Determination of DNA Conformational Features from Selective Two-Dimensional NMR Experiments

Lyndon Emsley, Tammy J. Dwyer,† H. Peter Spielmann, and David E. Wemmer*

Contribution from the Materials Sciences Division and Structural Biology Division, Lawrence Berkeley Laboratory, and Department of Chemistry, University of California, Berkeley, California 94720

Received March 9, 1993

Abstract: Selective two-dimensional NMR correlation experiments are demonstrated to be particularly well suited to the problem of obtaining quantitative three-bond coupling constants in samples of DNA. The techniques were demonstrated by obtaining high-resolution spectra of cross peaks between sugar protons, which were then iteratively fitted to simulations derived from trial coupling topologies to determine the coupling constants which contribute to the cross-peak patterns. These coupling constants were then used to determine structural parameters such as the phase angle of the deoxyribose pseudorotation and the backbone torsion angles δ and ϵ. Examples are given for both single-stranded and duplex DNA. We have studied both the unmodified 8-mer duplex $d(GCGTACGC)_2$ and the furan-side monoadduct formed between the thymidine of one of the strands and 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen (HMT) of the same DNA sequence. We comment on the structural differences between the two.

Introduction

The structural and conformational characterization of duplex DNA oligonucleotides presents a continuing challenge to modern NMR spectroscopy. Whereas the determination of protein structures has reached a stage where molecules of up to 20-kD molecular weight are nearly routine,[1-4] the determination of three-dimensional nucleic acid structures is still much more difficult. Compared to 2D NMR analysis in proteins using NOESY-derived restraints, structure determination in nucleic acids suffers from the absence of any long-range distance information. Another reason for the dearth of reliable NMR structures of DNA oligomers stems from the inherent difficulty in the measurement and interpretation of sugar ring conformations due to both signal overlap in the proton spectra and difficult to interpret ring dynamics. The structural character of a DNA double helix is highly dependent on the conformations adopted by the phosphodiester backbone. Determination of the torsion angles that define the phosphate backbone structure, via measurement of heteronuclear scalar couplings $^3J_{HP}$, is often difficult.

Given the small number of interproton distances available for structure determination in nucleic acids, it is essential to also obtain the three-bond coupling constants relevant to derive the torsion angles that determine the geometry of the backbone and sugar rings. Thus far, measurement of these coupling constants has been feasible only from an analysis of the fine structure of cross peaks in 2D COSY or TOCSY-type spectra.[5-11] Over the past several years, the use of such spectra to extract scalar couplings has resulted in the development of programs such as SPHINX and LINSHA.[7] Although these approaches have met with reasonable success, a serious drawback is that, due to the lack of digital resolution inherent in conventional 2D NMR experiments, the accuracy of the determination of the coupling constants contributing to the structure of a particular cross peak is usually rather poor.[9]

In this article we show that many of the problems associated with the determination of coupling constants from two-dimensional spectra can be eliminated by using selective correlation experiments. We show that these techniques are particularly useful to study nucleic acid oligomers and that the data obtained can be used to accurately determine the conformations of the sugar rings and, in favorable cases, the conformation of the backbone. Examples for both a tetramer of single-stranded DNA and two larger octamer duplex DNA fragments are presented.

Selective Two-Dimensional NMR and Coupling Topology

Figure 1 shows the two pulse sequences that have been used in this work. The basic selective COSY sequence[12-14] of Figure 1 involves the application of three 270° Gaussian pulses.[15] As shown in the operator evolution graph[16] beneath the sequence, the first two pulses applied at the precession frequency Ω_{A1}, separated by the incremented delay t_1, serve to label the magnetization with the frequency of the spin A_1 during t_1. In-phase coherence of A_1 evolves during t_1 into anti-phase coherence between A_1 and A_2 which is converted by the second pulse into longitudinal two-spin order. Finally, the last pulse, applied at a frequency Ω_{A2} converts this into anti-phase coherence of A_2 which evolves into an observable in-phase coherence of A_2 during t_2. The result of this experiment is a cross peak between the two *active* spins A_1 and A_2.

Detailed discussions of soft-COSY can be found in refs 12–15, 22, and 23, so we merely recapitulate the main features of the

* Author to whom correspondence should be addressed.
† Present address: College of Arts and Sciences, California State University San Marcos, San Marcos, CA 92096.

(1) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632.
(2) Akke, M.; Drakenberg, T.; Chazin, W. J. *Biochemistry* **1992**, *31*, 1011.
(3) Knegtel, R. M. A.; Boelens, R.; Ganadu, M. L.; Kaptein, R. *Eur. J. Biochem.* **1991**, *202*, 447.
(4) Kallen, J.; Spitzfaden, C.; Zurini, M. G. M.; Wider, G.; Widmer, H.; Wüthrich, K.; Walkinshaw, M. D. *Nature* **1991**, *353*, 276.
(5) Ernst, R. R.; Bodenhausen, G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Clarendon Press: Oxford, England, 1987.
(6) Kessler, H.; Gehrke, M.; Griesinger, C. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 490.
(7) Widmer, H.; Wüthrich, K. *J. Magn. Reson.* **1990**, *70*, 270.
(8) Titman, J. J.; Keeler, J. *J. Magn. Reson.* **1990**, *89*, 640.
(9) Majumdar, A.; Hosur, R. V. *Prog. NMR Spectrosc.* **1992**, *24*, 109.
(10) Macaya, R. F.; Schultze, P.; Feigon, J. *J. Am. Chem. Soc.* **1992**, *114*, 781.
(11) Schmitz, U.; Zon, G.; James, T. L. *Biochemistry* **1990**, *29*, 2357.
(12) Brüschweiler, R.; Madsen, J. C.; Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Magn. Reson.* **1987**, *73*, 380.
(13) Cavanagh, J.; Waltho, J. P.; Keeler, J. *J. Magn. Reson.* **1987**, *74*, 386.
(14) Emsley, L.; Bodenhausen, G. In *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy*; Hoch, J. C., Ed.; Plenum Press: New York, 1991.
(15) Emsley, L.; Bodenhausen, G. *J. Magn. Reson.* **1989**, *82*, 211.
(16) Eggenberger, U.; Bodenhausen, G. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 374.

