Pergamon

Tetrahedron 54 (1998) 14515–14528

TETRAHEDRON

Partially-Deuterated oligo-DNA Reduces Overcrowding, Enhances Resolution and Sensitivity and Provide Improved NMR Constraints for Structure Elucidation of Oligo-DNA

Tatiana V. Maltseva, András Földesi & Jyoti Chattopadhyaya*

Department of Bioorganic Chemistry, Box 581, Biomedical Centre,
University of Uppsala, S-751 23 Uppsala, Sweden

E-mail: tania@bioorgchem.uu.se or jyoti@bioorgchem.uu.se

Received 30 March 1998; revised 26 September 1998; accepted 30 September 1998

Abstract: The usefulness of the "NMR-window III" approach (see ref. 2e) has been demonstrated through the use of the deuterio isotopomeric 12mer oligo-DNA duplex obtained from protected partially deuterated 2'(R/S),3',5'(R/S)-$^{2}H_3$-2'-deoxyribonucleoside blocks [see preceding paper). The 2D-NMR spectra of this deuterio isotopomeric DNA duplex **I** were compared with those of the natural counterpart (duplex **II**) and deuterium-labelled 12mer duplex **III** (prepared earlier by the "NMR-window II" approach, ref. 2e), and the results can be summarised as follows: (i) The simplification of the crosspeak pattern in H1'-H2'/2" and H4'-H5'/5" areas in the DQF-COSY spectra reduces the spectral crowding, thereby allowing a precise extraction of the coupling constants. (ii) The deuterio isotopomeric mixture of DNA duplex **I** provide hitherto unavailable nOe data sets which are sensitive to the conformational changes of the sugar moiety or the backbone. (iii) The extractable number of nOe constraints based on fully resolved crosspeaks and free from spin diffusion for deuterio isotopomeric dodecamer **I** are 308, whereas they are only 188 for the natural counterpart, suggesting the usefulness of the nOe crosspeaks in the semi-quantitative approach for the NMR-constrained structure refinement.

Introduction

Specific isotope labelling technologies[1] with ^{13}C and ^{15}N has proved to be an important tool to understand how the structure and dynamics of a biopolymer and its complexes dictate specific biological functions. Recently, the development of both chemoselective and chemospecific approaches for incorporation of ^{2}H in nucleoside residues have paved the way for selective non-uniform labelling of oligo-DNA and RNA with these building blocks at specific sites of interest[2]. The non-uniform deuterium labelling (*i.e.* the "NMR-window I and II" approaches)[2a-g] has been developed in our laboratory to aid the assignment problem of the overlapping NMR resonances in large oligo-DNA and RNA, which is one of the most serious problem encountered in the elucidation of their NMR structures. Based on this concept, the solution structures of 21mer[2f] and 31mer oligo-RNAs[2g] have been solved.

We here illustrate the application of our new "NMR-window III" concept to demonstrate the improvement of the number of NMR constraints through a comparative study on the deuterio isotopomeric Dickerson-Drew DNA dodecamer[3] [d(C^{1}G^{2}C^{3}G^{4}A^{5}A^{6}T^{7}T^{8}C^{9}G^{10}C^{11}G^{12})]$_2$ (duplex **I**) (double-underlined nucleotides are of type **A** and **B** deuterated blocks according to the "NMR-window III" concept as schematically shown in Fig. 1 in the preceding paper[4]) in comparison with its natural counterpart (**II**) as well as with an analogous duplex

0040-4020/98/$ - see front matter
PII: S0040-4020(98)00906-5

[d($C^1G^2\underline{C}^3G^4\underline{A}^5A^6T^7\underline{T}^8C^9\underline{G}^{10}C^{11}G^{12}$)]$_2$ (**III**) deuterium labelled according to the "NMR-window II" concept[2e] (the specific site of incorporations of 2'(R[~15%]),S[~85%]),3',4'[~65%],5',5"-$^2H_5$-β-D-2'-deoxyribonucleoside blocks[2e] are shown by bold-underlined letters). It has been shown that (i) an unambiguous determination of the homo- and heteronuclear vicinal coupling constants has been achieved by elimination of large H2'-H2" and H5'-H5" geminal coupling constants, and that (ii) the deuterio isotopomeric mixture of DNA duplex **I** provides hitherto unavailable nOe data sets such as H5'/H5"-H2", H4'-H2", H4'-H5'/H5", H1'-H5', H1'-H5", H1'$_i$-H5'$_{i+1}$, H1'$_i$-H5"$_{i+1}$, H1'$_i$-H4'$_i$, H1'$_i$-H4'$_{i+1}$, arom-H5'/H5", which are sensitive to the conformational changes of the sugar moiety or the backbone. The extractable number of nOe constraints based on fully resolved crosspeaks for deuterio isotopomeric dodecamer **I** are 308, whereas they are only 188 for the natural counterpart, suggesting the usefulness of the "NMR-window III" approach in the semi-quantitative 3D-structure refinement procedure[11].

Results and Discussion

(A) The conformational analysis of torsion angles.

In oligonucleotides, accurate homonuclear $^3J_{HH}$ coupling constants provide information about sugar pucker and C4'-C5' torsion, whereas the heteronuclear $^3J_{HP}$ coupling constants give information about the phosphodiester backbone torsions. In this work, we show that an isotopomeric mixture of partially-deuterated nucleosides provides more unambiguous homonuclear $^3J_{H1'H2'}$, $^3J_{H1'H2"}$, $^3J_{H4'H5'}$ and $^3J_{H4'H5"}$ and heteronuclear $^3J_{PH4'}$, $J_{PH5'}$ and $^3J_{PH5"}$ based on routine DQF-COSY experiment compared to the natural counterpart[2c,e].

(i) Influence of deuterium on the determination of proton-proton coupling constants. We have not observed any geminal H/D coupling in any of our partially-deuterated building blocks or in the 12mer duplex in 1D experiments even at higher temperature. This is apparently owing to the strong effect of scalar relaxation of the second kind[22] of ^2H, which decouples the splitting of ^1H by ^2H. It is known that the contribution of the scalar relaxation of the second kind to the transverse relaxation could be eliminated by the deuterium decoupling with large field strengths[21]. Indeed, in 1D ^1H spectra for partially-deuterated nucleoside the narrowing (25% of linewidth, about 0.5-0.8 Hz) of 2' or 5' protons (*i.e.* geminal to deuterium) in 2',5'-partially-deuterated nucleoside blocks (see the preceding synthetic paper) has been observed upon application of deuterium decoupling. Nevertheless, for the same deuterated nucleosides we did not observe any noticeable difference in the linewidth of the 1'-2' or 4'-5' crosspeaks in the DQF-COSY or E-COSY spectra collected with 4Kx1K data sets (obtained with and without deuterium decoupling). The same result was also obtained in DQF-COSY experiments for duplex (**I**). Hence, we conclude that proton-deuterium interaction is negligible for the determination of vicinal proton -proton coupling constants. Nevertheless, all experiments performed in this paper (see below) have been carried out with the deuterium decoupling on, which is also beneficial for the suppression of HDO resonance.

