J|A|C|S
ARTICLES
Published on Web 06/24/2006

Quantification of DNA BI/BII Backbone States in Solution. Implications for DNA Overall Structure and Recognition

Brahim Heddi,[§] Nicolas Foloppe,*[,†] Nadia Bouchemal,[‡] Edith Hantz,[‡] and Brigitte Hartmann*[,§]

Contribution from the Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie Physico-chimique, 13 rue Pierre et Marie Curie, Paris 75005, France, and BioMoCeTi, Université Paris 13, 74 rue Marcel Cachin, Bobigny 93017, France

Received March 10, 2006; E-mail: Brigitte.Hartmann@ibpc.fr; nf_research@hotmail.com

Abstract: The backbone states of B-DNA influence its helical parameters, groove dimensions, and overall curvature. Therefore, detection and fine characterization of these conformational states are desirable. Using routine NMR experiments on a nonlabeled B-DNA oligomer and analyzing high-resolution X-ray structures, we investigated the relationship between interproton distances and backbone conformational states. The three $H2'_i-H6/8_{i+1}$, $H2''_i-H6/8_{i+1}$, and $H6/8_i-H6/8_{i+1}$ sequential distances were found cross-correlated and linearly coupled to $\epsilon-\zeta$ values in X-ray structures and ^{31}P chemical shifts (δP) in NMR that reflect the interconversion between the backbone BI ($\epsilon-\zeta < 0°$) and BII ($\epsilon-\zeta > 0°$) states. These relationships provide a detailed check of the NMR data consistency and the possibility to extend the set of restraints for structural refinement through various extrapolations. Furthermore, they allow translation of δP in terms of BI/BII ratios. Also, comparison of many published δP in solution to crystal data shows that the impact of sequence on the BI/BII propensities is similar in both environments and is therefore an intrinsic and general property of B-DNA. This quantification of the populations of BI and BII is of general interest because these sequence-dependent backbone states act on DNA overall structure, a key feature for DNA−protein-specific recognition.

Introduction

A crucial, but still largely unresolved, question in molecular biology is that of protein−DNA recognition, which governs many biological processes. Recently, the role of DNA plasticity in such mechanisms has attracted a lot of attention. In addition to direct recognition elements such as hydrogen bonds, the overall DNA intrinsic structure and flexibility as a function of base sequence also contribute to its specific recognition by cognate proteins. This indirect readout can originate from backbone equilibria between canonical BI and alternative BII states that depend on the base sequence and are associated with dramatic helical changes.[1−6] The notion that the sequence-dependent DNA backbone BI/BII flexibility assists in selective binding of proteins is supported by numerous results relevant to important biological processes.[7−20] Thus, a detailed description and quantification of DNA backbone conformations are necessary to understand protein−DNA recognition. However, obtaining a reliable characterization of the DNA backbone in solution remains a challenge.

BII backbone conformation differs from its canonical BI alternative in the torsion angles ϵ and ζ (Figure 1), which are, respectively, *trans/g−* in BI ($\epsilon-\zeta = -90°$) and *g−/trans* in BII ($\epsilon-\zeta = +90°$). The first report of BII came from crystallographic studies.[21] Later, NMR studies provided evidence for the BI↔BII equilibrium in solution.[22−26] Thus far, detection

† Present address: 51 Natal Road, Cambridge CB1 3NY, UK.
‡ BioMoCeTi.
§ Institut de Biologie Physico-chimique.

(1) Hartmann, B.; Piazzola, D.; Lavery, R. *Nucleic Acids Res.* **1993**, *21*, 561−568.
(2) Srinivasan, A. R.; Olson, W. K. *J. Biomol. Struct. Dyn.* **1987**, *4*, 895−938.
(3) van Dam, L.; Levitt, M. H. *J. Mol. Biol.* **2000**, *304*, 541−561.
(4) Djuranovic, D.; Hartmann, B. *Biopolymers* **2004**, *73*, 356−368.
(5) Winger, R. H.; Liedl, K. R.; Pichler, A.; Hallbrucker, A.; Mayer, E. *J. Biomol. Struct. Dyn.* **1999**, *17*, 223−235.
(6) Djuranovic, D.; Hartmann, B. *J. Biomol. Struct. Dyn.* **2003**, *20*, 771−788.
(7) Schroeder, S. A.; Roongta, V.; Fu, J. M.; Jones, C. R.; Gorenstein, D. G. *Biochemistry* **1989**, *28*, 8292−8303.
(8) Karslake, C.; Botuyan, M. V.; Gorenstein, D. G. *Biochemistry* **1992**, *31*, 1849−1858.
(9) Botuyan, M. V.; Keire, D. A.; Kroen, C.; Gorenstein, D. G. *Biochemistry* **1993**, *32*, 6863−6874.
(10) Tisne, C.; Delepierre, M.; Hartmann, B. *J. Mol. Biol.* **1999**, *293*, 139−150.
(11) Wellenzohn, B.; Flader, W.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R. *Biochemistry* **2002**, *41*, 4088−4095.
(12) Wecker, K.; Bonnet, M. C.; Meurs, E. F.; Delepierre, M. *Nucleic Acids Res.* **2002**, *30*, 4452−4459.
(13) Hartmann, B.; Sullivan, M. R.; Harris, L. F. *Biopolymers* **2003**, *68*, 250−264.
(14) Reddy, S. Y.; Obika, S.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15475−15480.
(15) Djuranovic, D.; Hartmann, B. *Biophys. J.* **2005**, *89*, 2542−2551.
(16) Huang, D. B.; Phelps, C. B.; Fusco, A. J.; Ghosh, G. *J. Mol. Biol.* **2005**, *346*, 147−160.
(17) Wellenzohn, B.; Flader, W.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R. *J. Am. Chem. Soc.* **2001**, *123*, 5044−5049.
(18) Banyay, M.; Graslund, A. *J. Mol. Biol.* **2002**, *324*, 667−676.
(19) Geahigan, K. B.; Meints, G. A.; Hatcher, M. E.; Orban, J.; Drobny, G. P. *Biochemistry* **2000**, *39*, 4939−4946.
(20) van Dam, L.; Korolev, N.; Nordenskiold, L. *Nucleic Acids Res.* **2002**, *30*, 419−428.
(21) Fratini, A. V.; Kopka, M. L.; Drew, H. R.; Dickerson, R. E. *J. Biol. Chem.* **1982**, *257*, 14686−14707.
(22) Gorenstein, D. G. *Phosphorus-31 NMR: Principles and Applications*; Academic Press: New York, 1984.
(23) Chou, S. H.; Cheng, J. W.; Reid, B. R. *J. Mol. Biol.* **1992**, *228*, 138−155.
(24) Lefebvre, A.; Mauffret, O.; Lescot, E.; Hartmann, B.; Fermandjian, S. *Biochemistry* **1996**, *35*, 12560−12569.