0002-7863/93/1515-7765$04.00/0

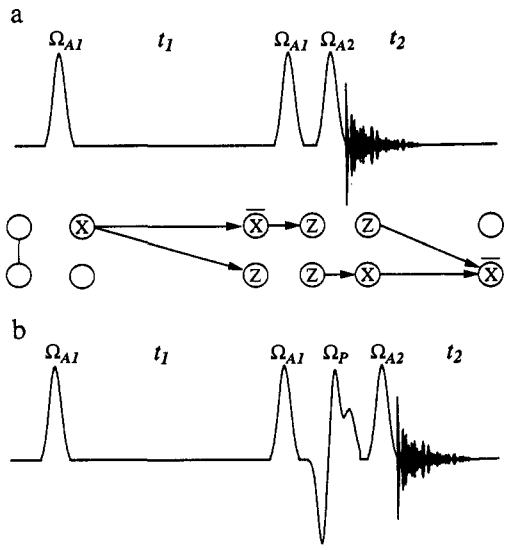


Figure 1. Pulse sequences used to obtain (a) soft-COSY spectra and (b) SIS-COSY spectra. The operator evolution graph shown below the sequence in part a illustrates how coherence transfer is achieved via anti-phase coherence of the two spins involved in the cross peak.

experiment. Soft-COSY has at least two distinct advantages over normal nonselective double-quantum-filtered COSY (DQF-COSY), the experiment most commonly used to obtain quantitative coupling constants.

Firstly, the definition of the fine structure of cross peaks in DQF-COSY spectra is normally limited by poor digital resolution (typically 3–6 Hz/point in ω_1). The use of selective pulses allows one to obtain spectra with small spectral widths (typically around 100 Hz), so that with the acquisition of a relatively small number of points one can obtain a very high digital resolution in a short time (typically 10 min is required to obtain a resolution of <1 Hz/point in ω_1). In this case, the definition of the fine structure is limited purely by constraints of the natural line widths. Secondly, because of the selective nature of the experiments, the fine structure of the cross peaks is simpler than the DQF-COSY cross-peak patterns. In fact, the cross-peak structure is the same as that observed in E.COSY-[17] or *z*-COSY[18]-type spectra. These spectra are characterized by anti-phase square patterns with an active splitting $J_{A_1A_2}$ and displacements representing couplings $J_{A_1P_i}$ to *passive* spins in t_1 and $J_{A_2P_i}$ in t_2. Passive spins are coupled to A_1 and/or A_2 but are not directly observed in the experiment. Such patterns have already been shown to be particularly suitable for automated analysis in terms of all the couplings that contribute to the pattern.[10,19,20] However, measurement of *J*-couplings by E.COSY is limited in accuracy by broad lines and digital resolution.[21] To obtain accurate coupling-constant information, one must apply methods which do not suffer from poor digital resolution.

The second sequence in Figure 1 is a selectively inverted soft-COSY or SIS-COSY sequence[22,23] which is used to assign the passive couplings observed in a particular cross peak (see the example of the TATA 1′2′ and 2′2″ cross peaks shown in Figure 2). By inverting a passive spin P_i between the two last pulses with a selective 180° pulse (in this case a Q^3 Gaussian pulse cascade) applied at the frequency Ω_{P_i}, the apparent sign of the coupling $J_{A_2P_i}$ is reversed in ω_2, allowing simple assignment of the different passive couplings.

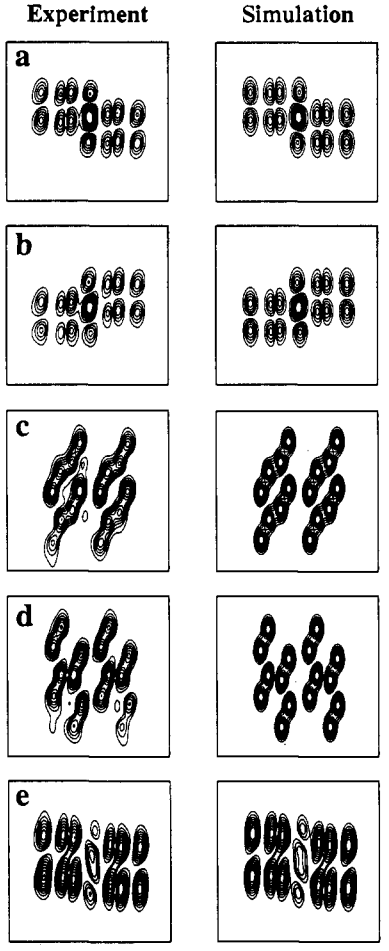


Figure 2. Comparison between experimental soft-COSY and SIS-COSY spectra and their "best fit" simulations. Spectra a, c, and e are soft-COSY spectra of the 1′2′, 2′2″, and 2′3′ cross peaks in the sugar ring of T1 in TATA. Spectrum b is a SIS-COSY spectrum of the 1′2′ cross peak in which a 180° pulse has been applied to the 2″ resonance between the last two pulses (see Figure 1). In this way the 2′2″ coupling appears reversed in sign in ω_2, allowing easy discrimination between the 2″ and 3′ passive couplings. Similarly spectrum d is a SIS-COSY spectrum of the 2′2″ cross peak, in which the 1′ spin has been inverted. All spectra were recorded under the same conditions with 8 scans for each of 64 points in t_1 and 1024 points in t_2 zero filled to 256 × 1024 before being apodized with a Lorentz to Gauss function (LB1 = −3, GB1 = 0.1, LB2 = −3, GB2 = 0.1) and Fourier transformed. We used 270° Gaussian pulses to ensure in-phase excitation. The pulses were all 24 ms long, giving a bandwidth of ≈40 Hz. 50-ms Q^3 inversion pulses, also with a bandwidth of ≈40 Hz, were used to invert the passive spins in spectra b and d. Spectral widths were 100 Hz in ω_1 and 500 Hz in ω_2, of which 50 × 50 Hz sections are shown in the figure.

