures

	10mer				14mer			
refinement constraints[a]								
NOEs	*	*	*	*	*	*	*	*
sugar pucker torsions	*	*	*	*	*	*	*	*
backbone torsions			*	*			*	*
dipolar couplings		*		*		*		*
energies (kcal mol^{-1})[b]								
total	−222	−200	−242	−237	−295	−252	−333	−317
bonds	4.2	5.7	4.7	4.9	6.5	8.2	7.4	7.3
angles	30	36	26	28	50	56	42	41
impropers	5.4	6.1	5.5	5.3	8.0	9.0	7.9	8.0
vdw	−262	−251	−278	−275	−360	−332	−391	−376
NOE	0.46	0.68	0.05	0.02	0.97	1.25	0.06	0.20
torsion angle	0.009	0.015	0.006	0.008	0.11	0.05	0.04	0.02
dipolar couplings		2.5		0.8		4.9		2.3
rmsd from restraints								
NOE (Å)	0.0025	0.0027	0.0008	0.0005	0.0031	0.0078	0.0008	0.0030
torsion angles (°)	0.031	0.028	0.011	0.013	0.11	0.057	0.035	0.0261
dipolar couplings (Hz)[c]		0.75		0.68		0.77		0.75
number of violations								
NOE (>0.1 Å)	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0
dihedral (>3°)	0.0	0.0	0.0	0.0	0.0	0	0.0	0.0
dipolar coupling (>1.5 Hz)		5.0		3.3		7.3		6.9
dipolar coupling (>3.0 Hz)		0.5		0.4		0.6		0.6

[a] All values are averaged for the 15 final structures. An asterisk (*) indicates this constraint was included in the refinement. [b] The energies were calculated using the following force constants for the experimental-type constraints: NOE-distance, 50 kcal mol^{-1} Å^{-2}; torsion angle, 500 kcal mol^{-1} rad^{-2}; dipolar coupling, 0.05−0.15 kcal mol^{-1} Hz^{-2}. Energy constants for bonds, angles, and impropers are defined in the X-PLOR parameter file, dna-rna-allatom.param.[50] [c] Dipolar coupling rmsds are calculated from the input dipolar couplings.

and sugar pucker constraints. For the grid search on the 10mer, D_a was varied from −30 to −12 Hz (2 Hz steps), and R, from 0 to 0.4 (0.1 step size). For each D_a and R combination, 2 ps of MD was performed, for which the temperature was reduced from 350 to 100 K, and the dipolar coupling force constant was increased from 0.01 to 0.3 kcal mol^{-1} Hz^{-2}. This was followed by a final Powell minimization. The lowest energy in the grid search were for values around $D_a = -20$ Hz and $R = 0$, which are the values of D_a and R that are used for generating the dipolar coupling constraints. This demonstrates that it is possible to determine the correct values of D_a and R for the DNA using a grid search procedure in the refinement. For the rest of the simulations, the actual values of D_a and R that were employed in calculating the dipolar couplings were used as input for X-PLOR, except for refinements on the 10mer that used incorrect values for D_a and R, as discussed below.

Refinement with Dipolar Couplings. The dipolar coupling refinements[15] started from the structures generated with NOEs and the different sets of torsion angle constraints. A quadratic harmonic function was used as the dipolar coupling energy term.[15] The force constant for dipolar couplings was calibrated so that the rmsd between the input and predicted dipolar couplings is approximately 0.7 Hz. In the final structures, most of the deviations between the final and target (Table S1 and S2) dipolar couplings are less than ±1.5 Hz, and there are no deviations larger than 5 Hz.

For the 10mer, the first round of refinement consisted of a 2.5-ps MD at 400 K, followed by a 7-ps MD for which the temperature was reduced to 100 K. During cooling, the dipolar coupling force constant was gradually increased from 0.01 to 0.1 kcal mol^{-1} Hz^{-2}. The second round of refinement included a 0.5-ps MD at 400 K, followed by a 7-ps MD where the temperature was reduced to 100 K, but the dipolar coupling force constant was kept at 0.1 kcal mol^{-1} Hz^{-2}. The molecular dynamics was followed by 200 steps of Powell minimization with a Lennard−Jones potential. The refinement procedure for the 14mer structure was similar to the 10mer, except that the 14mer required longer MD periods for which the force constant for dipolar couplings was gradually increased to 0.15 kcal mol^{-1} Hz^{-2}. This refinement was completed using 200 steps of Powell minimization with a Lennard−Jones potential.

The 10mer and 14mer dipolar coupling refinements that started from the structures generated with NOEs, sugar pucker, and backbone torsion angle constraints used a similar protocol, except that they required shorter MD periods and lower dipolar coupling force constants of 0.05 and 0.07 kcal mol^{-1} Hz^{-2} to achieve low energies. The refinements were completed using 500 steps of Powell minimization with a Lennard−Jones potential.

Refinements with dipolar coupling constraints using incorrect values for D_a and R were performed for the 10mer to examine how errors in the determination of these alignment parameters affect the results. The dipolar coupling constraints were generated with $D_a = -20$ Hz and $R = 0$ as described above, and the following incorrect alignment values were used for the structure calculations: −20, 0.2; −20, 0.1; −20, 0.05; −25, 0; and −21, 0 for D_a and R, respectively. The refinement procedure was the same as described above for the 10mer refinement using NOE, torsion angle, and dipolar coupling constraints.