10.1021/ja061686j CCC: $33.50

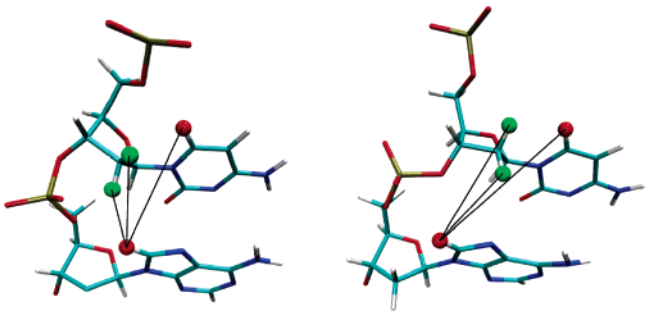


Figure 1. BI (left) and BII (right) phosphate linkage conformations with a CpA dinucleotide. The dashed lines depict the sequential distances between H8 of the adenine (red) and H2′/H2″ (green) or H6 (red) of the cytosine.

of BII states in NMR on nonlabeled oligonucleotides has relied upon measurements of $^3J_{H3'-P}$ spin−spin coupling constants and/or ^{31}P chemical shifts, δP.[27] $^3J_{H3'-P}$ is directly related to the torsion angle ϵ via the "modified Karplus equation".[28] However, the $^3J_{H3'-P}$ measurements suffer from numerous H3′ overlaps in spectra of nonlabeled DNA double helices. Furthermore, $^3J_{H3'-P}$ coupling constants depend on the type of experiment performed, so that the phosphate belonging to the same DNA sequence can be associated with very different $^3J_{H3'-P}$,[29] probably due to interferences with different passive coupling constants, difficult to neutralize or to estimate. In contrast, δP are easily assigned and very accurately measured. It is postulated and commonly accepted that δP values are shifted downfield when ζ changes from $g-$ (BI) to *trans* (BII),[22] but no simple and reliable relationship between δP and the BI/BII ratio has been presented. One could possibly obtain accurate information about ϵ with ^{13}C-labeled oligomers, from the $^3J_{C4'-P}$ and $^3J_{C2'-P}$ coupling constants that distinguish ϵ(*trans*) from ϵ($g+$ or $g-$) and ϵ($g-$) from ϵ($g+$ or *trans*), respectively.[30,31] We are, however, not aware of any such recently reported experiment, presumably because ^{13}C labeling of DNA for this sole purpose is perceived as far too tedious.

The present work offers an approach to quantify the proportions of BI and BII states for every phosphodiester link without the limiting requirement for labeled material. To tackle this issue, we build on previous work,[32] which found strong correlations between some proton−proton distances and helicoidal parameters in B-DNA, but did not consider the backbone states. Since the helicoidal parameters are sensitive to the BI/BII states, one expects a direct link between these proton−proton distances and the ^{31}P chemical shifts, as alluded to previously,[24,25] and thus between proton−proton distances and the BI/BII ratio.

To investigate such relationships by NMR, we selected a 14bp oligomer that captures the salient features of the DNA target of the JunFos transcription factor. The eukaryotic oncoproteins Jun and Fos belong to the group of dimeric AP-1 proteins, which bind to a common AP-1 DNA site.[33] These factors influence the control of cell proliferation, neoplasic transformation, and apoptosis.[34] Our NMR study was combined with an analysis of high-resolution X-ray structures of B-DNA, which strengthens and generalizes the quantitative link between the phosphate conformational states and their NMR signature presented here. The results highlight numerous correlations between three interproton sequential distances, δP and the BI/BII ratios, that allow us to greatly improve the detection and quantification of the BI/BII populations in solution.

In addition, this report demonstrates how these relationships can be used to establish the sequence effect on the percentages of BI/BII and illustrates the role of BII rich steps for modulating the overall structural properties of DNA. This will provide new information that can be exploited practically in structure determination protocols, yielding insights into the role of BII substates in influencing the intrinsic structure of B-DNA and their specific recognition by regulatory proteins.

Results

NMR Analysis of Sugars and Backbone of the JunFos Oligomer. The $^3J_{H1'-H2'}$ scalar couplings, measured by DQF−COSY experiments on 21 sugars out of 24, vary from 8 to 15 Hz, indicating a predominance of sugars in the south conformation.[30] In addition, the 12 available intranucleotide distances H4′$_i$−H1′$_i$, between 2.4 and 3.5 Å, exclude significant populations of the east domain.[35,36] Thus, the sugars in the JunFos oligomer are mostly in south, a signature of DNA in the B form.

For nonlabeled samples, the values of the torsional angles γ and β can be extracted from various coupling constants:[30] $^3J_{H5'-P}$ and $^3J_{H5''-P}$ for β, $^3J_{H4'-H5'}$ and $^3J_{H4'-H5''}$ for γ, and $^4J_{H4'-P}$ for both β and γ. Unfortunately, in DNA double helices, $^3J_{H4'-H5'}$ and $^3J_{H4'-H5''}$ are not measurable as the corresponding protons suffer from many overlaps and are too close to the diagonal. Because of severe proton overlaps, $^4J_{H4'-P}$, $^3J_{H5'-P}$, and $^3J_{H5''-P}$ cannot be accurately measured; however, the presence or absence of the corresponding cross-peaks provides qualitative information. The absence or the very low intensity of both H5′−P and H5″−P cross-peaks in our ^{31}P−^{1}H spectra matches with β torsions in canonical *trans*. Also, the H4′−P cross-peaks are clearly observed for all the phosphate linkages. The $^4J_{H4'-P}$ coupling constants are large enough only if β/γ fall in the *trans*/$g+$ region; for other conformations much smaller values (<1.5 Hz) are expected, which would not appear in the spectra. The unambiguous presence of all H4′−P cross-peaks confirms β in *trans* and shows that γ is $g+$ throughout. Finally, let us examine the intranucleotide H5′$_i$−H6/8$_i$ distances, which decrease markedly when γ changes from $g+$ (d(H5′$_i$ −H6/8$_i$) around 5.5 Å) to *trans* (d(H5′$_i$ −H6/8$_i$) between 2.5 and 2.9 Å).[30] That no intranucleotide cross-correlation is observed in the H5′−H6/8 region, either in the short or long mixing times (50−200 ms) NOESY spectra, confirms that γ torsions are in canonical $g+$.