Once a high-resolution soft-COSY cross peak has been obtained, it can be analyzed to yield a fragment of the coupling network to which it belongs. In the terminology of Pfändler *et al.*,[24] a network represents all the interconnected couplings in a particular compound, while a fragment is that part of the network which can be derived from a single cross peak. In order to determine fragments from cross peaks, we have adopted the approach of iteratively fitting the data to different fragment topologies and coupling constants. In practice this means using the couplings to generate a simulated cross peak and performing a least squares fitting procedure of the simulation to the experimental spectrum.

This analysis is particularly suitable for cross peaks between sugar protons in DNA. By recording only two different cross

(17) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1985**, *107*, 6394.
(18) Oschkinat, H.; Pastore, A.; Pfändler, P.; Bodenhausen, G. *J. Magn. Reson.* **1986**, *69*, 559.
(19) Mádi, Z.; Meier, B. U.; Ernst, R. R. *J. Magn. Reson.* **1987**, *72*, 584.
(20) Pfändler, P.; Bodenhausen, G. *J. Magn. Reson.* **1988**, *79*, 99.
(21) Gochim, M.; Zon, G.; James, T. L. *Biochemistry* **1990**, *29*, 11161.
(22) Emsley, L.; Huber, P.; Bodenhausen, G. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 517.
(23) Emsley, L.; Bodenhausen, G. *J. Am. Chem. Soc.* **1991**, *113*, 3309.
(24) Pfändler, P.; Bodenhausen, G. *J. Magn. Reson.* **1990**, *87*, 26.

Table I. ^3J Coupling Constants Which Can Be Measured from Cross Peaks between Protons in the Sugar Rings of DNA Fragments[a]

	^3J coupling constant					
cross peak	1′2′	1′2″	2′3′	2″3′	3′4′	3′P
1′2′	A	P	P	–	–	–
1′2″	P	A	–	P	–	–
2′2″	P	P	P	P	–	–
2′3′	P	–	A	P	P	P
2″3′	–	P	P	A	P	P
3′4′	–	–	P	P	A	P

[a] A indicates that the coupling appears in the cross peak as an anti-phase *active* coupling, P indicates that the coupling is a *passive* in-phase displacement, and a dash indicates that the coupling does not contribute to the structure of the cross peak.

peaks, such as the 1′2′ and 2′3′ peaks, we can determine the whole coupling network by combining the two fragments we obtain. As shown in Table I, each cross peak between protons in a sugar ring of DNA contains information about three or four three-bond coupling constants. As can be appreciated from the table, measurement of all the cross peaks would provide four or five *independent* estimates of each coupling constant. Any two of the three bond coupling constants can be used to determine the geometry of the sugar ring.[9,25] The overdetermined nature of these measurements is extremely useful. While the error of the fitting procedure determines the *precision* of the fit to the couplings (which can be in the mHz range), this is in fact misleading and only represents how well the algorithm can locate a minimum for the error function. It does not reflect the effect of systematic and random errors such as distortions in the line shapes and amplitudes caused by experimental artifacts. The spread of repeated independent measurements is more likely to tell us about the *accuracy* of the measured couplings, which for our experimental conditions varies from 0.2 to 1.0 Hz. It is the accuracy of the determination, and not the precision, which defines the range of torsion angles that are compatible with the data.

In order to exemplify the usefulness of selective correlation techniques for characterizing structure and conformation in DNA, we have used the techniques outlined above to record spectra from both the single-stranded DNA oligomer 5′-TATA-3′ and duplex DNA oligomers of the sequence d(GCGTACGC)$_2$. We have recorded spectra of both the unmodified 8-mer duplex (UM) and a 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen (HMT) monoadducted 8-mer heteroduplex molecule (MAf).[26] Because TATA is small and single stranded under the experimental conditions, it has narrow natural line widths (<1 Hz). Both the unmodified 8-mer and the HMT monoadducted molecule form duplexes under the experimental conditions, leading to much broader lines in the spectrum (approximately 10–20 Hz).

Experimental Section

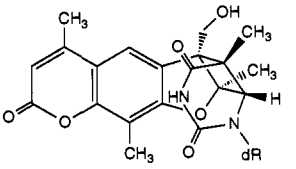
The DNA oligomers were synthesized by using the phosphite triester method[27,28] on an Applied Biosystems Inc. automated DNA synthesizer. The DNA oligomer samples were cleaved from the solid support and deprotected using standard methods. The TATA was purified on a Sephadex G-10 column and dissolved in a 10 mM sodium phosphate buffer, pH 7.0. The sample was lyophilized several times from D_2O and finally dissolved in 99.96% D_2O. The resulting solution was 10 mM in single-stranded tetramer. The unmodified 8-mer and MAf were synthesized and purified as described in ref 26. The unmodified 8-mer was dissolved in 200 mM NaCl, 10 mM deuterated Tris-HCl (d_{11}), pH 7.0, and lyophilized from D_2O repeatedly. The resulting solution was 1.5 mM in UM duplex.