Results and Discussion

Solution Structures of DNA Using Distance and Torsion Angle Constraints. The quality of structures produced by standard solution NMR techniques has been the subject of various investigations, and the general consensus is that some of the local structural features, but not the global structure, of DNA can be determined from standard NOE and *J*-coupling data.[2−6] The goal of this study was to extend these earlier studies to determine how the addition of dipolar coupling constraints affects the quality of nucleic acid structures generated by solution NMR-type data. However, we first needed to perform simulations with the standard NMR constraints. Thus, control calculations were performed that employed NOE distance constraints and sugar pucker and/or backbone constraints that mimic a conservative set of the most easily measured NMR data (see Methods). As seen in Table 1, the different sets of simulations used approximately 10 NOE, 1 or 3 torsion angle, and 0 or 4 dipolar coupling constraints per residue. The structural statistics for the control calculations on the 10mer and 14mer are given in Table 2 and show low-energy structures with essentially no NOE or torsion angle violations. For the 10mer and 14mer, the calculations with only the NOE and sugar pucker constraints had an average rmsd to the target structure of 2.37 and 3.31 Å, respectively (see Table 3). The rmsds for the

Table 3. Average Root Mean Square Deviations for the Structures from Different Refinements (in Å)

	target structure[a]		average structure[b]
	all[c]	middle[d]	all[c]
10mer			
control[e]	2.37	1.47	1.63
dipolar coupling[f]	1.29	0.95	0.78
control w/ backbone[g]	1.67	0.90	1.23
dipolar coupling w/backbone[h]	0.83	0.60	0.41
perpendicular alignment[i]	1.31	1.04	0.94
rhombic alignment[j]	1.16	0.84	0.73
14mer			
control[e]	3.31	2.06	2.52
dipolar coupling[f]	1.55	1.20	0.99
control w/backbone[g]	2.73	1.62	2.29
dipolar coupling w/backbone[h]	1.15	0.75	0.72
perpendicular alignment[i]	2.37	1.73	1.75

[a] Average rmsd between the 15 low-energy structures and the target structure. [b] Average rmsd between the 15 low-energy structures and the unminimized average structure. [c] Rmsd values were calculated for all non-hydrogen atoms for all residues. [d] Rmsd values were calculated for non-hydrogen atoms of residues 3−8, 13−18 (10mer) and 3−12, 17−26 (14mer). [e] Refinement included NOE and sugar pucker constraints. [f] Refinement included NOE, sugar pucker, and dipolar coupling constraints. [g] Refinement included NOE, sugar pucker, and backbone torsion angle constraints. [h] Refinement included NOE, sugar pucker, backbone torsion angle, and dipolar coupling constraints. [i] Refinement with the principal axis of alignment perpendicular to DNA axis. Constraints included NOE, sugar pucker, and dipolar coupling. [j] Refinement with the rhombic component in alignment ($D_a = -20$ Hz, $R = 0.2$). Refinement included NOE, sugar pucker, and dipolar coupling constraints.

middle 6 base pairs in the 10mer (Table 3) were similar to what was previously observed in computer simulations on DNA or RNA duplexes using NOE constraints.[4,6] Figures 2a and 3a show the superpositions of two nonterminal base pairs on their respective target structures for the 10mer and 14mer structures that employed the NOE and sugar pucker constraints. Superimposing only two base pairs helps illustrate that the local structure is reasonably well-defined by conventional NMR constraints, but the global structure is poorly defined. Thus, the larger rmsd for superposition of the full structures (Table 3) simply reflects that NOEs and torsion angle constraints do not precisely define the global structure. This is especially true for the 14mer, which emphasizes that the global structure becomes more poorly defined as nucleic acid length increases (see Figures 2a and 3a).

The rmsds for these simulations are somewhat higher than what is seen in experimental NMR structure determinations of this length DNA. This is because the rmsds calculated here are for the superimposition on the target structure; however, in real NMR structure calculations, the true structure is not known and, therefore, the rmsds are reported to the average structures. Thus, rmsds to the average structures are also included in Table 3. These rmsds are always smaller than rmsds to the target and will only give the same values if the average structure is exactly the same as the target structure.

Although the NOE-distance and *J*-coupling sugar pucker constraints are the most common types of constraints in NMR solution structure determinations of DNA, it is possible to generate a more extensive set of experimental constraints for the β, γ, and ϵ torsion angles by the analysis of various homonuclear and heteronuclear *J*-couplings.[34] Thus, we also performed calculations that included constraints for the β, γ, and ϵ backbone torsion angles (see Methods). The inclusion of these backbone torsion angle constraints improved the precision and accuracy of the structures. As seen in Table 3, the average rmsd to the target is reduced to 1.67 and 2.73 Å for the 10mer and 14mer duplexes, respectively.

Improved Global and Local Structure by Inclusion of Dipolar Coupling Data. The goal of this study was to determine how the addition of dipolar coupling constraints affects the quality of nucleic acid structures generated by solution NMR data. We chose to include the one-bond $^1H-^{13}C$ dipolar coupling constraints in these simulations because they can be readily measured in nucleic acids at natural abundance. Although techniques for uniform $^{13}C/^{15}N$ labeling of RNA are now routine,[38,39] and labeling of DNA is now possible,[40−43] the use of $^1H-^{13}C$ dipolar couplings at natural abundance makes the approach employed here applicable to a wide range of nucleic acids, not only isotopically labeled systems.