(25) Tisne, C.; Hantz, E.; Hartmann, B.; Delepierre, M. *J. Mol. Biol.* **1998**, *279*, 127−142.
(26) Tisne, C.; Hartmann, B.; Delepierre, M. *Biochemistry* **1999**, *38*, 3883−3894.
(27) Gorenstein, D. G. *Methods Enzymol.* **1992**, *211*, 254−286.
(28) Lankhorst, P. P.; Haasnoot, C. A.; Erkelens, C.; Altona, C. *J. Biomol. Struct. Dyn.* **1984**, *1*, 1387−1405.
(29) Clore, G. M.; Murphy, E. C.; Gronenborn, A. M.; Bax, A. *J. Magn. Reson.* **1998**, *134*, 164−167.
(30) Wijmenga, S. S.; Van Buuren, B. N. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287−387.
(31) Szyperski, T.; Ono, A.; Fernandez, C.; Iwai, H.; Tate, S.; Wüthrich, K.; Kainosho, M. *J. Am. Chem. Soc.* **1997**, *119*, 9901−9902.
(32) Lefebvre, A.; Fermandjian, S.; Hartmann, B. *Nucleic Acids Res.* **1997**, *25*, 3855−3862.
(33) Rauscher, F. J., 3rd; Sambucetti, L. C.; Curran, T.; Distel, R. J.; Spiegelman, B. M. *Cell* **1988**, *52*, 471−480.
(34) Shaulian, E.; Karin, M. *Nat. Cell Biol.* **2002**, *4*, E131−136.
(35) Chuprina, V. P.; Nerdal, W.; Sletten, E.; Poltev, V. I.; Fedoroff, O. *J. Biomol. Struct. Dyn.* **1993**, *11*, 671−683.
(36) Wüthrich, K. *NMR of proteins and nucleic acids*; Wiley & Sons: New York, 1986.

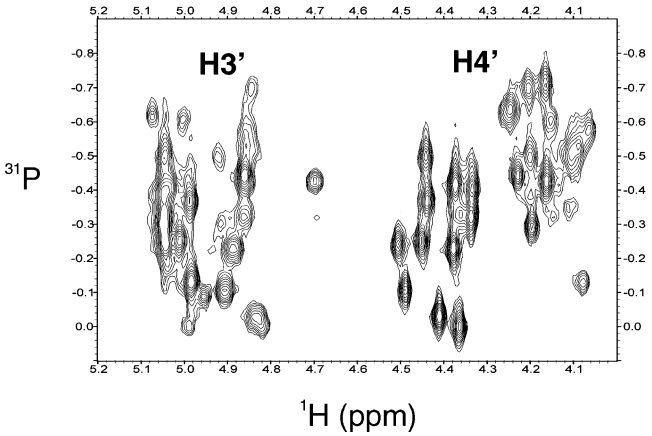


Figure 2. Spectrum from a $^{31}P-^{1}H$ HSQC experiment showing the correlation between P_i and the H3′$_i$ and H4′$_{i+1}$ sugar protons.

Table 1. NMR Characterization of the Phosphate Groups in the JunFos Oligomer[a]

base steps	δP	$^{3}J_{H3-P}$	%BII
5′-G_1pC_2	−0.44		22
C_2pA_3	−0.11		69
A_3pT_4	−0.62	4.2	0
T_4pT_5	−0.44	5.3	22
T_5pC_6	−0.50		13
C_6pT_7	−0.50	5.3	13
T_7pG_8	−0.33	6.9	37
G_8pA_9	−0.25	7.2	49
A_9pG_{10}	−0.41		26
$G_{10}pT_{11}$	−0.70	4.2	0
$T_{11}pC_{12}$	−0.53		9
$C_{12}pA_{13}$	−0.03		80
$A_{13}pG_{14}$-3′	−0.30		42
5′-$C_{15}pT_{16}$	−0.43		23
$T_{16}pG_{17}$	−0.23	7.3	52
$G_{17}pA_{18}$	−0.24	7	50
$A_{18}pC_{19}$	−0.39		29
$C_{19}pT_{20}$	−0.70		0
$T_{20}pC_{21}$	−0.58		2
$C_{21}pA_{22}$	0.00		85
$A_{22}pG_{23}$	−0.40		28
$G_{23}pA_{24}$	−0.37		32
$A_{24}pA_{25}$	−0.50	4.6	13
$A_{25}pT_{26}$	−0.61	4.2	0
$T_{26}pG_{27}$	−0.41		26
$G_{27}pC_{28}$-3′	−0.13		66

[a] δP (ppm): ^{31}P chemical shifts; $^{3}J_{H3'-P}$ (Hz): only coupling constants devoid of any overlap are presented; % BII: estimated BII percentage using eq 1, see the section "Quantification of the BI/BII Ratios in Solution".

These data on β and γ torsions, consistent with a canonical B structure, reflect general knowledge about DNA structure, with 95% of the γ and β conformers being canonical in X-ray structures of B-DNA devoid of any ligand,[6] consistent with α/γ flippings found to be highly unlikely in free-energy calculations.[37] These studies highlighted a systematic correlation between the α and γ angles, strengthening further the likelihood of the $g-$ conformation for the α torsions in the JunFos oligomer. All these observations show that the JunFos oligomer adopts a B form with no detectable unusual conformations of α, β, and γ.

The $^{31}P-^{1}H$ HSQC spectrum is shown in Figure 2, and the ^{31}P chemical shift values (δP) are reported in Table 1. The dispersion reaches 0.7 ppm and, while remaining in a standard

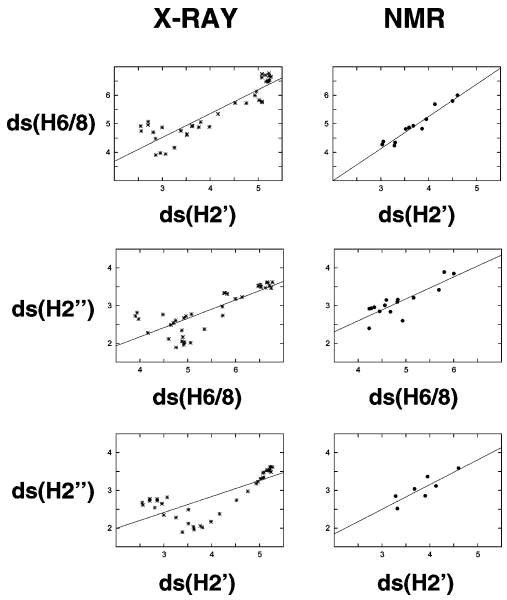


Figure 3. Cross-correlations between the ds(H2″), ds(H2′), and ds(H6/8) sequential interproton distances (Å) in X-ray structures (left) and from the NMR data on the JunFos oligomer (right).

