The Octamer Motif in Immunoglobulin Genes: Extraction of Structural Constraints from Two-Dimensional NMR Studies†

Klaus Weisz,‡ Richard H. Shafer,‡ William Egan,§ and Thomas L. James*,‡

Departments of Pharmaceutical Chemistry and Radiology, University of California, San Francisco, California 94143-0446, and Biophysics Laboratory, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Maryland 20892

Received March 25, 1992; Revised Manuscript Received May 26, 1992

ABSTRACT: Phase-sensitive two-dimensional nuclear Overhauser enhancement (2D NOE) and double-quantum-filtered correlated (2QF-COSY) spectra were recorded at 500 MHz for the DNA duplex d(CATTTGCATC)·d(GATGCAAATG), which contains the octamer element of immunoglobulin genes. Exchangeable and nonexchangeable proton resonances including those of the H5′ and H5″ protons were assigned. Overall, the decamer duplex adopts a B-type DNA conformation. Scalar coupling constants for the sugar protons were determined by quantitative simulations of 2QF-COSY cross-peaks. These couplings are consistent with a two-state dynamic equilibrium between a minor N- and a major S-type conformer for all residues. The pseudorotation phase angle P of the major conformer is in the range 117–135° for nonterminal pyrimidine nucleotides and 153–162° for nonterminal purine nucleotides. Except for the terminal residues, the minor conformer comprises less than 25% of the population. Distance constraints obtained by a complete relaxation matrix analysis of the 2D NOE intensities with the MARDIGRAS algorithm confirm the dependence of the sugar pucker on pyrimidine and purine bases. Averaging by fast local motions has at most small effects on the NOE-derived interproton distances.

Immunoglobulin (Ig) genes are expressed exclusively in B lymphocytes. Two regulatory DNA fragments consisting of the same octamer motif d(ATGCAAAT) have been described as major determinants of the cell type-specific transcription of immunoglobulin heavy-chain (IgH) genes (Gillies et al., 1983; Mason et al., 1985). One element is located upstream of the TATA box in the promoter region. Another consensus sequence is part of the enhancer region, which resides in the large intron between the variable and constant segments of the IgH gene. The same structural features apply to the genes of immunoglobulin κ light chains but with the octamer motif replaced by its inverted and complementary sequence d(ATTTGCAT) (Falkner & Zachau, 1984). Deletion of either element in the promoter or enhancer region results in significant reductions of transcriptional activity (Perez-Mutul et al., 1988; Kemler et al., 1989). The conserved octamer sequence constitutes a binding site for both cell type-specific and non-tissue-specific nuclear proteins called octamer transcription factors. Consequently, the B cell type-specific expression of Ig genes may be attributed to differences in the interaction of various transcription factors with the DNA octamer fragment (Kemler & Schaffner, 1990).

The rapid advance of new NMR techniques over the past years has led to significant progress in the application of high-resolution NMR to proteins and DNA oligomers in solution. Two-dimensional NMR spectroscopy has made major contributions in elucidating the solution structures of a variety of biologically important DNA sequences and their interaction with drugs and proteins (Patel et al., 1987; Reid, 1987; Shafer & Brown, 1990). These studies rely almost exclusively on the quantitative or semiquantitative evaluation of sugar torsion angles and interproton distances from COSY and 2D NOE data. The NMR-derived structural parameters may subsequently be used as constraints in refinement methods such as restrained molecular dynamics, molecular mechanics, or distance geometry calculations to obtain a high-resolution structure. Of course, the quality of the final structure thus determined depends heavily on the number and accuracy of the NMR-derived constraints, and much effort has been directed toward a proper interpretation and quantitation of the experimental data (Kaluarachchi et al., 1991).

In principle, the conformation of the deoxyribose sugars can be obtained from the homonuclear proton–proton coupling constants between the various sugar protons. Optimized Karplus equations for deoxyribose have been derived and permit the straightforward determination of dihedral angles and thus sugar pucker from the vicinal couplings (De Leeuw & Altona, 1982). However, direct extraction of scalar couplings from J-resolved 2D[1] NMR spectra is often limited by the severe overlap of multiplet components due to large inherent line widths. E.COSY (Griesinger et al., 1985), P.E.COSY (Müller, 1987; Bax & Lerner, 1988), and the DISCO technique (Kessler et al., 1985) have all been proposed as methods to circumvent part of these difficulties. Recently, a quantitative simulation procedure has been developed that calculates theoretical COSY spectra and allows the extraction of J couplings by comparison with the experiment (Widmer & Wüthrich, 1986). Since the effects of line width and limited digital resolution are reproduced in the simulated spectra, accurate coupling constants may even be obtainable in the presence of broad and poorly resolved cross-peak components by matching the simulated and experimental cross-peak patterns (Schmitz et al., 1990).

More structural information in terms of interproton distances is provided by 2D NOE intensities. An important factor to consider in the determination of distances from NOE data

† This work has been supported by USPHS Grants CA27343 and Gm39247 awarded by the National Institutes of Health.

* Author to whom correspondence should be addressed.
‡ University of California.
§ Food and Drug Administration.

[1] Abbreviations: 1D, one-dimensional; 2D, two-dimensional; NOE, nuclear Overhauser enhancement; COSY, correlated spectroscopy; 2QF-COSY, double-quantum-filtered COSY; FID, free induction decay; EDTA, ethylenediaminetetraacetic acid; rms, root mean square.

0006-2960/92/0431-7477$03.00/0

is the presence of spin diffusion. A widely used approach, the isolated spin-pair approximation (ISPA), ignores the effect of neighboring spins on the cross-peak intensity for a particular proton pair, resulting in systematic distance errors especially at longer mixing times (Borgias & James, 1988). A more rigorous, self-consistent approach which accounts for multispin effects is based on a complete relaxation matrix analysis (Keepers & James, 1984; James et al., 1991). Several iterative matrix methods for determining internuclear distances have been described recently and include COMATOSE (Borgias & James, 1988), IRMA (Boelens et al., 1988, 1989), and MARDIGRAS (Borgias & James, 1990). A common feature of these methods is the calculation of a complete cross-relaxation rate matrix for a starting structure which is subsequently solved for the NOE intensities.

In the present study, structural information for the regulatory octamer sequence in Ig genes is obtained by a detailed analysis of double-quantum-filtered COSY (2QF-COSY) and 2D NOE data of the DNA duplex d(CATTTGCATC)·d(GATGCAAATG). Having unambiguously assigned all resonances, scalar couplings for the deoxyribose ring are determined by the quantitative simulation of 2QF-COSY cross-peaks and interpreted in terms of sugar conformations. In addition, a large number of distance constraints are obtained by analyzing 2D NOE intensities with the MARDIGRAS algorithm. These are compared with the J coupling-derived structural parameters.

MATERIALS AND METHODS

Sample Preparation. The DNA duplex for the NMR measurements was prepared by combining equimolar amounts of the decadeoxynucleotides d(CATTTGCATC) and d(GATGCAAATG) in a phosphate buffer, pH 7.0. The stoichiometry of duplex formation was previously determined by adding aliquots of one decamer solution to a similar solution of the complementary single-strand oligomer, the titration being monitored by UV spectroscopy. No peaks corresponding to single-strand oligonucleotide were found for the final DNA solution in the NMR spectrum. The NMR sample was lyophilized several times from ^2H_2O and finally dissolved in 0.4 mL of 99.996% ^2H_2O. The resulting buffer solution (20 mM sodium phosphate, 100 mM NaCl, 0.1 mM EDTA, 1 mM NaN_3, pH 7.0) was about 1.3 mM in duplex. For experiments with exchangeable protons, a second sample of the same buffer and duplex concentration was prepared in 90% H_2O–10% ^2H_2O.

NMR Spectroscopy. ^1H NMR experiments were run at 500 MHz on a GE GN500 spectrometer equipped with an Oxford Instruments magnet and a Nicolet 1280 computer. All 2D NMR measurements were carried out at 25 °C except for a 2D NOE experiment in H_2O which was run at 15 °C. Pure absorption 2D NOE spectra were obtained with the States method of phase cycling and alternating block acquisition (States et al., 1982). 2D NOE spectra with mixing times of 80, 140, and 200 ms were recorded in ^2H_2O. The delay between scans was 10.5 s, thus allowing for sufficient relaxation of the magnetization. A spectral width of 4400 Hz was used with the carrier frequency set to the HDO resonance frequency. A total of 400 FIDs of 2K complex data points were collected in t_1 with 16 scans at each t_1 value. All data were subsequently transferred to a Sun workstation for processing with locally written software. In both dimensions, a Gaussian window function with −7-Hz line broadening and a shift of 20% was used for resolution enhancement. Prior to Fourier transformation, the t_1 and t_2 FIDs were zero-filled to give a 2K × 1K data set with a digital resolution of 2.1 Hz/point in ω_2 and 4.3 Hz/point in ω_1.