The addition of dipolar coupling constraints dramatically improved the global structure of these DNA duplexes. The average rmsd to the target structure for the 10mer was reduced from 2.37 to 1.29 Å in the simulations with NOE, sugar pucker, and dipolar coupling constraints (Table 3). The average rmsd was also lower for the dipolar coupling calculations that included the backbone torsion angle constraints, for which the average rmsd to the target structure for the 10mer was reduced from 1.67 to 0.83 Å. An even more dramatic improvement was observed in the 14mer structure for which the average rmsd to the target was reduced from 3.31 to 1.55 Å by including dipolar coupling constraints in simulations with NOE and sugar pucker constraints and from 2.73 to 1.15 Å for the calculations that included the backbone torsion angles. Figures 2 and 3 show the structures of the 10mer and 14mer generated by simulations that included the dipolar couplings with the NOE and sugar pucker and/or backbone constraints. These superpositions illustrate that addition of these readily measured $^1H-^{13}C$ dipolar couplings provides much better definition of the global structure, as compared to the simulations with standard NOEs and torsion angle constraints.

The dipolar coupling constraints also significantly improve the local structure of the DNA. Figure 4a,b shows the rmsds for superposition of each two neighboring base pairs in the 10mer and 14 mer duplexes. In the 10mer and 14mer, the average rmsds were 1.0 and 0.88 Å for the calculations with NOE and sugar pucker constraints, and addition of the dipolar coupling constraints reduced these average rmsds to 0.69 and 0.72 Å, respectively. The local structure is equally well-defined in both the 10mer and the 14mer duplex; thus, as expected, duplex length has little effect on the local structure. Many helical parameters are also much better defined by additional dipolar coupling constraints (Tables S3 and S4). For example, the standard deviation for the helical rise in the 10mer structures with NOE and sugar pucker constraints was reduced from 0.4 to 0.2 Å (see Table S3). There were significant reductions in the standard deviations for most helical parameters: 4.2 to 2.1° for helical tilt, 8.3 to 3.5° for roll, 5.0 to 2.9° for the helical twist, and 11.5 to 4.6° for propeller twist. Thus, the inclusion of these $^1H-^{13}C$ dipolar coupling constraints leads to a dramatic

(38) Nikonowicz, E. P.; Sirr, A.; Legault, P.; Jucker, F. M.; Baer, L. M.; Pardi, A. *Nucleic Acids Res.* **1992**, *20*, 4507−4513.
(39) Batey, R. D.; Inada, M.; Kujawinski, E.; Puglisi, J. D.; Williamson, J. R. *Nucleic Acids Res.* **1992**, *20*, 4515−4523.
(40) Zimmer, D. P.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3091−3095.
(41) Smith, D. E.; Su, J. Y.; Jucker, F. M. *J. Biomol. NMR* **1997**, *10*, 245−253.
(42) Kainosho, M. *Nat. Struct. Biol.* **1997**, *4*, 858−861.
(43) Louis, J. M.; Martin, R. G.; Clore, G. M.; Gronenborn, A. M. *J. Biol. Chem.* **1998**, *273*, 2374−2378.

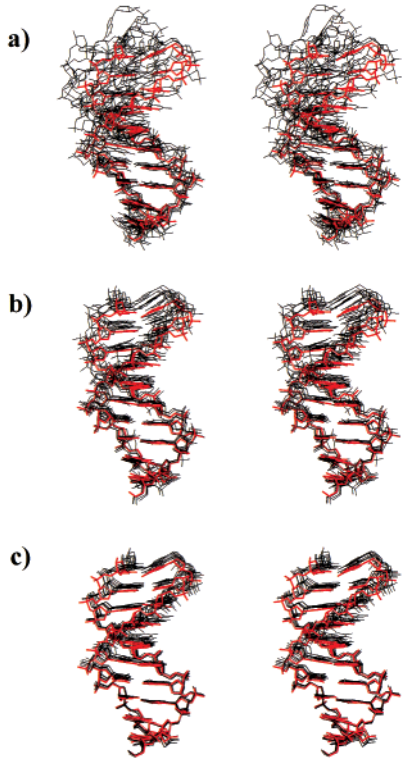

Figure 2. Stereoviews of the superposition of the refined 10mer structures (black) on their target structure (red) with (a) only NOE and sugar pucker constraints, (b) NOE, sugar pucker, and dipolar coupling constraints, and (c) NOE, sugar pucker, backbone torsion angle, and dipolar coupling constraints. Only two nonterminal base pairs (residues 2, 3, 18, and 19) were superimposed in all the structures to help illustrate that the global bend is not defined from traditional NMR structural data. For clarity, the heavy atoms are only shown for 6 of the 15 low-energy structures, but these 6 span the full range of rmsds for that refinement.

improvement in the precision of both the local and global structure determinations of DNA duplexes.