B-DNA range,[38] reveals a variety of behaviors for the phosphate groups. In the absence of unusual conformations of α, β, and γ, these δP can only reflect the ϵ and ζ motion (i.e., the BI↔BII equilibrium). This is reinforced by the $^{3}J_{H3'-P}$ coupling constants extracted from the $^{31}P-^{1}H$ COSY spectrum, which are sensitive to ϵ. Only 10 of them could be obtained, but a close correspondence is found between them and the corresponding δP values (Table 1). This confirms that the different δP values reflect different BI/BII ratios, but more information is needed for a quantitative interpretation.

Correlations between Sequential Interproton Distances in Solution and Crystals. Strong linear cross-correlations (correlation coefficients between 0.85 and 0.94) are found between the three sequential distances H2′$_i$−H6/8$_{i+1}$ [ds(H2′)], H2″$_i$−H6/8$_{i+1}$ [ds(H2″)], and H6/8$_i$−H6/8$_{i+1}$ [ds(H6/8)], both by NMR in the JunFos oligomer and in X-ray structures (Figure 3). In the X-ray data, we excluded the rare steps containing north sugars since they are not detected in our NMR experiments. Also, north sugars are associated with short ds(H2′) and ds(H2″) distances that obscure the correlations, decreasing and increasing the ds(H2′) and ds(H2″) distances, respectively, by 0.5 Å on average. In the NMR data, no distance deviates significantly from the linear correlation (Figure 3), as expected for sugars predominantly in south. Despite a widening in the region of short distances extracted from the X-ray structures, in particular for the ds(H2′) versus ds(H2″) plot, the correlation slopes for NMR or X-ray data are comparable (Figure 3). Nevertheless, we note a systematic shift of 0.2 Å between X-ray and NMR H2″$_i$ −H6/8$_{i+1}$ distances. This is because in NMR these distances are systematically overestimated and often need to be recalibrated.[25,39−41] The strong similarity between the

(37) Varnai, P.; Djuranovic, D.; Lavery, R.; Hartmann, B. *Nucleic Acids Res.* **2002**, *30*, 5398−5406.
(38) Gorenstein, D. G. *Chem. Rev.* **1994**, *94*, 1315−1358.
(39) Cuniasse, P.; Sowers, L. C.; Eritja, R.; Kaplan, B.; Goodman, M. F.; Cognet, J. A.; LeBret, M.; Guschlbauer, W.; Fazakerley, G. V. *Nucleic Acids Res.* **1987**, *15*, 8003−8022.
(40) Mauffret, O.; Hartmann, B.; Convert, O.; Lavery, R.; Fermandjian, S. *J. Mol. Biol.* **1992**, *227*, 852−875.
(41) Lefebvre, A.; Mauffret, O.; Hartmann, B.; Lescot, E.; Fermandjian, S. *Biochemistry* **1995**, *34*, 12019−12028.

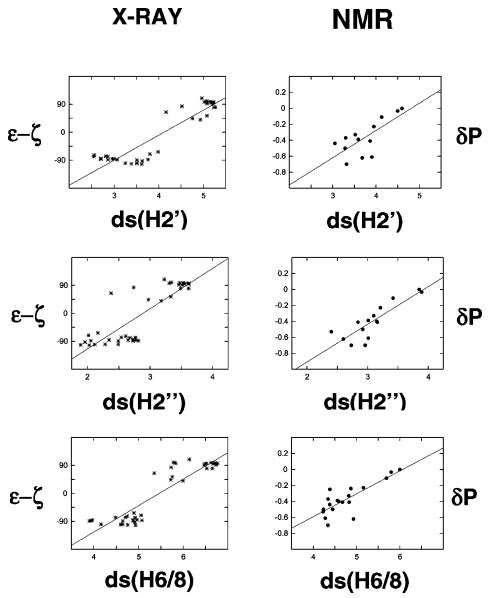


Figure 4. Correlations between the ds(H2″), ds(H2′), and ds(H6/8) sequential interproton distances (Å) and either ($\epsilon-\zeta$) (deg) in X-ray structures (left) or δP (ppm) in NMR (right).

correlations found by NMR on the JunFos oligomer and in X-ray structures indicates that these correlations are general to B-DNA.

These experimental correlations indicate that ds(H2′), ds(H2″), and ds(H6/8) are sensitive to a same conformational parameter. Previous theoretical results[32] did not consider the backbone states but showed that the interdependent $H2'_i - H6/8_{i+1}$ and $H6/8_i - H6/8_{i+1}$ distances vary with the roll and the twist. The well-known relationship between these helical parameters and the BI/BII equilibrium[1−6] prompted us to investigate a quantitative correlation between the BI/BII backbone conformers and the interproton distances.

Quantification of the BI/BII Ratios in Solution. In NMR ds(H2′), ds(H2″), and ds(H6/8) are correlated to δP values (Figure 4, correlation coefficients between 0.72 and 0.85), themselves linked to ϵ and ζ conformations, as seen above. The solid-state data represent structures that are essentially static and capture either the BI or BII conformer. Consequently, the X-ray structure analysis yields a bimodal BI/BII distribution. In contrast, DNA is free to undergo BI/BII transitions in solution NMR experiments. The time scale of these experiments with respect to that of the BI/BII transitions results in an average phosphate chemical shift (δP) per step that reflects the dynamical equilibrium between the BI and BII states and is directly related to the BI/BII ratio. Thus, the δP distribution from NMR appears continuous rather than bimodal. Nevertheless, the correlations are equivalent with both types of data, δP in NMR being substituted by $\epsilon-\zeta$ in the crystals (Figure 4). With both NMR and X-ray, the distributions associated with large distances (BII region) are clearly less scattered than those coupled to short distances (BI region), consistent with the known local stiffening of the double helix with BII conformers. Indeed, while BI phosphate groups tolerate a relatively large range of sequence-dependent helical parameter values, the BII backbone imposes rather specific values of high twist and negative roll, whatever the sequence.[4,6] Overall, the present relationships precisely recapitulate known intrinsic mechanical properties of B-DNA.