2D NOE experiments in 90% H_2O–10% ^2H_2O were acquired with a $13\bar{3}\bar{1}$ excitation pulse centered on H_2O and a short homospoil pulse during the mixing time to suppress the water signal. The $13\bar{3}\bar{1}$ reading pulse results in a sharper null with less water suppression than does the standard $1\bar{3}\bar{1}$ sequence (Hore, 1983) but produces spectra that are more easily phase-corrected. Data were collected for a spectral width of 10 kHz with the excitation maximum set to the imino proton region at 13.2 ppm. Spectra in H_2O were recorded with a 140- and 200-ms mixing time at 25 °C and with a 200-ms mixing time at 15 °C. After zero-filling the FID twice, a final data set of 4K × 2K was obtained with a corresponding digital resolution of 2.4 Hz/point in ω_2 and 4.9 Hz/point in ω_1.

A pure absorption double-quantum-filtered COSY spectrum was recorded with a time-proportional phase incrementation scheme (Marion & Wüthrich, 1983). A spectral width of 4000 Hz and a repetition time of 2.5 s were employed. FIDs (4K complex data points) were acquired with 64 scans at each of 720 t_1 values, resulting in acquisition times of 90 and 512 ms in t_1 and t_2, respectively. After apodization in both dimensions with a squared sine bell shifted by 30°, zero-filling was used to obtain a final data set of 4K × 1K corresponding to a digital resolution of 1.0 Hz/point in ω_2 and 3.9 Hz/point in ω_1.

Measurement of the spin–lattice relaxation time T_1 was performed by the inversion–recovery method employing a 180° composite pulse (Freeman et al., 1980). Spin–spin relaxation times T_2 were determined using the Hahn spin-echo method.

2QF-COSY Simulation. 2QF-COSY cross-peaks for the deoxyribose sugar protons were simulated with the programs SPHINX and LINSHA (Widmer & Wüthrich, 1986). SPHINX (Simulation of Pulse sequences in HIgh-resolution Nmr eXperiments) calculates stick spectra for the spin system which comprises H1′, H2′, H2″, H3′, and H4′ protons as well as ^{31}P bound to the 3′-position. Except between H2′ and H2″, weak coupling between the sugar protons was assumed. Line shapes are subsequently simulated and added by LINSHA (LINe SHApe simulation program) taking into account experimental and processing parameters, i.e., acquisition times, digital resolution, and apodization functions.

Proton coupling constants for specific deoxyribose conformers were taken from Rinkel and Altona (1987).

Measurement of NOE Cross-Peak Intensities. 2D NOE intensities were determined by a line fit of the cross-peaks to a Gaussian and the subsequent integration of the theoretical curve using the locally written NMR-processing program package Sparky. In contrast to boxing methods which simply add up the signal inside a boxed region (Zhou et al., 1987), this line-fitting procedure enabled the deconvolution and reliable extraction of volumes for even strongly overlapped peaks. If the intensity of corresponding cross-peaks on both sides of the diagonal differed significantly from each other (>50%), the more reliable value as judged by the goodness of the line fit was taken. Otherwise, the volume was discarded. Cross-peaks with apparent non-Gaussian line shapes and cross-peaks of low intensity, i.e., those represented by only a few data points, yielded less reliable volumes when the line-fitting procedure was used. In these cases, nonoverlapped peaks were preferably integrated by applying the boxing method.

Intensities of cross-peaks from the labile imino protons were obtained from the 2D NOE spectra in H_2O and appropriately scaled by taking into account the excitation profile of the $13\bar{3}\bar{1}$ reading pulse. Cross-peaks of imino protons from the

Table I: Proton Chemical Shifts (ppm) of d(CATTTGCATC)·d(GATGCAAATG) at 25 °C[a]

position	C1	A2	T3	T4	T5	G6	C7	A8	T9	C10
H1′	5.71	6.38	6.02	6.13	5.845	5.815	5.62	6.245	6.005	6.25
H2′	1.995	2.81	2.10	2.175	2.09	2.625	2.10	2.66	2.025	2.255
H2″	2.45	3.03	2.62	2.615	2.48	2.685	2.41	2.905	2.445	2.265
H3′	4.72	5.05	4.875	4.905	4.895	4.985	4.82	4.985	4.84	4.56
H4′	4.08	4.465	4.28	4.22	4.10	4.38	4.19	4.385	4.145	4.01
H5′	3.76	4.17	4.21	4.15	4.11–4.21	4.15	4.15	4.15	4.125	4.04/4.16
H5″	3.76	4.06	4.34	4.225	4.10–4.22	4.11	4.22	4.10	4.285	4.04/4.16
H2		7.82						7.675		
H5	5.95						5.36			5.685
H6	7.71		7.265	7.46	7.295		7.36		7.215	7.58
H8		8.45				7.865		8.29		
CH_3			1.425	1.59	1.66				1.45	
NH[b]			13.75	13.94	13.69	12.56			13.74	
$NH_2(1)$[b,c]	6.98	6.62 br				*d*	6.40	6.36 br		6.99
$NH_2(2)$[b,c]	8.04	7.67 br				*d*	8.28	7.78 br		8.18

position	G11	A12	T13	G14	C15	A16	A17	A18	T19	G20
H1′	5.67	6.32	5.77	5.83	5.43	5.795	5.90	6.115	5.805	6.145
H2′	2.575	2.745	2.035	2.55	1.855	2.715	2.66	2.495	1.84	2.58
H2″	2.76	2.98	2.415	2.66	2.24	2.865	2.87	2.855	2.32	2.345
H3′	4.855	5.055	4.875	4.965	4.795	5.035	5.055	4.97	4.825	4.675
H4′	4.20	4.47	4.185	4.35	4.105	4.36	4.445	4.44	4.10	4.18
H5′	3.705	4.22	4.17	4.15	4.14–4.18	4.11	4.20	4.23	4.08	4.095
H5″	3.705	4.14	4.28	4.08	4.12–4.22	3.99	4.20	4.275	4.21	4.07
H2		7.955				7.115	7.11	7.67		
H5					5.335					
H6			7.12		7.285				7.02	
H8	7.905	8.32		7.80		8.19	8.125	8.095		7.845
CH_3			1.395						1.35	
NH[b]	12.57/12.71 br		13.51	12.53					13.67	12.57/12.71 br
$NH_2(1)$[b,c]	*d*	*d*		*d*	6.31	6.10 br	6.10 br	6.06 br		*d*
$NH_2(2)$[b,c]	*d*	*d*		*d*	8.31	7.66 br	7.35 br	7.25 br		*d*

[a] Uncertainty ±0.005 ppm; br, broad. [b] At T = 15 °C. [c] $NH_2(1)$ and $NH_2(2)$ denote non-hydrogen and hydrogen-bonded amino protons. [d] Not assigned.

terminal and penultimate residues of the duplex were not considered due to considerable proton exchange with water.

Model Structures. All-atom models with all hydrogen atoms explicitly included were established for A DNA (Arnott & Hukins, 1972) and standard B DNA (Arnott & Hukins, 1973) by use of the NUCGEN and EDIT modules of the modeling program package AMBER Version 4.0 (Pearlman et al., 1991). Hexahydrated Na^+ counterions were added at a distance of 0.5 nm to the phosphorus atoms of the phosphate groups to neutralize any negative charges. Molecular mechanics calculations were carried out on these models with AMBER using a distance-dependent dielectric constant to mimic bulk solvent effects. The force field parameters of the AMBER force field have been reported previously (Weiner et al., 1986).

Energy minimization was performed by steepest descent for the initial 200 cycles followed by conjugate gradient methods. The structures were refined until the energy gradient with respect to atomic coordinates was lower than 4 $kJ \cdot mol^{-1} \cdot nm^{-1}$. The final energy-minimized A- and B-form structures were subsequently used as starting models in the MARDIGRAS calculations.

RESULTS AND DISCUSSION

In the following, residues of the decamer duplex were labeled and numbered according to

```
      1 2 3 4 5 6 7 8 9 10
5'-C A T T T G C A T C-3'
3'-G T A A A C G T A G-5'
     20 19 18 17 16 15 14 13 12 11
```

Proton Resonance Assignments of d(CATTTGCATC)·d(GATGCAAATG). (a) Nonexchangeable Protons. Nonexchangeable protons of the duplex were assigned by analyzing 2QF-COSY and 2D NOE spectra (mixing times of 80, 140, and 200 ms) by following established strategies (Feigon et al., 1982, 1983; Scheek et al., 1983, 1984). The chemical shifts of the individual protons are given in Table I.

