Journal of Biomolecular NMR, **26:** 297–315, 2003.
KLUWER/ESCOM

Overall structure and sugar dynamics of a DNA dodecamer from homo- and heteronuclear dipolar couplings and ^{31}P chemical shift anisotropy

Zhengrong Wu[a], Frank Delaglio[a], Nico Tjandra[b], Victor B. Zhurkin[c] & Ad Bax[a]
[a]*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases,* [b]*Laboratory of Biophysical Chemistry, National Heart, Lung and Blood Institute, and* [c]*Laboratory of Experimental and Computational Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, U.S.A.*

Received 20 November 2002; Accepted 27 March 2003

Key words: chemical shift anisotropy, Dickerson dodecamer, dipolar coupling, liquid crystal, NMR, sugar pucker

Abstract

The solution structure of d(CGCGAATTCGCG)$_2$ has been determined on the basis of an exceptionally large set of residual dipolar couplings. In addition to the heteronuclear ^{13}C-^{1}H and ^{15}N-^{1}H and qualitative homonuclear ^{1}H-^{1}H dipolar couplings, previously measured in bicelle medium, more than 300 quantitative ^{1}H-^{1}H and 22 ^{31}P-^{1}H dipolar restraints were obtained in liquid crystalline Pf1 medium, and 22 ^{31}P chemical shift anisotropy restraints. High quality DNA structures can be obtained solely on the basis of these new restraints, and these structures are in close agreement with those calculated previously on the basis of ^{13}C-^{1}H and ^{15}N-^{1}H dipolar couplings. In the newly calculated structures, ^{31}P-^{1}H dipolar and $^{3}J_{H3'P}$ couplings and ^{31}P CSA data restrain the phosphodiester backbone torsion angles. The final structure represents a quite regular B-form helix with a modest bending of $\sim$10°, which is essentially independent of whether or not electrostatic terms are used in the calculation. Combined, the number of homo- and heteronuclear dipolar couplings significantly exceeds the number of degrees of freedom in the system. Results indicate that the dipolar coupling data cannot be fit by a single structure, but are compatible with the presence of rapid equilibria between C2′-endo and C3′-endo deoxyribose puckers (sugar switching). The C2′-H2′/H2″ dipolar couplings in B-form DNA are particularly sensitive to sugar pucker and yield the largest discrepancies when fit to a single structure. To resolve these discrepancies, we suggest a simplified dipolar coupling analysis that yields N/S equilibria for the ribose sugar puckers, which are in good agreement with previous analyses of NMR J_{HH} couplings, with a population of the minor C3′-endo form higher for pyrimidines than for purines.

Introduction

Although for nucleic acids the secondary structure is usually clear from their primary sequence, the paucity of long-range NOE restraints limits the accuracy of their structure determination by NMR (Metzler et al., 1990; Ulyanov et al., 1992; Varani et al., 1996; Allain and Varani, 1997). Other experimental parameters, such as chemical shifts and scalar couplings, only depend on local geometry and their conversion to structural restraints is limited both by measurement error and by intrinsic uncertainties in their empirical parametrizations. Small errors in local geometry, derived from these parameters, tend to accumulate when evaluating more global properties such as, for example, DNA bending. More recently, measurement of residual dipolar couplings (RDCs) in weakly aligned macromolecules has been introduced as a remedy to solve this inherent weakness in NMR structure determination. RDCs restrain the time-averaged orientation of a given internuclear vector relative to the magnetic field. This magnetic field direction therefore serves as a common reference axis, relative to which all internuclear vector orientations are restrained. RDCs have proven particularly useful for defining relative orientations in protein-protein complexes (Clore, 2000), or relative domain orientations in multi-domain proteins (Fischer et al., 1999; Bewley and Clore, 2000) (Skrynnikov et al., 2000; Mollova et al., 2000; Goto et al., 2001).

Weak alignment of the molecule relative to the magnetic field is a prerequisite for measurement of RDCs. It can result either from its own intrinsic magnetic susceptibility anisotropy (Tolman et al., 1995; Kung et al., 1995; Tjandra et al., 1996) or from the use of a dilute liquid crystalline medium (Tjandra and Bax, 1997; Hansen et al., 1998; Clore et al., 1998). Numerous different liquid crystals suitable for macromolecular alignment have been proposed to date (Tjandra and Bax, 1997; Hansen et al., 1998; Clore et al., 1998; Ruckert and Otting, 2000) and, with the exception of the positively charged cetyl pyridinium based liquid crystals (Prosser et al., 1998; Barrientos et al., 2000), all appear suitable for the study of nucleic acids. Despite the obvious potential of dipolar couplings in nucleic acid structure determination, it has taken several years before their use started taking root (Bayer et al., 1999; Mollova et al., 2000; Trantirek et al., 2000; Warren and Moore, 2001). Recent studies using experimental and modeled dipolar couplings confirmed their potential to study subtle features such as bending of the helical axis in relatively short oligomers (Tjandra et al., 2000b,c; Vermeulen et al., 2000; MacDonald et al., 2001; Murphy et al., 2001).

Because dipolar couplings can increase the accuracy of both the local and global features of oligonucleotide structures determined by NMR, there has been renewed interest in capturing structural details such as the average values and cross-correlations between base pair roll, tilt, and twist. These details form the basis for DNA deformations, such as bending, which in turn are critical for recognition, packaging and regulation of DNA transcription (Olson and Zhurkin, 1995; Crothers, 1998). In the crystalline state, these correlations can be masked by DNA deformations induced by packing forces (Dickerson et al., 1987; DiGabriele et al., 1989), and access to the precise values of these parameters in solution is therefore of fundamental interest.

Much of our current understanding of the details underlying DNA structure and its sequence-dependent variability derives from experimental and modeling studies of the so-called Dickerson dodecamer, d(CGCGAATTCGCG)$_2$. Its atomic resolution structure was first determined by Dickerson et al. (Wing et al., 1980; Drew et al., 1981) using single crystal X-ray diffraction and more recently, at higher resolution, by Williams and co-workers (Shui et al., 1998a,b). These structures have provided invaluable insight in the relations between DNA sequence and structure (Dickerson and Drew, 1981; Dickerson, 1983), backbone flexibility and bending (Fratini et al., 1982), and the effect of hydration and counterion condensation (Drew and Dickerson, 1981; Shui et al., 1998a,b; Young and Beveridge, 1998; Chiu et al., 1999; Tereshko et al., 1999a,b; Sines et al., 2000). Remarkably, all these independent X-ray studies show distinct asymmetric features for this palindromic DNA, including some unusual backbone torsion angles and sugar puckers (Dickerson and Drew, 1981; Shui et al., 1998a). The presence of some highly ordered hydration water molecules and localized counterions was first observed by Dickerson et al. (Drew et al., 1981; Drew and Dickerson, 1981), and their role recently has been discussed by Williams and co-workers (Shui et al., 1998a,b; Sines et al., 2000) and by Tereshko et al. (1999a,b). Although the presence of monovalent cations in the tightly coordinated hydration shell is still subject to debate (Chiu et al., 1999; Tereshko et al., 1999b; Woods et al., 2000), it appears clear that hydration is closely coupled to the commonly observed narrowing of the minor groove in A-tracts. These observations have led to the current consensus that DNA conformation is influenced intrinsically by its sequence and externally by the environment in which it is studied.

The same dodecamer has also been studied extensively by molecular dynamics calculations (Young et al., 1997; McConnell and Beveridge, 2000) and was also the subject of many early NMR studies (Hare et al., 1983; Nerdal et al., 1989; Lane et al., 1991). However, due to the above mentioned intrinsic problems in determining global features from local NMR restraints, these NMR studies were ill suited for studying the oligomer's overall structural parameters, such as the bend angle. In addition to these short-range restraints, a recent study also included a nearly complete set of one bond ^{1}H-^{13}C and ^{1}H-^{15}N dipolar couplings (Tjandra et al., 2000c). The resulting structure is highly regular and exhibits only minimal helical bending. None of the unusual features present in the X-ray structures are seen in this NMR structure. However, considering that there are seven variable torsion angles per nucleotide, the number of observed dipolar couplings in this study was not much larger than the number of degrees of freedom. Together with the non-uniform distribution of the dipolar interactions, this made it impossible to cross validate the resulting structure in the regular rigorous manner (Tjandra et al., 2000c). Furthermore, as very few of the restraints have any direct relation to the torsion angles surrounding

the backbone phosphate groups, these values remained rather ill determined.

After completion of the dipolar coupling NMR study of the Dickerson dodecamer, a number of new methods have been developed that provide access to additional parameters. These include methods for measuring ^{1}H-^{1}H (Tian et al., 1999, 2000; Wu and Bax, 2001; Delaglio et al., 2001) and ^{1}H-^{31}P dipolar couplings (Wu et al., 2001a; Henning et al., 2001), as well as the orientation of the phosphate group from the effect of ^{31}P chemical shift anisotropy (CSA) (Wu et al., 2001b). ^{1}H-^{1}H and ^{1}H-^{31}P dipolar couplings provide information on both the internuclear distance and orientation. ^{31}P CSA and ^{1}H-^{31}P dipolar coupling also provide a handle on the conformation of the phosphodiester linkage. The present study combines all these parameters, such that its structure becomes overdetermined. However, we find that for many of the nucleotides, in particular cytosines, the measured dipolar couplings are incompatible with a single static structure. Locally, this fit can be improved by assuming a two-state model, where the sugar is in a dynamic equilibrium between the N- and S-state puckers.

Materials and methods

Sample preparation

The unlabeled DNA dodecamer, d(CGCGAATTCGCG)$_2$ (Midland, Texas), was dialyzed against sterilized H_2O using a 2500 MW cut-off dialysis cassette and subsequently lyophilized to dryness. The resulting DNA powder was then dissolved in D_2O containing 10 mM sodium phosphate (pH = 6.8), 50 mM KCl, and 6 mM NaN_3. Two different NMR samples were used for measuring ^{1}H-^{1}H coupling constants at the DNA concentration of approximately 2 mM duplex, one isotropic sample and the other one containing an additional 20 mg ml^{-1} Pf1 phage (ASLA Labs, http://130.237.129.141//asla/asla-phage.htm) (Hansen et al., 1998), which was first exchanged into the same buffer in D_2O using a 30 kD cut-off centricon concentrator. For the measurement of ^{31}P-^{1}H couplings and ^{31}P-CSA, two different alignment media were employed, one with 20 mg ml^{-1} Pf1 phage and the other with 50 mg ml^{-1} phospholipids in a 3.1:1 molar ratio of dimyristoyl phosphatidylcholine (DMPC) and dihexanoyl phosphatidylcholine (DHPC). Along with each liquid crystalline sample, an isotropic sample was also prepared with the same buffer condition. Experiments were performed on Bruker DRX-800 and DMX-600 spectrometers equipped with a triple resonance, three-axis pulsed field gradient probehead and a broadband inverse probehead. Data were processed using the NMRPipe software package (Delaglio et al., 1995).

Measurement of ^{1}H-^{1}H couplings

Proton-proton couplings of the DNA dodecamer were derived from 800 MHz 2D phase-sensitive COSY spectra of the isotropic sample and the liquid crystalline (Pf1) sample by amplitude-constrained multiplet evaluation (ACME), a newly developed analysis program that obtains quantitative J-coupling information from a direct analysis of the intensity and shape of cross peak multiplets (Delaglio et al., 2001). The 2D COSY data set was processed twice, once in such a way that the cross peaks are phased absorptive and the dispersive diagonal is suppressed using a routine described by Delaglio et al. (2001), and the second time with only the absorptive diagonal while cross peaks are being suppressed. An important value, referred to as the intrinsic intensity, was first extracted from several isolated diagonal peaks by fitting them with zero active coupling in the diagonal-only subspectrum. After this intrinsic intensity has been established it is kept fixed for the cross peak analysis, where each individual peak is best fitted with variable active coupling, line width, peak position, and passive couplings. Dipolar couplings are then derived from the difference between the active couplings measured for the isotropic and the aligned sample.