To form the heteroduplex, the single-stranded monoadducted oligomer was taken up in 80 μL of 1 M NaCl and the unmodified oligomer was taken up in 200 μL of 20 mM deuterated Tris-HCl (d_{11}), pH 7.0, 400 mM NaCl and lyophilized. The samples were lyophilized several times from D_2O, and finally the 8-mer MAf was dissolved in 400 μL and the unmodified 8-mer was dissolved in 70 μL of 99.996% D atom D_2O. The final Na^+ and deuterated Tris concentrations in these two solutions were 200 mM and 10 mM, respectively. The single-stranded 8-mer MAf was then titrated with the more concentrated unmodified 8-mer to give duplex DNA. The titration was monitored by observing the appearance of two resonances assigned to the adenine C8-H protons in the ^1H NMR at 600 MHz. Aliquots of UM were added until the peaks corresponding to single-stranded MAf disappeared. The resulting solution was 2 mM in MAf heteroduplex.

The unmodified 8-mer duplex is self-complementary, and we use the following numbering scheme:

5′-G1 C2 G3 T4 A5 C6 G7 C8-3′
3′-C16 G15 C14 A13 T12 G11 C10 G9-5′

Because of its self-complementary nature, there are only eight spectroscopically distinct residues. The sugar protons were assigned by working outward from the intraresidue aromatic proton to the H1′ and H2′, H2″ cross peaks to find the corresponding H3′ and H4′ resonances. In the MAf, the psoralen derivative (**1**) is covalently attached to T12 through



1
dT-HMT furanside monoadduct, *cis-syn* conformation
dR is deoxyribose

a cyclobutane ring formed between the 5,6 double bond of the thymidine and the 4′,5′ double bond of the psoralen. The twofold symmetry of the d(GCGTACGC)$_2$ self-complementary duplex is broken upon HMT adduct formation because only one of the two strands in the heteroduplex is covalently modified by the drug. The MAf heteroduplex is more stable than UM or the MAf homoduplex.[29] The strand-specific assignment of the now unsymmetric DNA base sequence was resolved by observation of NOE connectivities between the psoralen protons and the thymidine H6 and methyl protons and the adenosine H8 protons. The sequential aromatic to H1′ NOE assignments of the two strands are interrupted at the 5′-TpA-3′ step in both strands of the duplex. The resonances of T4 can be connected to those of A5, and those of T12 can be connected to those of A13 through the HMT proton resonances in the MAf. These results indicate that the HMT is intercalated in the central 5′-TpA-3′ site in the MAf. In the heteroduplex, T4 is unmodified.

Two-dimensional soft-COSY and SIS-COSY spectra were recorded using the sequences described above on a 600-MHz Bruker AM-600X spectrometer at a temperature of 303 K. Spectral widths ranged from 100 to 150 Hz in t_1 and from 250 to 500 Hz in t_2. Data sets were zero filled before being weighted with a Lorentz to Gauss apodization function before Fourier transform and phase correction. Specific parameter values are given in the figure captions. Typically, digital resolution was around 1.5 and 0.4 Hz/point in ω_1 before and after Fourier transform and 1.0 and 0.24 Hz/point in ω_2. Spectra were recorded in times ranging from around 10 min for high-concentration samples of TATA to about 3 h for the low-concentration samples of d(GCGTACGC)$_2$. After Fourier transform the spectra were transferred to a Stardent 3000 Titan mini-supercomputer where the data were fitted to simulations using the MIGRAD algorithm from the MINUIT package in the CERNLIB program library available from CERN. The algorithm uses a variable metric method incorporating inexact line search, a stable metric updating scheme, and checks for positive definiteness. The fitting process typically took around 30 min of CPU time to converge, using around 300 calls to optimize the 10 parameters in a fragment with two passive couplings. The optimization takes into account the active coupling, the passive couplings, the Gaussian and Lorentzian contributions to the linewidths in each dimension, the offsets from zero in each dimension, and the absolute intensity.

(25) Altona, C.; Sundralingham, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.
(26) Spielmann, H. P.; Sastry, S. S.; Hearst, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4515.
(27) Letsinger, R. L.; Lunsford, W. B. *J. Am. Chem. Soc.* **1976**, *98*, 3655.
(28) Matteucci, M. D.; Caruthers, M. H. *J. Am. Chem. Soc.* **1981**, *103*, 3185.
(29) Shi, Y. B.; Hearst, J. E. *Biochemistry* **1986**, *25*, 5895.

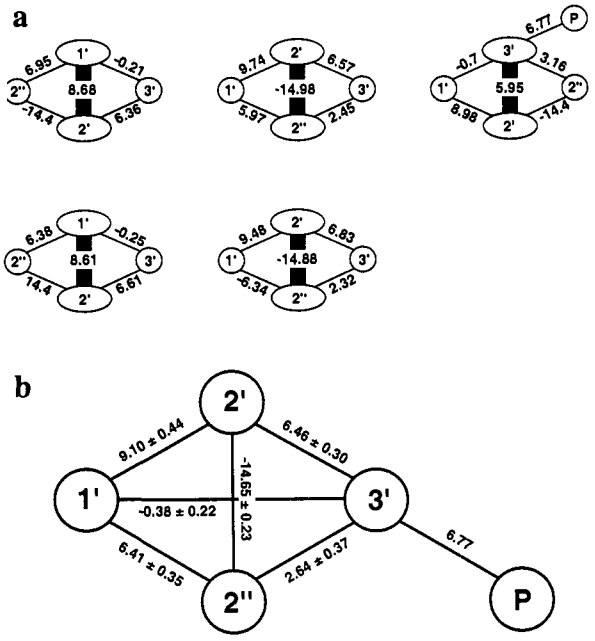


Figure 3. (a) Fragments of the coupling network obtained from the simulations of Figure 2. The solid lines represent couplings observed in the experiment, with the derived numerical values. Ovals show active spins, with A_1 above A_2, and circles show passive spins. The global coupling network derived using these fragments is shown in part b. The numerical values are the mean observed coupling and the one standard deviation error. Note that no error is given for the 3'P coupling, since we only measure it once.