The overall pattern of NOE cross-peaks and their relative intensities are consistent with a right-handed double helix, a glycosidic torsion angle in the anti range, and a S-type sugar pucker characteristic of B DNA (Wüthrich, 1984). The well-resolved 2D NOE spectra allowed the sequential assignments of all base H8 and H6 and the H1′, H2′, H2″, and H3′ sugar resonances via continuous base-to-sugar proton coupling networks at longer mixing times. This is illustrated in Figure 1 for the region of cross-peaks between the aromatic and H1′ and H3′ protons. H4′, H5′, and H5″ sugar protons show intranucleotide NOE connectivities to base (weak) and H3′ as well as intra- and internucleotide connectivities to H1′ protons. Assignments of H4′ resonances are based on their strong intranucleotide cross-peak to H1′ and were further confirmed by scalar connectivities to H3′ in a 2QF-COSY spectrum. Distinction between H5′ and H5″ was possible for most of the nucleotides due to their differential intensities for cross-peaks to H1′ and H3′ protons at shorter mixing times: while H5′ exhibits stronger intranucleotide contacts to H3′, internucleotide cross-peaks between H5′ and H1′ of the 5′-linked nucleotide are expected to be of lower intensity in B DNA when compared to corresponding contacts to H5″ (Wüthrich, 1984).

Adenine H2 resonances are easily identified by their comparatively long spin–lattice relaxation times with T_1's as long as 6.1 s at 25 °C for the rather isolated H2 protons of adenines 8 and 12 (see Table II). They exhibit various weak intranucleotide, sequential, and interstrand contacts to H1′

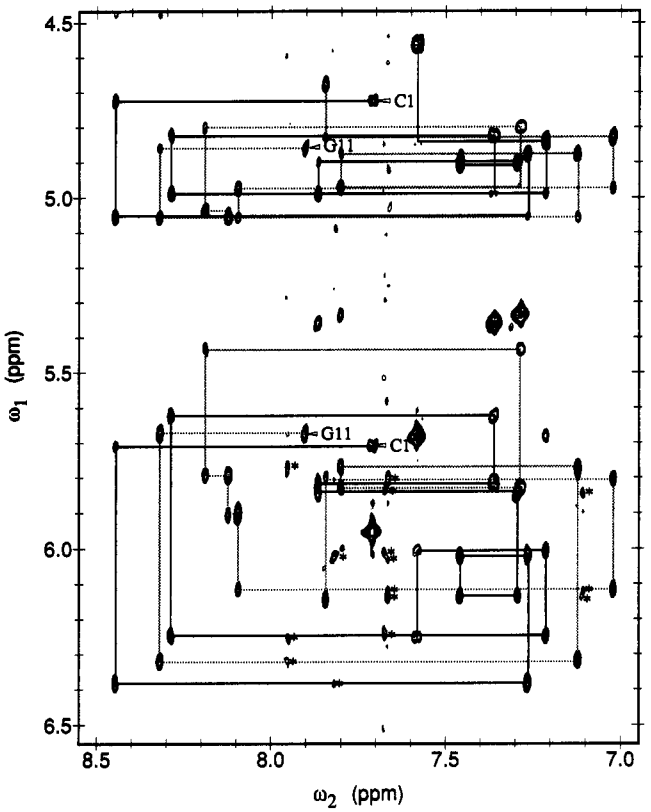

FIGURE 1: Part of the 140-ms 2D NOE spectrum of d(CATTTGCATC)·d(GATGCAAATG) showing cross-peaks between aromatic and H1′ and H3′ protons. Sequential assignments for the two strands starting at the 5′-terminal nucleotides are indicated by the black and grey lines. Also shown are adenine H2–H1′ cross-peaks, both cross-strand and intrastrand (marked by *), as well as intrabase and sequential cytosine H5–H6/H8 contacts.

Table II: Spin–Lattice Relaxation Times T_1 of Protons in the DNA Duplex d(CATTTGCATC)·d(GATGCAAATG) at 25 °C[a]

proton	T_1 (s)	proton	T_1 (s)
all H1′	2.0–2.3	A18 H2	5.5
all H2′/H2″	1.1–1.8	C1 H5	3.5
all H3′	2.0–2.3	C7 H5	2.9
all H4′	2.0–2.3	C15 H5	2.9
C1 H5′/H5″	1.0	T9 H6	1.8
G11 H5′/H5″	1.0	A17 H8	2.5
A2 H2	5.8	A18 H8	2.5
A8 H2	6.1	other H6/H8	2.0–2.3
A12 H2	6.1	T4 CH_3	2.0
A17 H2	5.6	other CH_3	2.2

[a] Uncertainty ±0.2 s.

sugar protons, particularly at longer mixing times (see Figure 1). These include cross-peaks to their own H1′, to H1′ of the 3′-linked nucleotide, to H1′ of the base-paired nucleotide, and to H1′ of the nucleotide base-paired to their 5′-linked neighbor. In addition, sequential and interstrand contacts are found between A17 H2 and A18 H2 as well as between A8 H2 and A12 H2. Expected cross-peaks between A2 H2 and A18 H2 and A16 H2 and A17 H2 are too close to the diagonal to be observed.

(*b*) *Exchangeable Protons.* Assignment schemes of labile protons based on their connectivities to other exchangeable and nonexchangeable protons have been described previously (Boelens et al., 1985; Rajagopal et al., 1988). A continuous train of connectivities links the neighboring imino protons of nonterminal base pairs. With guanine imino protons resonating at around 12.5 ppm, distinctly upfield from the thymine iminos, NH protons of the central nucleotides can be sequentially assigned. The absence of marked imino–imino contacts to the outer base pairs can be attributed to enhanced proton exchange with water toward the fraying end of the DNA duplex, making detection of corresponding cross-peaks increasingly difficult. Thus, imino protons of the terminal guanines give rise to very broad signals just above baseline in a 1D spectrum at 15 °C and exhibit no detectable cross-peaks to other DNA protons in the 2D NOE spectrum. However, they could be identified by their very strong exchange cross-peaks with the water signal at ω_1 = 4.90 ppm.

Except for those from the terminal residues, imino protons of guanines and thymines show additional contacts to various base protons. Strong cross-peaks are expected and observed to adenine H2 and cytosine H4 amino protons within a base pair while medium- and low-intensity cross-peaks are found to A H2 and C H4 aminos of 5′- and 3′-neighboring base pairs, respectively. Also, intrabase-pair connectivities are observed between T H3 imino and adenine H6 amino protons as are weak NOEs from thymine and guanine iminos to methyl and C H5 protons within the same base pair. Amino protons of A12 could not be assigned due to the absence of obvious cross-peaks with T9 H3. The C H4 protons of the terminal cytosines also exhibit no detectable NOEs in the imino proton region but were assigned from their intrabase contact to C H5. The amino protons of C10 show additional weak cross-peaks to the T9 methyl resonance.

Under the conditions employed, the cytosine H4 amino protons are slowly exchanging through rotation about the C–N bond and give rise to a pair of sharp cross-peaks in the imino proton region, the hydrogen-bonded proton shifted to lower field. In contrast, the two adenine H6 amino protons are significantly exchange-broadened, and amino protons of the guanines, probably exchanging at intermediate rates, are not observed at 15 °C. These findings are compatible with previous observations and can be related to the varying double-bond character of the C–N bond in the nucleobases (Patel, 1976; Rajagopal et al., 1988). In 2D NOE spectra taken at 25 °C, cross-peaks between imino and amino protons of adenines have disappeared while guanine amino protons start to exhibit a very broad intrabase NOE to G H1 (data not shown). At both temperatures a corresponding broad resonance along the diagonal, which is centered around 6.7 ppm and shows an exchange cross-peak with water, is observed and can be attributed to guanine amino protons.

Analysis of 2QF-COSY Cross-Peaks. Regions of the 2QF-COSY spectrum which were used for the determination of coupling constants include H2′/H2″(ω_1)–H1′(ω_2), H1′(ω_1)–H2′/H2″(ω_2), and H3′(ω_1)–H2′/H2″(ω_2) cross-peaks. The H2′/H2″(ω_1)–H3′(ω_2) cross-peaks were not included in the analysis, since t_1 noise from the water signal partially interferes in this region. Despite some peak overlap, cross-peaks are generally well-resolved, as shown in Figure 2 for the H2′/H2″(ω_1)–H1′(ω_2) region.