(ii) The extraction of $^3J_{H1'H2'}$ and $^3J_{H1'H2"}$ for determination of sugar conformation in DNA. We have shown[2e] that the incorporation of a 2'/2"-deuterio isotopomeric mixture of building blocks into an oligo-DNA allows extraction of both $^3J_{H1'H2'}$ and $^3J_{H1'H2"}$ from the DQF-COSY spectra (Fig. 1) in a straightforward manner with

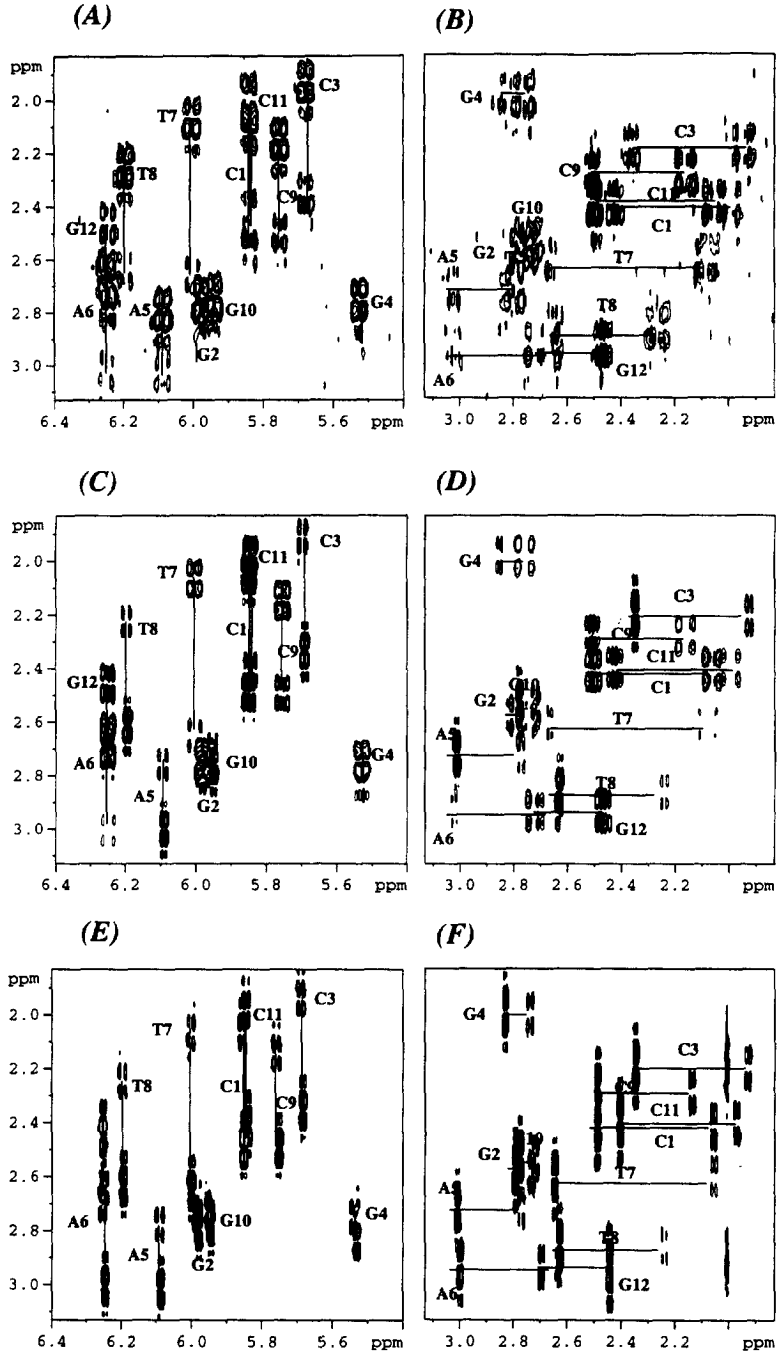

Figure 1: Expanded H1' - H2'/H2" [located on up diagonal: panels (A), (C) and (E)], and H2'/H2" - H1' regions [located on down diagonal: panels (B), (D) and (F)] of the DQF-COSY spectra of the natural Dickerson's dodecamer **II** [panels (A), (B)], its deuterated analogue **III**, [d($C^1G^2\underline{C}^3G^4\underline{A}^5A^6T^7\underline{T}^8C^9\underline{G}^{10}C^{11}G^{12}$)]$_2$ [panels (C), (D)] and the deuterium labelled duplex **I** [panels (E), (F)]. The assignments of the crosspeaks are labelled by name of nucleoside and its sequential number in duplexes. The crosspeaks H1'-H2' and H1'-H2" are connected by solid lines.

improved precision. This was possible due to two factors: (i) The elimination of the geminal passive $^{2}J_{H2'H2''}$, and (ii) the increase of T_2 of H1' and H2" in the partially-deuterated residue, which decreases the linewidth with an overall increase of the crosspeak intensity owing to the elimination of a number of relaxation pathways [compare Fig. 1(B) for natural dodecamer **II** with Fig. 1(F) for the deuterated counterpart **I**]. The reduction of the linewidth (Δ) in deuterated compounds (**I** and **III**) compared to the natural counterpart [compare Fig. 1(A), (C) and (E)] shows that $^{3}J_{1',2} > \Delta$ and $^{3}J_{1',2''} \geq \Delta$, thereby allowing more accurate estimation of these coupling constants through the cross-section of the crosspeaks in the F2 dimension. Thus the method of Kim and Prestegard[5] could easily be applied for the determination of coupling constants from AX type spectra in which the splittings of the absorptive and dispersive signal are used, not the knowledge of the linewidth[5,6], which is impossible to obtain in case of natural counterpart because of the complex coupling pattern. The simplification of crosspeak pattern resulting from $^{3}J_{H1'H2''}$ doublet causes the increase of the intensity of the crosspeak that, in turn, allows a quantitative determination of $^{3}J_{H1'H2''}$ through a simulation procedure with higher precision (~0.2 Hz) of the fit to the coupling. Moreover, the equally well resolved H1'-H2' and H1'-H2" crosspeaks located above and below the diagonal in DQF-COSY spectra [Fig. 1(E) and 1(F)] allow a straightforward estimation of the experimental error in the determination of the J-coupling, which has been a problem in case of the natural oligonucleotides where the crosspeaks located above the diagonal could be reliably used for the determination of $^{3}J_{H1'H2'}$ and $^{3}J_{H1'H2''}$ (below the diagonal, the crosspeaks are complicated by the passive couplings). Obviously the simplification of the crosspeak pattern in the H1'-H2'/2" area of DQF-COSY spectra reduces the spectral crowding in this area which in turn allows an unambiguous extraction of the $^{3}J_{H1'H2'}$ and $^{3}J_{H1'H2''}$ for large DNA molecule to elucidate the conformation of the sugar moiety.