The ACME fitting method provides only the absolute value for $J_{HH} + D_{HH}$. However, it is also desirable that the sign of the dipolar couplings be determined prior to structure calculation. If the dipolar coupling superimposes on a relatively large $^{2}J_{HH}$ or $^{3}J_{HH}$ coupling, data obtained for a more dilute phage sample ($\sim$10 mg ml^{-1}) were used for determining the sign of the dipolar couplings, using the fact that dipolar couplings scale linearly with respect to Pf1 concentration. If no significant J coupling is present, signed values for D_{HH} usually can be determined on the basis of preliminary structures, calculated without inclusion of these couplings (Tjandra et al., 2000c). For couplings whose sign could not be determined, structure calculations optimized agreement between the absolute values of the measured and calculated couplings (Tjandra et al., 2000a).

^{1}H-^{31}P dipolar couplings

2D ^{31}P-^{1}H CT-NOESY difference experiments were collected at 600 MHz ^{1}H frequency with a NOE mixing time of 300 ms for both the isotropic and the two aligned samples (Wu et al., 2001a). Three-bond ^{1}H3′-^{31}P J couplings were derived from the ratio between cross peak intensities in constant-time NOESY spectra recorded for the isotropic sample with and without $^{3}J_{HP}$ evolution during the constant time evolution period. With a Karplus parameterization available for this interaction (Lankhorst et al., 1984), these $^{3}J_{HP}$ values couplings were used as direct coupling restraints in the structure calculations. The dipolar couplings, $D_{H3'P}$, were derived from the difference between the couplings measured in the aligned and isotropic samples. These $D_{H3'P}$ measurements were carried out both in Pf1 phage and in liquid crystalline bicelle media.

^{31}P CSA measurement

^{31}P CSA-induced chemical shift differences, $\Delta\delta(^{31}P)$, between isotropic and aligned samples were measured using selective 2D ^{1}H-^{31}P HSQC correlation spectra (Wu et al., 2001b), recorded at 600 MHz using a Bruker broadband inverse probehead. In order to minimize the broadening of ^{31}P signal by trace paramagnetics, 1 mM EDTA was added to all four samples used for measuring the ^{31}P CSA effect. For the Pf1 containing sample, CSA restraints were derived from the difference in ^{31}P chemical shift between the isotropic and phage sample (Wu et al., 2001b). The CSA restraint in bicelle medium was obtained from the chemical shift difference above (35 °C) and below (25 °C) the temperature threshold where the bicelle sample switches from an isotropic to a nematic liquid crystalline phase (Cornilescu and Bax, 2000). The contribution of temperature to the ^{31}P chemical shift was derived using an isotropic sample with the same buffer composition, and subtracted from the above chemical shift difference of the bicelle sample. Conformational exchange causes increased ^{31}P transverse relaxation at temperatures below 25 °C, and in contrast to an analogous CSA study in a protein, the temperature dependence of the chemical shift could not be measured accurately at temperatures much below 25 °C (Cornilescu and Bax, 2000). Owing to the collinearity between the alignment tensors in Pf1 and bicelle media, the $\Delta\delta(^{31}P)$ values obtained from the two measurements were highly correlated (Pearson's correlation coefficient, $R_P = 0.98$).

Structure calculations

Structure calculation was carried out with a total of 162 NOEs, 22 $^{3}J_{H3'P}$ couplings, 452 orientational restraints from ∼5% DMPC/DHPC bicelle medium (including 198 D_{CH} couplings, 10 D_{NH} values, 200 approximate D_{HH} values, 22 $D_{H3'P}$ couplings and 22 ^{31}P CSA values) and 390 restraints from Pf1 phage medium (including 346 D_{HH} quantitative values, 22 $D_{H3'P}$ couplings (Wu et al., 2001a) and 22 ^{31}P CSA values (Wu et al., 2001b)), using restrained molecular dynamics calculations with two different alignment tensors, using XPLOR version 3.84 (Brunger, 1993) supplemented with home-written routines for inclusion of dipolar coupling and CSA restraints (Schwieters et al., 2003). Except for the backbone angle ε, all torsion angles were left floating. The ε angles are restrained by previously reported $^{3}J_{H3'P}$ values (Sklenar and Bax, 1987), with a loose margin of $\pm 20^{o}$. These loose restraints are only for better convergence and have zero violations in the final set of refined structures. In contrast to our earlier study of this dodecamer (Tjandra et al., 2000c), no 'convergence restraints' were used for any of the remaining torsion angles, and a high rate of convergence was obtained when D_{HH}, $D_{H3'P}$ and CSA restraints were included.

The restrained molecular dynamics (MD) protocol followed in the present study is similar to that described earlier (Tjandra et al., 2000c). Calculations were started from random initial structures, generated by 15 ps restrained MD at 1000 K with only energy terms for bonds, angles and improper torsions left on, followed by MD with the inclusion of only NOE and H-bond restraints in order to obtain roughly folded structures. These were then further minimized with NOEs, ε-torsion angle, and J coupling restraints. Subsequently, all $^{1}D_{CH}$, $^{1}D_{NH}$, and D_{HH} dipolar couplings were included and slowly ramped up to their final values through a 15 ps MD run. The final set of structures was calculated after adding $^{3}D_{H3'P}$ and ^{31}P CSA restraints, which were also slowly ramped up during 50-cycle steps, followed by another 25 ps of restrained MD at a constant temperature of 300 K. The representative final structures were generated from the minimized average of the last 10-ps trajectory of this 25-ps period. The force constants for $J_{H3'P}$, $D_{H3'P}$ and CSA were adjusted empirically to the values of 10 kcal Hz^{-2}, 20 kcal Hz^{-2} and 0.02 kcal ppb^{-2}, such that the agreements between the observed and predicted values are comparable to the uncertainties

of each measurement. In a similar manner, the force constants for NOE and D_{CH} were obtained to be 50 kcal Å^{-2}, 0.2 kcal Hz^{-2}. In cases of partial overlap, experimental errors in D_{CH} were larger, and a smaller force constant was used. When $^{3}D_{H3'P}$ and ^{31}P CSA restraints were introduced at the same time as all $^{1}D_{CH}$, $^{1}D_{NH}$, and D_{HH} dipolar couplings, few structures converged to the final low energy conformation, and these restraints were therefore introduced sequentially.

Two additional calculations were also carried out: one without any heteronuclear $^{1}D_{CH}$ and $^{1}D_{NH}$ dipolar coupling restraints, and the other one with the electrostatic term in XPLOR turned off. The first of these calculations is used to evaluate whether a high-resolution structure can be obtained without heteronuclear labeling and specific deuteration. The second evaluates the effect of the empirically parameterized electrostatic force on the structure determination process.

Results

Measurement of dipolar couplings and CSA

Figure 1A shows the 2D phase sensitive COSY spectrum of the DNA dodecamer in 20 mg ml^{-1} Pf1 phage liquid crystalline medium, with its diagonal being suppressed by digital means (Delaglio et al., 2001). Numerous cross peaks between pairs of protons that lack a J_{HH} interaction are visible in this spectrum, e.g. between H2′/H2″ and H8/H6, indicating the presence of significant dipolar couplings. An expansion of the boxed region in Figure 1A shows several of the H6/H8-H2′ cross peaks (Figure 1B) and Figure 1C represents the spectrum best-fitted by ACME (Delaglio et al., 2001). Although it is well known that very similar looking cross peaks can be generated using a wide range of different coupling values, the ACME program restrains the amplitude of the time domain signals to values derived from a few resolved diagonal peaks, and thereby permits an accurate determination of the active coupling constant. For each measured coupling, both cross peaks above and below the diagonal were fitted, yielding a pairwise root-mean-square difference (rmsd) of less than 0.8 Hz. Assuming the errors are random and uncorrelated, this indicates an error of less than 0.6 Hz in values derived from individual cross peaks, and only 0.4 Hz in cases where averaging over values derived from the two mirror-image cross peaks is used. Slightly better reproducibility is obtained for the isotropic sample due to the much simpler spin systems in the absence of dipolar interactions. Isotropic $^{3}J_{HH}$ values were found to be in excellent agreement with those reported previously using E.COSY methods (Bax and Lerner, 1988; Yang et al., 2000). Although weak alignment in the liquid crystalline phase introduces large numbers of unresolved passive couplings, the value of an active coupling measured by ACME is not much affected by passive couplings (Delaglio et al., 2001). Proton-proton dipolar couplings were derived from the difference between isotropic and aligned samples. The uncertainty of those dipolar couplings was set to be ±1 Hz for cases when cross peaks are measurable in both samples, ±2 Hz for cases when cross peaks are below the noise level in either the isotropic or aligned sample, and a value of 0 ± 4 Hz was used for three-bond couplings if cross peaks were below the noise level in both the isotropic and the aligned state. A total of 120 ^{1}H-^{1}H dipolar couplings was measured at high accuracy (±1 Hz), 46 couplings at medium accuracy (±2 Hz), and for 3 couplings only an upper limit (<4 Hz) could be established. Owing to the palindromic nature of the dodecamer, the number of resulting restraints equals twice the number of couplings.

As described in the experimental section, sequential ^{1}H3′(i)-^{31}P(i+1) scalar and dipolar couplings were obtained from the H3′ → H4′/H5′/H5″ and H2′/H2″ regions of the 2D selective CT-NOESY difference experiment, recorded with and without ^{31}P decoupling (Wu et al., 2001a). The homonuclear decoupling obtained in the F_1 dimension of the spectrum results in excellent resolution, even in the presence of substantial size dipolar interactions, and couplings for all 11 phosphates were measured, both in bicelles and in Pf1. With a Pearson's correlation coefficient R_P = 0.91, dipolar couplings in the two media are strongly correlated, which is expected on the basis of their similar alignment tensor orientation and rhombicities (see below).