As illustrated in Figure 3 for some representative residues, there are significant variations in cross-peak fine structure among individual nucleotides, in particular for the H1′(ω_1)–H2′(ω_2) and H3′(ω_1)–H2″(ω_2) multiplet. Due to smaller line widths, the terminal cytosine C1 exhibits well-resolved cross-peak patterns, and its H3′(ω_1)–H2″(ω_2) fine structure shows the complete eight-component multiplet for H2″ theoretically expected for a proton coupled to three other nondegenerate spins. In contrast, the number of observable components is reduced together with their associated intensities for the other three nucleotides, and only two very weak peaks just above noise level are found for T5. All of the H1′(ω_1)–H2′(ω_2)

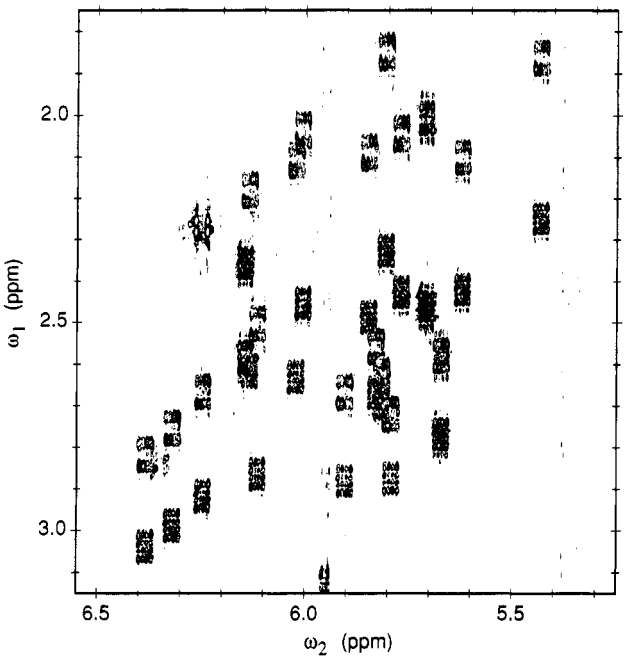

FIGURE 2: Part of the 2QF-COSY spectrum of d(CATTTGCATC)·d(GATGCAAATG) showing the H2′/H2″(ω_1)–H1′(ω_2) region.

cross-peaks exhibit two outer components along ω_2 with a partially resolved in-phase splitting due to passive coupling. However, the central multiplet pattern markedly differs among the residues. Two intense pairs of in-phase components are seen for C1 while four central fine-structure components of decreased intensity and alternating phase are observed for C15. Also, the two central components of corresponding cross-peaks from T5 and A8 significantly differ in phase and frequency position. These large variations in fine structure cannot be solely attributed to different resonance line widths but apparently reflect very different scalar couplings and hence sugar conformations for the various nucleotides.

Evaluation of the Line Width. As was found previously (Schmitz et al., 1990), establishing the proper line width for individual protons is crucial for any accurate determination of coupling constants. In Figure 4, the effect of line width is illustrated on simulated H1′(ω_1)–H2′(ω_2) and H3′(ω_1)–H2″(ω_2) cross-peaks. Changing the line width of protons along ω_2 not only significantly affects the appearance of the fine structure but may in some cases also lead to deceptive similarities of cross-peak patterns calculated with different J couplings. Correspondingly, establishing the appropriate line width for protons involved in the particular scalar coupling is prerequisite for any reliable extraction of coupling constants by the evaluation of COSY cross-peaks.

The H2′(ω_1)–H1′(ω_2) cross-peak was found to be especially useful for establishing the line width of the H1′ proton. In the experimental spectra, the pair of in-phase components along ω_2 due to passive coupling are well-resolved, pointing to a H1′ line width below 7 Hz. Careful inspection of experimental and simulated cross-peaks suggests a value for the H1′ line width of 5 Hz. Since variations in coupling constants may affect the estimation of the actual line width, line widths were verified after the J couplings for the individual nucleotides were extracted. However, only minor line-width adjustments were necessary, and these in turn did not affect the determined scalar couplings within experimental error.

Similarly, H1′(ω_1)–H2′(ω_2) cross-peaks proved to be the most useful ones for establishing the line width of the H2′ proton. Again, resolution of the outer pair of in-phase components along ω_2 which arise from passive coupling between H2′ and H3′ protons strongly depends on the H2′ line width employed. Careful comparison between simulated and experimental cross-peaks for each individual residue led to line widths for the H2′ protons ranging from 7.5 to 8.5 Hz, depending upon the particular nucleotide.

Wiggles in the experimental spectra resulting from truncation of the FID are influenced by the transverse relaxation and hence by the actual line width of individual protons as well as by acquisition times and apodization functions employed. Since experimental and processing parameters are known and their effects reproduced by the simulation program, truncation effects may contain valuable information about the line width of specific protons. Because no reliable evaluation of the H2″ line width could be achieved by looking at specific COSY cross-peak patterns, truncation effects along ω_1 were used to estimate the line width for H2″ protons. Apparently, these were smaller than those determined for H2′ protons, and H2″ line widths of 6–7 Hz best reproduced the experimental data.

The terminal nucleotides were found to have significantly reduced proton line widths, indicative of considerable fraying at the end of the duplex. The line widths determined for these terminal residues were 3.5–4.5 Hz for the H1′ proton, 5–6 Hz for the H2″ proton, and 6–7 Hz for the H2′ proton.

Determination of Coupling Constants. The simulation of COSY cross-peaks requires the adjustment of all coupling constants in the deoxyribose spin system which includes H1′, H2′, H2″, H3′, H4′, and ^{31}P bound to the 3′-position. To avoid the cumbersome simultaneous variation of several coupling constants, a strategy was employed which focuses on the better resolved cross-peak fine structure along the ω_2 axis. H2′(ω_1)–H1′(ω_2) and H2″(ω_1)–H1′(ω_2) cross-peaks exhibit a relatively simple four-peak pattern along ω_2, as is expected for a proton scalar coupled to two other spins. The pairs of in-phase and antiphase peaks correspond to the passive and active coupling, respectively. In cases where the peak-to-peak separations are not significantly below the line width of the H1′ protons, the individual couplings $^3J_{1'2'}$ and $^3J_{1'2''}$ can be directly determined by a first-order measurement of the respective peak-to-peak separations along ω_2. However, in most cases, amalgamation and cancellation effects will reduce or increase the separation of peak maxima, and measured distances commonly differ from the true values. For the present H1′ line width, distances between peaks less than 4 Hz introduced significant errors; therefore, computer simulations had to be applied for obtaining proper coupling constants. Note, however, that only $^3J_{1'2'}$ and $^3J_{1'2''}$ are to be independently varied within a reasonable range, while other coupling constants only affect spectral features along the poorly resolved ω_1 axis but have no direct impact on the position of the peak maxima along ω_2. The uncertainty in the coupling constants $^3J_{1'2'}$ and $^3J_{1'2''}$ determined by the simulation of the H1′ multiplet of H2′(ω_1)–H1′(ω_2) and H2″(ω_1)–H1′(ω_2) cross-peaks was estimated not to exceed ±0.2 Hz. Obtaining accurate values for these couplings is of particular importance, since initial errors may impair the proper evaluation of the remaining J couplings.

Once $^3J_{1'2'}$ and $^3J_{1'2''}$ have been extracted, the H1′(ω_1)–H2′(ω_2) and H3′(ω_1)–H2′(ω_2) cross-peak pattern along ω_2 solely depends on the coupling constants $^2J_{2'2''}$ and $^3J_{2'3'}$. That is, fitting of these experimental cross-peaks along ω_2 is accomplished by only adjusting two variables and provides the coupling constants between H2′ and H2″ as well as between H2′ and H3′. Since $^2J_{2'2''}$ is independent of the sugar pucker