(iii) The extraction of $^{3}J_{H4'H5'}$, $^{3}J_{H4'H5''}$, $^{3}J_{PH4'}$, $J_{PH5'}$ and $^{3}J_{PH5''}$ to estimate the backbone conformation. The determination of homonuclear $^{3}J_{H4'H5'}$ and $^{3}J_{H4'H5''}$ in an oligo-DNA molecule with middle to large size has posed a considerable challenge because of the severe increase of the resonance overlap of H4' to H5'/5" compared with H1' to H2'/2". The use of ^{13}C labelling at C-5' of an oligo- RNA[7,8] or DNA[9,10] is one of the ways to resolve this problem. In this case, the HCCH-E.COSY type experiments are normally used[9,11] and the J-couplings are derived from the crosspeak pattern that correlates carbon with a two- bond away proton. The multiplet patterns of crosspeaks from C2'-H1' and C5'-H4' give the sum[9] of the $^{3}J_{H1'H2'} + {}^{3}J_{H1'H2''}$ and $^{3}J_{H4'H5'} + {}^{3}J_{H4'H5''}$ couplings, respectively. To extract information from these crosspeaks, thegeminal H5'- H5" and/or H2'-H2" J-couplings should be taken into account. In the present study, we demonstrate an alternative approach where a deuterio isotopomeric mixture of building blocks at C5'(H/D) and C5"(H/D) was used in the construction of an oligo-DNA. Here, we have shown that the homonuclear $^{3}J_{H4'H5'}$ and $^{3}J_{H4'H5''}$ coupling constants could be determined from routine DQF-COSY experiment. In Fig. 2, the expanded parts of H4'-H5'/5" area of the DQF-COSY spectra with phosphorus decoupling [Fig. 2(A) and (B)] and without phosphorus decoupling [Fig. 2(C) and (D)] are presented. It is noteworthy that for purine nucleotide residues ($\underline{\underline{A}}^{5}$, $\underline{\underline{A}}^{6}$, $\underline{\underline{G}}^{2}$, $\underline{\underline{G}}^{4}$, $\underline{\underline{G}}^{10}$ and $\underline{\underline{G}}^{12}$) in duplex (**I**), the H4'-H5' and H4'-H5" crosspeaks are well dispersed because of the large downfield shift of H-4'

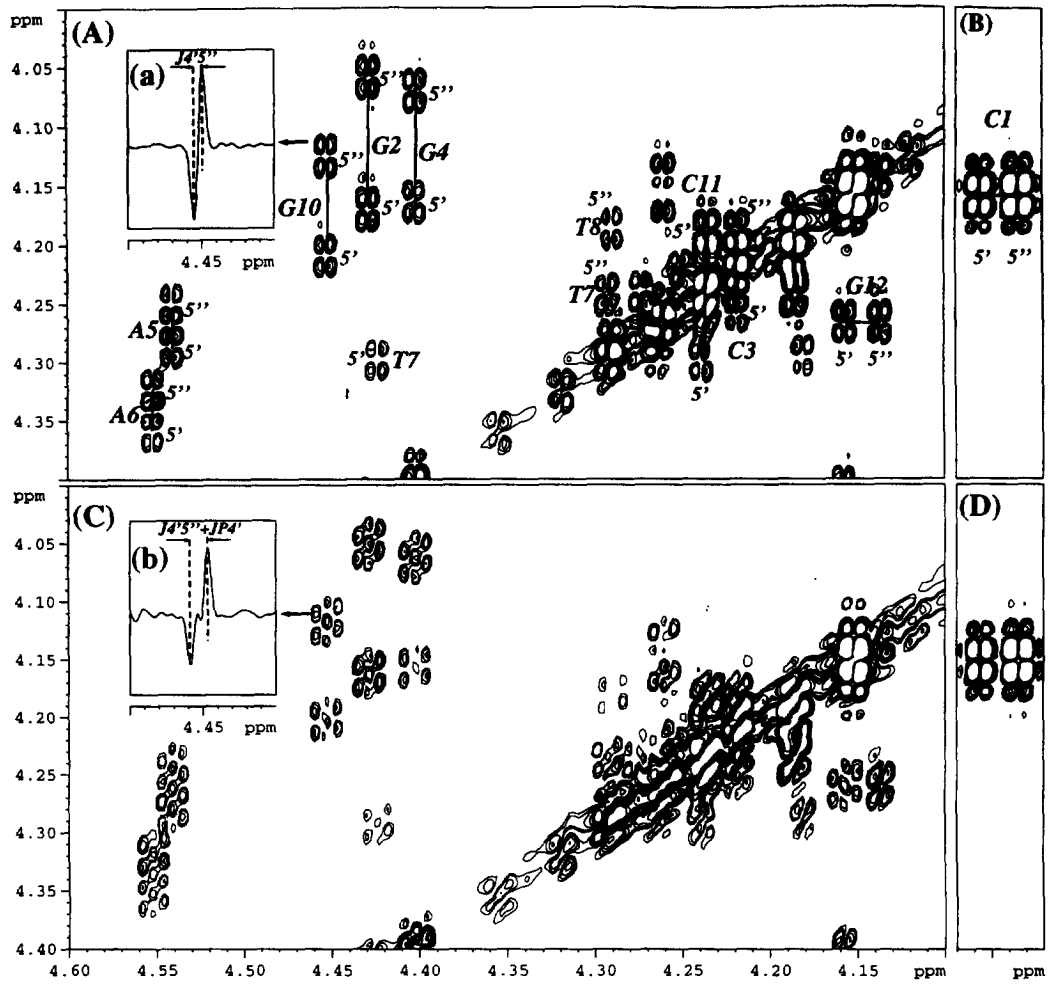

Figure 2: Expanded H4'-H5'/H5" regions of the DQF-COSY spectra of duplex **I** with ^{31}P decoupling [panels (A), (B)] and without [panels (C), (D)]. The crosspeaks are labelled by name of nucleoside and its sequential number. The H4'-H5' and H4'-H5" crosspeaks in panels (A) are connected by solid lines. In boxes (a) and (b) in panels (A) and (C), the cross sections through the crosspeaks at the arrows near the boxes are presented. In boxes (a) and (b), the splitting between positive and negative components is indicated by dashed lines and the corresponding coupling constants responsible for this splitting are indicated.

resonances in comparison with H5' and H5". For pyrimidine residues ($\underline{\underline{T}}^7$, $\underline{\underline{T}}^8$, $\underline{\underline{C}}^3$, $\underline{\underline{C}}^9$ and $\underline{\underline{C}}^{11}$), except for the terminal residue $\underline{\underline{C}}^1$, the situation is more severe because the H4's are located close to H5' and H5", and the crosspeaks have appeared close to the diagonal. Nevertheless, in 2D DQF-COSY experiment with 1K increments in F1 dimension, these peaks also could be resolved [Fig. 2(A), (B)] successfully. The pattern of H4'-H5' and H4'-H5" crosspeaks in the DQF-COSY spectra with phosphorus decoupling is expected to be a simple doublet with antiphase component because of the elimination of the geminal H5'-H5" coupling constant [see Fig. 2(A), the cross-section through crosspeak is presented in the panel (a) of Fig. 2(A)]. These crosspeaks carry straight information about $^3J_{H4'H5'}$ and $^3J_{H4'H5"}$ which could be extracted by either simulation of the crosspeak or through

the analysis of the cross-section of the crosspeak. The pattern of the H4'-H5' and H4'-H5" crosspeaks in the DQF-COSY spectra without phosphorus decoupling is more complicated and looks like an E.COSY type pattern [see Fig 2(C)] because of the presence of heteronuclear passive $^3J_{PH4'}$, $^3J_{PH5'}$, $^3J_{PH5"}$ for crosspeaks located above and below the diagonal areas, respectively. The cross-section through this crosspeak in the F2 dimension is presented in the panel (b) of Fig. 2(C). The difference between outer antiphase lines relates the information about the sum of the $^3J_{H4'H5'} + {}^3J_{PH4'}$ and $^3J_{H4'H5"} + {}^3J_{PH4'}$ [for the crosspeaks above the diagonal] and $^3J_{H4'H5'} + {}^3J_{PH5'}$ and $^3J_{H4'H5"} + {}^3J_{PH5"}$[for the crosspeaks below the diagonal].