^{31}P CSA measurement is less straightforward in bicelle medium than in Pf1 phage (Wu et al., 2001b). The ^{31}P chemical shift difference between 35 °C (aligned phase) and 24 °C (isotropic phase) represents the sum of the CSA contribution and the temperature dependence of the ^{31}P chemical shift. Previously, the temperature contribution was measured by taking the difference between 24 °C and 13 °C for the same sample, and assuming this temperature dependence remains constant when comparing chemical shifts at 24 and 35 °C (Cornilescu and Bax, 2000). However,

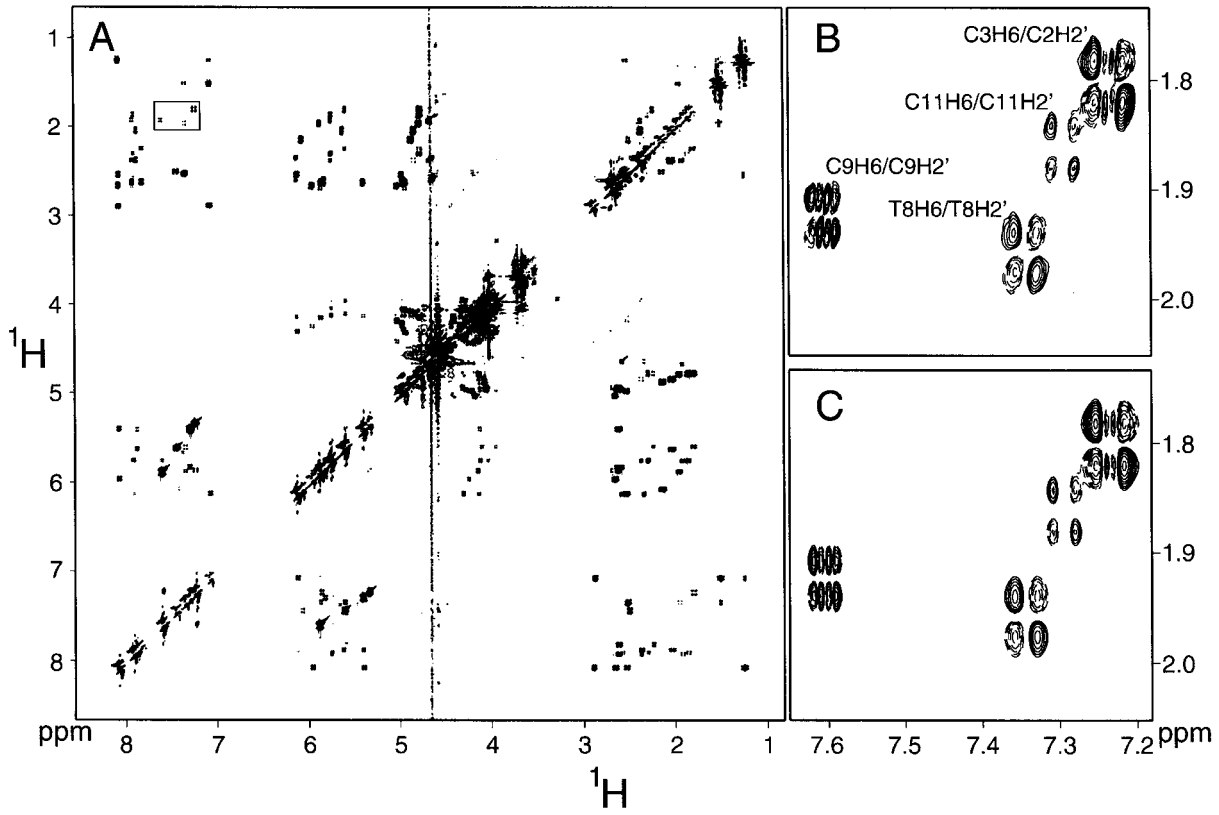

Figure 1. 800-MHz phase-sensitive COSY spectrum of the DNA dodecamer in 20 mg ml^{-1} Pf1 phage. (A) Full spectrum, with the dispersive diagonal suppressed by digital means (Delaglio et al., 2001). The spectrum was acquired as a 512* × 2048* data matrix with 9 kHz spectral width for both dimensions. (B) Expansion of the region boxed in (A). (C) ACME best-fitted spectrum of the multiplets in (B), restraining the amplitude in this fit to a value derived from the average over eight resolved diagonal resonances.

presumably as a result of intermediate rate conformational exchange processes, the line width of ^{31}P signals increases strongly at lower temperature, impeding accurate determination of the ^{31}P frequencies. Instead, the temperature dependence was measured separately for an isotropic sample with the same buffer composition. $\Delta\delta(^{31}P)$ values were measured for all 11 phosphates in the molecule. The good agreement between $\Delta\delta(^{31}P)$ values measured in bicelle and in Pf_1 media ($R_P = 0.98$) again testifies to the accuracy of the two measurements, but also indicates that the two sets of data are highly correlated.

Determination of alignment tensor

In our earlier study, it was shown that several of the final parameters defining the characteristics of the dodecamer, such as helical bending and minor groove width, are sensitive to the magnitude (D_a) and rhombicity (R) of the alignment tensor. Here, we determine these values for Pf1 medium by means of a grid search, starting from initial values derived by fitting the couplings measured in Pf1 data to the previously determined structure (PDB code 1DUF) (Tjandra et al., 2000c). The grid search aims to find the D_a and R values that yield structures with the lowest energies. Similarly, the D_a and R of the alignment tensor for samples used for $^3D_{HP}$ and ^{31}P CSA measurement were also first estimated by fitting to the 1DUF structure, and subsequently refined by a two-dimensional grid search. Although different magnitudes for D_a^{CH} were obtained for the various Pf1-containing samples, in perfect linear agreement with the observed ^{2}H solvent signal splitting, the rhombicity R was found to be the same for all Pf1-containing samples and have the very small value of 0.06. Remarkably, in bicelle medium we find an R value of only 0.09, instead of $R = 0.26$ found previously (Tjandra et al., 2000c). This difference is likely the result of a different batch of phospholipids used for the present bicelle preparation. Predictions of alignment tensor magni-

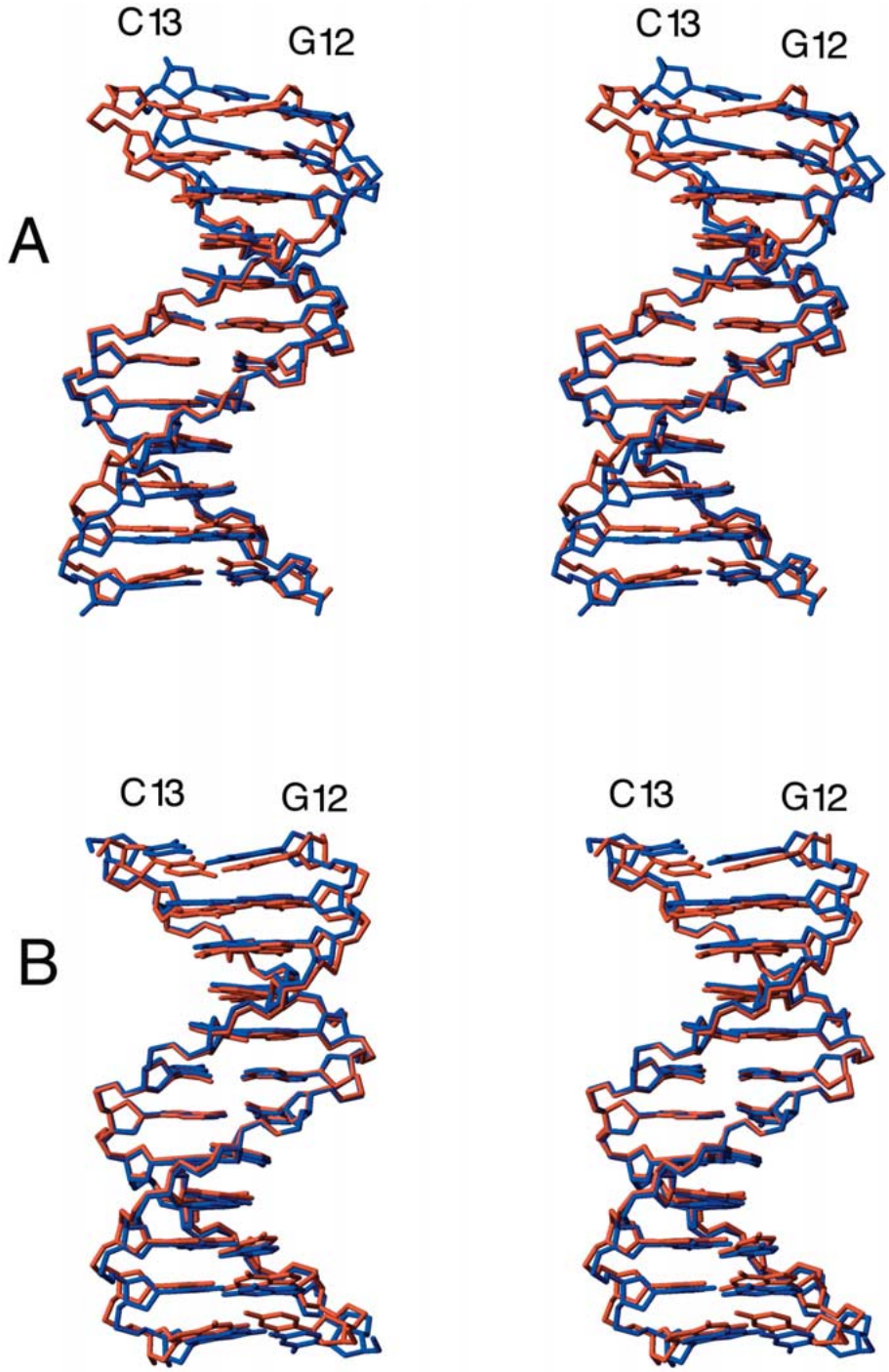

Figure 2. Stereo representation of the superposition of the newly calculated d(CGCGAATTCGCG)$_2$ structure, NMR-all (red), with (A) the original X-ray structure (blue), PDB code 1BNA, and (B) with the most recent NMR structure of this dodecamer, 1DUF (B).

tude and rhombicity on the basis of shape and charge of the solute, using a modified version of the program PALES (Zweckstetter and Bax, 2000), indicate that the rhombicity for the Dickerson dodecamer is quite sensitive to the surface charge density of the nematogen. For bicelles, this surface charge density is a function of the fraction of hydrolyzed lipids present in the sample.

Description of NMR structure

Figure 2 compares the newly calculated d(CGCGAATTCGCG)$_2$ structure with the original X-ray structure (PDB code 1BNA) (Dickerson and Drew, 1981) and with the most recent NMR structure of this dodecamer (1DUF) (Tjandra et al., 2000c). Only a single set of resonances is observed for the two strands of this palindromic oligomer. Not surprisingly, therefore, the resulting NMR structure shows nearly perfect inversion symmetry. This contrasts with the X-ray structures, which all exhibit considerable asymmetry induced by crystal contacts.

As was found by Tjandra et al. (2000), the solution structure differs considerably from structures observed in the crystalline state. The bending of the helical axis (10°, as defined by the program CURVES (Lavery and Sklenar, 1988)) toward the minor groove in the center of the dodecamer is somewhat larger than the 7° found in the previous NMR study, but still within the error bounds (Tjandra et al., 2000c). However, the NMR structure lacks the pronounced kinks around base pairs 3 and 4, seen in the crystalline state. As a result, the rmsd of the new NMR structure relative to the original X-ray structure (1BNA) is rather large (1.3 Å), and only slightly smaller (1.2 Å) when compared to the more recent high resolution structure (PDB code 355D). When only the center six base pairs are considered, much better agreement is obtained, an rmsd of 0.55 Å relative to 1BNA and 0.58 Å relative to 355D. With an rmsd of 0.68 Å for the full length and 0.35 Å for the center six base pairs, the earlier and new NMR structures differ considerably more from one another than the uncertainty suggested by the spread observed in each of the two studies. Closer inspection shows that most of the difference can be attributed to a slight but systematic decrease in the rise between base pairs, i.e., a translational effect that determines the overall length of the oligonucleotide, and which is relatively insensitive to orientations of individual bond vectors, i.e., insensitive to the one-bond heteronuclear dipolar couplings used in the previous study.

Table 1 lists the backbone and glycosyl torsion angles, as well as the pseudorotation phase angles and amplitudes calculated from the lowest energy structure. As can be seen from this table, the torsion angles α, β, γ, ε, ζ and χ are consistent with a regular B-DNA conformation, as expected on the basis of the narrow chemical shift range observed in the ^{31}P NMR spectrum. The δ angles are in remarkable agreement with values reported previously from an analysis of $^3J_{H3'H4'}$ couplings (Bax and Lerner, 1988), with a Pearson's correlation coefficient, R_P, of 0.92 when the outliner C9 is excluded (with C9 included, $R_P = 0.82$). All torsion angles associated with the center six base pairs fall within the range of $\pm 20°$ (rmsd = 11.3°) relative to those seen in the 1BNA crystal structure. Angles that differ substantially between the two structures are all located in the flanking GC tracts. The same is found when comparing our NMR structure with the more recent higher resolution crystal structure, 355D. This again confirms that the main difference between the NMR and crystal structures stems from the distortion in the GC tracts, apparently caused by the presence of the tightly coordinated, hydrated Mg^{2+} and crystal packing, which give rise to unusual backbone torsion angles. As expected, our new NMR structure shows somewhat better agreement with the center six base pairs in the crystal structure than the earlier NMR structure (1DUF), which lacked explicit experimental restraints impacting the phosphodiester linkages.