These correlations can be used to improve the structure determination of DNA by NMR in solution. First, they allow

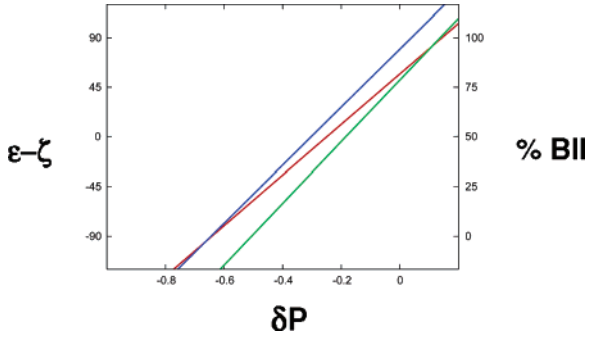


Figure 5. Relationships between δP (ppm) and either $\epsilon-\zeta$ values (deg) (left *y*-axis) or the BII percentage (right *y*-axis) established via the X-ray and NMR correlations found for the sequential distances (i) ds(H2″) recalibrated by 0.2 Å in blue or (ii) ds(H2′) in green or (iii) ds(H6/8) in red.

us to extrapolate some distances not directly measured from other, measured distances, or from the corresponding δP. This opportunity provides advantages illustrated below by an application to the JunFos oligomer.

More generally, the NMR and X-ray correlations establish a quantitative correspondence between δP, $\epsilon-\zeta$, and BII percentages. We first calculate by simple linear regressions the best fits for the correlations shown Figure 4, and we combine them to express δP in function of $\epsilon-\zeta$. Then, associating 0% and 100% of BII with $\epsilon-\zeta = -90°$ and $\epsilon-\zeta = +90°$, the two maxima of $\epsilon-\zeta$ distribution in free B-DNA X-ray structures,[6] any $\epsilon-\zeta$ intermediate value is converted in BII percentage by linear extrapolation between these extreme points. Figure 5 shows the δP versus either $\epsilon-\zeta$ or BII percentage relationships obtained from ds(H2′), ds(H6/8), and ds(H2″) correlations. The three straight lines converge remarkably and allow us to convert δP in terms of BII percentage ±8% according to eq 1:

$$\mathrm{BII}(\%) = 143\,\delta P + 85 \quad \text{for } \delta P \text{ referenced to phosphoric acid}$$

$$\mathrm{BII}(\%) = 143\,\delta P + 621 \quad \text{for } \delta P \text{ referenced to trimethyl phosphate}$$

BII(%) should be expressed from 0 to 100. We find that purely BI or BII states correspond to −0.60 and +0.11 ppm (δP referenced to phosphoric acid), respectively, in line with previous estimations made on the basis of $^3J_{H3'-P}$ measurements.[38] A BI/BII ratio of 1 gives a δP of −0.24 ppm. This is exactly the average from measurements of 64 phosphates in single-stranded DNA[42] in which one expects equal populations of BI and BII, following $^3J_{H3'-P}$ measurements carried out in other DNA above the melting temperature.[38]

In X-ray analyses[6] and molecular dynamics (MD) studies,[4] BI and BII correlate with two distinct zones in the roll/twist space. Thus, BI and BII are accompanied by typical helical parameters, BII being characterized by marked negative roll and high twist. In turn, these helical parameter values affect the interproton distances. Thus, our results reflect, and take advantage of, these mechanical couplings between base and backbone and the modulation of helical parameters by the BI or BII states.

Generalized Sequence Effects on BI/BII Propensities. We collected 241 δP values previously reported for various B-DNA

(42) Ho, C. N.; Lam, S. L. *J. Magn. Reson.* **2004**, *171*, 193−200.

Table 2. Influence of DNA Base Sequence on the BII Percentages in Solution and Crystals[a]

base steps		solution			crystal	
		N	δP_{av}	%BII	*N*	%BII
YpR	CpA	27	−0.23	52	18	87*
	CpG	14	−0.34	36	66	14
	TpA	10	−0.49	15	13	31
	TpG	27	−0.38	31	19	69*
RpY	ApC	21	−0.55	6	13	0
	ApT	15	−0.66	0	47	0
	GpC	16	−0.44	22	46	47*
	GpT	21	−0.70	0	9	11
RpR	ApA	10	−0.51	12	48	1
	ApG	16	−0.49	15	18	8
	GpA	17	−0.33	37	38	21*
	GpG	2	−0.26	48	18	35
YpY	CpC	2	−0.37	32	18	4
	CpT	16	−0.64	0	18	0
	TpC	17	−0.53	9	42	0
	TpT	10	−0.60	0	48	0

[a] R: purine; Y: pyrimidine. *N* is the total number of dinucleotide steps per sequence category. The BII percentages (%BII) in solution were deduced from the averaged ^{31}P chemical shifts (δP_{av}, ppm) referenced in the section "Generalized Sequence Effects on BI/ BII Propensities". The BII percentages in crystals were obtained by a previous X-ray data analysis.[6] In the solution data set, there are only two instances of GpG and CpC steps, and therefore the associated properties are not be statistically relevant. The asterisks indicate the percentages suspected to be overestimated (see main text). The %BII in solution is given with an uncertainty of ±8%.

sequences in solution[12,26,40,43−47] and the δP data obtained in this work, excluding the phosphate linkages at the sequences termini, submitted to end effects. These δP were converted to BII percentages with eq 1 to compare the effects of sequence on BI/BII in solution with those ascertained in crystals.[6] The influence of sequence was analyzed in terms of dinucleotide steps, showing that this influence is very similar in crystals and solution (Table 2). There are categories of base steps (YpY, ApR, and RpY, but not GpC; Y for any pyrimidine, R for any purine) in which the phosphate group is essentially trapped in BI. In contrast, YpR, GpR, and GpC are conducive to BII. Even more striking, each CpA phosphate is more than 50% in BII. This means that these propensities are intrinsic properties of the DNA molecule itself and are primarily controlled by the local base sequence.

In addition, we observe a sequence effect beyond the immediate dinucleotide base step, as previously suggested by NMR and modeling studies.[12,26,46,48,49] Considering a dinucleotide and its two bracketing nearest neighbor bases on the same strand, in the CpA·TpG NMR data set all the types of 3′- and 5′- nearest neighbors are equally represented in terms of purine/pyrimidine (Y--Y, R--R, Y--R, R--Y, with -- either CpA or TpG). The corresponding δP values show that RCpA(R/Y) tetranucleotides disfavor the CpA BII phosphates, while YCpAR sequences especially enhance them (Table 3). Although less marked, this effect is retrieved for TpG, with YTpGR exhibiting

Table 3. Tetranucleotide Sequence Effects around CpA Steps in Solution[a]

sequence	*N*	δP_{av}	%BII
RCpA(R/Y)	14	−0.37	32
YCpAY	6	−0.19	58
YCpAR	6	−0.06	76

[a] R: purine; Y: pyrimidine. *N* is the total number of CpA steps per sequence category. The BII percentages (%BII) in solution were deduced from the averaged ^{31}P chemical shifts (δP_{av}, ppm) referenced in the section "Generalized Sequence Effects on BI/BII Propensities".

the highest TpG BII percentages (between 36 and 55%). Consequently, TpG·CpA in the Y--R context can form BII·BII facing steps, which are otherwise known to enhance the distortions from a straight canonical B helix.[4,6,10,50] This tetranucleotide effect is echoed in the X-ray data, where 16 out of 18 CpA·TpG steps belong to the Y(CpA/TpG)R pattern, leading to BII percentages higher than those found in solution (Table 2). This secondary sequence effect may thus account for the differences observed between crystal and solution BII percentages (flagged by asterisks in Table 2). Unfortunately, in the cases of GpC, CpG, and GpA, the four possible sequence contexts are unequally populated in both solution and crystals, limiting the opportunity to analyze tetranucleotide effects, and more structural work will be needed to clarify these remaining ambiguities.