Factors That Need To Be Considered when Refining Solution Structures with Residual Dipolar Couplings. Residual dipolar couplings represent a very different type of structural constraint, as compared to standard NOE distance or *J*-coupling torsion angle constraints.[7,44] The distance and torsion angle constraints only give structural information between atoms that are close together in the structure; specifically, the atoms must be less than 5 Å apart or within three covalent bonds. In contrast, the residual dipolar couplings yield orientation information on individual internuclear vectors relative to the alignment axes of the molecule. Therefore, it is possible to obtain the relative orientation of two internuclear vectors that are far apart in the structure. This has important implications for the protocols used in solution structure refinements when residual dipolar couplings are included as constraints. Distance and torsion angle constraints primarily induce changes in the local geometry for which the global structure is only indirectly defined from a limited set of short-range distance constraints involving tertiary structure interactions. On the other hand, dipolar coupling constraints are readily satisfied by changing either the local or global conformation of the molecule. This is dramatically illustrated in the simulations on the 14-base-pair duplex, for which Figure 3a shows that the global bend of the molecule is not well-defined in the control calculations without dipolar couplings. Thus, even if the local structure is exactly the same

(44) Prestegard, J. H. *Nat. Struct. Biol.* **1998**, *5 Suppl*, 517−522.

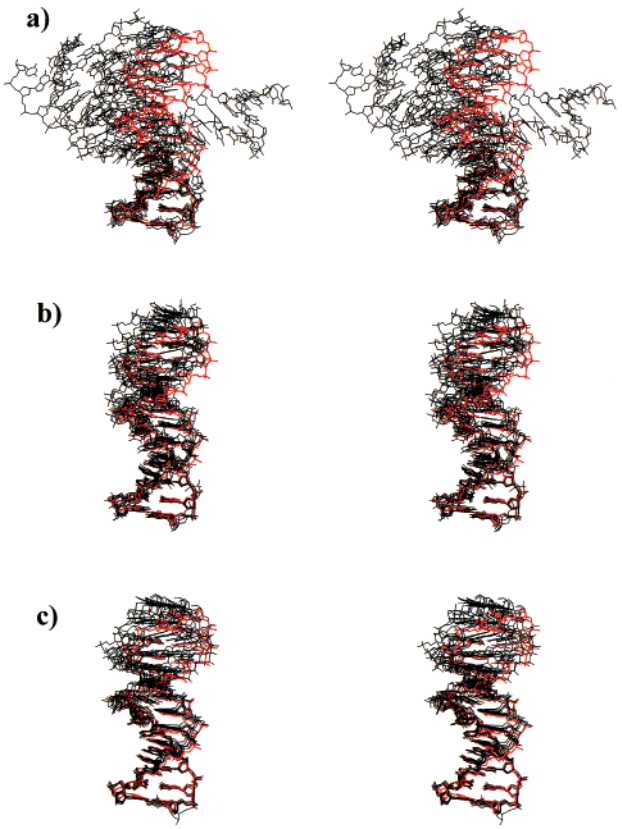

Figure 3. Stereoviews of the superposition of the refined structures (black) on the target structure (red) of the 14mer for the refinements with (a) only NOE and sugar pucker constraints, (b) NOE, sugar pucker, and dipolar couplings, and (c) NOE, sugar pucker, backbone torsion, and dipolar coupling constraints. Only two nonterminal base pairs (residues 2, 3, 26, and 27) were superimposed in all the structures to help illustrate that the global bend is not defined from traditional NMR structural data. For clarity, the heavy atoms are only shown for 6 of the 15 low-energy structures, but these 6 span the full range of rmsds for that refinement.

in all of these structures, the different global bends will lead to very different C−H bond vector orientations and, therefore, to different residual dipolar couplings. These differences in the global structure will lead to large differences between the predicted and observed dipolar couplings when the dipolar coupling constraints are initially included in the refinement.

An important property of MD calculations is that global structural changes only result from accumulation of local structural changes. The small time step required for MD calculations means there will be only very small distance changes for any atom or group of atoms at each step.[45] Thus, if the initial structure does not satisfy a particular $^{13}C-^{1}H$ dipolar coupling constraint, the calculations will try to rearrange the local structure to give an orientation of that C−H bond vector that is consistent with the input dipolar coupling constraints (and the NOE and torsion angle constraints). Because the dipolar couplings are a very sensitive function of angle,[7] a small change in position of the C−H bond vector can lead to a large change in angle. For example, if the orientation of a C−H bond vector on the terminal base pair of the 14mer needs to change by 20°, this can be achieved by moving the local position of this H atom by only 0.35 Å, and could be achieved with only a few steps in the MD calculation. However, if the angle change was instead achieved by inducing a 20° kink in the center of the

(45) McCammon, J. A.; Harvey, S. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987.