Results and Discussion

Single-Stranded DNA: TATA. In Figure 2 we show three experimental soft-COSY multiplets obtained from the T1 sugar protons in a sample of 5'-TATA-3'. For the 1'2' and 2'2'' cross peaks we also show the SIS-COSY multiplets that were used to determine which of the passive displacements corresponded to which passive coupling. The right-hand side of the figure shows simulations of spectra which are the "best fit" to the experiments. The reduced χ^2 is defined as

$$\chi^2 \equiv \frac{\sum_{i=1}^{n}(x_i - e_i)^2}{n\sigma^2} \qquad (1)$$

where n is the number of data points, $(x_i - e_i)$ is the difference between the observed and simulated points, and σ^2 is the variance of the noise. The values obtained for these fits are <5. This highlights the excellent agreement between theory and experiment. The precision of these fits, as represented by the one standard deviation error, is less than ±0.01 Hz. As mentioned above, this precision is misleading. The accuracy of the measurement can only be obtained by comparing the spread of values obtained from several *independent* determinations. In Figure 3 we show the fragments derived from each of the spectra in Figure 2. It can be appreciated that we have made five separate and independent measurements of the 1'2' coupling constant. These five values can be simply combined to give a mean of 9.10 Hz and a standard deviation of 0.44 Hz. It is this standard deviation which should reflect the accuracy of the measurement of $J_{1'2'}$, within the scope of the NMR technique. It is important to note that this is not equivalent to values obtained by repeatedly measuring the *same* cross peak. The contributions to line widths, line shapes, and cancellation effects are different for each individual cross peak, so that a measurement of a single coupling constant from several *different* cross peaks is immune to systematic errors that could be caused by these effects and reflects the true random error. Indeed, the cross-peak patterns observed in the spectra of Figure 2 are clearly very different from one peak to another (c.f. Figure 2a,c,e). Note that all our measured coupling constants, summarized in Figure 3b, agree well with those previously measured by Mellema *et al.*[30] In this way we have demonstrated that in a short amount of experimental time (all five spectra for T1 were obtained in a total time of less than 1 h) we are able to obtain reliable data which can be used to determine couplings to a high accuracy.

Derivation of geometry from three-bond coupling constants is a well-established procedure which can be found discussed in many articles.[31] We have used the modified Karplus equation,[32,33] provided by Altona and co-workers[34,35] for the protons in the sugar rings of DNA, to determine torsion angles from the coupling constants. The equation is

$$^3J_{HH} = P_1 \cos^2 \phi_{HH} + P_2 \cos \phi_{HH} + P_3 + \sum[\Delta\chi_i(P_4 + P_5 \cos^2(\xi\phi_{HH} + P_6|\Delta\chi_i|))] \qquad (2)$$

where the coefficients $P_1, \ldots, P_6$ are empirically derived constants and the $\Delta\chi_i$ are the electronegativities of the substituents. Of the four three-bond couplings we have measured, we present a representative derivation of the ring conformation using the 1'2' and 2'3' couplings, for which the values of Figure 3 used in eq 2 lead to torsion angles $\phi_{1'2'}$ and $\phi_{2'3'}$ of 160.7 ± 6.6° and −38.6 ± 2.4°, respectively.

The sugar rings in DNA can exist in at least two distinct solution conformations[25,30,34] referred to as N-type (characterized by a pseudorotation phase angle $P_s = 0 \pm 90°$), which is the predominant conformation in B-DNA, and S-type (characterized by a pseudorotation phase angle $P_s = 180 \pm 90°$), which is the predominant conformation in A-DNA. In reality a sugar ring need not exist in a static conformation, but rather can rapidly interconvert (on the NMR time scale) between the N-type and S-type conformers. Calculation of the pseudorotation phase angle, P_s, and the maximum amplitude of pucker ϕ_m for a residue affords an estimation of the geometry of a sugar ring in terms of N- and S-types.[34] Using this model for the DNA conformation, the torsion angles can be used to obtain a pseudorotation phase angle P_s and a maximum amplitude for the ring pucker ϕ_m using the following equations[25,34]

$$\phi_{1'2'} = 121.4 + 1.03\phi_m \cos(P_s - 144) \qquad (3)$$

$$\phi_{2'3'} = 2.4 + 1.06\phi_m \cos(P_s) \qquad (4)$$

$$\phi_{1'2''} = 0.9 + 1.05\phi_m \cos(P_s - 144) \qquad (5)$$

(similar equations exist for the other measurable torsion angles in the ring). Using these equations we obtain $P_s = 163 \pm 13°$ and $\phi_m = 40 \pm 4°$ for the sugar ring of T1.

We can determine the backbone torsion angle δ (the 4'3' torsion angle) from our measurement of P_s and ϕ_m. The determination of any two torsion angles completely defines the conformation of the deoxyribose ring. This is done using the equation[36]

$$\delta = 120.6 + 1.1\phi_m \cos(P_s + 145.3) \qquad (6)$$

which yields $\delta = 142 \pm 9°$. All of the above values are in agreement with the normal range found in DNA.

Finally, a second backbone torsion angle can be obtained from the measurement of the heteronuclear coupling to phosphorus

(30) Mellema, J. R.; Pieters, J. M. L.; Marel, G. A. v. d.; Boom, J. H. v.; Haasnoot, C. A. G.; Altona, C. *Eur. J. Biochem.* **1984**, *143*, 285.
(31) Reference 9 and references therein.
(32) Karplus, M. J. *J. Chem. Phys.* **1959**, *30*, 11.
(33) Karplus M. J. *J. Am. Chem. Soc.* **1963**, *85*, 2870.
(34) Rinkel, L. J.; Altona, C. *J. Biomol. Struct. Dyn.* **1987**, *4*, 621.
(35) Haasnoot, C. A. G.; Leeuw, F. A. A. M. d.; Altona, C. *Tetrahedron* **1980**, *36*, 2783.
(36) Leeuw, F. A. A. M. d.; Kampen, P. N. v.; Altona, C.; Diez, E.; Esteban, A. L. *J. Mol. Struct.* **1984**, *125*, 76.