Overall, the BI/BII propensities associated with dinucleotide steps in solution parallel sufficiently clearly their crystal counterpart to conclude that these propensities obey the same trends and are modulated by the same factors in both environments. In turn, this confirms the validity and the generality of the method developed in this work to quantify the BI/BII ratios in solution.

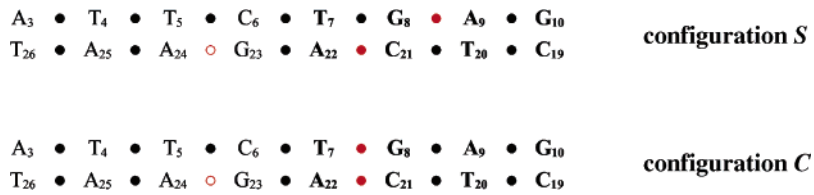
Implications for the Overall Structure of the JunFos Oligomer. The general interest of the above methodology to structural biology is illustrated by its application to a 14bp oligomer containing the 5′-TGACTCA-3′ sequence related to AP-1 sites,[33] the core element recognized by the AP-1 proteins, including Fos and Jun. Sequences outside the AP-1 site influence the affinity of Fos and Jun binding, determine the orientation of the JunFos heterodimer, and contribute to the differences in Fos and Jun recognition at different promoters.[51] This recognition is clearly governed by both direct and indirect readout, and the involved DNA sequences can therefore act as a paradigm to unravel the mechanisms of selective recognition of DNA by proteins.

To demonstrate the concrete use of the information extracted with the above NMR methodology, we present how the derived restraints can be implemented in a structure calculation of the JunFos oligomer, leading to new insights into its overall shape. A fully detailed structure refinement along with a discussion of the many implications for biological function is beyond the scope of this work and will be presented elsewhere. Rather, the aim is to illustrate the general potential and main strength of the method for structure determination of DNA.

An overall structure calculation was carried out with two sets of interproton restraints termed *Res_direct* and *Res_extra*. Set *Res_direct* contains the 82 sequential distances directly measured

(43) Castagne, C.; Murphy, E. C.; Gronenborn, A. M.; Delepierre, M. *Eur. J. Biochem.* **2000**, *267*, 1223−1229.
(44) Castagne, C.; Terenzi, H.; Zakin, M. M.; Delepierre, M. *Biochimie* **2000**, *82*, 739−748.
(45) Khandelwal, P.; Panchal, S. C.; Radha, P. K.; Hosur, R. V. *Nucleic Acids Res.* **2001**, *29*, 499−505.
(46) El Antri, S.; Bittoun, P.; Mauffret, O.; Monnot, M.; Convert, O.; Lescot, E.; Fermandjian, S. *Biochemistry* **1993**, *32*, 7079−7088.
(47) Cordier, C.; Marcourt, L.; Petitjean, M.; Dodin, G. *Eur. J. Biochem.* **1999**, *261*, 722−733.
(48) Bertrand, H.; Ha-Duong, T.; Fermandjian, S.; Hartmann, B. *Nucleic Acids Res.* **1998**, *26*, 1261−1267.
(49) Isaacs, R. J.; Spielmann, H. P. *J. Mol. Biol.* **2001**, *311*, 149−160.
(50) Djuranovic, D.; Oguey, C.; Hartmann, B. *J. Mol. Biol.* **2004**, *339*, 785−796.
(51) Leonard, D. A.; Kerppola, T. K. *Nat. Struct. Biol.* **1998**, *5*, 877−881.

Scheme 1 Central Part of the JunFos Oligomer, with the Bases Belonging to the AP-1 Site in Bold[a]



[a] The black ● and the red ● represent phosphates in BI and BII, respectively. The red ○ represents the phosphate either in BI or in BII. The two phosphate configurations (*S* and *C*) are in equilibrium in solution for the structures obtained by restrained MD, when $C_{21}pA_{22}$ is in BII.

from the NMR spectra. These constraints are heterogeneously distributed along the sequence and under constrain some dinucleotides, for instance with only three distances for the G_8-pA_9·$T_{20}pC_{21}$ facing steps. In set *Res_extra* 18 extrapolated sequential distances were added as compared to *Res_direct*. For instance, the G_8pA_9·$T_{20}pC_{21}$ steps gain five supplementary distances in *Res_extra*. These two sets of constraints were applied for structure calculation with molecular dynamics simulations in explicit solvent. With both *Res_direct* and *Res_extra* the constrained distances are well-respected in the resulting structures without hampering the motions of the phosphate group (no restraints are directly applied to the phosphate groups). Nevertheless, in the *Res_direct* trajectory, the BII percentages are generally not in agreement with those derived from experiment. The discrepancies are particularly evident on the GpA and CpA that are 90% and less than 20% in BII, respectively, whereas their δP values indicate high/moderate and very high percentages of BII, respectively (Table 1). The use of *Res_extra* in the second trajectory clearly improves the fit between δP and the BI/ BII ratios in the structural model, leading to around 35 and 50% of BII for GpA and CpA, respectively. This illustrates the validity and usefulness of the extrapolated restraints.

We focus our analysis on the center of the oligomer that is the junction between the outside sequence ATTC and the TGAG AP-1 site and on structures modeled with *Res_extra*. They display a significant minor groove curvature (>10° for 70% of the time), consistent with a key hypothesis underpinning the interpretation of previous studies.[51] A more precise analysis reveals that the curvature intensity results from a cooperative effect involving the relative position of the steps that are simultaneous in BII, in conjunction with negative rolls. Four steps (T_7pG_8, G_8pA_9, $C_{21}pA_{22}$, and $G_{23}pA_{24}$) are BII-rich in the middle of the oligomer. When $C_{21}pA_{22}$ is in BII (50% of the time in the MD trajectory) $G_{23}pA_{24}$ is in either BI or BII. However, as two successive BII phosphates on the same strand are not stable,[6,49] T_7pG_8 and G_8pA_9 are not in BII simultaneously. This results in an equilibrium between the two backbone configurations, summarized in Scheme 1.