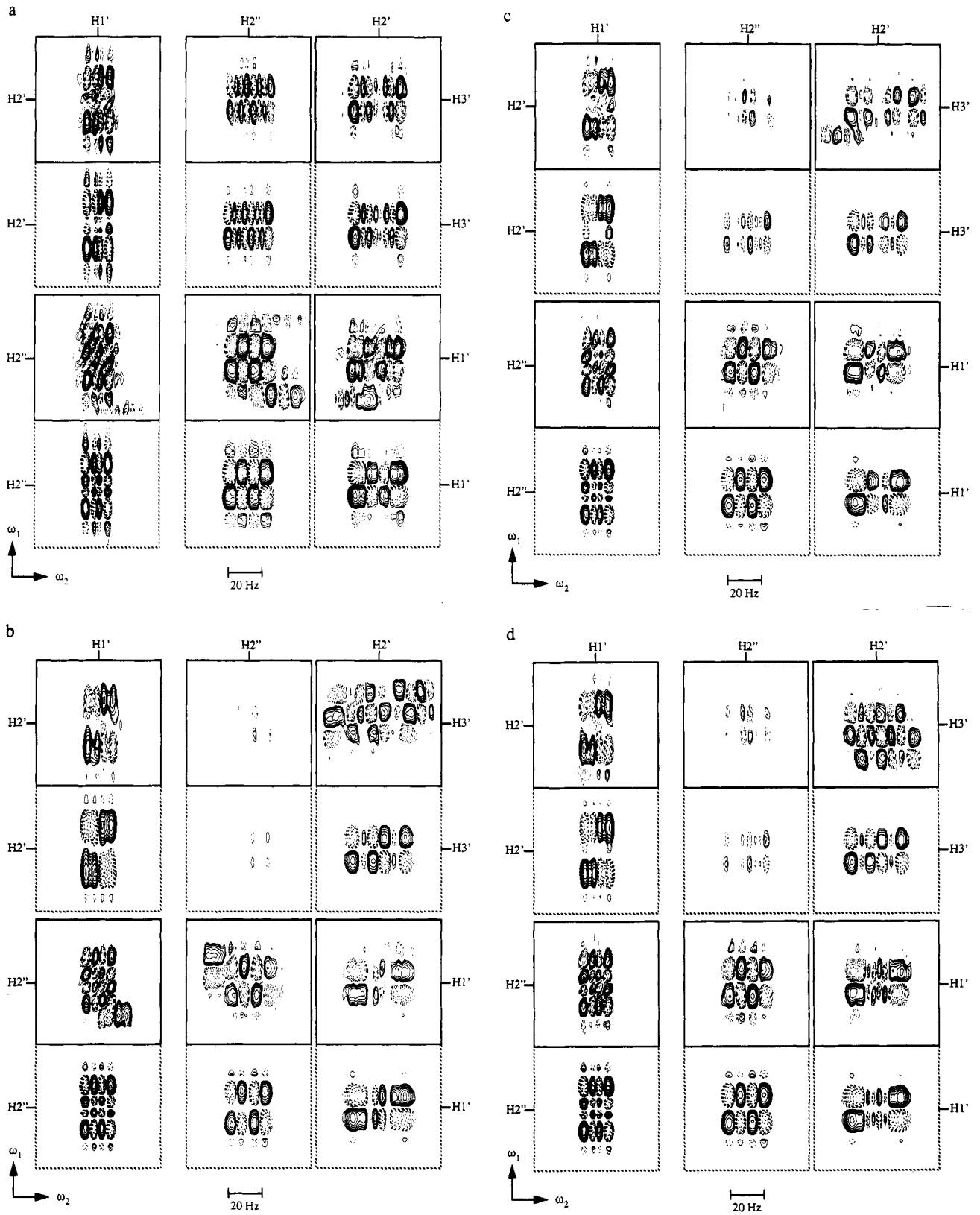

FIGURE 3: Experimental 2QF-COSY cross-peaks (solid boxes) and simulations using the coupling constants of Table III (dashed boxes) for C1 (a), T5 (b), A8 (c), and C15 (d). The line widths employed were 5.0 Hz for H1′ and H3′ (3.5 Hz for C1), 8.0 Hz for H2′ (6.0 Hz for C1), and 6.5 Hz for H2″ (5.0 Hz for C1). Negative peaks are drawn with dashed lines, and proton chemical shifts are indicated by the label positions.

and known to be approximately −14.0 Hz, variation of this coupling constant was limited to the range from −13.0 to −15.0 Hz. Values of −13.0 and −15.0 Hz were always found to yield worse fits than −14.0 Hz, whereas −13.5 or −14.5 Hz gave reasonable fits for some nucleotides. However, because in no cases were significant improvements achieved, $^2J_{2'2''}$ was set to −14.0 Hz with an uncertainty of ±0.5 Hz. The H1′(ω_1)–H2′(ω_2) cross-peaks are particularly sensitive to $^3J_{2'3'}$ (see Figure 4); therefore, this coupling constant could be reliably determined by simulating the corresponding cross-peak pattern along ω_2. On the other hand, H3′(ω_1)–H2′(ω_2) cross-peaks are less affected by $^3J_{2'3'}$, yet proved to be useful for verification

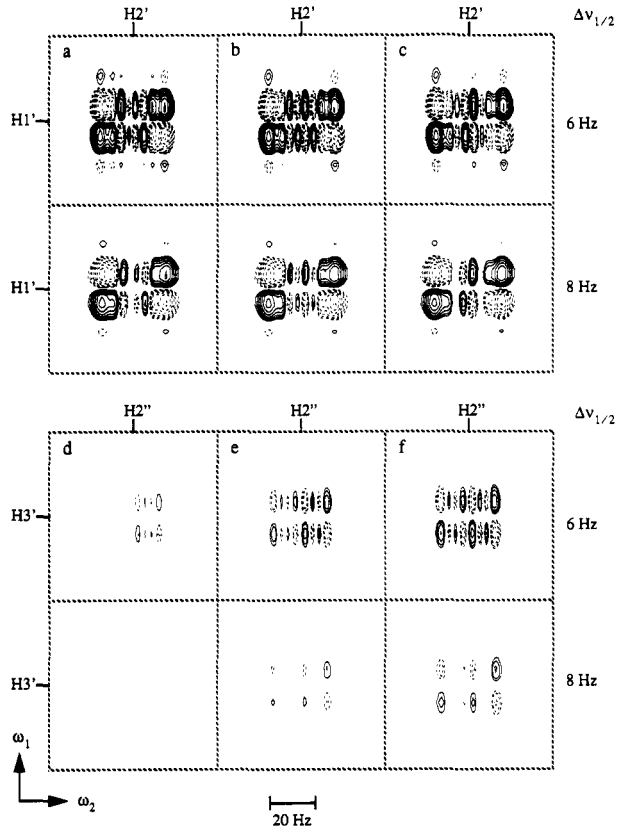

FIGURE 4: Effect of line width and coupling constants on the appearance of 2QF-COSY cross-peaks. Top: simulated H1′(ω_1)–H2′(ω_2) cross-peaks using coupling constants of $^3J_{2'3'}$ = 6.2 (a), 7.0 (b), and 7.8 Hz (c) and H2′ line widths of 6 and 8 Hz. Bottom: simulated H3′(ω_1)–H2″(ω_2) cross-peaks using coupling constants of $^3J_{2''3'}$ = 2.0 (d), 2.8 (e), and 3.6 Hz (f) and H2″ line widths of 6 and 8 Hz. Other parameters were those determined for cytosine C15. Negative peaks are indicated by dashed lines.

of the value determined from the H1′(ω_1)–H2′(ω_2) region. Also, in cases of ambiguity, e.g., peak overlap, H3′(ω_1)–H2′(ω_2) cross-peaks are important for establishing the proper coupling constant $^3J_{2'3'}$. The estimated error for this coupling constant was below ±0.4 Hz.

Similarly, the coupling constant $^3J_{2''3'}$ was determined by simulating H1′(ω_1)–H2″(ω_2) and H3′(ω_1)–H2″(ω_2) cross-peak patterns along ω_2. Again, only $^3J_{2''3'}$ remains as independent parameter for the simulations once $^3J_{1'2''}$ and $^2J_{2'2''}$ have been determined. While H1′(ω_1)–H2″(ω_2) cross-peaks are not very sensitive toward changes of coupling constants, H3′(ω_1)–H2″(ω_2) cross-peaks exhibit a high sensitivity for the $^3J_{2''3'}$ couplings (see Figure 4). Due to cancellation effects, small changes in this coupling constant result in the appearance or disappearance of detectable H3′(ω_1)–H2″(ω_2) cross-peaks. However, choice of an appropriate line width for H2″ is particularly important in evaluating $^3J_{2''3'}$.

Since only regions with H3′ proton resonances along the poorly resolved ω_1 dimension were used, coupling constants between H3′ and H4′ protons as well as H3′ and phosphorus could not be determined very well. A coupling constant $^3J_{3'P}$ = 5.8 Hz was therefore assumed for all residues. Variation of $^3J_{3'4'}$ within the range 2–5 Hz was found to have only marginal effects on the COSY cross-peak intensity and did not affect the overall H3′(ω_1)–H2′(ω_2) and H3′(ω_1)–H2″(ω_2) cross-peak pattern.

The terminal cytosine C10 exhibits a relatively intense, simple cross-peak pattern due to its near-isochronous H2′ and H2″ resonances. In this case, no individual couplings $^3J_{1'2'}$ and $^3J_{1'2''}$ are obtainable from peak-to-peak separations, and only the sum $^3J_{1'2'} + {}^3J_{1'2''}$ can be determined from the first-order measured distance between the outer peaks of the H2′/H2″(ω_1)–H1′(ω_2) cross-peak along ω_2. In addition, separation of the outer peaks from the H1′(ω_1)–H2′/H2″(ω_2) and H3′(ω_1)–H2′/H2″(ω_2) cross-peaks corresponds to $(^3J_{1'2'} + {}^3J_{1'2''} + {}^3J_{2'3'} + {}^3J_{2''3'})/2$, thus allowing for the determination of the sum $^3J_{2'3'} + {}^3J_{2''3'}$. This again leaves essentially two independent coupling constants to be adjusted in simulating the COSY cross-peaks.

Table III: Deoxyribose Coupling Constants of d(CATTTGCATC)·d(GATGCAAATG) from Quantitative Simulations of 2QF-COSY Cross-Peaks[a]

nucleotide	$^3J_{1'2'}$ (Hz)[b]	$^3J_{1'2''}$ (Hz)[b]	$^3J_{2'3'}$ (Hz)[c]	$^3J_{2''3'}$ (Hz)[c]
C1	7.4	6.0	6.0	4.0
A2	9.6	5.4	5.6	1.6
T3	9.8	5.6	8.0	2.2
T4	9.8	5.4	8.0	2.2
T5	9.8	5.4	8.0	2.2
G6	9.8	5.2	6.0	1.6
C7	8.2	5.8	6.8	3.2
A8	8.2	5.8	5.8	3.2
T9	8.6	6.0	8.0	3.2
C10	6.8	6.6	8.0	4.8
G11	9.0	6.0	5.2	2.4
A12	9.6	5.2	5.6	1.6
T13	9.6	5.4	7.2	1.8
G14	9.8	5.2	6.0	1.6
C15	8.8	6.0	7.0	2.8
A16	9.8	5.2	5.8	1.6
A17	9.8	5.2	5.6	1.6
A18	8.8	5.8	6.0	2.4
T19	8.6	6.0	7.4	3.0
G20	8.2	6.0	6.6	3.6

[a] $^2J_{2'2''}$ = −14.0 ± 0.5 Hz; $^3J_{3'P}$ = 5.8 Hz. [b] Uncertainty ±0.2 Hz. [c] Uncertainty ±0.4 Hz.