(iv) The extraction of $^3J_{31P\text{-}13C}$ coupling. Another application of deuterated nucleoside with natural abundance carbon is based on the analysis of the shape of ^1H-$^{13}C(2')$ crosspeak in the folding HSQC experiment to analyze $^3J_{31P\text{-}13C}$. In Fig 3(A), the folding HSQC experiment[12] with ^2H decoupling on duplex (**I**) is presented. The elimination of geminal coupling in the sugar benefits to the increasing of intensities of ^1H-$^{13}C(2')$ crosspeaks what is clearly observed on Fig 3(A). This experiment also provides an unique opportunity for the determination of $^3J_{31P\text{-}13C(2')}$ which is very important for distinguishing between B_I and B_{II} type of sugar conformation of DNA. Our data show that the $^3J_{31P\text{-}13C(2')}$ for 12mer investigated in this work are well below the linewidth of crosspeaks in F1 dimension (<2 Hz), which corresponds to the B_I type conformation of DNA for all residues[13].

(B) The dipole-dipole relaxation in the partially-deuterated DNA duplexes containing deuterated nucleotide blocks [2'/2"(R/S),3',5'/5"(R/S)-$^2H_3$].

The majority of constraints for the structural analyses of biomolecules is based on the interproton distance information derived from NOESY type experiments. As usual, the analysis of NOESY spectra can be performed using one of the following ways: (i) The first is based on an approximation in which it is assumed that the ratio of the intensity of the crosspeak (a_{ij}) to intensity of the corresponding diagonal peak (a_{ii}) is linearly proportional both to the mixing time (τ_m) in which the magnetization transfer is taken place and to the cross-relaxation rate (σ_{ij}) between observed spins i and j. This dependence is presented by the first term in equation (1) and only valid at very short mixing time for the fully protonated DNA:

$$\frac{a_{ij}}{a_{ii}} \approx -\sigma_{ij}\tau_m + \tfrac{1}{2}\sum_{k\neq i,j}\sigma_{ik}\sigma_{kj}\tau_m^2 \qquad \text{.....} \qquad \text{eqn (1)}$$

where, the (σ_{ij}) is related to the inverse sixth power of the distance (r_{ij}) between proton i and j, and proportional to the correlation time (τ_c). For the rigid body molecular tumbling approximation, it is assumed that τ_c is similar for all protons and r_{ij} is obtained from the ratio of two cross-relaxation rates, one of which is unknown but the other is obtained from a reference crosspeak with a fixed distance such as H5-H6. The main limitation of this method is the influence of the spin diffusion [*i.e.*, the second term in eqn (1)] involving protons in the spatial proximity. In this case, the direct dipolar interaction between spins i and j competes with indirect pathways where transfer takes place through a third spin which is located closer to spin i than the direct spin j. This situation is

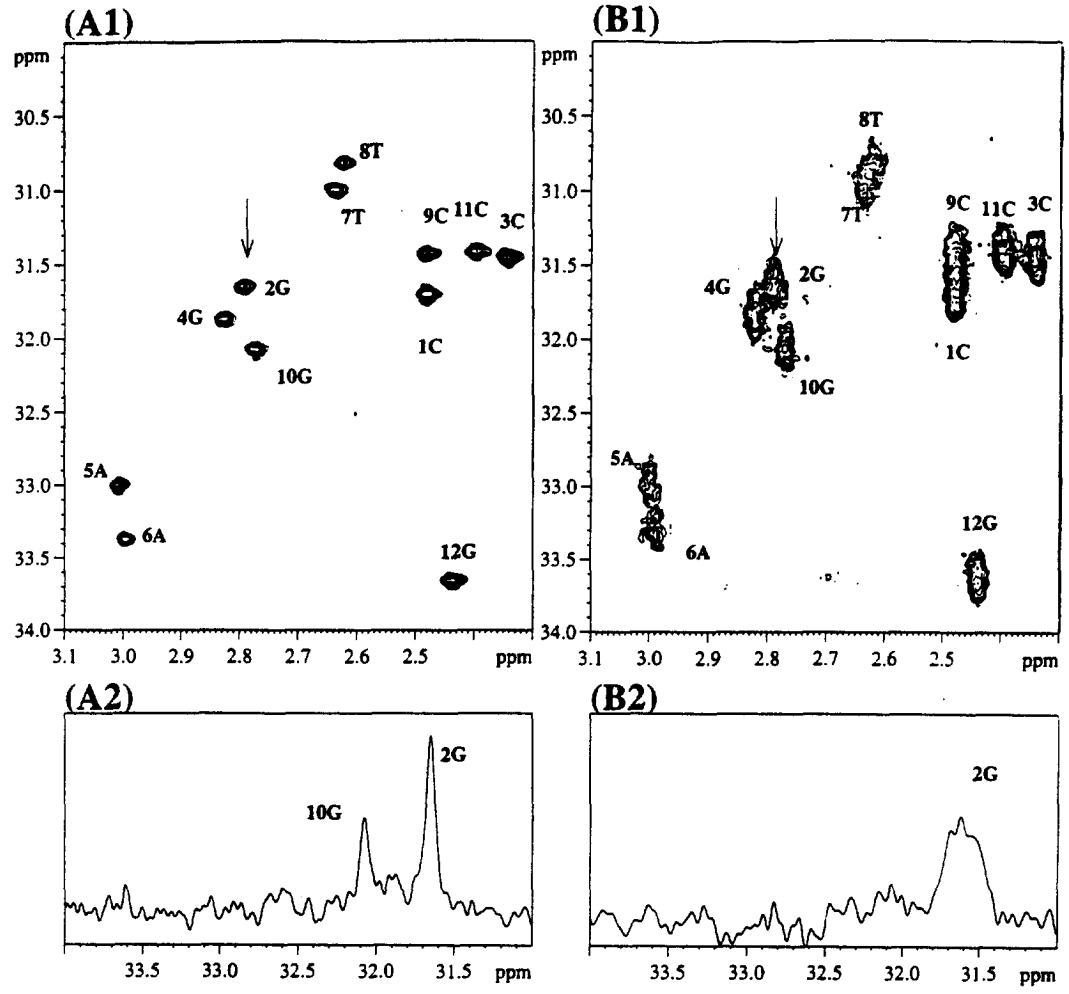

Figure 3: Comparison of two [^{1}H,^{13}C]-HSQC spectra with deuterium decoupling [panel A] and without [panel (B)] of duplex (I). Both panels A and B show the H2'/H2"-^{13}C(2') region, but only the H2"-^{13}C(2') crosspeaks are observed (marked by the residue number) because H2' has been replaced by ^{2}H in all 2'(R/S),3',5'(R/S)-$^2H_3$-2'-deoxynucleosides. The projection through the H2"-^{13}C(2') crosspeak in F1 dimension has been shown, only as an example, for ^{2}G residue (a part of of the signal of ^{10}G residue is also seen) in panel A1 in deuterium decoupled mode showing that the pseudosinglets for $^3J_{^{31}P-^{13}C(2')}$ are visible because this coupling constant is less than the linewidth of the resonance (<2 Hz). In panel B, there is only ^{31}P decoupling applied (no deuterium decoupling), thereby allowing us to observe the ^{13}C(2')-^{2}H splitting in the methylene triplet for ^{2}G residue.

readily observed for H2' and H2" and H5' and H5" in DNA molecule where the effect of the spin diffusion from the geminal proton partner is dramatic[2e,14]. Another limitation of the two-proton approximation is the assumption of a single isotropic correlation time. (ii) To overcome the above spin diffusion problem, a second method has been developed which relies on the solution of the relaxation matrix describing all dipolar interaction of a full spin system[15]. To obtain a well-defined structure by this method, accurate NOE distance constraints are preferable. (iii) The third approach to obtain distance constrains is semi-quantitative, in which a large number of nOes are

used in the form of crosspeak intensities which are categorised as strong (1.8-3.0 Å), medium (1.8-4.0 Å), week (1.8-5.0 Å) or very weak (1.8-7.0 Å)[2g,16].