Structure calculated without electrostatics

Calculation of the above structure includes Lennard-Jones, van der Waals, and electrostatic non-bonded contacts in the target function. These terms are commonly used in NMR structure calculations of nucleic acids (Gronenborn and Clore, 1989; Metzler et al., 1990; Ulyanov et al., 1992; Varani et al., 1996; Allain and Varani, 1997) and serve to balance the empirical energy function such that reasonable non-bonded contacts are obtained in the final structure (Gronenborn and Clore, 1989). However, without inclusion of counterions and explicit water, the electrostatic term is based on a number of relatively poor approximations and therefore has the potential to introduce substantial distortions in the final structure (Kuszewski et al., 2001).

Our present data set contains far more independent experimental restraints per nucleotide than have been used in any previous oligonucleotide structural study. Therefore, we investigated to what extent the elec-

Table 1. Backbone and glycosyl torsion angles, pseudorotation phase angles and amplitudes in d(CGCGAATTCGCG)$_2$ derived from 16 lowest energy structures[a]

Nucleotide	α (deg)	β (deg)	γ (deg)	δ (deg)	ε (deg)	ζ (deg)	χ (deg)	Phase P (deg)	Amplitude (deg)
C(1)			55 ± 17	118 ± 3	−158 ± 1	−93 ± 1	−123 ± 1	121 ± 5	32 ± 2
G(2)	−61 ± 2	170 ± 2	46 ± 1	132 ± 1	−173 ± 1	−114 ± 1	−104 ± 1	144 ± 1	37 ± 1
C(3)	−61 ± 3	166 ± 1	55 ± 2	120 ± 1	−164 ± 1	−97 ± 1	−121 ± 1	124 ± 2	32 ± 1
G(4)	−69 ± 1	176 ± 1	54 ± 1	134 ± 1	−168 ± 1	−122 ± 1	−102 ± 1	148 ± 1	38 ± 1
A(5)	−58 ± 1	164 ± 1	50 ± 1	131 ± 1	−175 ± 1	−103 ± 1	−115 ± 1	145 ± 1	34 ± 1
A(6)	−71 ± 1	177 ± 1	58 ± 1	125 ± 1	−178 ± 1	−93 ± 1	−112 ± 1	140 ± 1	31 ± 1
T(7)	−58 ± 1	172 ± 1	52 ± 1	109 ± 1	−178 ± 1	−88 ± 1	−122 ± 1	109 ± 1	34 ± 1
T(8)	−51 ± 1	175 ± 1	43 ± 1	125 ± 1	−174 ± 1	−92 ± 1	−111 ± 1	134 ± 1	31 ± 1
C(9)	−66 ± 1	176 ± 1	50 ± 1	133 ± 1	−165 ± 1	−114 ± 1	−107 ± 1	146 ± 1	36 ± 1
G(10)	−60 ± 1	177 ± 1	43 ± 1	139 ± 1	−170 ± 1	−114 ± 1	−99 ± 1	156 ± 1	35 ± 1
C(11)	−52 ± 1	161 ± 1	48 ± 1	111 ± 1	−164 ± 1	−96 ± 1	−125 ± 1	109 ± 1	31 ± 1
G(12)	−73 ± 1	174 ± 1	57 ± 1	115 ± 1			−109 ± 1	117 ± 3	26 ± 1

[a]Reported uncertainties represent the r.m.s. distributions in the ensemble of 16 structures and underestimate the true uncertainty.

trostatic term is actually needed when using a large number of homo- and heteronuclear dipolar couplings. As expected, turning off this term in the XPLOR dynamics run has a minimal effect: The two sets of structures calculated with and without the electrostatic term differ by only 0.3 Å when considering the center 10-base pairs. In contrast, when only the NOE and J-coupling experimental restraints are used, turning off the electrostatic term results in a substantially larger rmsd of >1.0 Å. Our finding that the electrostatic term can be turned off when working with large numbers of dipolar restraints is encouraging because it indicates that the calculated NMR structures no longer rely on the oversimplified electrostatic energy term.

Cross-validation of NMR structure

In analogy to the free R-factor (Brunger, 1992), commonly used in crystallography, the accuracy of an NMR structure can be evaluated by the agreement between the structure and dipolar couplings that were not used in the structure calculation process (Cornilescu et al., 1998; Clore and Garrett, 1999). The so-called Q factor is defined by:

$$Q = \text{r.m.s}(D_{obs} - D_{pred})/\text{r.m.s.}(D_{obs}), \quad (1)$$

where r.m.s. is the root-mean-square function, and D_{obs} and D_{pred} are the observed and predicted dipolar couplings, respectively. Equation 1 applies only to fixed-distance interactions, such as $^1D_{CH}$ or $^1D_{NH}$. Defining a quality factor on the basis of variable distance ^{1}H-^{1}H interactions in principle is also possible but requires correction factors to ensure that small fractional errors in D_{HH} for very proximate pairs of protons do not outweigh larger errors in orientation for less proximate pairs. Here, calculation of Q factors is restricted to $^1D_{CH}$ and $^1D_{NH}$ interactions. In Equation 1, the denominator rms(D_{obs}) accounts for the magnitude of the alignment tensor, which ideally requires that the bond vectors are uniformly distributed in space. However, in the case of B-form DNA this distribution is quite non-uniform. Instead, r.m.s.(D_{obs}) can be derived from (Clore et al., 1998):

$$\text{r.m.s.}(D_{obs}) = \{(D_a^2[4 + 3(D_r/D_a)^2]/5\}^{1/2}, \quad (2)$$

which corresponds to the case where the bond vectors are uniformly distributed.

Typically, the Q-factor is determined by randomly omitting a small fraction of dipolar couplings, and then by calculating the agreement for this omitted fraction with the calculated structure, and repeating this procedure many times. In the earlier NMR structure (Tjandra et al., 2000c), it was found that the number of dipolar couplings available for calculating the structure was too small for applying a Q factor analysis; removal of even a single dipolar restraint resulted locally in an underdetermined structure. Our present experimental restraint set is much larger, and this previous limitation no longer applies.

Cross-validation of the DNA dodecamer was carried out by omitting various subsets of dipolar couplings during the structure calculation, followed by application of Equation 1, using the denominator of

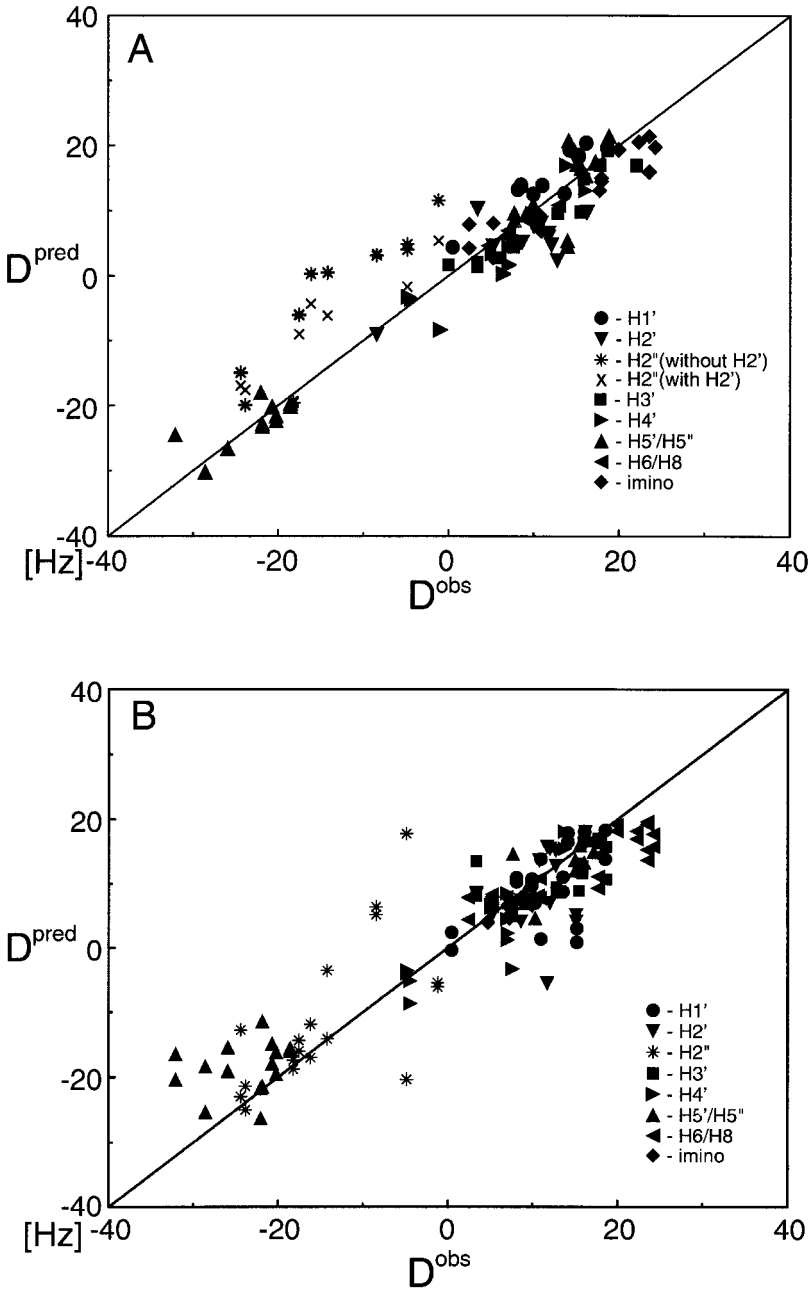

Figure 3. Agreement between dipolar couplings and (A) NMR and (B) X-ray structures. (A) Cross validation results for the NMR-all structure. The figure is a composite of nine structure calculations, each time leaving out one particular type of heteronuclear one bond coupling, and the correlation shown corresponds to the measured coupling, versus the coupling predicted by the structure which did not include that coupling in its input restraints. Cross validation results shown are restricted to the center 10 base pairs. (B) Correlation shown is simply that between experimental couplings and those predicted after best fitting these couplings to the X-ray structure (1BNA), using singular value decomposition.