Once all individual coupling constants have been determined, simulated cross-peaks have to reproduce their experimental counterparts in both dimensions. Figure 3 illustrates the quality of the fit for several residues. As expected for a non-self-complementary decamer, cross-peaks are often partially overlapped. Fortunately, except for the G6 residue, peak overlap did not substantially impair extraction of individual J couplings. In the case of G6, values for $^3J_{2'3'}$ and $^3J_{2''3'}$ could not be reasonably well determined due to severe peak overlap but were assigned on the basis of close similarity in the cross-peak patterns to residue G14. The scalar coupling constants for the various nucleotides are summarized in Table III.

Evaluation of Sugar Pucker. In order to obtain the pseudorotational parameters which describe the geometry of the five-membered sugar ring, coupling constants for different sugar puckers were compared with the previously determined scalar couplings. No reasonable fit could be achieved for any of the nucleotides when one single conformer was used. Therefore, an analysis based on a two-state model with rapid interconversion between a N- and S-type sugar pucker was employed (Altona & Sundaralingam, 1973). In this case, the observed J coupling is averaged according to

$$^3J_{obs} = (1 - f_S)^3J_N + f_S{}^3J_S \quad (1)$$

with f_S representing the fraction of S-type conformer and ^3J_N and ^3J_S being the J-coupling constants of the pure N- and S-type conformers, respectively. The phase angle of pseudorotation for the minor N-type conformer was set to 9°, and equal pucker amplitudes for N and S conformers were assumed. The remaining parameters, i.e., the phase angle of pseudorotation for the major S-type conformer P_S, the fraction

Table IV: Pseudorotation Phase Angle P_S, Pucker Amplitude Φ_S, and Fraction f_S of the S-Type Conformer for the Deoxyribose Rings in d(CATTTGCATC)·d(GATGCAAATG)[a]

nucleotide	P_S (deg)	Φ_S (deg)	f_S (%)	rms
C1	171	37	70	0.15
A2	162	37	95	0.12
T3	117	40	95	0.16
T4	117	40	95	0.19
T5	117	40	95	0.19
G6	153	37	95	0.15
C7	135	37	75	0.17
A8	162	37	76	0.15
T9	117	37	84	0.17
C10	117	37	62	0.10
G11	171	37	88	0.18
A12	162	37	95	0.19
T13	135	37	92	0.16
G14	153	37	95	0.15
C15	135	37	80	0.09
A16	153	37	95	0.15
A17	162	37	95	0.16
A18	153	37	82	0.20
T19	126	37	82	0.12
G20	144	37	75	0.10

[a] $P_N = 9°$ and $\Phi_N = \Phi_S$ were assumed in the analysis; rms refers to the root-mean-square deviation between the coupling constants of Table III and those of the two-state model.

of the S-type conformer f_S, and the pucker amplitude Φ, were varied to match the observed coupling constants. The search for appropriate parameters is greatly facilitated by the different correlations existing between specific couplings and pseudorotational parameters. While there is a strong dependence of $^3J_{2'3'}$ on P_S, $^3J_{1'2'}$ is mostly affected by the fraction of S-type conformer f_S. Similarly, the coupling $^3J_{1'2''}$ proved to be particularly useful in establishing the pucker amplitude Φ. Hence, only a limited number of different combinations had to be considered. Table IV lists the pseudorotational parameters which were found to closely match the experimental data. In all cases, the root-mean-square (rms) deviation between the couplings calculated for the particular sugar geometry and the previously determined coupling constants are well within experimental error (rms $\leq$ 0.2 Hz). The uniqueness of the set of pseudorotational parameters depends on the particular nucleotide. However, in order to adequately reproduce the observed coupling constants within a two-state model, changes in P_S, Φ, and f_S are confined for all nucleotides to within 9°, 2°, and 5%, respectively.

Deoxyribose rings of most nucleotides are best described by a pucker amplitude of 37°. However, an amplitude of 40° was prerequisite for a reasonable description of the scalar couplings for the three thymidines T3, T4, and T5. Pseudorotational phase angles P_S for nonterminal purine nucleotides vary in the range 153–162°, characteristic of a C2′-endo conformer. In contrast, nonterminal pyrimidine residues exhibit phase angles of pseudorotation between 117° and 135°, significantly deviating from the standard B-DNA structure. Similar sugar puckers corresponding to a C1′-exo conformation were also found for other DNA duplexes (Celda et al., 1989; Schmitz et al., 1990).

Apparently, the attached base affects the deoxyribose ring conformation, with purine residues preferring larger pseudorotation phase angles P_S. Corresponding findings were observed for the octamer duplex $d(AC)_4$·$d(GT)_4$ (Gochin et al., 1990). Assuming a rigid conformer in this study, experimental 2QF-COSY cross-peaks were best reproduced by pseudorotation phase angles of 144° and 180° for pyrimidine and purine nucleotides, respectively. It is interesting to note that the analysis of several data sets from duplex DNA fragments also revealed conformational differences in purine and pyrimidine nucleotides, thus corroborating the present findings (Rinkel & Altona, 1987). However, these differences were attributed to a lower population of the major S-type conformer in a two-state equilibrium rather than to a lower pseudorotation phase angle for pyrimidine residues, implying less motional flexibility in purine sugars.

Except for terminal nucleotides, the fraction of minor N-type conformer amounts to less than 25%. Higher populations of the N-type conformer for some terminal nucleotides again indicate their high conformational mobility due to partial fraying at the duplex termini.

Analysis of 2D NOE Intensities. Internuclear distances were calculated from experimental NOE cross-peak intensities by use of the program MARDIGRAS (Borgias & James, 1990). MARDIGRAS (Matrix Analysis of Relaxation for DIscerning GeometRy of an Aqueous Structure) is based on the complete relaxation matrix and incorporates all network relaxation and multispin effects. This offers the opportunity of determining proton–proton distances with greater accuracy. Starting from a model structure, a complete relaxation rate matrix is set up and solved for the theoretical 2D NOE intensities by the CORMA algorithm (Keepers & James, 1984). CORMA calculates the matrix **a** of cross-peak intensities via diagonalization of the relaxation rate matrix **R**

$$\mathbf{a}(\tau_m) = \exp(-\mathbf{R}\tau_m) \tag{2}$$

Those elements for which experimental data are available are subsequently replaced by the normalized experimental intensities, resulting in a hybrid matrix. The normalization factor s was determined from the ratio of calculated and experimental intensities for fixed proton distances, i.e., cytosine H5-H6 and thymidine methyl-H6

$$s = \sum a_{ij,\mathrm{calc}} / \sum a_{ij,\mathrm{exp}} \tag{3}$$

The hybrid intensity matrix is back-transformed to a new relaxation rate matrix corresponding to a new set of internuclear distances. For internal consistency, certain elements of the new relaxation matrix are reset. A new cycle of intensity calculation, substitution and back-transformation is then performed until calculated NOE intensities converge.

The MARDIGRAS algorithm employed explicitly incorporates methyl rotors which are modeled with a three-site jump motion (Liu et al., 1992). Distances are calculated iteratively to a pseudoatom at the geometric center of the three methyl protons. For a globular molecule such as a decamer duplex with a hydrated length-to-diameter ratio of about 1.4, overall molecular tumbling is effectively isotropic in terms of influence on relaxation parameters and is well-approximated by a single correlation time τ_c. It can be evaluated from the spin–spin and spin–lattice relaxation of individual protons. Assuming $\omega^2\tau_c^2 \gg 1$, the correlation time of isotropic motion τ_c and the dipolar relaxation times T_1 and T_2 are related for an isolated homonuclear two-spin system according to

$$\tau_c = \frac{2}{\omega}\sqrt{\frac{T_1}{3T_2}} \tag{4}$$

where ω is the proton Larmor frequency. The correlation time thus obtained for different base protons of the decamer duplex at 25 °C was 1.6 ns, a value subsequently used in the MARDIGRAS calculations. Due to the various approximations made in deriving eq 4, uncertainties in the calculated tumbling time must be considered. To check the effect of τ_c on the

results, final distances obtained with MARDIGRAS runs for correlation times of 1 and 3 ns were compared to each other. However, only minor changes, averaging about 5%, were found.