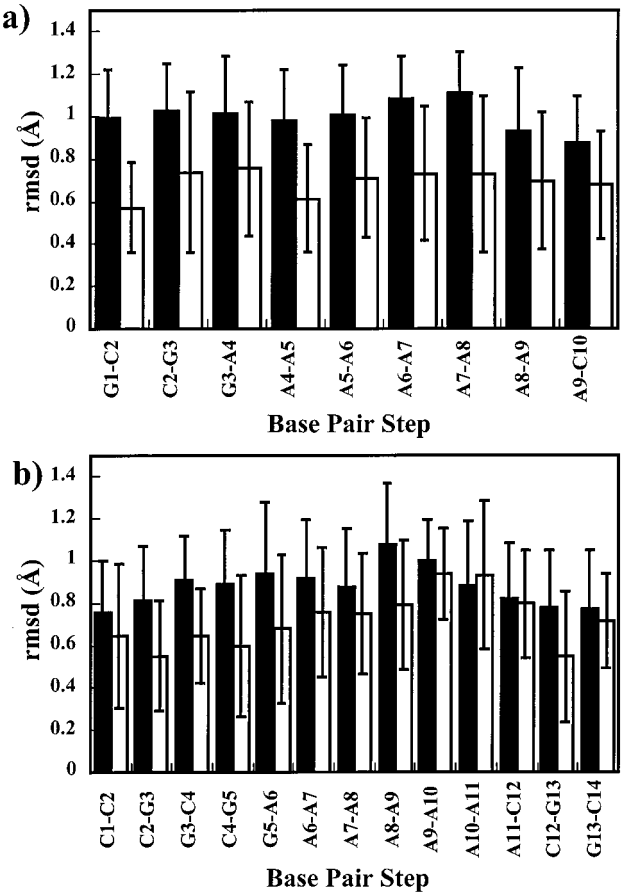

Figure 4. Average rmsds for superposition of each neighboring base-pair step is shown for (a) the 10mer and (b) the 14mer structures. The black bars are for the refinement with NOEs and sugar pucker constraints, and the open bars are for the refinement that also included the dipolar coupling constraints. Error bars are the standard deviations from the average.

helix, where the local structure of the terminal base pair is unchanged, this would lead to an ~8 Å change in the position of the H atom (along with a large change in the position of the rest of the helix). Such a large concerted conformational change would require a large number of MD steps and, therefore, will only happen if it is impossible to achieve the correct orientation by changes in local structure (without violating NOE or torsion angle constraints). An implication of this is that current molecular dynamics protocols are highly biased to inducing local structural modifications. This needs to be considered when interpreting structures generated by restrained molecular dynamics with residual dipolar couplings.

This biasing toward changes in local structures has also been observed in refinements of a protein−DNA complex with residual dipolar coupling by Wright and co-workers.[46] Their calculations included additional nonexperimental short internuclear distance constraints having high weighting factors to help fix the local structure of the molecule at the beginning of the refinement with the dipolar couplings. In the refinement protocol employed here, we started with a very small force constant for the dipolar coupling constraints (relative to the NOE or torsion angle constraints) similar to the procedure used in protein refinements.[37] By combining this small force constant with a large number of steps of low temperature molecular dynamics, the local structure established by the conventional constraints could be maintained while also allowing movement of the entire helix. This principle is especially important for longer DNA helices. As seen in Figure 3a, there is a significant curvature in the helical axis for many of the 14mer structures generated without the dipolar coupling constraints. Consequently, there were high violations of the dipolar coupling constraints in these starting structures and longer MD periods were required to refine the 14mer structures when including dipolar coupling constraints. Using a small force constant in the early stages of the dipolar coupling refinement allows the molecule to change the C−H bond orientation by gradually modifying the curvature of the helix axis instead of inducing incorrect stacking or incorrect local orientations for the base pairs. This type of incorrect stacking of base pairs was observed when we initially started the structure calculations with a force constant that was too high for the dipolar couplings (data not shown).

Another issue that needs to be considered when refining with dipolar couplings is that the quality of the structures defined by the conventional constraints can have a significant effect on how the dipolar coupling constraints are satisfied during the refinement. If local structure is well-defined by NOEs and torsion angles, the only way that dipolar coupling constraints can be accommodated is by changing global structure. In cases in which the local structure is not well-defined by distance and torsion angle constraints, the dipolar coupling constraints are easily satisfied by adjusting local structure without any changes in the global conformation. As discussed above, MD calculations bring about global structural changes through a large number of small local structural changes. This will not necessarily be true of other methods of structure refinement, such as minimization in torsion angle space,[47,48] in which it is possible to induce large changes in the global structure by changing only a few torsion angles while keeping most of the other torsion angles fixed. It will be interesting to compare the efficiency of these various methods for refining solution structures of macromolecules when residual dipolar couplings are included as constraints. However, the results here demonstrate that by carefully ramping up the force constant for the dipolar couplings, standard MD simulations can be efficiently used to generate DNA structures from experimental NMR data.

Do Variations in the Alignment Tensor Affect the Results of Refinements with Dipolar Couplings? As shown for several protein systems, it is sometimes possible to obtain good estimates for the magnitudes of the alignment tensor without a structural model if there is a relatively random orientation of the internuclear vectors of the measured dipolar couplings.[23] However, the regular structure of DNA and RNA helices means that there will generally not be a random angular distribution of H−C and/or H−N bond vectors in the molecule. For example, in an A- or B-form helix, there are many H−C and H−N bonds oriented perpendicular to the helix axis but relatively few bonds oriented parallel to the helix axis. Thus, the axial and rhombic components of the alignment tensor must be determined as part of the structure refinement.

For the simulations here, we incorporated a grid search algorithm into the calculations to test the possibilities of finding the correct D_a and R (see Methods).[37] Figure 5 shows the energies of the structures produced in a set of MD refinements

(46) Tsui, V.; Zhu, L.; Huang, T. H.; Wright, P. E.; Case, D. A. *J. Biomol. NMR* **2000**, *16*, 9−21.

(47) Guntert, P.; Mumenthaler, C.; Wuthrich, K. *J. Mol. Biol.* **1997**, *273*, 283−298.

(48) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr., Sect. D* **1998**, *54*, 905−921.