All these approaches have been used in the determination of the structure of DNA and RNA. It has been however shown[11] for RNA that structures calculated from a dataset of large number of unprecise distances are more accurate than those based on fewer and precise distances.

Recently[2e] we have attempted to solve some of these issues by developing new deuterium isotope labelling techniques for non-uniform labelling of the oligo-DNA or RNA. We have also shown[2e] that in a partially-deuterated isotopomeric oligo-DNA with ^2H in 2' and 2" positions, the intraresidual spin-diffusion and the relaxation pathways can be eliminated allowing extraction of accurate intraresidual distances $(H1')_i$-$(H4')_i$ and $(H2")_i$-$(H4')_i$ which are very sensitive to the sugar conformation. In our "NMR-window II" approach, relaxation pathways between H2'-Aromatic (Ar), H1'-H2', H3'-Ar, H1'- H3', H2'-H3' and H2'-H2" can be eliminated, enabling[2e] us to quantitatively estimate the build-up curves for the $(H2"\text{-}Ar)_i$ crosspeaks depending upon the mixing time (τ_m), which are known to be affected by spin-diffusion through the H2'-Ar and H2'-H2" pathways. This data allowed us to obtain the nOe distance constraints with improved precision, and could be readily used either by two-proton approximation method or by full matrix relaxation analysis. Nevertheless, the most serious disadvantage of this deuteration approach is that it reduces the number of nOe constraints for natural protonated DNA which already has a low density of protons relative to proteins.

In this work, we have studied the possibility to use the isotopomeric mixture of deuterated sugar moiety to obtain more number of nOe constraints. Following conclusions can be drawn from a comparative analysis of NOESY spectra of two types of partially-deuterated DNA duplexes **I** and **III** with their natural counterpart **II**:

(i) Fig. 4 shows the panels with expanded part of the NOESY spectra in Ar-H1', Ar-H3', Ar-H4', Ar-5',5" [Figs. 4(A), 4(C) and 4(E)] and Ar-H2', 2" [Figs. 4(B), 4(D) and 4(F)] regions for all three types of dodecamers. As expected, the Ar-H1' area of the three duplexes are very similar, the largest deviation of chemical shifts of aromatic and H1' protons was not more than 0.01 ppm for the terminal residues suggesting that isotope-induced chemical shift is negligible. This means that the information from one deuterated analogue to the other (or to natural) can be easily extrapolated or superimposed. In the Ar-H2', H2" area, there is a clear change of the intensity of crosspeaks in the deuterated residues compared to the natural counterpart. Note that in the deuterated residues of duplexes **I** and **II** the intensity of peaks increases and the linewidth decreases because of the elimination of the geminal coupling constant splitting of H2' and H2" resonances. The more dramatic change has been observed for Ar-H3', Ar-H4', Ar-H5'/5". Since ^1H at C3' is substituted by ^2H for all deuterated building blocks, the crosspeaks for all nucleosides in Ar-H3' area in dodecamer **I** [Fig. 4(E)] and for four nucleosides in dodecamer **III** [Fig. 4(C)] are absent. Instead, the dramatic improvement in the resolution of Ar-H4', Ar-H5'/5" crosspeaks [Figs. 4(A), 4(C) and 4(E)] is observed. Indeed, for duplex **I**, all Ar-H4' and Ar-H5'/5" crosspeaks are clearly resolved [Figs. 4(E)] compared with the natural counterpart [Fig. 4(A)]. It can be also noted that for the nondeuterated residues in duplex **III** the Ar-H4' and Ar-H5'/5" areas are very poorly resolved, and hardly could be used, partly because of broadening of Ar-H5'/5" by geminal H5'-H5" coupling compared to the deuterated residues which have no H5'/5" protons [Fig. 4(B)].

(ii) In Fig. 5, the panels with H1'-H3', H1'-H4', H1'-H5', H1'-H5" [Figs. 5(A), 5(C), 5(E)] and H1'-H2', H1'-H2" [Figs. 5(B), 5(D), 5(F)] areas of the spectra are presented. In the H1'-H2'/H2" areas the crosspeak intensities of deuterated residues are increased in a similar way as for those in Ar-H5'/H5" areas