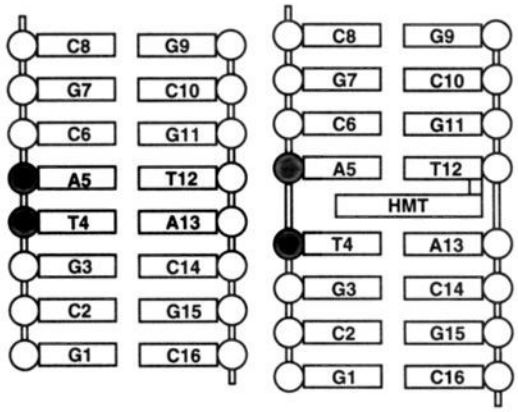


Figure 4. Schematic diagram of the unmodified and HMT adducted DNA 8-mer duplexes. The bases are represented by rectangles, and the sugars are represented by circles. The sugars that were studied are shaded. The sugar in T4 of the unmodified DNA 8-mer has a conformational purity greater than 80% C2′ endo. A5 in the unmodified DNA 8-mer also appears in the normal DNA conformation with greater than 70% C2′ endo. In contrast we find that the sugar of the HMT photomodified base in residue T4 of the MAf is in neither a C2′ endo or C3′ endo conformation and can be best described as a dynamically averaged fluxional mixture of the two states. A5 in the MAf can be best described as having a conformational purity of greater than 60% C2′ endo.

$^3J_{3'P}$. As is indicated in Table I, this coupling appears as a passive displacement in 2′3′, 2″3′, and 3′4′ cross peaks. In T1 we have measured the 2′3′ cross peak, which yields $^3J_{3'P} = 6.77$ Hz. This can be used to derive the ϵ torsion angle (the H3′–C3′–O–P angle) by finding the roots of the equation[37]

$$^3J_{HP} = 15.3 \cos^2 \epsilon - 6.1 \cos \epsilon + 1.6 \qquad (7)$$

Using the 3′P coupling in T1, we obtain roots for $\epsilon = 245°$ and 115°. One should note that neither of these roots is close to the values typically found for ϵ (210° or 270°)[9] in normal duplex DNA. We are probably observing a dynamic average value due to the rapid rotation about the C–O bond in this terminal residue of the single-stranded molecule.

Duplex DNA. The example of single-stranded TATA shows that soft-COSY provides a reliable way of determining conformation and geometry of sugar rings. However TATA is not a particularly convincing "real" example. As mentioned above, the TATA spectrum has narrower lines than those of most biopolymers of interest. More realistic DNA samples that are double stranded often have line widths an order of magnitude larger. To show that soft-COSY techniques are useful in these systems, we have applied them to the study of d(GCGTACGC)$_2$. The interest in this case is to compare the structure and dynamics of the unmodified duplex with that of the modified HMT furan-side monoadduct on T12, as shown in Figure 4.[26] Since psoralens are photomutagens and suspected carcinogens, an important structural question is how much and in what way does the drug adduct distort the DNA.

Toward this end we have obtained soft-COSY cross peaks for residues T4 and A5 in both the unmodified duplex and the monoadduct. The spectra of the 1′2′, 1′2″, and 2′2″ cross peaks are shown in Figures 5 and 6, respectively. The spectra of the 1′2′ and 1′2″ cross peaks result in best fits that are both accurate and precise. These fits are not much less accurate than those for the TATA data, even though the broader lines lead to multiplets in these larger compounds that are not as aesthetically pleasing as those in Figure 2. The resulting simulations are shown on the right-hand side of the figures. However, one should note that large line widths do cause problems. Even though both types of peak provide "good" fits (i.e. precise fits in the sense discussed above), the quality of the data for the 2′2″ cross peaks resulted in best fits that were much less reliable than those obtained with 1′2′ and 1′2″ data. The structure of the resonances in the 2′2″ cross peaks allowed for multiple solutions for the values of active and passive couplings, some of which are clearly unphysical. Even when values for coupling constants determined for the 1′2′ and 1′2″ cross peaks were used as starting points in the fitting routine for the 2′2″ cross peaks, widely divergent solutions were produced. This observation is not particularly surprising, as the 2′ and 2″ resonances have significantly larger line widths (24 Hz) than the 1′ resonances (13 Hz).

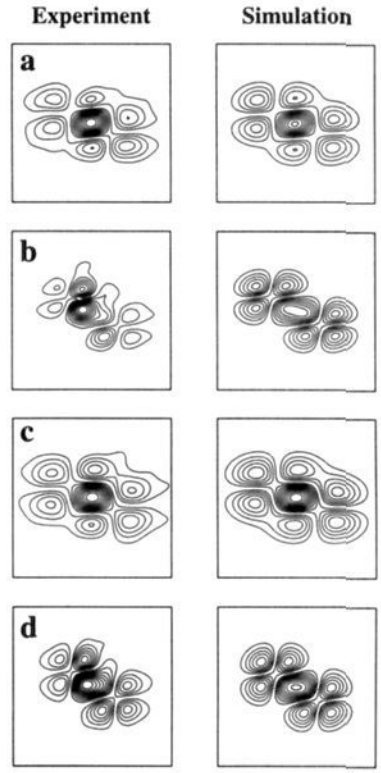


Figure 5. Comparison between experimental soft-COSY spectra of d(GCGTACGC)$_2$ and their "best fit" simulations. In parts a and b we show the 1′2′ and 1′2″ cross peaks, respectively, from T4, and in parts c and d we show the same for A5. The spectra were recorded with between 64 and 96 scans for each of 40–60 points in t_1 and 512 points in t_2 zero filled to 512 × 1024 before being apodized with a Lorentz to Gauss function (LB1 = −3, GB1 = 0.1, LB2 = −3, GB2 = 0.1) and Fourier transformed. Pulse lengths were 17.6 ms for the 270° Gaussian pulses. Spectral widths were 200 Hz in ω_1 and 500 Hz in ω_2, of which 60 × 60 Hz sections are shown in the figure.