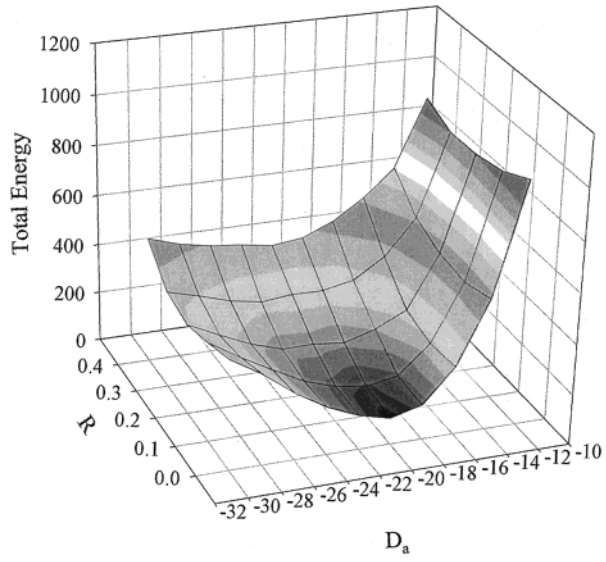

Figure 5. Average final energies as a function of D_a and R from the grid search procedure used on the 10mer. The lowest energies correspond to refinements with D_a and R around −20 Hz and 0, respectively, which are the actual values used in calculating the dipolar coupling constants (see text). These refinements were performed using the X-PLOR repel function instead of the Lennard−Jones potential, which is why the absolute energies are different from those in the final structures (Table 2), which employed the Lennard−Jones potential.

in which D_a was varied from −30 to −12 Hz, and R, from 0 to 0.4. The lowest energy was found at $D_a = -20$ Hz and $R = 0$, which were the values used for generating the dipolar coupling constraints in these simulations. To further illustrate the validity of this approach in determining the proper D_a and R values to be used in the dipolar coupling refinement, we carried out several extended simulations with incorrect D_a and/or R values. The results from these simulations are given in Tables S6 and S7. The simulations were carried out for the 10mer and with dipolar couplings data generated with $D_a = -20$ Hz and $R = 0$. It is clear that the use of incorrect D_a and R values significantly increased the number of violations in dipolar coupling constraints and the overall energy. For instance, without backbone torsion constraints, the number of violations larger than 3 Hz from the correct dipolar couplings increased to an average of 12 if $D_a = -25$ Hz and $R = 0.0$ were used and to an average of 10 if $D_a = -20$ Hz and $R = 0.2$ were used, as compared to an average of 0.5 when the correct values were used. With the backbone torsion constraints, an even larger number of severe violations for the dipolar constraints were seen. It is interesting to note that although the structures refined with incorrect D_a and R had higher energies than with the correct values, they still had improved precision when compared to the control refinement (Table S7). For actual experimental data, the results of the grid-search method in determining D_a and R will depend on the accuracy and precision of the measured dipolar couplings, whereas in these simulations, all of the input constraints were perfectly accurate (but not perfectly precise).

Another important question of dipolar coupling refinement is whether the orientation of the alignment tensor significantly influences the quality of the structures generated. We, therefore, performed identical refinements on the 10mer and 14mer for which, instead of assuming an axially symmetric alignment tensor with the alignment axis parallel to the helix axis, the axially symmetric alignment tensor was oriented perpendicular to the helix axis (see Figure 1c). Although it would be unlikely to have this orientation for an alignment tensor in an isolated double helix, it is possible to have such an orientation when a DNA or RNA helix is bound in a larger protein-nucleic complex.

As seen in Table 3 for the 10mer, small increases were seen in the rmsds from the target and average structures with the perpendicular alignment compared to the parallel alignment. Larger increases were observed for the 14mer with perpendicular alignment. This means that the structures are not so precisely defined with the perpendicular alignment. This difference indicates that data from the perpendicular alignment are not so effective as those from parallel alignment in defining the global structure of nucleic acids. One explanation is that in an axially symmetric system ($R = 0$) in which the helix axis is parallel to the principal axis of the alignment tensor, rotations about the helix axis have no effect on the alignment angle of individual bond vectors. However, for the same system in which the principal axis is perpendicular to the helix axis, rotations about the DNA helix axis will affect the alignment angle of individual bond vectors. Thus, there are more orientations of a vector which can give the correct dipolar coupling. Overall, this leads to a small, but significant, increase in the rmsds for structures generated with the principal axis of the alignment tensor perpendicular to the helix axis, as compared to the parallel case (Table 3).

To test the effect of nonaxial symmetry on refinement with dipolar couplings, an identical set of structure calculations were performed for the 10mer for which, instead of having an axially symmetric alignment tensor ($R = 0$), a rhombic term ($R = 0.2$) was introduced to determine if there is a difference in the quality of the generated structures. The average rmsds for the 10mer with dipolar couplings were very similar for these two different alignment tensors, 1.29 versus 1.16 Å for refinements with $R = 0$ and 0.2, respectively, as seen in Table 3. This indicates that for the set of dipolar couplings used here, whether the system is axially symmetric or has a significant rhombic component does not affect the precision or accuracy of the refined structures.