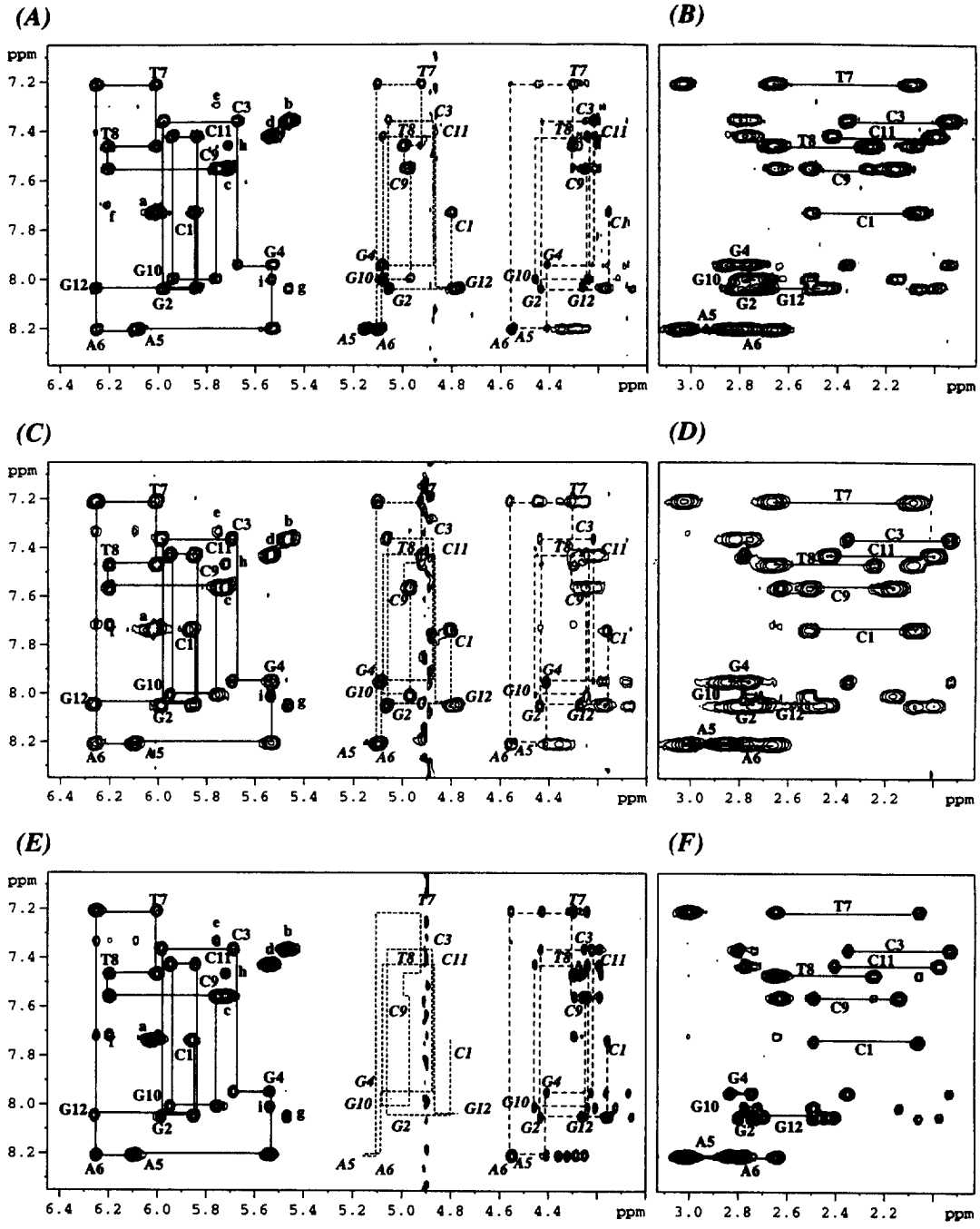

Figure 4: Expanded Ar-H1'/H3'/H4'/H5'/H5" regions [panels (A), (C) and (E)] and Ar-H2'/H2" [panels (B), (D) and (F)] of the NOESY spectra of the natural dodecamer II [panels (A) and (B)], and its deuterated analogues duplex III [panels (C), (D)] and I [panels (E), (F)]. The connectivity pathways from the first (C^1) to the last (G^{12}) nucleosides are presented by solid lines for 5'Ar(i)-H1'(i) → H1'(i) - Ar(i+1) connectivities and by dushed lines for 5'Ar(i)-H3'(i) → H3'(i) - Ar(i+1) and 5'Ar(i)-H4'(i) → H4'(i) - Ar(i+1) connectivities. The intraresidual crosspeaks are labelled by name of nucleoside and its sequential number in duplexes. The Ar-H2' and Ar-H2" crosspeaks are connected by solid lines.

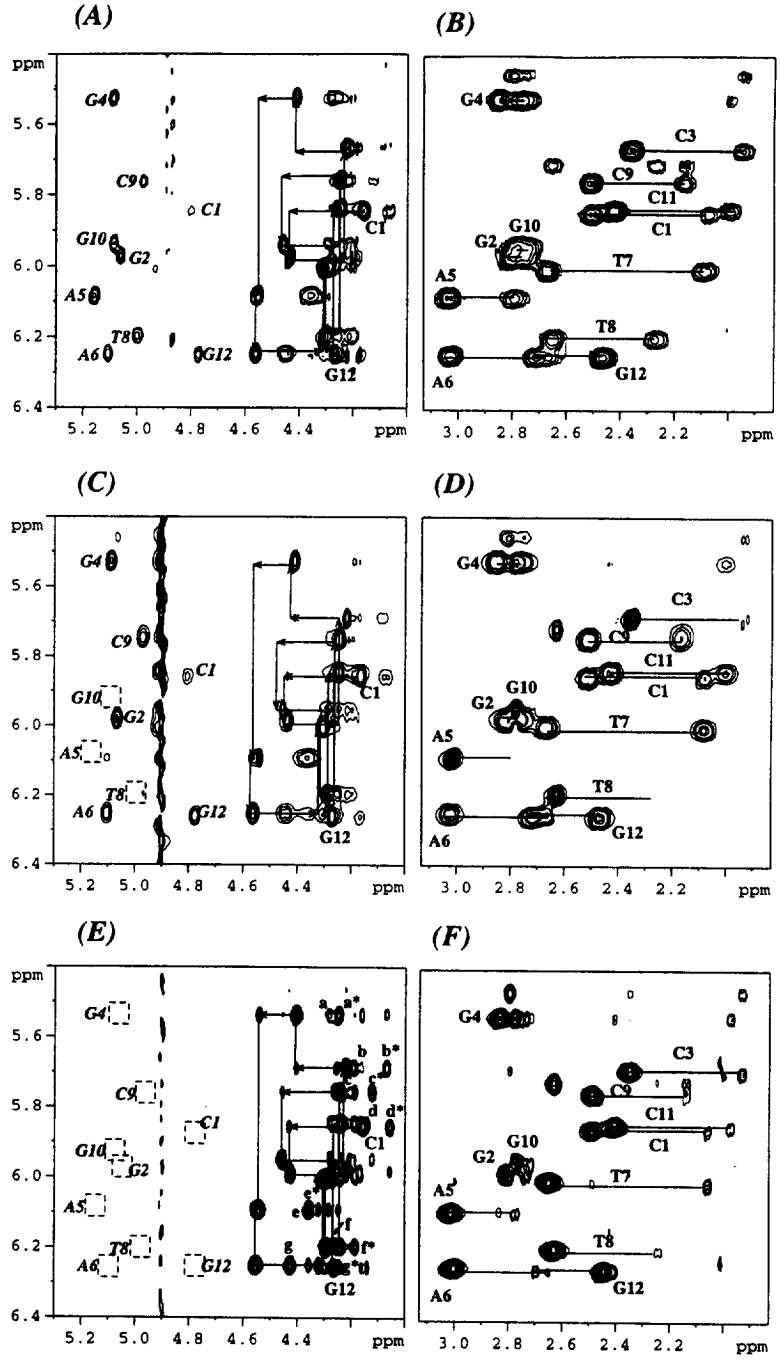

Figure 5: Expanded H1'-H3'/H4'/H5'/H5" [panels (A), (C) and (E)] and H1'-H2'/H2" regions [panels (B), (D) and (F)] of NOESY spectra of duplex **II** [panels (A), (B)] and its deuterated analogues duplex **III** [panels (C), (D)] and **I** [panels (E), (F)]. The connectivity pathways from C^1 till G^{12} nucleosides are presented by solid lines for 5'H1'-H4'(i) → H1'(i) - H4(i+1). The H1'-H2' and H1'-H2" crosspeaks are connected by solid lines and are labelled by name of nucleoside. The intraresidual H1'-H3' crosspeaks are labelled by name of nucleoside and its sequential number in duplexes. The boxes inside panels (C) and (E) show the absence of these H1'-H3' crosspeaks for the deuterated duplex **III** and duplex **I**, respectively. In panel (E) the interresidual H1'(i)-H5'(i+1) (labelled with small alphabetic latters) and H1'(i)-H5"(i+1) crosspeaks (labelled with small alphabetic latter with *) are indicated.

because of elimination of geminal H2' and H2" coupling constants. The dramatic improvement is achieved in H1'-H3', H1'-H4', H1'-H5', H1'-H5" areas. For deuterated nucleotides in duplex **I**, it is clearly possible to trace the connectivity walk, $H1'_i$-$H4'_i \rightarrow H1'_i$-$H4'_{i+1}$ [compare Fig. 5(A) with 5(E)] starting from the terminal nucleotide G^{12} or C^1. Moreover, mostly all crosspeaks, H1'-H5', H1'-H5" and $H1'_i$-$H5'_{i+1}$ and $H1'_i$-$H5"_{i+1}$ are clearly distinguishable in the spectra shown in Fig. 5(E) for duplex **I** but certainly not in duplex **II** [Fig. 5(A)] and **III** [Fig. 5(C)].

(iii) In Fig. 6, the panels with H3'/ H4'/ H5'/ H5" - H3'/ H4'/ H5'/ H5" [Figs. 6(A), 6(C), 6(E)] and H3'/ H4'/ H5'/ H5" - H2'/ H2" [Figs. 6(B), 6(D), 6(F)] areas of the spectra are presented. Note that even though the H3'-H4', H3'-H5'/5" crosspeaks are missing in H3'/ H4'/ H5'/ H5" - H3'/ H4'/ H5'/ H5" areas for deuterated duplex **I**, this causes no harm because the information from those crosspeaks could not be extracted anyway in the natural counterpart, duplex **II**, or from the nondeuterated part of duplex **III** because of the line broadening and severe spectral crowding in that region of the spectra. Instead, in the H5'/H5"/H4'-H2'/H2" area of duplex **I** the *conformation-sensitive* H4'-H2" crosspeaks are clearly observable and could be used for unambiguous integration of their nOe volumes. Moreover, H4'-H5'/H5" crosspeaks are also distinguishable, and those volumes could be easily extracted which, to the best of our knowledge, was impossible for natural oligo-DNA molecule of middle to larger size.