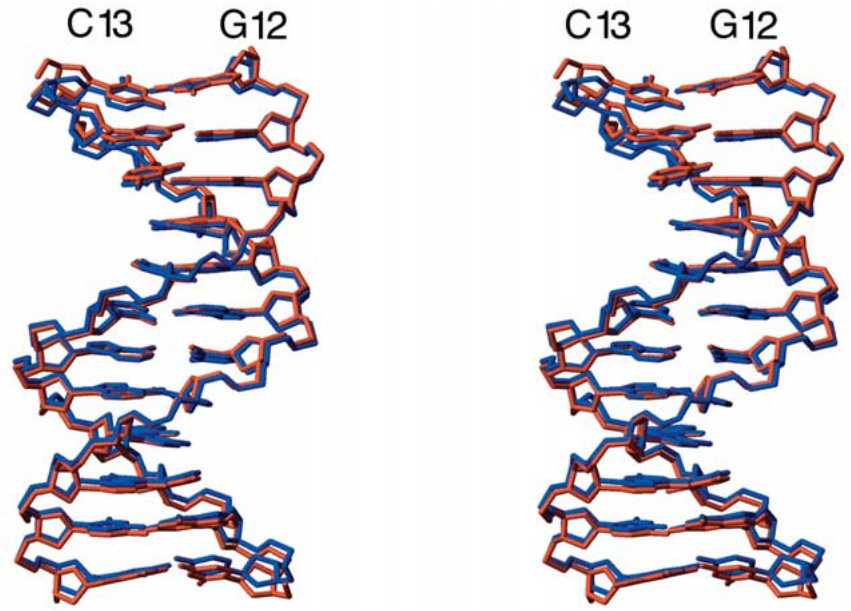

Figure 4. Superposition of the NMR-all (red) and NMR-HH (blue) structures in stereo representation. When excluding the terminal base pairs, the rmsd between these two structures is 0.40 Å. When only the center six base pairs are considered, the rmsd is 0.24 Å.

Equation 2. For N1/N3-H1/H3, C1′-H1′, C2′-H2′, C3′-H3′, C4′-H4′, C6/C8-H6/H8 and C5′-H5′/H5″ we find an average Q-factor of 25% (considering only the center 10 base pairs; the terminal base pairs are subject to extensive internal dynamics). Figure 3A shows the correlation between the observed and predicted dipolar couplings of each omitted data set. As can be seen, the agreement is reasonable for all dipolar couplings with the exception of $D_{C2'\text{-}H2''}$ (asterisks in Figure 3A). As discussed in more detail below, the likely reason for the poor cross validation of $D_{C2'\text{-}H2''}$ values stems from their exquisite sensitivity to the deoxyribose ring pucker, and dynamic effects are believed to be responsible for our inability to obtain Q factors lower than 25%. As shown in Figure 3B, the same validation procedure applied to the original crystal structure, 1BNA, yields considerably worse agreement (Q = 0.36) whereas a similar analysis of the more recent high resolution X-ray structure (355D) is dominated by several outliers, resulting in Q = 0.44 (data not shown). So, although our new NMR structure cross validates better than the crystal structures, Q values remain considerably higher than for high-resolution protein structures.

As mentioned by Tjandra et al. (2000c), the accuracy of the NMR structure can be evaluated qualitatively by comparing the agreement between the structures of two dimeric fragments. Our new structure exhibits an rmsd of 0.35 Å between the G2-C23, C3-G22 and G10-C15, C11-G14 fragments, which is identical to what was seen in the earlier structure. The close structural similarity for these two-base pair fragments is consistent with the very similar ^{1}H and ^{13}C chemical shifts observed for G2 and G10, and for C3 and C11 (Hare et al., 1983; Tjandra et al., 2000c). A much larger difference (0.92 and 0.82 Å for 1BNA and 355D, respectively) when comparing the same two-base pair fragments in the crystalline state is caused by the asymmetric Mg^{2+} metal binding (Chiu et al., 1999; Tereshko et al., 1999a) and crystal contacts (Dickerson et al., 1987; DiGabriele et al., 1989).

Structure without ^{1}H-^{13}C and ^{1}H-^{15}N dipolar couplings

In order to evaluate whether a high quality structure can be determined in the absence of heteronuclear dipolar couplings, two parallel calculations were carried out: One with all the restraints (NMR-all) and one without $^{1}D_{CH}$ and $^{1}D_{NH}$ restraints (NMR-HH). Calculations of the NMR-HH set of structures converged remarkably well. For the 25 lowest energy structures, the width of the bundle for NMR-all and NMR-HH is 0.04 Å and 0.08 Å, respectively. However, as discussed previously (Tjandra et al., 2000c), when using dipolar couplings the width of the bundle is not representative of the accuracy of the structure.

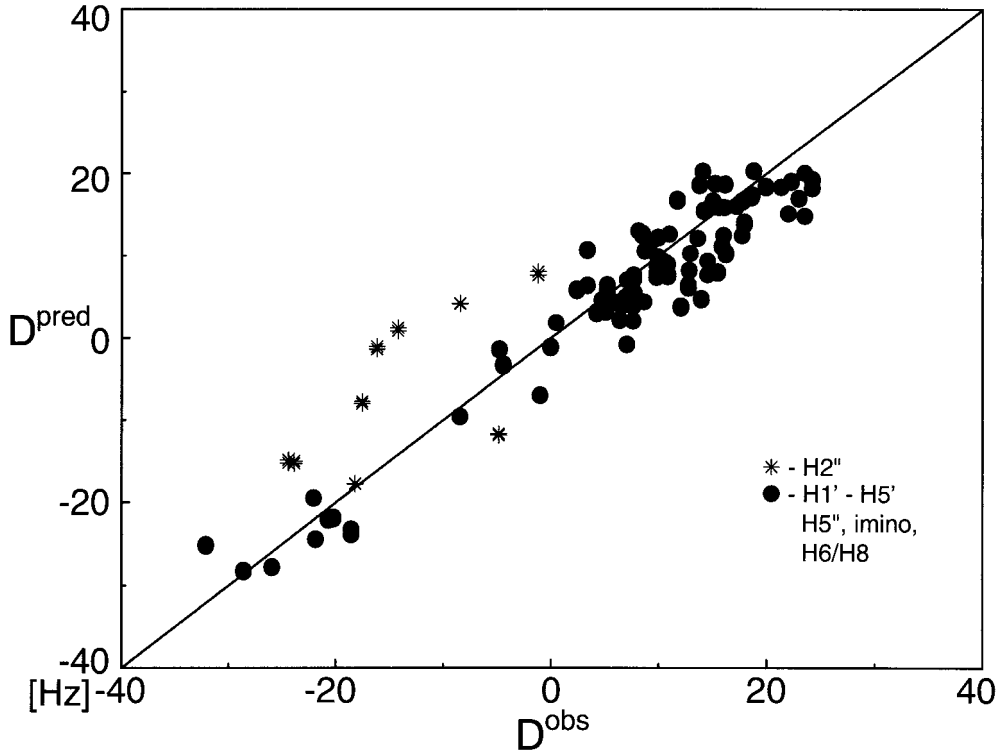

Figure 5. Correlation between measured and predicted D_{CH} and D_{NH} values for the NMR-HH structure, with most of the outliers corresponding to C2′-H2″ couplings.

Comparison of the two sets of structures confirms that they are very similar (Figure 4). When excluding the terminal base pairs, the rmsd between NMR-HH and NMR-all is 0.40 Å. When only the center six base pairs are considered, the rmsd decreases to 0.24 Å.

Another indicator of the quality of NMR-HH is provided by the above mentioned cross validation procedure. When the D_{CH} and D_{NH} dipolar couplings, previously measured in bicelle medium, are fitted to the NMR-HH structure, a good correlation is observed, again with the exception of the C2′-H2″ coupling (Figure 5). When excluding two terminal base C1 and G12 and all $D_{C2'-H2''}$, the Q-factor for the center 10-base pairs is 26%, which is almost indistinguishable from the above reported cross validation result (Q = 25%) in which all heteronuclear dipolar couplings were included. As mentioned above, the ^{1}H-^{1}H dipolar couplings were measured in Pf1 medium, whereas the heteronuclear couplings used in the structure calculation were measured in bicelles. However, when the proton-only structure is best-fitted to additional $^1D_{CH}$ and $^1D_{NH}$ couplings (associated with six nucleotides: C1, G2, A5, T7, C9 and G10, using a selectively labeled oligomer) measured in Pf1 phage, a comparable Q-factor of 23% was obtained. These results indicate that the fact that the fit remains imperfect is not caused by a difference in the liquid crystalline media used. It also indicates that for small oligomers, where most ^{1}H-^{1}H couplings can be measured at high field, the quality of the structure attainable without measurement of heteronuclear couplings is comparable to what can be obtained with isotopic labeling. However, we also point out that in order to calculate the NMR-HH structures, a large number of ^{1}H-^{1}H dipolar couplings (more than 29 per base pair in the present study) were used.

Both NMR-HH and NMR-all were generated with $^1D_{HP}$ and $\Delta\delta(^{31}P)$ restraints included. When both of these two restraints were omitted from the NMR-HH and NMR-all calculation, the deviation between NMR-all and NMR-HH becomes much more pronounced (rmsd of 1.09 Å for the center 10 base pairs, 0.52 Å for the center six base pairs. Therefore, these D_{HP} and $\Delta\delta(^{31}P)$ restraints are particularly important when lacking heteronuclear dipolar couplings.

Ring pucker dynamics

Previous NMR studies of deoxyribose conformation have been based largely on ^{1}H-^{1}H scalar couplings (van Wijk et al., 1992; Rinkel et al., 1987; Iwahara et al., 2001), but also on NOEs (Tonelli and James, 1998), ^{13}C chemical shifts (LaPlante et al., 1994), and most recently on cross-correlated relaxation (Felli et al., 1999; Boisbouvier et al., 2000). It is now well established that many of the deoxyriboses in DNA oligomers are subject to rapid dynamic averaging between C2′-endo and C3′-endo ring conformations, also denoted S and N, respectively (van Wijk et al., 1992; Tonelli and James, 1998; Ojha et al., 1999; Kamath et al., 2000). This dynamic behavior is mostly invisible in X-ray crystal structures, where electron density can be fit satisfactorily to a single conformer which, depending on hydration conditions and crystal packing interactions, may not be the predominant one in solution (van Wijk et al., 1992). Although the C2′-endo conformation predominates in canonical B-DNA, the C2′-endo and C3′-endo conformations are close in energy. Therefore, at the cryogenic temperatures used for the high resolution X-ray structures, the Boltzmann population of the slightly higher energy conformer is decreased considerably relative to room temperature. This may make it particularly hard to observe the minor population by X-ray crystallography.

The data presented below indicate that in B DNA, with the z axis of the alignment tensor along the helical axis, the $^{1}D_{C2'H2''}$ coupling is particularly sensitive to the deoxyribose pucker. As an example, Figure 6 shows the changes of the various deoxyribose ^{13}C-^{1}H dipolar couplings of cytosine C9 as a function of the pseudorotation phase P. Clearly, $D_{C2'\text{-}H2''}$ and $D_{C2'\text{-}H2'}$ change sharply when P shifts from C2′-endo to C3′-endo, whereas the other couplings are relatively insensitive to this change. The dependence of the $D_{C2'\text{-}H2''}$ and $D_{C2'\text{-}H2'}$ on P is highly nonlinear, and this has important consequences for the calculated structures if the data are fit to a single, average structure. For example, if the sugar exists as a 75% C2′-endo conformation and 25% C3′-endo, this will increase the observed $D_{C2'\text{-}H2''}$ by about 8 Hz relative to a pure C2′-endo pucker, and decrease $D_{C2'\text{-}H2'}$ by roughly the same amount. However, there is no sugar conformation corresponding to these intermediate values. For example, a decrease of 30° in P from C2′-endo increases the predicted $D_{C2'\text{-}H2''}$ by about 18 Hz, while decreasing the predicted $D_{C2'H2'}$ by less than 5 Hz (Figure 6). This numerical example illustrates the likely reason for the poor cross validation of the $D_{C2'H2''}$ couplings in Figure 3A.

The above example points to the presence of a dynamic C2′-endo/C3′-endo equilibrium as the main reason for the failure to fit all dipolar couplings simultaneously to a single structure. As can be seen in Figure 7, this effect is most pronounced for the cytosines, whereas the purine data fit reasonably well with a single conformer. This agrees with both computational and experimental studies, which indicate that pyrimidines, especially cytosines, tend to have a higher C3′-endo population than purines (Gorin et al., 1990; van Wijk et al., 1992).