Local Kinks and Bending in DNA Duplexes. A-tract sequences have been shown to bend DNAs in solution, but the molecular mechanism for the bending is the subject of much debate.[25,27−30] Unfortunately, X-ray structures cannot be directly used to address this problem because it is not clear to what extent the crystal packing forces may be perturbing the global (and local) structure of short DNA duplexes in high-resolution X-ray structures.[25,26] The simulations here demonstrate that both the local and the global structure of the DNA duplex are much better defined by inclusion of dipolar coupling constraints. Thus, one interesting question is whether the dipolar coupling constraints will make it possible to define DNA bending in solution. The A-tract sequence used as the target structure shows a bend in the X-ray structure, but the regularization procedure employed here (see Methods) leads to changes in the local and global structure, which means the kinks and bends in the target structure no longer accurately reproduce those observed in the X-ray structure. However, it is still possible to analyze the curvature of the structures produced here to see how accurately and precisely they reproduce the bends in the target structure.

The CURVES program[32] was used to compare local kinks and the global bend in the calculated and target structures for the 10mer and 14mer. To calculate these parameters, CURVES first determines an average linear DNA axis, and then a separate axis is drawn for each base-pair step in the molecule. Finally, the angle between the axis for each base-pair step and the average linear DNA axis is calculated, which CURVES calls the "global axis curvature angle". Figure 6 shows the curvature

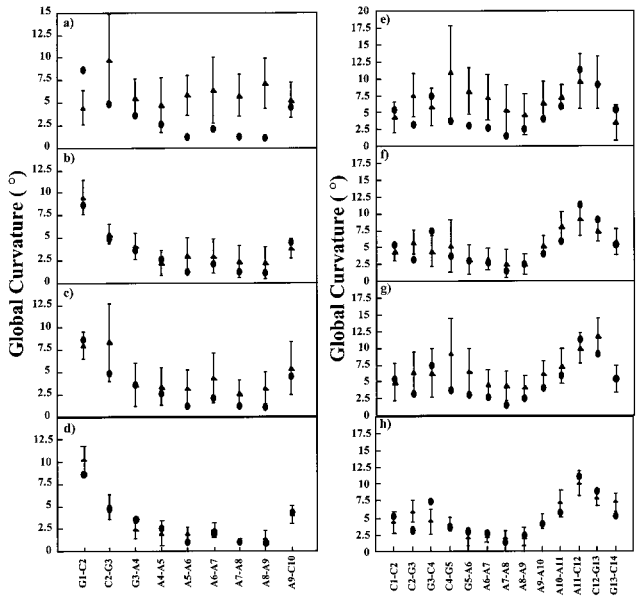

Figure 6. Global axis curvature is shown for each base-pair step in the DNA structures of (a) the 10mer with only NOE and sugar pucker constraints, (b) the 10mer with NOE, sugar pucker, and dipolar coupling constraints, (c) the 10mer with NOE, sugar pucker, and backbone torsion angle constraints, (d) the 10mer with NOE, sugar pucker, backbone torsion angle, and dipolar coupling constraints, (e) the 14mer with only NOE, and sugar pucker constraints, (f) the 14mer with NOE, sugar pucker, and dipolar coupling constraints, (g) the 14mer with NOE, sugar pucker, and backbone torsion angle constraints, and (h) the 14mer with NOE, sugar pucker, backbone torsion angle, and dipolar coupling constraints. These structures were calculated with the axially symmetric tensor with the principal axis aligned parallel to the helix axis. Circles are the angles in the target structure, triangles are the average angles for the 15 structures generated in this refinement, and error bars are the standard deviations from the average.

angle data for each base-pair step in the 10mer and 14mer target structures. The curvatures for most base-pair steps are between 0 and 5°, but there are sections in both molecules in which there are larger angles, indicating a kink, particularly at one end of the 14mer structure.

For the refinement with distance and sugar pucker torsion angles, the curvature angle at each base-pair step has a relatively large range with standard deviations of up to ±5° for the 10mer and ±7° for the 14mer (Figure 6a,e), which means the global axis curvature is not precisely or accurately defined by these constraints. The addition of backbone torsion angle constraints improves the precision of global axis curvature for these base-pair steps, with the largest standard deviation being approximately ±4 and ±5° for the 10mer and 14mer structures, respectively (Figure 6c,g). However, the average curvature values for a particular base-pair step are often not very close to the target structure curvature values.

Both the accuracy and the precision of these global axis curvature values are much better defined by the addition of dipolar coupling constraints for the 10mer and 14mer (Figure 6b,d,f,h). For almost every base-pair step the average global axis curvatures in the 10mer and 14mer simulations with dipolar couplings are very close to the values in the target structure. Exceptions are steps C2−G3 and G3−C4 in the 14mer structures, refined with backbone torsion angle and dipolar coupling constraints (Figure 6h) for which the simulation did not give accurate values for the global axis curvature. We do not understand the source of the deviation at this point, but overall it appears that the dipolar coupling constraints used here make it possible to define kinks for the individual base-pair steps in a DNA duplex.

The results of the simulations also clearly demonstrate that the global bending of the DNA cannot be accurately determined from traditional NMR constraints. For the 10mer, the target structure has an overall bend from CURVES of 4°. However, the simulations with the NOE and all the torsion angle constraints give an average bend of 15°, with a very large standard deviation of ±9° (Figure 2a and Table S5 in Supporting Information). There is dramatic improvement by the addition of dipolar couplings for which the average overall bend is 6 ± 3°. The 10mer and 14mer target structures have different degrees of bending, with the 14mer having a larger overall bend of 25° (Table S5). The simulations with NOE and backbone torsion angle, but no dipolar coupling constraints, give bends of 45 ± 14° (Figure 3a and Table S5 in Supporting Information). Addition of dipolar coupling constraints again leads to a much more accurate and precise definition of the overall bending of the DNA, with an average bend of 21 ± 5°.