The above data show that the observed nOe crosspeaks in the diasteremeric mixture of oligo-DNA (duplex **I**), although can not be used for full-relaxed NOESY matrix calculations[15], will be however immensely useful in setting up qualitative nOe constraints[11] as strong-medium-week basing on their relative intensities.

Conclusions

Following are the conclusions of our analysis of DQF-COSY and 2D NOESY spectra of deuteroisotopomeric mixture of duplex **I** [~85% of the total diastereomeric mixture for each sugar residue has natural H2" in the α–face and ~15% natural H2' at the β-face, and the rest is ^2H at C2'; the C5' has a 1:1 diastereomeric mixture of deuterio isotopomers; *i.e.* the total number of isotopomers: 12^4]:

(1) The advantages to perform the J-coupling analysis of isotopomers for reasonably large DNA duplex is evident from the simplification of the crosspeak pattern through elimination of the large germinal 5'/5" and 2'/2" coupling constants. It is noteworthy that we have not observed any geminal H/D coupling in any of our partially-deuterated building blocks or in the 12mer duplex in 1D experiments even at higher temperature. This is apparently owing to the strong effect of scalar relaxation of the second kind[22] of ^2H, which decouples the splitting of ^1H by ^2H, thereby allowing us to determine the vicinal homonuclear coupling constants with higher accuracy than in the natural counterpart.

(2) These deuterio isotopomeric mixtures of DNA provide more complete nOe information (with large dataset) compared with the natural counterpart, and they can be used as distance constraints in a semi-quantitative approach[11] as strong-medium-week nOes, basing on their relative intensities. The new nOe dataset obtainable from the our partially-deuterated 12mer compared to the native counterpart includes the nOes which are indeed sensitive to the conformation of sugar moiety or for the backbone conformation such as H5'/H5"-H2", H4'-H2", H4'-H5'/H5", H1'-H5', H1'-H5", $H1'_i$-$H5'_{i+1}$, $H1'_i$-$H5"_{i+1}$, $H1'_i$-$H4'_i$, $H1'_i$-$H4'_{i+1}$, Ar-H5'/H5". It is

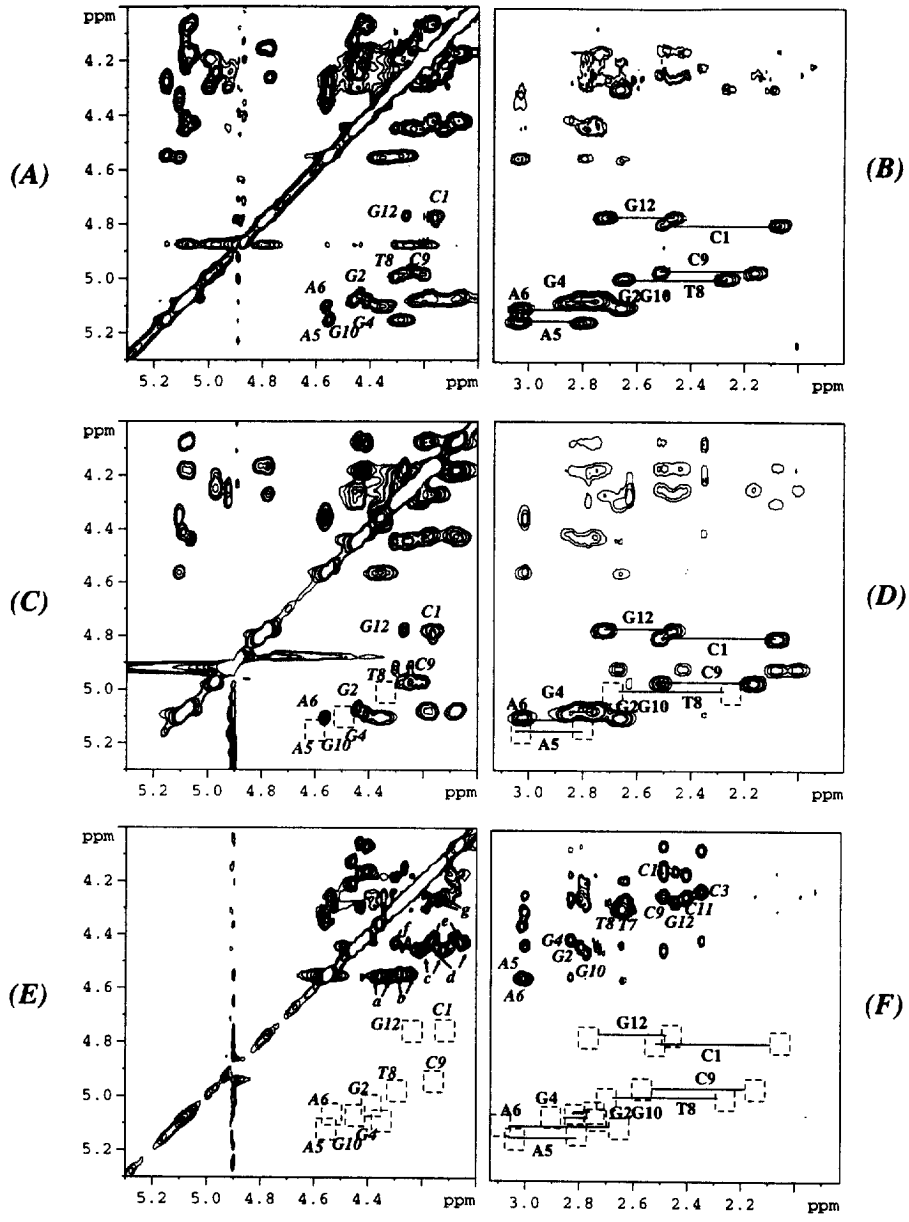

Figure 6: Expanded H3'/H4'/H5'/H5" - H3'/H4'/H5'/H5" [panels (A), (C) and (E)] and H3'/H4'/H5'/H5" - H2'/H2" regions [panels (B), (D) and (F)] of NOESY spectra of the natural duplex **II** [panels (A), (B)], and its deuterated analogue **III** [panels (C), (D)] and **I** [panels (E), (F)]. The H3'-H2' and H3'-H2" crosspeaks are connected by solid lines and labelled by name of nucleoside. The intraresidual H1'-H3' crosspeaks are labelled by name of nucleoside and its sequential number in duplexes. The boxes inside panels (D) and (F) show the absence of these H3'-H2', H3'-H2" crosspeaks in the spectra of duplex **III** and duplex **I**, respectively. In panel (F) there are well resolved H4'-H2" crosspeaks labelled by name and number of residue which could not be distiguished in corresponding panels (B) and (D) of natural duplex **II** and labelled **III**, respectively. The boxes inside panels (C) and (E) show the absence of H3'-H4' crosspeaks in the spectra of deuterated duplex **III** and duplex **I**, respectively. In panel (E) belonging to the spectrum of duplex **I**, the crosspeaks H4'-H5' and H4'-H5" are indicated by small alphabetic latters and picked by arrows.

noteworthy that these nOes are simply not accessible in the NOESY spectra of the fully protonated DNA because of severe spectral overlap. Indeed, the observed crosspeaks in duplex **I** are free from one of the strongest pathway of spin diffusion, *i.e.* through geminal H2' and H2" and H5' and H5" protons, and their intensities are linear upto the 300 ms of mixing time. The trustworthiness of this approach is higher with increasing of number of nOe constraints[11]. The estimation of number of possible nOe constraints based on fully resolved crosspeaks (Figs. 4-6) for natural Dickerson-Drew dodecamer **II** and the analogous deuterium modified dodecamer **I** are 188 and 308, respectively.