The average curvature of configuration *S* ("straight") is less than 10°. In contrast, in configuration *C* ("curved"), the facing T_7pG_8·$C_{21}pA_{22}$ phosphates create a strong negative roll of −10°. As the roll is the major parameter controlling curvature, this configuration enhances bending toward the minor groove, which reaches 20° on average.

The T_7pG_8·$C_{21}pA_{22}$ steps belong to the AP-1 site, but, as seen above, their capability to adopt a BII·BII combination strongly depends on their neighbors, which are partially (C_6 and G_{23}) outside the recognition region, strictly defined. A similar feature has been described for the NF-κB targets,[10,12,14,25,26] introducing the notion that "extended" recognition sequences, via their nonlocal, unusual deformability, are determinant in selective DNA−protein recognition indirect readout. Although the detailed structure and dynamics of this oligomer necessitate further work, the present structure calculations already indicate that the new type of constraints used in the present approach are instrumental in obtaining an overall DNA structure that can be reconciled with, and maybe explain, the selective recognition of the JunFos oligomer.

Discussion and Conclusion

We first recapitulate on the key methodological elements developed in this work. At its core, the method simply relies on a careful extraction of distance restraints from NOE cross-peaks and assignment of all the δP with unlabeled B-DNA. This highlighted strong linear cross-correlations between the three sequential distances ds(H2′), ds(H2″), and ds(H6/8), themselves correlated to δP. These correlations are also present in very high-resolution X-ray structures where δP can be substituted by $\epsilon-\zeta$ values. The key point is that these correlations establish a quantitative relation between δP and $\epsilon-\zeta$, strictly representative of the BI/BII ratios. The results are general because they were derived from analyzing a large number of different dinucleotide steps in diverse environments.

The interproton and δP correlations allow us to easily and rapidly assess the consistency of the NMR data. The distances that do not obey the correlations may reflect inconsistencies in interpretation of the spectra or may signal unusual backbone conformations ($\alpha/\beta/\gamma$ transitions). Nevertheless, it should be kept in mind that the analysis excluded B-DNA oligonucleotides with bulges or mismatches, which sometimes significantly alter both proton−proton distances and backbone torsion angles. In addition, these correlations allow us to extend and homogenize the set of constraints from one step to the next during the structure calculation. Due to overlaps, one can rarely measure all the sequential distances potentially accessible to NMR along the whole oligomer. This results in an unequal distribution of restraints, mixing well-controlled steps and poorly constrained dinucleotides almost entirely under the influence of the force field. The correlations allow us to extrapolate the distances not directly measured from other, measured distances or from the corresponding δP. This yields a better-balanced set of restraints over the dinucleotides across the oligomer, as illustrated here with the JunFos oligomer. In addition, the well-defined and simple correspondence that we establish between δP and BII percentages in solution will go a long way toward a correct representation of the backbone motions in the DNA structures in solution. Apart from an obvious role to better describe and understand DNA structure, this will provide new and specific tests of the performance of molecular mechanics force fields aimed at representing DNA in solution.

Our results allowed us to infer quantitative BI/BII ratios in

solution from δP published for diverse sequences over many years and to extract the dependency of these ratios on their local sequence environment. We showed that (i) the BII propensities are clearly influenced by sequence in solution and (ii) the effects of sequence in solution on the BI/BII ratios are very similar to those observed in crystals. This second point is an additional and very strong validation of the general character of the method. But more importantly, it shows that there is a limited, and yet well-defined, repertoire of sequences that significantly favor BII in solution. This will equip molecular biologists with simple guides to recognize the sequences of interest, in relationship with structural consequences of biological relevance, discussed below. We anticipate that this will ultimately be a valuable addition to the tools being developed to extract biological information from the immense amount of sequence data produced by genomics projects.

Finally, we used a B-DNA oligomer of special interest for the structural biology of selective protein−DNA recognition to illustrate the quantification of the DNA BI/BII states in solution. We first established the BI/BII populations for JunFos oligomer in solution and their sensitivity to base sequence. Second, we presented how the new method was instrumental in deriving NMR restraints, which, implemented in combination with MD, yielded a structure of JunFos oligomer with BI/BII populations in agreement with experiment and therefore a credible overall structure, especially with respect to its helical curvature. Indeed, the BII states can markedly influence the base displacement, roll, and twist that influence the base accessibility in the grooves and the intensity and direction of the helical curvature. We showed here how the BII steps, modulated by base sequence, contribute to mediate the curvature intensity of JunFos oligomer and thereby would intervene in its selective binding by proteins, as in other DNA transcription factor recognition mechanisms.[10−16,38] Hence, characterizing the BI/BII ratios is a prerequisite to a general and complete description and understanding of protein−DNA recognition mechanisms.

Materials and Methods

Sample Preparation. The JunFos oligonucleotide 5′-d($G_1C_2A_3T_4T_5C_6T_7G_8A_9G_{10}T_{11}C_{12}A_{13}G_{14}$)-3′∞5′-d($C_{15}T_{16}G_{17}A_{18}C_{19}$-$T_{20}C_{21}A_{22}G_{23}A_{24}A_{25}T_{26}G_{27}C_{28}$)-3′ was synthesized by Eurogentec, Inc. (Belgium). The sample was dissolved in an aqueous 10 mM sodium−phosphate buffer with 0.15 M NaCl, 0.1 mM EDTA, and 0.01 mM NaN_3, pH 7.0. The duplex was made by mixing the two complementary strands in a 1:1 ratio. The NMR sample was obtained by lyophilizing the duplex twice and dissolving it in 500 μL of 99.9% ^2H_2O to a final concentration of 5.3 mM.

NMR Spectroscopy. All NMR experiments were carried out on a Varian Unity Inova spectrometer (11.7 T) operating at a proton frequency of 500 MHz and at a phosphorus frequency of 202 MHz with a 5-mm gradient indirect probe. One-dimensional ^1H spectra were collected from 0 to 60 °C by steps of 5 °C to check the stability of the duplex. The experiments were performed at 27 °C, apart from the proton-detected selective heteronuclear experiment, which was performed at 47 °C. A spectral width of 1000 Hz for phosphorus and of 5000 Hz for proton was used.