The fragments which result from the 1′2′ and 1′2″ cross peaks are shown in Figure 7. (Note that in order to minimize the number of variables the geminal $^2J_{H2'-H2''}$ was fixed at −14.5 Hz in the fitting routine used to analyze the cross peaks for the UM and MAf 8-mers.) Using the methods of Rinkel and Altona,[34] analysis of the coupling constants allows us to determine the percent C2′ endo and phase angle P_s assuming a sugar pucker amplitude ϕ_m of 35°. Given that we estimate the values of the coupling constants determined for these DNA oligomers to have an accuracy of ±0.5 Hz, we determine the T4 in the unmodified DNA 8-mer to be in the normal DNA conformation with a P_s of between 171° and 189°. The conformational purity of this residue is greater than 80% C2′ endo. A5 in the unmodified DNA 8-mer also appears in the normal DNA conformation with a P_s of between 144° and 162°. The conformational purity of this residue is greater than 70% C2′ endo. In contrast we find that the sugar of the HMT photomodified base in residue T4 of the MAf is in neither a C2′ endo or C3′ endo conformation, and the simple two-state model is therefore not appropriate to describe this residue. A self-consistent fit of the data to a percent C2′ endo confor-

(37) Lankhorst, P. P.; Haasnoot, C. A. G.; Erkelens, C.; Altona, C. *J. Biomol. Struct. Dyn.* **1984**, *1*, 1387.

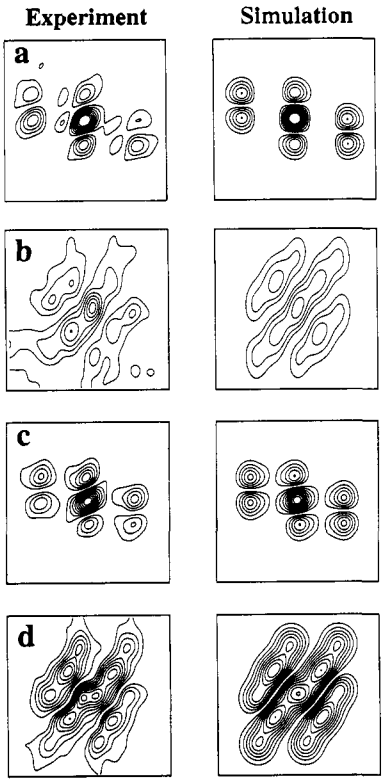


Figure 6. Comparison between experimental soft-COSY spectra of MAf and their "best fit" simulations. In parts a and b we show the 1′2′ and 2′2″ cross peaks, respectively, from T4 and in parts c and d we show the same for A5. The spectra were recorded with 64 scans for each of 64 points in t_1 and 256 points in t_2 zero filled to 256 × 1024 before being apodized with a Lorentz to Gauss function (LB1 = −3, GB1 = 0.1, LB2 = −3, GB2 = 0.1) and Fourier transformed. The 270° Gaussian pulses were 24 ms long. Spectral widths were 100 Hz in ω_1 and 250 Hz in ω_2, of which 50 × 50 Hz sections are shown in the figure. As discussed in the text, physically unreasonable solutions were obtained for fits of the 2′2″ peaks due to their broad lines.

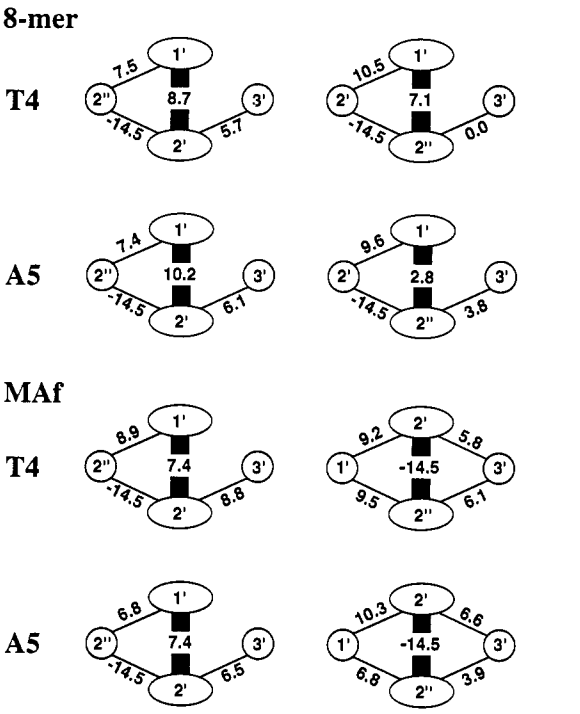


Figure 7. Fragments of the coupling networks derived from the spectra of Figures 5 and 6. The notation is the same as that of Figure 3. The upper two fragments correspond to the T4 and A5 cross peaks of the unmodified 8-mer, and the lower two correspond to the T4 and A5 cross peaks from the monoadduct. As discussed in the text, physically unreasonable solutions were obtained for fits of the 2′2″ peaks due to their broad lines.

mation was not possible. This sugar conformation can be best described as a dynamically averaged fluxional mixture of the two states. A5 in the MAf can be best described as having a P_s of between 90° and 108° and a conformational purity of approximately 60% C2′ endo. When these residues are analyzed in terms of a single major conformer, the measured coupling constants lead to mutually inconsistent values for the torsion angles. A more detailed analysis of the conformational equilibrium is beyond the scope of this work.