Multiple conformer fit of dipolar couplings

The above discussion indicates that ideally the data must be fit to an ensemble in which each of the sugars samples two different conformations. A change in sugar pucker for one sugar is expected to have a (presumably small) effect on the orientation of other nucleotides relative to the average alignment tensor of the dodecamer. Assuming that changes in ring pucker occur independently of one another, a rigorous analysis of the dodecamer therefore would require 2^{24} conformers, which is clearly not feasible. However, if changes in ring pucker are assumed to have no effect on the remainder of the molecule, a simplified analysis is possible in which first the average conformation is determined in the manner described by Tjandra et al. (2000), followed by two-conformer refinement in which a single nucleotide is treated as a linear combination of two conformers. Analogous multiple-conformer analyses have previously been proposed for NOE-based structures (Torda et al., 1990; Bonvin and Brunger, 1995; Gorler et al., 2000). However, such analyses increase the degrees of freedom and with the small number of dipolar couplings affected by the change in sugar pucker, this makes it difficult to evaluate whether the observed improvement in the fit is statistically meaningful. Therefore, we resort to a third approach, where the sugar is simply assumed to be a linear combination of C2′-endo and C3′-endo.

In this procedure, two separate oligomer structures, NMR-2′ and NMR-3′ are generated for each nucleotide investigated. These two structures are calculated starting from the NMR-all structure, but removing all experimental restraints involving any of the protons of the deoxyribose in question, i.e., excluding $^{1}D_{CH}$, D_{PH}, and D_{HH} couplings, whereas those corre-

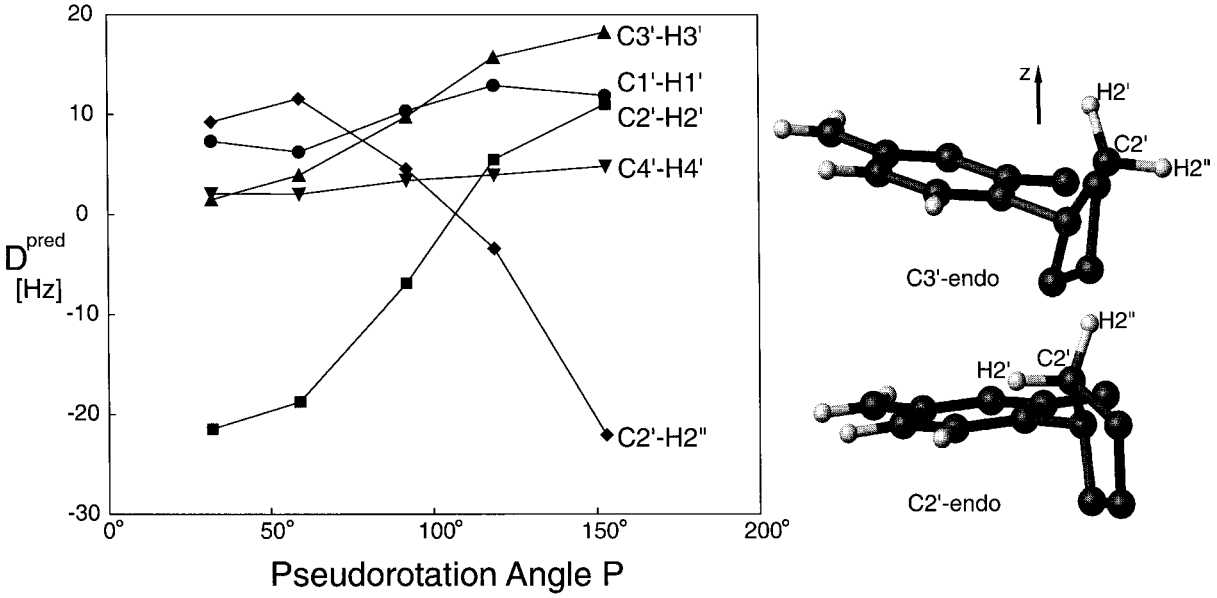

Figure 6. Predicted one-bond dipolar couplings as a function of the deoxyribose pseudorotation angle, P, for cytosine C9. Structures were generated in the same way as described for generating the pure C2′-endo and C3′-endo conformers. Very similar patterns are obtained for the other nucleotides (data not shown). The inset shows the C2′-endo and C3′-endo conformers of C9 and their orientation relative to the *z* axis of the alignment tensor.

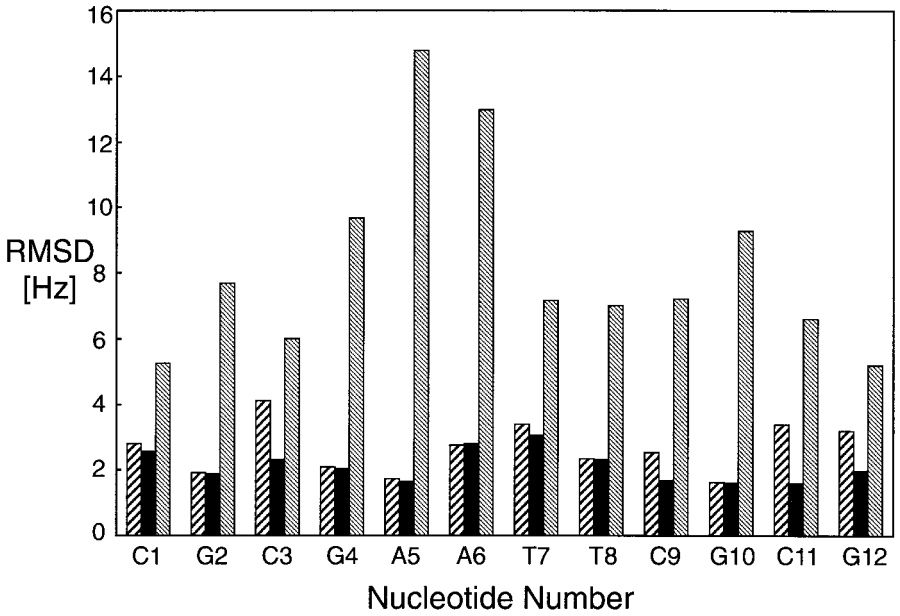

Figure 7. Residual rmsd of all dipolar couplings, involving at least one deoxyribose proton, as a function of nucleotide number, when the particular nucleotide is forced into C2′-endo (hashed bars) or C3′-endo (shaded) conformations, or when fitting a linear combination of these two conformers (solid bars).

sponding to the base and the hydrogen bonds were left intact. For all other nucleotides, the coordinates are fixed to their positions in NMR-all. The geometry of the deoxyribose in question is forced to be either C2′-endo or C3′-endo by introducing the corresponding torsion angle restraints (Markley et al., 1998) in the XPLOR structure calculation. Subsequently, a best fit to the dipolar couplings is carried out which minimizes

$$F(\alpha) = \Sigma_i\{[\alpha D_i^{pred,1} + (1-\alpha)D_i^{pred,2}] - D_i^{meas}\}^2 \quad (3)$$

with respect to the fraction α, where $D_i^{pred,1}$ and $D_i^{pred,2}$ are the dipolar couplings predicted for interaction i in the NMR-3′ and NMR-2′ structures, respectively, and the summation extends over all couplings left out of the structure calculation, i.e., all couplings involving any of the protons of the deoxyribose in question. The predicted values, D_i^{pred}, for both structures are calculated for the alignment tensors used in deriving the NMR-all structure, which differ for the ^{13}C-^{1}H and ^{1}H-^{1}H couplings.

Figure 7 displays the rms difference between measured and predicted dipolar couplings obtained from Equation 3 as a function of nucleotide number. It is clear from this figure that the C3′-endo conformations (shaded bars) result in very high rmsds, whereas C2′-endo conformations (hashed bars) fit considerably better. However, a linear combination of the two conformations (solid bars) yields even lower rmsds, which approach the experimental uncertainty in the data. The largest improvement in the fit is obtained for the cytosines, whereas little or no improvement is seen for most purines.

The optimal ratio, α, obtained from the two-conformer fit provides an estimate for the percentages of C2′- and C3′-endo conformers for each nucleotide in the dodecamer. As shown in Table 2, population of the C3′-endo conformer is negligible for most purines whereas for pyrimidines it can reach as high as 36%. For T(7) and T(8), no separate $D_{C2'\text{-}H2''}$ and $D_{C2'\text{-}H2'}$ values were available due to the absence of specific deuteration at their C2′ position. Considering that $D_{C2'\text{-}H2''}$ and $D_{C2'\text{-}H2'}$ values are the most sensitive to pucker change, and consequently dominate the rmsd for most other nucleotides, the fraction α of C2′-endo conformer derived from the minimization of Equation 3 is less reliable for T(7) and T(8) than for the remainder of the nucleotides.

Table 2. C3′-endo percentage from fitting dipolar couplings to a two-state equilibrium[a]

Residue	% C3′endo
C(1)	20
G(2)	3
C(3)	32
G(4)	0
A(5)	1
A(6)	4
T(7)	24
T(8)	2
C(9)	22
G(10)	3
C(11)	36
G(12)	33

[a]Values correspond to a linear combination of a nearly C2′-endo conformer (P = 155°, pucker amplitude 35°) and a C3′-endo conformer (P = 18°, amplitude 35°), when fitting all ^{1}H-^{1}H and ^{1}H-^{13}C dipolar couplings, and restraining the alignment tensor to the value obtained from the global fit of the full oligomer.

Discussion and conclusions

As pointed out above, it is clear from the NMR data that the sugar conformations are dynamically averaging. Nevertheless, the cross validation results (Q = 0.25) obtained when considering the structure as a single, average conformer indicate that such a single conformer provides a reasonable description for its time-averaged conformation. The helical parameters of this structure are found to be highly regular (Table 3), even more so than in the previous NMR study by Tjandra et al. (2000).

Comparison with crystal structures

Perhaps the most striking local feature of the current solution structure is the high absolute value of propeller twist for the center AT-tract compared with the flanking GC-tracts (an average of −24° *versus* −17°). This trend is similar to that observed in the crystalline state, despite the substantial differences in global structure in the two states. However, the solution structure does not exhibit the sudden drop in propeller twist at base pairs C3-G22 and C10-C15, observed in the 1BNA (Dickerson and Drew, 1981) and 355D (Shui et al., 1998a) crystal structures and, to a much lesser extent, in the earlier solution structure (Tjandra et al., 2000).

It has been commonly known that A-tract DNA exhibits a narrow minor groove and large propeller twist (Berman, 1997). This feature has been frequently observed in many cases of oligo(dA)-oligo(dT) containing structures, both by X-ray crystallography and NMR methods (Nelson et al., 1987; Coll et al., 1987: DiGabriele et al., 1989; Shatzky-Schwartz et al., 1997; MacDonald et al., 2001). Since Watson–Crick AT base pairs lack the third hydrogen bond, present in GC, twisting around the long axis of the AT base pair is less restricted by H-bonding (Dickerson and Drew, 1981). The unusually large propeller twist shortens the distance between N6 of A5 and O4 of T7 on the complementary strand, facilitating formation of a bifurcated hydrogen bond (Taylor et al., 1984; Nelson et al., 1987). In the present structure, the length of this potential hydrogen bond is 3.2 Å, which is comparable to the values reported for other A-tract structures (Nelson et al., 1987; DiGabriele et al., 1989; Shatzky-Schwartz et al., 1997; MacDonald et al., 2001).