So far, a rigid molecule with the absence of any internal motion except methyl rotation has been assumed. However, analysis of scalar coupling constants suggests rapid interconversion of at least two sugar conformers in the duplex, and there is ample evidence of local mobility within the DNA molecule (Keepers & James, 1982; Lane & Forster, 1989). Fast internal motions in the picosecond range will average the dipolar interaction, thus reducing some of the cross-relaxation rate constants and, consequently, NOE intensities. Recently, local motions have been incorporated into the MARDIGRAS algorithm (Kumar et al., 1992), employing the model-free analysis of Lipari and Szabo (Lipari & Szabo, 1982a,b). This approach describes internal rearrangements in terms of an internal correlation time τ_i and an order parameter S^2 which varies between 0 and 1, the latter corresponding to the absence of any local motion. Order parameters for DNA fragments have been estimated from experiment (Lane & Forster, 1989) or extracted from molecular dynamics simulations (Koning et al., 1991) and shown to be in the range $0.6 \leq S^2 \leq 1.0$, depending on the particular interproton vector. However, fraying effects at the end of the duplex are expected to increase the mobility of terminal residues, giving rise to order parameters as low as $S^2 = 0.4$ (Lane & Forster, 1989; Koning et al., 1991). This may explain the anomalously low intensities observed for some experimental NOE cross-peaks of terminal nucleotides, particularly those involving H2′ and H2″ protons. In this case, intensities of experimental cross-peaks, especially when they are insensitive to the actual conformation like H1′–H2″ or H1′–H3′, should indicate the relative flexibility of the DNA residues. Thus, comparison of intensities for different nucleotides suggests, that (1) the 5′-terminus is more mobile than the 3′-terminus and (2) the 5′-terminal cytosine experiences more flexibility than the 5′-terminal guanine. Similar findings in other sequences (Koning et al., 1991; Lane et al., 1991) suggest that this may be a general feature of DNA fragments.

To assess the effect of fast internal motions, model calculations were performed with order parameters for some residues arbitrarily set to $S^2 \sim 0.8$. Changes in final distances were found to be marginal, corroborating previous results that low-amplitude internal fluctuations will have little effect on the NOE intensities (Lane, 1990). Due to the difficulty of assigning individual order parameters to the various interproton vectors without any additional information, local motions were consequently ignored in the present MARDIGRAS calculations. However, neglecting internal motions for the more flexible 5′-terminal nucleotides resulted in unreasonable distances for some of the proton pairs, and corresponding intensities were therefore discarded.

MARDIGRAS runs were performed for the three mixing times of 80, 140, and 200 ms using standard B DNA, energy-minimized B DNA, and energy-minimized A DNA as starting structures. Depending upon the mixing time, between 308 and 366 distances were obtained corresponding to about 15–18 distance constraints per nucleotide. Generally, the accuracy of the derived distances for the experimentally observed cross-peaks depends on the number of available experimental intensities, the signal-to-noise ratio, and the accuracy of the initial model structure. To evaluate the latter, corresponding distances for energy-minimized A and B DNA are compared in Figure 5 before and after MARDIGRAS for the 80-ms NOE intensities. As can be seen, MARDIGRAS-derived distances are

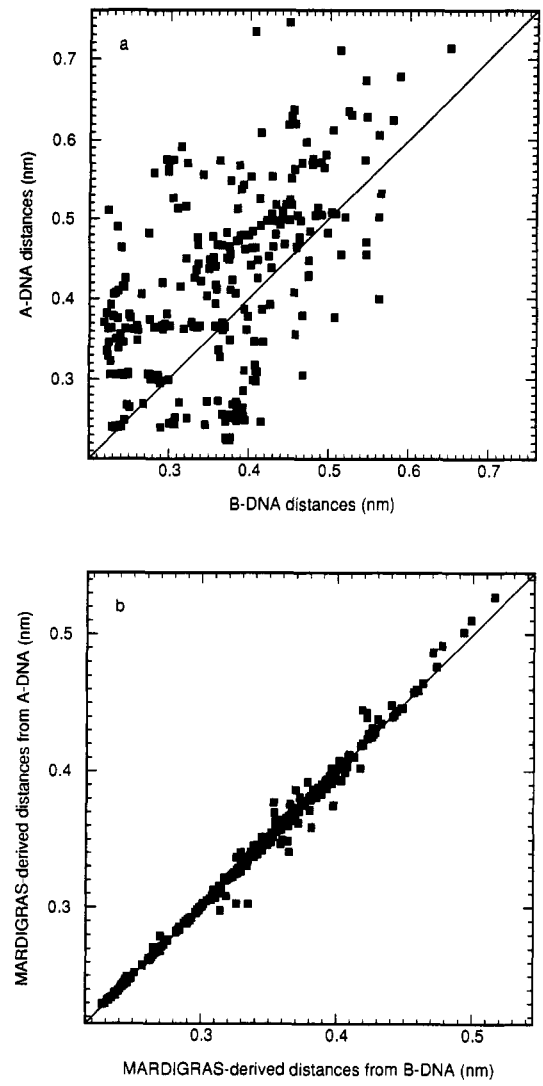

FIGURE 5: Comparison of interproton distances for d(CATTTGCATC)·d(GATGCAAATG) in energy-minimized A- and B-DNA conformations before (a) and after (b) refinement with MARDIGRAS. 2D NOE intensities with a mixing time of 80 ms were used for the MARDIGRAS calculations.

essentially invariant to the initial model and finally converge for the two different starting structures after four cycles of MARDIGRAS.

To judge the quality of the starting models, the absolute rms error between final distances d_o and distances derived for the three initial structures d_c were calculated according to

$$\text{rms} = \left(\frac{1}{N} \sum_{\tau_m} \sum_{i,j} (d_{o,ij}(\tau_m) - d_{c,ij})^2 \right)^{1/2} \quad (5)$$

where the summation runs over all of the reported distances and each of the three mixing times. Values thus obtained for standard B, energy-minimized B, and energy-minimized A DNA were 0.60, 0.59, and 1.26 Å, respectively. Not surprisingly, A DNA is the poorest initial approximation to the actual structure yielding the largest rms error. The best starting model appears to be energy-minimized B DNA, which was not only found to have the lowest rms value but also showed the best convergence in the MARDIGRAS calculations. This is consistent with previous results showing that distances extracted with MARDIGRAS exhibit essentially no systematic bias dependent on the starting structure. However, the accuracy of the resulting distances is improved with a better initial structural model.

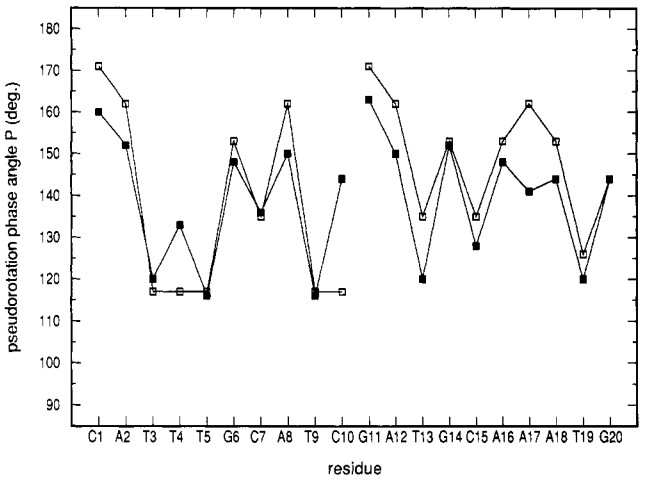

FIGURE 6: Plot of the pseudorotation phase angle P for the various nucleotides as determined from 2D NOE-derived H1′–H4′ distances, assuming $\Phi = 35°$ (filled squares). Also shown are the corresponding pseudorotation angles P of the major conformer from the coupling constant analysis (open squares).

Comparison of 2QF-COSY- and 2D NOE-Derived Structural Constraints. Do the 2QF-COSY- and 2D NOE-derived constraints reflect the same conformational behavior? Analysis of scalar coupling constants clearly indicates a dynamic structure of the sugar ring with rapid interconversion of at least two conformational states. In contrast, evaluation of the NOE data set is based on the assumption of a rigid DNA molecule. Fortunately, sugar repuckering should have a relatively small impact on the calculated distances for the nonterminal nucleotides whose minor conformer was found to comprise less than 25% of the population. However, a high population of N-type sugar conformers in terminal residues is expected to result in significant deviations between 2QF-COSY- and NOE-derived sugar puckers, and care must be taken when such data are interpreted.

Generally, the pseudorotational phase angle P is not well-determined by NOE data. There are three NOEs that are particularly sensitive to the sugar pucker, namely, H1′–H4′, H2″–H4′, and H6/H8–H3′, the latter also varying with the glycosidic torsion angle. The H6/H8–H3′ NOE strongly depends on the sugar pucker in the range $0° < P < 90°$ but is mostly invariant for $P > 90°$. On the other hand, H2″–H4′ NOE cross-peaks are difficult to use because they are weak for $P > 90°$ and often strongly overlapped. This essentially leaves the H1′–H4′ NOE for defining the pseudorotational phase angle. In Figure 6, the angle P determined from the intrasugar H1′–H4′ distances under the assumption of a single conformer is plotted as a function of the residue. Also shown are the corresponding pseudorotational angles for the major S conformer as determined from the J-coupling analysis. Considering the uncertainties in the NOE-derived values, results from J-coupling and NOE analysis favorably agree. In particular, the NOE-derived pseudorotational angles verify the differential sugar pucker for the nonterminal purine and pyrimidine nucleotides with $P_{Pyr} < 140° < P_{Pur}$. The average deviation in P for nonterminal residues is less than 8°, reflecting the good match between 2QF-COSY and 2D NOE data.