The present simulations demonstrate that both the local kinks and overall helical bends in target DNA duplexes are accurately and precisely reproduced using solution NMR-type constraints that incorporate $^{1}H-^{13}C$ dipolar couplings. Thus, the inclusion of the orientational information available by measurement of dipolar coupling data on DNA in liquid crystalline-type media should prove extremely valuable in elucidating the molecular mechanisms for the bending of A-tract sequences in DNA duplexes. However, it should be noted that the simulations here were for a rigid structure. In solution, the DNA helix will be somewhat flexible and, thus, the average bend will result from an ensemble of interconverting structures with various bend angles. The effects of this type of global dynamics on the NMR structure determinations of DNA will need to be considered when including dipolar coupling constraints in refinements of DNA helices.

Conclusions

Dipolar coupling restraints are now being routinely used to improve the NMR solution structure determination of proteins. However, these methods have not yet been widely applied to nucleic acids. Thus, we performed a series of simulations on different length DNA duplexes using distance, torsion angle, and dipolar coupling constraints that mimic experimental NMR data. The molecular dynamics calculations show that inclusion of dipolar coupling constraints dramatically improves both the local and global structure of the DNA duplexes. The addition of dipolar coupling constraints in a 10-base-pair duplex reduced the rmsd for the final structures from 2.4 to 1.3 Å for structures that included NOEs and sugar pucker constraints, and the improvement in rmsd was from 1.7 to 0.8 Å for structures that also included β, γ, and ϵ backbone torsion angles constraints. There was an even larger improvement in the average rmsd to the target structure for the 14-base-pair duplex by including dipolar couplings, from 3.3 to 1.6 Å for NOEs and sugar pucker constraints, and from 2.7 to 1.2 Å when also including backbone torsion angle constraints. The local structure was also significantly improved, as judged by the rmsds for superposition of neighboring base pairs and by analysis of various helical parameters in the DNA structures.

The dipolar couplings give information on the orientations of internuclear vectors relative to the alignment axes in the molecule. In these simulations, similar improvements in the precision and accuracy of the final structures were obtained for both axially symmetric and nonaxially symmetric alignment

tensors. The simulations also illustrate an important feature of the molecular dynamics calculations: that global structural changes occur as a result of many small steps of local structural adjustments. Therefore, it is more likely that dipolar coupling will be accommodated by local structural changes before global structural changes. The 10mer and 14mer target structures had very different overall helical bends, and the simulations here demonstrate that without dipolar couplings, solution NMR-type data cannot define helical bends in DNA. However, including $^{1}H-^{13}C$ dipolar coupling constraints in the calculations makes it possible to accurately and precisely reproduce the overall helical bends in the DNA target structures. Thus, these studies demonstrate that incorporating a relatively easily measured set of $^{1}H-^{13}C$ residual dipolar coupling constraints in nucleic acids leads to dramatic improvement in the precision and accuracy of both the local and global structure of the helical regions. Such dipolar coupling data should help address important questions such as the relative angle between helical regions in various RNAs and the mechanisms for bending in A-tract sequences in DNA duplexes.

After this paper was submitted, Tjandra, Bax, and co-workers published a refinement of a DNA dodecamer that included experimental residual dipolar couplings.[49] Their study employed specific labeling which allowed for a very complete set of $^{1}H-^{13}C$ dipolar couplings and included $^{1}H-^{15}N$ and $^{1}H-^{1}H$ dipolar couplings as well as "artificial" constraints for the α and ζ backbone torsion angles. Thus, their refinement of this DNA duplex included more dipolar coupling and torsion angle constraints, but somewhat fewer NOE-distance constraints, than were employed in the simulations here. Overall, their refinement with experimental data gave similar results to what is observed here in the model simulations. Although the accuracy of their experimental structures cannot be determined, the rmsd for the center 10 base pairs of their dodecamer to the average structure (0.33 Å) is very similar to the rmsd of the average obtained for the 10mer in this work (0.41 Å). Thus, both the simulation and experimental studies demonstrate that refinements with residual dipolar couplings substantially improve the solution structure determinations of nucleic acids.

Acknowledgment. This work was supported by NIH grant AI33098 to A.P. We also thank Dr. Fiona Jucker for helpful discussions, Dr. R. A. Byrd for use of computer resources at NCI-Frederick, and Professor M. D. Tsai for providing computer resources and lab space to A.V.

Supporting Information Available: Tables containing the $^{1}H-^{13}C$ dipolar coupling constraints, average helical parameters, average helical curvature for the 10mer and 14mer DNA duplexes, and the analyses of the refinements of the 10mer with incorrect values of D_a and R (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA001919L

(49) Tjandra, N.; Tate, S.; Ono, A.; Kainosho, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 6190−6200.

(50) Rife, J. P.; Stallings, S. C.; Correll, C. C.; Dallas, A.; Steitz, T. A.; Moore, P. B. *Biophys. J.* **1999**, *76*, 65−75.