Strong propeller twist is correlated with a decreased rise and increased helical twist for the AT-tract in d(CGCGAATTCGCG)$_2$ (Table 3). The smaller rise in the center AT-tract makes the current structure slightly more compressed than the crystal structures (1BNA and 355D) and the previous NMR structure (1DUF). The overall length in the present study is 35.7 Å, which is very close to the 355D crystal structure (36.2 Å) but 1.5 Å shorter than 1BNA. The structure calculated without the electrostatic 'elec' term in XPLOR shows the same total length, indicating that the slight shortening is not caused by the empirical non-bonded functions (Kuszewski et al., 2001). The helical twist for the AT-tract corresponds to 9.9 base pairs per turn, essentially identical to what was found by Dickerson and Drew (1981). On the other hand, the twist angle for the AA:TT dimer, 38° (Table 3), is somewhat larger than the 'canonical' value ∼36° for poly(dA):poly(dT) in solution, obtained by the electrophoresis band-shift method (Wang, 1979). Also, similar to what is seen in the 1BNA crystal structure, the average helical twist angle for pyrimidine-purine (CG) steps is smaller than that for purine-pyrimidine steps (GC), which has been attributed to optimization of stacking interactions (Dickerson and Drew, 1981). A distinctive feature of the NMR structure is a relatively small sequence dependence of the twisting compared to the crystalline phase, where the sequence specificity is likely to be emphasized by the packing interactions. For example, the absolute difference in twisting between the GC and CG steps in our structure is ∼2° (Table 3), whereas in the crystal structures this difference varies from 7° (1BNA) to 17° (355D).

The inter-base-pair tilt and roll also show less variation than in the crystals, especially in the center AT-tract where it exhibits almost zero tilt and small negative roll angles (−0.8° to −2.6°). These roll values are practically the same as observed in the two crystal structures mentioned above (0° to −4°). In the GC region, our structure does not show the swinging opposite signed roll angle at C1/G2/C3 and C9/G10/C11, seen in the X-ray structures. However, a decrease in the roll angle for the GC step (1°) is clearly evident compared to the relatively larger roll for the CG step, 5–8° (Table 3). This tendency for the purine-pyrimidine steps to open their base planes toward the major groove, and for the pyrimidine-purine step to open towards the minor groove, was also noted by Dickerson and Drew (1981) and by Calladine (1982). Again, as in the case of the helical twist, the discrepancy between the NMR and X-ray measurements of the roll angle is substantially higher in the GC region, reaching its maximum for the GC step (1° in Table 3 versus −11° in 1BNA, and −13° in 355D).

Overall, we see that the NMR and X-ray data together give a fairly consistent description of the average structure of the AT-tract, with small absolute values of the roll and tilt angles, strong propeller twist and narrow minor groove. In contrast, the GC-rich ends of the dodecamer have different local geometries in the two environments (solution and crystal), this geometry probably being affected by intermolecular interactions in the crystal. As to the global description of the DNA curvature in solution, the crystal-based model of the AT-tracts (Dickerson et al., 1996) agrees well with our NMR data, whereas the structures of the GC-containing DNA fragments demonstrate a profound variability, and more measurements are necessary to clarify this issue. One of the principal difficulties in solving the solution structure of the GC-rich DNA is the inherent flexibility of the cytosine deoxyribose, fluctuating between the C2′-endo and the C3′-endo conformations (Table 2). In turn, this sugar switching influences the mutual orientation of the bases, and as a result, the overall configuration of the duplex (Kamath et al., 2000). In this respect, it is important to point out that the structure reported here represents the average structure, from which considerable deviations are occurring over time.

Table 3. Average helical parameters for the ensemble of 16 lowest energy structures

	Propeller twist (deg)	Rise[a] (Å)	Helical twist[a] (deg)	Tilt[a] (deg)	Roll[a] (deg)	Slide[a] (pm)	Shift[a] (pm)
C(1)-G(24)	-21 ± 1[b]	3.5 ± 0.1	33.3 ± 0.4	-3.4 ± 0.4	8.1 ± 0.7	-19 ± 5	45 ± 9
G(2)-C(23)	-16 ± 1	3.2 ± 0.1	36.2 ± 0.4	2.4 ± 0.5	1.0 ± 0.1	-95 ± 6	-22 ± 4
C(3)-G(22)	-15 ± 2	3.5 ± 0.1	34.5 ± 0.5	1.8 ± 0.3	4.9 ± 0.4	-48 ± 7	$3 + 6$
G(4)-C(21)	-15 ± 1	3.2 ± 0.1	36.1 ± 0.1	-0.5 ± 0.3	6.0 ± 0.3	-68 ± 3	-37 ± 3
A(5)-T(20)	-23 ± 1	3.1 ± 0.1	38.0 ± 0.2	-0.6 ± 0.2	-0.8 ± 0.2	-68 ± 2	-41 ± 1
A(6)-T(19)	-27 ± 1	3.0 ± 0.1	34.3 ± 0.2	0 ± 0.1	-2.6 ± 0.3	-103 ± 1	0 ± 2

[a]Defined relative to the next base pair (Lavery and Sklenar, 1988). Due to the dyad symmetry of the dodecamer in solution, the parameters are given for only one half of the oligomer.
[b]Reported uncertainties represent the rms distributions in the ensemble of 16 structures and underestimate the true uncertainty.

Effect of counterions

Crystallographic data have clearly shown that the presence of divalent metal ions such as Mg^{2+} can affect DNA structure (Chiu and Dickerson, 2000). Metal ions localized in both the major and minor grooves of DNA have been observed by X-ray crystallography and NMR (Drew and Dickerson, 1981; Hud and Feigon, 1997; Shui et al., 1998a, b; Tereshko et al., 1998a, b; Hud et al., 1999; Sines et al., 2000). Chemical shifts and RDC values are exquisitely sensitive reporters of minor changes in structure, and we therefore have monitored these parameters also as a function of Mg^{2+} and K^+ concentration. Similar to what was reported previously for this dodecamer (Tjandra et al., 2000), ^{1}H chemical shift changes induced by Mg^{2+} at near physiological Mg^{2+} concentration (0.5–1.0 mM) are negligible for most of the nucleotides, with a maximum change of about -0.01 ppm for H8 of G4. Even in the presence of 10 mM Mg^{2+} (and 0.5 mM duplex DNA) the largest chemical shift change occurs for G4-H8 and is only -0.02 ppm. Effects on the ^{31}P chemical shifts are considerably larger and range from -0.03 ppm for G12, to -0.2 ppm for G4 and A5 at 1 mM Mg^{2+}, increasing to -0.6 ppm for G4 at 10 mM Mg^{2+}. Differences in ^{1}H-^{13}C RDCs in the absence and presence of 1.5 mM Mg^{2+} (0.3 mM duplex) are within the experimental error (correlation coefficient $R = 0.994$), except for a uniform 7% decrease in alignment induced by the Mg^{2+} (which results from its effect on the Pf1 medium, as evidenced by a similar reduction in ^{2}H lock solvent splitting). In these latter experiments the effective Mg^{2+} concentration was about 1 mM, as monitored by its effect on the ^{31}P shifts; the effective Mg^{2+} concentration accounts for traces of EDTA, present in the washed Pf1 medium, and for binding of Mg^{2+} to the negatively charged phage.

Together, the absence of significant chemical shift and RDC changes upon increasing the Mg^{2+} concentration from zero to the physiological range (0.5–1 mM), in the presence of moderate ionic strength (50 mM KCl; 20 mM phosphate buffer), indicates there is no significant perturbation of the dodecamer structure by such low levels of Mg^{2+}. The effect of changing the concentration of monovalent K^+ and Na^+ ions over the 40–140 mM range on ^{1}H chemical shifts and ^{1}H-^{13}C dipolar couplings was found to be comparably small as the above mentioned effect of 1 mM Mg^{2+}.

Effect of internal dynamics

Our study demonstrates that even in the absence of heteronuclear labeling, measurement of ^{1}H-^{1}H and ^{1}H-^{31}P dipolar couplings, together with ^{31}P CSA induced chemical shift changes in liquid crystalline media, provide sufficient information for deriving the three-dimensional structure of small DNA oligomers. When both homonuclear and one-bond heteronuclear couplings are available, cross validation of the structural quality is feasible. Results for the Dickerson dodecamer indicate that no single structure can accurately fit the entire ensemble of experimental data, and this imperfect fit is most notable for the cytosines. A significant improvement in the fit is obtained, particularly for the C2′-H2″ dipolar interactions, when considering the structure as a dynamic average of C2′-endo and C3′-endo conformers. The presence of such dynamic equilibria has previously been proposed on the basis of $^3J_{HH}$-based NMR studies (van Wijk et al., 1992; Tonelli and James, 1998; Kamath et al.,

2000). With a Pearson's correlation coefficient of 0.9, the fraction of C3′-endo conformer obtained from the dipolar coupling analysis (Table 3) agrees remarkably well with that of a previous J_{HH} analysis of the same dodecamer (Bax and Lerner, 1988; Yang et al., 2000).

There has been considerable discussion in recent years on the effect of internal dynamics on dipolar couplings, and the inability to perfectly fit dipolar couplings to a structure has been attributed to such effects (Tolman et al., 1997, 2001). Because dipolar couplings are relatively insensitive to small angular fluctuations around an average orientation, large amplitude dynamics are required if modest discrepancies between experimental values and couplings predicted by a static structure are attributed to internal motion. Deoxyribose ring pucker changes are just one example of such large fluctuations, particularly when considering the C2′-H2″, which changes orientation by more than 75° between the two states.

Acknowledgements

The authors acknowledge insightful discussions with Marius Clore, Shin-Ichi Tate and Masatsune Kainosho and wish to thank Jerome Boisbouvier for useful comments during the preparation of the manuscript.