CONCLUSIONS

A strategy which involves the simulation of 2QF-COSY cross-peaks enabled the straightforward extraction of the scalar couplings between the deoxyribose protons. By concentrating on the fine-structure components and their separation along the well-resolved ω_2 axis, we have been able to determine coupling constants without relying primarily on the often subjective visual comparison between overall experimental and simulated cross-peak patterns. Of course, this method must fail in cases where broad resonances result in extensive amalgamation and cancellation of multiplet components. Also, care must be taken to ensure proper adjustment of the line width for individual resonances. Within the deoxyribose system, proton resonances vary significantly in line width. Also, terminal nucleotides were found to have smaller line widths when compared to nonterminal residues due to their increased mobility. The detailed analysis of coupling constants indicated that the deoxyriboses in the DNA fragment exist as an equilibrium mixture of at least two conformations. A minor conformer corresponds to C3′ endo whereas the major conformer exhibits pseudorotational angles in the range 117–135° (C1′ exo) for nonterminal pyrimidine nucleotides and 153–162° (C2′ endo) for nonterminal purine nucleotides.

MARDIGRAS was employed for the evaluation of interproton distances from 2D NOE intensities. Incorporation of low-amplitude internal motions was found to have at most small effects on the final distances, and NOE data were consequently interpreted in terms of one rigid conformer. Although there is excellent agreement between data derived from 2QF-COSY and 2D NOE experiments, it should be remembered that NOE intensities and thus distance constraints strictly reflect time-averaged values and may not correspond to a low-energy conformer. This is particularly true for the sugar ring, where energy differences between conformers are small and NMR coupling constants indicate rapid interconversion between different conformations.

The large number of structural parameters obtained from 2QF-COSY and 2D NOE experiments, i.e., bond torsion angles and internuclear distances, will be used as constraints in subsequent refinement methods. Restrained molecular dynamics calculations may hopefully give more insight into the dynamic solution structure of the Ig gene regulatory octamer motif and ultimately into the molecular basis of Ig gene expression.

ACKNOWLEDGMENT

We thank Dr. D. Kneller and M. Day for developing and improving the NMR-processing software and Drs. U. Schmitz and A. Kumar for helpful discussions.

SUPPLEMENTARY MATERIAL AVAILABLE

Four figures showing part of the imino proton region of a 200-ms 2D NOE spectrum recorded in H_2O at 15 °C, the H1′(ω_1)–H2′/H2″(ω_2) and the H3′(ω_1)–H2′/H2″(ω_2) regions of a 2QF-COSY spectrum, and the effect of line width on simulated 2QF-COSY cross-peaks (4 pages). Ordering information is given on any current masthead page.

REFERENCES

Altona, C., & Sundaralingam, M. (1973) *J. Am. Chem. Soc. 95*, 2333.

Arnott, S., & Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun. 47*, 1504.

Arnott, S., & Hukins, D. W. L. (1973) *J. Mol. Biol. 81*, 93.

Bax, A., & Lerner, L. (1988) *J. Magn. Reson. 79*, 429.

Boelens, R., Scheek, R. M., Dijkstra, K., & Kaptein, R. (1985) *J. Magn. Reson. 62*, 378.

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

Boelens, R., Koning, T. M. G., van der Marel, G. A., van Boom, J. H., & Kaptein, R. (1989) *J. Magn. Reson. 82*, 290.
Borgias, B. A., & James, T. L. (1988) *J. Magn. Reson. 79*, 493.
Borgias, B. A., & James, T. L. (1990) *J. Magn. Reson. 87*, 475.
Celda, B., Widmer, H., Leupin, W., Chazin, W. J., Denny, W. A., & Wüthrich, K. (1989) *Biochemistry 28*, 1462.
De Leeuw, F. A. A. M., & Altona, C. (1982) *J. Chem. Soc., Perkin Trans 2*, 375.
Falkner, F. G., & Zachau, H. G. (1984) *Nature (London) 310*, 71.
Feigon, J., Wright, J. M., Leupin, W., Denny, W. A., & Kearns, D. R. (1982) *J. Am. Chem. Soc. 104*, 5540.
Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1983) *Biochemistry 22*, 5943.
Freeman, R., Kempsell, S. P., & Levitt, M. H. (1980) *J. Magn. Reson. 38*, 453.
Gillies, S. D., Morrison, S. L., Oi, V. T., & Tonegawa, S. (1983) *Cell 33*, 717.
Gochin, M., Zon, G., & James, T. L. (1990) *Biochemistry 29*, 11161.
Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1985) *J. Am. Chem. Soc. 107*, 6394.
Hore, P. J. (1983) *J. Magn. Reson. 55*, 283.
James, T. L., Borgias, B., Bianucci, A. M., & Zhou, N. (1991) in *NMR and Biomolecular Structure* (Bertini, I., Molinari, H., & Niccolai, N., Eds.) pp 87–111, VCH Verlagsgesellschaft, Weinheim, Germany.
Kaluarachchi, K., Meadows, R. P., & Gorenstein, D. G. (1991) *Biochemistry 30*, 8785.
Keepers, J. W., & James, T. L. (1982) *J. Am. Chem. Soc. 104*, 929.
Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson. 57*, 404.
Kemler, I., & Schaffner, W. (1990) *FASEB J. 4*, 1444.
Kemler, I., Schreiber, E., Müller, M. M., Matthias, P., & Schaffner, W. (1989) *EMBO J. 8*, 2001.
Kessler, H., Müller, A., & Oschkinat, H. (1985) *Magn. Reson. Chem. 23*, 844.
Koning, T. M. G., Boelens, R., van der Marel, G. A., van Boom, J. H., & Kaptein, R. (1991) *Biochemistry 30*, 3787.
Kumar, A., James, T. L., & Levy, G. C. (1992) *Isr. J. Chem.* (in press).
Lane, A. N. (1990) *Biochim. Biophys. Acta 1049*, 189.
Lane, A. N., & Forster, M. J. (1989) *Eur. Biophys. J. 17*, 221.
Lane, A. N., Jenkins, T. C., Brown, T., & Neidle, S. (1991) *Biochemistry 30*, 1372.
Lipari, G., & Szabo, A. (1982a) *J. Am. Chem. Soc. 104*, 4546.
Lipari, G., & Szabo, A. (1982b) *J. Am. Chem. Soc. 104*, 4559.
Liu, H., Thomas, P. D., & James, T. L. (1992) *J. Magn. Reson. 98*, 163.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967.
Mason, J. O., Williams, G. T., & Neuberger, M. S. (1985) *Cell 41*, 479.
Müller, L. (1987) *J. Magn. Reson. 72*, 191.
Patel, D. J. (1976) *Biopolymers 15*, 533.
Patel, D. J., Shapiro, L., & Hare, D. (1987) *Q. Rev. Biophys. 20*, 35.
Pearlman, D. A., Case, D. A., Caldwell, J. C., Seibel, G. L., Singh, U. C., Weiner, P., & Kollman, P. A. (1991) AMBER 4.0, University of California, San Francisco.
Perez-Mutul, J., Macchi, M., & Wasylyk, B. (1988) *Nucleic Acids Res. 16*, 6085.
Rajagopal, P., Gilbert, D. E., van der Marel, G. A., van Boom, J. H., & Feigon, J. (1988) *J. Magn. Reson. 78*, 526.
Reid, B. R. (1987) *Q. Rev. Biophys. 20*, 1.
Rinkel, L. J., & Altona, C. (1987) *J. Biomol. Struct. Dyn. 4*, 621.
Scheek, R. M., Russo, N., Boelens, R., Kaptein, R., & van Boom, J. H. (1983) *J. Am. Chem. Soc. 105*, 2914.
Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371.
Schmitz, U., Zon, G., & James, T. L. (1990) *Biochemistry 29*, 2357.
Shafer, R. H., & Brown, S. C. (1990) in *Chemistry & Physics of DNA–Ligand Interactions* (Kallenbach, N. R., Ed.) pp 109–142, Adenine Press, New York.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286.
Weiner, S. J., Kollman, P. A., Nguyen, D. T., & Case, D. A. (1986) *J. Comput. Chem. 7*, 230.
Widmer, H., & Wüthrich, K. (1986) *J. Magn. Reson. 70*, 270.
Wüthrich, K. (1984) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Zhou, N., Bianucci, A. M., Pattabiraman, N., & James, T. L. (1987) *Biochemistry 26*, 7905.