An important factor in structure determination by NMR is the measurement of how dynamic a system is and how any dynamics may effect the interpretation of the data. If there are structural isomers of the molecule that interconvert at rates that are either comparable to or faster than the NMR time scale, not only does this have obvious importance in terms of the biological activity of the molecule but steps must be taken to incorporate the dynamics into the interpretation of both nuclear Overhauser effect (NOE) and torsional restraints in molecular modeling simulations. Comparison of the coupling constants of the UM sugars with those of their MAf counterparts indicates a substantial conformational change in the helix segment at the site of photoalkylation and gives information on the magnitude of the local changes in helix dynamics (Figure 7). The sugar residues go from relatively pure C2′ endo conformations in the unmodified duplex to dynamic, rapidly averaging conformations at the site of damage by the drug. The significant population of both C2′ endo and C2′ exo conformations of the sugar residues has implications for the interpretation of interproton NOEs used in structure determination. Analysis of the NOEs in terms of a single sugar conformation can lead to incorrect and inconsistent distance restraints. Because the conformations of the residues in question cannot be analyzed in sufficient detail by the two-state model for deoxyribose conformations, the magnitudes of the changes in dynamics are interpreted in terms requiring the use of conservative bounds for NOE-derived distance constraints in structure determinations of this molecule. The soft-COSY method allows for the more accurate determination of both structural and dynamic parameters for DNA oligomers of biological interest and can be used to set upper and lower distance limits derived from NOEs that are observed for protons involved with the protons of the dynamic sugar residues. The range of interproton distances in the case of deoxyribose sugars in a duplex can be interpreted as being bounded by the changes in geometry between A-form C2′ exo and B-form C2′ endo conformations. The recognition of DNA damage by cellular enzyme systems has been proposed to take place by the detection of structural changes and/or the alteration of the dynamic behavior of the polymer.[38-40] The soft-COSY method allows for the accurate determination of both structural and dynamic parameters for DNA oligomers of biological interest.

Experimental Limitations. In the examples we have given the measurement and subsequent extraction of accurate coupling constants from soft-COSY experiments enable a complete analysis of the conformation of the deoxyribose rings in the DNA oligomers. In less favorable cases, where an analysis of coupling constants alone does not provide a sufficiently detailed picture, we suggest that these data be incorporated into the structure determination process using distance restraints determined from Overhauser effect measurements. In cases such as the MAf where an analysis of coupling constants indicates that a sugar residue cannot be described as being primarily one conformation, the

(38) Pu, W. T.; Kahn, R.; Munn, M. M.; Rupp, W. D. *J. Biol. Chem.* **1989**, *264*, 20697.
(39) Lin, J. J.; Sancar, A. *Biochemistry* **1989**, *28*, 7979.
(40) Sancar, A.; Sancar, G. B. *Annu. Rev. Biochem.* **1988**, *57*, 29.

upper and lower limits given to NOE-derived restraints used in structure determinations need to be modified to allow for the differences in interproton distances of the conformations available to the sugar. Interproton distances derived from relaxation matrix analysis of NOE volumes, such as IRMA[41] and MARDIGRAS,[42] must be interpreted with caution in these systems because the methods assume a structure that is primarily one conformation. The combination of constraints generated from NOEs and torsional angles that incorporate information about dynamics should provide increasingly accurate structures.

As we have seen, the soft-COSY technique is ultimately limited by the line width of the system under study. In the case of particularly large fragments, the experiment suffers from the dual drawbacks of short transverse relaxation times and cancellation due to anti-phase line shapes (note that this latter is shared with any of the nonselective techniques which generate anti-phase patterns). Transverse relaxation becomes a problem for the sensitivity of the experiments when the T_2 is on the order of the duration of the selective pulses (typically tens of milliseconds). Thus, the choice of pulse lengths in larger fragments represents a compromise between selectivity and sensitivity and results in an effective upper limit on the selectivity of the experiment.

Conclusions

We have demonstrated that soft-COSY and allied methods provide excellent high-resolution spectra for small DNA oligomers and that they allow for the extraction of useful structural information. Using the techniques outlined above it is possible to record high-resolution spectra in a relatively short time. These spectra are then simply analyzed in terms of the couplings that contribute to the multiplet patterns. In turn the couplings can be interpreted using a modified Karplus relation to yield structural parameters such as backbone torsion angles and the deoxyribose pseudorotation phase angle and ring pucker amplitudes. The experiments appear to be well suited to the type of spectra usually found in small DNA oligomers, and we have shown that they are useful in addressing real structural problems to extract useful modeling constraints. Note that because the technique relies on selective excitation, large impurity peaks (i.e. H_2O) can be tolerated, as long as they are not too close to the cross peak of interest. The main limitation of the technique is that signal to noise is degraded in larger fragments due to cancellation effects of the active anti-phase coupling, and we are currently investigating TOCSY-type experiments to try to avoid these problems.[43] The broad lines observed in some of the resonances lead to cancellation of the anti-phase multiplets. Despite this minor limitation, we predict that these experiments will be of value in studying structural features of small to medium-sized DNA fragments, such as those currently used in many DNA–drug binding studies.[44]

Acknowledgment. This work was supported by the National Institutes of Health Grant GM-43219 (D.E.W.) and postdoctoral fellowship GM-14966 (H.P.S.) and through instrumentation grants from the U.S. Department of Energy, DE FG05-86ER75281, and the National Science Foundation, DMB 86-09305 and BBS 87-20134, and by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. L.E. is a fellow of the Miller Institute for Basic Research in Science.

(41) Boelens, R.; Koning, T. M. G.; Kaptein, R. *J. Mol. Struct.* **1988**, *173*, 299.

(42) Borgias, M.; Gochin, M.; Kerwood, D. J.; James, T. L. *Prog. NMR Spectrosc.* **1990**, *22*, 83.

(43) Vincent, S. J. F.; Zwahlen, C.; Bodenhausen, G. *J. Am. Chem. Soc.*, in press.

(44) Gilbert, D. E.; Feigon, J. *Curr. Opin. Struct. Biol.* **1991**, *1*, 439.