References

Allain, F.H.T. and Varani, G. (1997) *J. Mol. Biol.*, **267**, 338–351.
Barrientos, L.G., Dolan, C. and Gronenborn, A.M. (2000) *J. Biomol. NMR*, **16**, 329–337.
Bax, A. and Lerner, L. (1988) *J. Magn. Reson.*, **79**, 429–438.
Bayer, P., Varani, L. and Varani, G. (1999) *J. Biomol. NMR*, **14**, 149–155.
Berman, H.M. (1997) *Biopolymers*, **44**, 23–44.
Bewley, C.A. and Clore, G.M. (2000) *J. Am. Chem. Soc.*, **122**, 6009–6016.
Boisbouvier, J., Brutscher, B., Pardi, A., Marion, D. and Simorre, J.P. (2000) *J. Am. Chem. Soc.*, **122**, 6779–6780.
Bonvin, A. and Brunger, A.T. (1995) *J. Mol. Biol.*, **250**, 80–93.
Brunger, A.T. (1992) *Nature*, **355**, 472–475.
Brunger, A.T. (1993) *XPLOR Manual Version 3.1*, Yale University, New Haven, CT.
Calladine, C.R. (1982) *J. Mol. Biol.*, **161**, 343–52.
Chiu, T.K. and Dickerson, R.E. (2000) *J. Mol. Biol.*, **301**, 915–945.
Chiu, T.K., Kaczor-Grezeskowiak, M. and Dickerson, R.E. (1999) *J. Mol. Biol.*, **292**, 589–608.
Clore, G.M. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 9021–9025.
Clore, G.M. and Garrett, D.S. (1999) *J. Am. Chem. Soc.*, **121**, 9008–9012.
Clore, G.M., Starich, M.R. and Gronenborn, A.M. (1998) *J. Am. Chem. Soc.*, **120**, 10571–10572.
Coll, M., Frederick, C.A., Wang, A.H.J. and Rich, A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8385–8389.
Cornilescu, G. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 10143–10154.
Cornilescu, G., Marquardt, J.L., Ottiger, M. and Bax, A. (1998) *J. Am. Chem. Soc.*, **120**, 6836–6837.
Crothers, D.M. (1998) *Proc. Natl. Acad. Sci. USA*, **95**.
Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) **6**, 277–293.
Delaglio, F., Wu, Z.R. and Bax, A. (2001) *J. Magn. Reson.*, **149**, 276–281.
Denisov, V.P. and Halle, B. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 629–633.
Dickerson, R., Goodsell, D., Kopka, M. and Pjura, P. (1987) *J. Biomol. Struct. Dyn.*, **5**, 557–579.
Dickerson, R.E. (1983) *J. Mol. Biol.*, **166**, 419–441.
Dickerson, R.E. and Drew, H.R. (1981) *J. Mol. Biol.*, **149**, 761–786.
Dickerson, R.E., Goodsell, D. and Kopka, M.L. (1996) *J. Mol. Biol.*, **256**, 108–125.
DiGabriele, A.D., Sanderson, M.R. and Steitz, T.A. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 1816–1820.
Drew, H.R. and Dickerson, R.E. (1981) *J. Mol. Biol.*, **151**, 535–556.
Drew, H.R., Wing, R.M., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2179–2183.
Felli, I.C., Richter, C., Griesinger, C. and Schwalbe, H. (1999) *J. Am. Chem. Soc.*, **121**, 1956–1957.
Fischer, M., Losonczi, J., Weaver, J. and Prestegard, J. (1999) *Biochemistry*, **38**, 9013–9022.
Fratini, A.V., Kopka, M.L., Drew, H.R. and Dickerson, R.E. (1982) *J. Biol. Chem.*, **257**, 14686–14707.
Gorin, A.A., Ulyanov, N.B. and Zhurkin, V.B. (1990) *Mol. Biol.*, **24**, 1036–1047.
Gorler, A., Ulyanov, N.B. and James, T.L. (2000) *J. Biomol. NMR*, **16**, 147–164.
Goto, N., Skrynnikov, N., Dahlquist, F. and Kay, L. (2001) *J. Mol. Biol.*, **308**, 745–764.
Gronenborn, A. and Clore, G. (1989) *Biochemistry*, **28**, 5978–5984.
Hansen, M.R., Mueller, L. and Pardi, A. (1998) *Nat. Struct. Biol.*, **5**, 1065–1074.
Hare, D.R., Wemmer, D.E., Chou, S.H., Drobny, G. and Reid, B.R. (1983) *J. Mol. Biol.*, **171**, 319–336.
Henning, M., Carlomagno, T. and Williamson, J.R. (2001) *J. Am. Chem. Soc.*, **123**, 3395–3396.
Hud, N.V. and Feigon, J. (1997) *J. Am. Chem. Soc.*, **119**, 5756–5757.
Hud, N.V., Sklenar, V. and Feigon, J. (1999) *J. Mol. Biol.*, **286**, 651–660.
Iwahara, J., Wojciak, J.M. and Clubb, R.T. (2001) *J. Magn. Reson.*, **153**, 262–266.
Kamath, S., Sarma, M.H., Zhurkin, V.B., Turner, C.J. and Sarma, R.H. (2000) *J. Biomol. Struct. Dyn.*, **S2**, 317–325.
Kung, H.C., Wang, K.Y., Goljer, I. and Bolton, P.H. (1995) *J. Magn. Reson. Ser.*, **B109**, 323–325.
Kuszewski, J., Schieters, C. and Clore, M.G. (2001) *J. Am. Chem. Soc.*, **123**, 3903–3918.
Lane, A.N., Jenkins, T.C., Brown, T. and Neidle, S. (1991) *Biochemistry*, **30**, 1372–1385.
Lankhorst, P.P., Haasnoot, C.A.G., Erkelens, C. and Altona, C. (1984) *J. Biomol. Struct. Dyn.*, **1**, 1387–1405.
LaPlante, S.R., Zanatta, N., Hakkinen, A.W., A.H.-J. and Borer, P.N. (1994) *Biochemie*, **33**, 2430–2440.
Lavery, R. and Sklenar, H. (1988) *J. Biomol. Struct. Dyn.*, **6**, 63–91.
MacDonald, D., Herbert, K., Zhang, X.L., Polgruto, T. and Lu, P. (2001) *J. Mol. Biol.*, **306**, 1081–1098.

Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright, P.E. and Wüthrich, K. (1998) *J. Biomol. NMR*, **12**, 1–23.

McConnell, K.J. and Beveridge, D.L. (2000) *J. Mol. Biol.*, **304**, 803–820.

Metzler, W., Wang, C., Kitchen, D., Levy, R. and Pardi, A. (1990) *J. Mol. Biol.*, **214**, 711–736.

Mollova, E.T., Hansen, M.R. and Pardi, A. (2000) *J. Am. Chem. Soc.*, **122**, 11561–11562.

Murphy, E.C., Zhurkin, V.B., Louis, J.M., Cornilescu, G. and Clore, G.M. (2001) *J. Mol. Biol.*, **312**, 481–499.

Nelson, H.C., Finch, J.T., Luisi, B.F. and Klug, A. (1987) *Nature*, **330**, 221–226.

Nerdal, W., Hare, D.R. and Reid, B.R. (1989) *Biochemistry*, **28**, 10008–10021.

Ojha, R.P., Dhingra, M.M., Sarma, M.H., Shibata, M., Farrar, M., Turner, C.J. and Sarma, R.H. (1999) *Eur. J. Biochem.*, **265**, 35–53.

Olson, W.K. and Zhurkin, V.B. (1995) In *Biological Structure and Dynamics*, Sarma, R.H. and Sarma, M.H. (Eds.), Adenine Press, Albany, NY, pp. 341–370.

Prosser, R.S., Losonczi, J.A. and Shiyanovskaya, I.V. (1998) *J. Am. Chem. Soc.*, **120**, 11010–11011.

Rinkel, L.J., van der marel, G.A., vanBoom, J.H. and Altona, C. (1987) *Eur. J. Biochem.*, **166**, 87–101.

Ruckert, M. and Otting, G. (2000) *J. Am. Chem. Soc.*, **122**, 7793–7797.

Schwieters, C.D., Kuszewski, J., Tjandra, N. and Clore, G.M. (2003) *J. Magn. Reson.*, **160**, 65–73.

Shatzky-Schwartz, M., Arbuckle, N.D., Eisenstein, M., Rabinovich, D., Bareket-Samish, A., Haran, T.E., Luisi, B.F. and Shakked, Z. (1997) *J. Mol. Biol.*, **267**, 595–623.

Shui, X., Mcfail-Isom, L., Hu, G.G. and Williams, L.D. (1998a) *Biochemistry*, **37**, 8341–8355.

Shui, X., Sines, C.C., Mcfail-Isom, L., VanDerveer, D. and Williams, L.D. (1998b) *Biochemistry*, **37**, 16877–16887.

Sines, C.C., McFail-Isom, L., Howerton, S.B., VanDerveer, D. and Williams, L.D. (2000) *J. Am. Chem. Soc.*, **122**, 11048–11056.

Sklenar, V. and Bax, A. (1987) *J. Am. Chem. Soc.*, **109**, 7525–7526.

Skrynnikov, N., Goto, N., Yang, D., Choy, W., Tolman, J., Mueller, G. and Kay, L. (2000) *J. Mol. Biol.*, **295**, 1265–1273.

Taylor, R., Kennard, O. and Versichel, W. (1984) *J. Am. Chem. Soc.*, **106**, 244–248.

Tereshko, V., Minasov, G. and Egli, M. (1999a) *J. Am. Chem Soc.*, **121**, 470–471.

Tereshko, V., Minasov, G. and Egli, M. (1999b) *J. Am. Chem. Soc.*, **121**, 3590–3595.

Tian, F., Bolon, P.J. and Prestegard, J.H. (1999) *J. Am. Chem. Soc.*, **121**, 7712–7713.

Tian, F., Fowler, C.A., Zartler, E.R., Jenney, F.A., Adams, M.W. and Prestegard, J.H. (2000) *J. Biomol. NMR*, **18**, 23–31.

Tjandra, N. and Bax, A. (1997) *Science*, **278**, 1111–1114.

Tjandra, N., Grzesiek, S. and Bax, A. (1996) *J. Am. Chem. Soc.*, **118**, 6264–6272.

Tjandra, N., Marquardt, J. and Clore, G.M. (2000a) *J. Magn. Reson.*, **142**, 393–396.

Tjandra, N., Tate, S., Ono, A., Kainosho, M. and Bax, A. (2000b) *J. Am. Chem. Soc.*, **122**, 6190–6200.

Tjandra, N., Tate, S.-i., Ono, A., Kainosho, M. and Bax, A. (2000c) *J. Am. Chem. Soc.*, **122**, 6190–6200.

Tolman, J.R., Al-Hashimi, H.M., Kay, L.E. and Prestegard, J.H. (2001) *J. Am. Chem. Soc.*, **123**, 1416–1424.

Tolman, J.R., Flanagan, J.M., Kennedy, M.A. and Prestegard, J.H. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 9279–9283.

Tolman, J.R., Flanagan, J.M., Kennedy, M.A. and Prestegard, J.H. (1997) *Nat. Struct. Biol.*, **4**, 292–297.

Tonelli, M. and James, T.L. (1998) *Biochemistry*, **37**, 11478–11487.

Torda, A.E., Scheek, R.M. and van Gunsteren, W.F. (1990) *J. Mol. Biol.*, **214**, 223–235.

Trantirek, L., Urbasek, M., Stefl, R., Feigon, J. and Sklenar, V. (2000) *J. Am. Chem. Soc.*, **122**, 10454–10455.

Ulyanov, N.B., Gorin, A.A., Zhurkin, V.B., Chen, B.C., Sarma, M.H. and Sarma, R.H. (1992) *Biochemistry*, **31**, 3918–3930.

van Wijk, J., Huckriede, B.D., Ippel, J.H. and Altona, C. (1992) *Meth. Enzymol.*, **211**, 286–307.

Varani, G., Aboul-ela, F. and Allain, F.H.-T. (1996) *Prog. Nucl. Magn. Reson. Spectrosc.*, **29**, 51–127.

Vermeulen, A., Zhou, H. and Pardi, A. (2000) *J. Am. Chem. Soc.*, **122**, 9638–9647.

Wang, J.C. (1979) *Nucl. Acids Res.*, **76**, 200–203.

Warren, J.J. and Moore, P.B. (2001) *J. Biomol. NMR*, **20**, 311–323.

Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1980) *Nature*, **287**, 755–758.

Woods, K.K., McFail-Isom, L., Sines, C.C., Howerton, S.B., Stephens, R.K. and Williams, L.D. (2000) *J. Am. Chem. Soc.*, **122**, 1546–1547.

Wu, Z. and Bax, A. (2001) *J. Magn. Reson.*, **151**, 242–252.

Wu, Z., Tjandra, N. and Bax, A. (2001a) *J. Biomol. NMR*, **19**, 367–370.

Wu, Z.R., Tjandra, N. and Bax, A. (2001b) *J. Am. Chem. Soc.*, **123**, 3617–3618.

Yang, J., McAteer, K., Silks, L.A., Wu, R., Isern, N.G., Unkefer, C.J. and Kennedy, M.A. (2000) *J. Magn. Reson.*, **146**, 260–276.

Young, M.A. and Beveridge, D.L. (1998) *J. Mol. Biol.*, **281**, 675–687.

Young, M.A., Ravishanker, G. and Beveridge, D.L. (1997) *Biophys. J.*, **73**, 2313–2336.

Zweckstetter, M. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 3791–3792.