All 2D proton NMR experiments (COSY−TOCSY−NOESY) were recorded with the phase-sensitive mode using the hypercomplex scheme.[52] Mixing times of 50, 75, 120, 200, and 300 ms were used for NOESY experiments. For the TOCSY experiments, a homonuclear proton spin-lock with MLEV-17[53] pulse sequence was applied with a mixing time of 120 ms. For the ^{31}P-decoupled DQF−COSY experiment, the spectrum was acquired with 8000 data points in the t2 dimension, 256 t1 increments, and a recycle delay of 2 s, and the matrix was zero filled to 16k*4k to obtain a final resolution of 0.6 Hz/points in the F2 dimension.

^{31}P chemical shifts are referred to external phosphoric acid (85%) at 0 ppm. The $^{31}P-^1H$ HSQC and $^{31}P-^1H$ HSQC TOCSY (mixing time of 60 ms) experiments[54−56] were performed with 96 scans in the t2 dimension and 180 t1 increments, and $^{31}P-^1H$ COSY experiments were performed with 80 scans in the t2 dimension and 128 t1 increments. For the HSQC experiments, a delay t/2 was set to 15 ms for optimal polarization transfer.

Strong correlations between the ^{31}P nucleus and H3′ sugar proton of the 5′-adjacent nucleotide and between ^{31}P and H4′ of the 3′-adjacent nucleotide enabled us to assign all the ^{31}P chemical shifts. In addition, every chemical shift was checked by employing a $^{31}P-^1H$ HSQC TOCSY experiment, which allows us to connect the phosphorus nucleus to the H1′ and H2′−H2″ protons of the 5′-adjacent nucleotide.

To measure the $^3J_{H3'-P}$ more accurately, we performed a proton-detected selective heteronuclear experiment with a selective pulse on the H3′ protons.[57,58] The transmitter offset was set on the water signal. The soft pulse was obtained with an i-SNOB-3[59] shape with a flip angle of 180° to invert the H3′ spin population. The bandwidth covered by the soft pulse was 300 Hz with a power level of 21dB for a pulse length of 9 ms. The data were collected with 2496 data points and 320 t1 increments for a ^{31}P spectral width of 600 Hz; the resolution of this experiment was 0.5 Hz/points.

Distance restraints for nonexchangeable protons were derived from well-resolved NOESY cross-peaks in the ^2H_2O spectra, which were integrated over 50-, 75-, 120-, 200-, and 300-ms mixing times. The volume integration of each resolved cross-peak on both sides of the diagonal was measured by summation of intensities in an area around the center of the peak using Felix Software (Accelrys, Inc.). All NOE distances were extracted with particular care: restraints of all cross-peaks were based on visual inspection of the build-up rates in the five different mixing times. For example, we used only the 50-, 75-, and 120-ms mixing time spectra for the short distances such as the sequential $H2''_i-H6/8_{i+1}$. Distances were obtained by scaling the intensities to the average rates of cross-relaxation for the cytosine H5−H6 proton pairs ($r = 2.5$ Å). With this careful protocol, we collected and measured 106 intranucleotide distances and 82 internucleotide distances.

Crystallographic Data. The crystal structures used in the section "Correlations between Sequential Interproton Distances in Solution and Crystals" included the six B-DNA oligomers (Nucleic Acid Data Bank[60] codes bd0023, bd0033, bd0034, bd0035, bd0036, and bd0037) of 1 Å resolution or better. The protons were added with Yasara.[61] For 61 dinucleotides steps, we measured sequential proton−proton distances that must be as precise as possible, given the drastic resolution cutoff. The analysis excluded the terminal base pairs, which are subject to end effects.

Backbone and Helicoidal Parameters. Torsion angles were defined as follows ϵ: $C_{4'}-C_{3'}-O_{3'}-P$; ζ: $C_{3'}-O_{3'}-P-O_{5'}$; α: $O_{3'}-P-O_{5'}-C_{5'}$; β: $P-O_{5'}-C_{5'}-C_{4'}$; and γ: $O_{5'}-C_{5'}-C_{4'}-C_{3'}$. The conventional

(52) States, D. J.; Haberkon, R. A.; Ruben, D. J. *J. Magn. Reson.* **1982**, *48*, 286−292.
(53) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355−360.
(54) Bodenhausen, G.; Ruben, D. J. *Chem. Phys. Lett.* **1980**, *69*, 185−189.
(55) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393−399.
(56) Norwood, T. J.; Boyd, J.; Heritage, J. E.; Soffe, N.; Campbell, I. D. *J. Magn. Reson.* **1990**, *87*, 488−501.
(57) Sklenar, V. B. A. *J. Am. Chem. Soc.* **1987**, *109*, 7525−7526.
(58) Tisne, C.; Simenel, C.; Hantz, E.; Schaeffer F.; Delepierre M. *Magn. Reson. Chem.* **1996**, *34*, s115−s124.
(59) Kupce, E.; Boyd, J.; Campbell, I. D. *J. Magn. Reson. B*, **1995**, *106*, 300−303.
(60) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S. H.; Srinivasan, A. R.; Schneider, B. *Biophys. J.* **1992**, *63*, 751−759.
(61) Krieger, E.; Vriend, G. *Bioinformatics* **2002**, *18*, 315−318.

3-fold staggered notation of the torsions is used: $g+$ (60° ± 40°), *trans* (180° ± 40°), and $g-$ (300° ± 40°). The sugar conformations are categorized following the pseudo-rotation angle (P): "*north*" (P = 300°−60°), "*east*" (P = 60°−115°), and "*south*" (P = 115°−220°). All DNA conformational and helicoidal parameters were obtained with the program CURVES.[62]

Molecular Dynamics Simulations. We performed MD simulations of the JunFos oligomer for 8 ns at 300 K using the program AMBER 8.0,[63] with the Parm98 force field, following a previously described protocol.[4,15,37] The system was neutralized by one Na+ for each DNA phosphate (26 Na^+) and solvated with 6770 TIP3P[64] water molecules in a truncated octahedral box. After 1.5 ns of free MD, NMR constraints were applied for the remaining 6.5 ns. Only the last 6 ns of this second segment of the trajectory was exploited for structural analysis. The constraints consisted only of sequential distances extracted from the NMR spectra. The distance restraints were applied via parabolic potentials of force constant of 10 kcal/mol/Å^2, around a central flat-bottomed well. The flat-bottomed well covered the experimental range of the distances, including experimental errors ($d_{measured}$ ± 10%).

Acknowledgment. We thank Patrick Ladam, Fatiha Kateb, and Christophe Oguey for helpful advice.

JA061686J

(62) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63−91.
(63) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(64) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.