(3) It thus clearly emerges from our above work that deuterio isotopomeric mixture of oligo-DNA provides reliable data set both for determination of accurate torsional angles from the COSY-type experiments as well as for the determination of spatial distances using the semi-quantitative approach[11]. Clearly, the use of the larger set of experimental NMR constraints from deuterium modified oligo-DNA (compared to the native counterpart) would enable structure elucidation of DNA molecules that would reflect more experimental NMR structure in solution than the algorithm of the molecular mechanics forcefield.

Experimental Section

The NMR experiments were carried out on a Bruker DRX spectrometer at magnetic field strength 14.1T, operating at 600.13 MHz for ^{1}H, 242.93 MHz for ^{31}P, 150.91 MHz for ^{13}C and 92.12 MHz for ^{2}H. The spectrometer was equipped with a Bruker digital lock and with a switching ^{2}H lock - ^{2}H pulse device and with inverse triple-resonance probehead for ^{1}H, ^{13}C and ^{31}P (TXI). ^{1}H, ^{13}C and ^{31}P pulses on this probehead were applied with a 31.7 kHz, 17.2 kHz and 15.4 kHz field strengths, respectively. ^{13}C and ^{31}P decoupling were performed using GARP[17a] with a 4.17 kHz and 1.79 kHz field strengths. For the 90° and 180° ^{2}H pulses the probe power after the switching block was 6.4 W, which corresponds to 2.08 kHz applied field. ^{2}H decoupling utilised a WALTZ16[17b] sequence using a 625 Hz field.

For HSQC[12,18] experiment , the total measurement time was ~18 h: 1292 x 4K complex t1 and t2 points were acquired with acquisition times of 0.077 and 0.170 s, respectively. The number of scans were 128 or 256. The experiments were performed with extensive folding[11] in F1 dimension: the sweep width of ^{13}C dimension was 28 ppm with the carrier for ^{13}C at 38 ppm in $^{13}C(2')$ region. The apodization, zero-filling and Fourier transformation led to a digital resolution of 1.03 Hz/point in F1 and 2.93 Hz/point in the F2 dimension.

Phase-sensitive NOESY experiments[19] were performed at 20 °C using the following parameters: mixing time 0.25 s, 4K complex data points in t_2, 512 complex data points in t_1, a relaxation delay of 3 s, a sweep width of 10 ppm in both dimensions, acquisitions per FID was 64; a shifted squared sine-bell apodization function was applied for both dimensions. The data were zero-filled in t_1 to give 2K x 2K complex data points. The residual water resonance was saturated during the relaxation delay.

Two-dimensional data sets for DQF-COSY[20] spectra were collected in the phase-sensitive mode with the time-proportional phase incrementation with and without phosphorus decoupling, with deuterium decoupling[21]. Typically 4096 data points were collected for each 512 t_1 values in DQF-COSY experiments. The 4096 x 512 data points were resolution enhanced by a shifted squared sine-bell window function in both the t_1 and t_2

directions, then Fourier transformed and phase adjusted, a relaxation delay of 3 s has been used. The data were collected with the nonspinning sample to avoid t_1 noise.

Acknowledgements

Thanks are due to the Swedish Board for Technical Development (NUTEK) to JC, the Swedish Natural Science Research Council (NFR contract # K-KU 12067-300 to AF & K-AA/Ku04626-321 to JC) and the Swedish Research Council for Engineereng Sciences (TFR) to JC for generous financial support. Authors also thank Wallenbergsstiftelsen, University of Uppsala and Forskningsrådsnämnden (FRN) for their financial supports for the purchase of the 500 and 600 MHz NMR spectrometers.

References

1. (a) Batey, R. T.; Battiste, J. L.; Williamson, J. R. *Methods Enzymol.* **1995**, *261*, 300. (b) Pardi, A. *Methods Enzymol.* **1995**, *261*, 350.
2. (a) Földesi, A.; Nilson, F. P. R.; Glemarec, C.; Gioeli, C.; Chattopadhyaya, J. *Tetrahedron* **1992**, *48*, 9033. (b) Földesi, A.; Nilson, F. P. R.; Glemarec, C.; Gioeli, C.; Chattopadhyaya, J. *J. Biochem. Biophys. Methods* **1993**, *26*, 1. (c) Yamakage, S.-i.; Maltseva, T. V.; Nilson, F. P. R.; Földesi, A.; Chattopadhyaya, J. *Nucleic Acids Res.* **1993**, *21*, 5005. (d) Agback, P.; Maltseva, T. V.; Yamakage, S.-i.; Nilson, F. P. R.; Földesi, A.; Chattopadhyaya, J. *Nucleic Acids Res.* **1994**, *22*, 1404. (e) Földesi, A.; Yamakage, S.-i.; Maltseva, T. V.; Nilson, F. P.; Agback, P.; Chattopadhyaya, J. *Tetrahedron* **1995**, *51*, 10065. (f) Földesi, A.; Yamakage, S.-I.; Nilsson, F. P. R.; Maltseva, T. V.; Chattopadhyaya, J. *Nucleic Acids Res.*, **1996**, *24*, 1187. (g) Glemarec, C.; Kufel, J.; Földesi, A.; Maltseva, T.; Sandström, A.; Kirsebom, L. A.; Chattopadhyaya, J. *Nucleic Acids Res.*, **1996**, *24*, 2022. (h) Kawashima, E.; Toyama, K.; Ohshima, K.; Kainosho, M.; Kyogoku, Y.; Ishido, Y. *Tetrahedron Lett.* **1995**, *36*, 6699. (i) Oogo, Y.; Ono, M. A.; Tate, S.-i.; Ono, A. S.; Kainosho, M. *Nucleic Acids Symp. Ser. 37* **1997**, 35. (j) Ono, A.; Makita, T.; Tate, S.-i.; Kawashima, E.; Ishido, Y.; Kainosho, M. *Magn. Reson. Chem.* **1996**, *34*, S40. (k) Tate, S.-i.; Kubo, Y.; Ono, A.; Kainosho, M. *J. Am. Chem. Soc.* **1995**, *117*, 7277. (l) Kojima, C; Kawashima, E.; Toyama, K.; Ohshima, K.; Ishido, Y; Kainosho, M.; Kyogoku, Y. *J. Biomol. NMR 11*, **1998**, 103. (m) Kawashima, E.; Aoyama, Y.; Sekine, T.; Miyahara, M.; Radwan, M. F.; Nakamura, E.; Kainosho, M.; Kyogoku, Y.; Ishido, Y. *J. Org. Chem.* **1995**, *60*, 6980. (n) Ono, A. (M.); Ono, A.; Kainosho, M. *Tetrahedron Lett.* **1997**, *38*, 395. (o) Ono, A. (M.); Shiina, T.; Ono, A.; Kainosho, M. *Tetrahedron Lett.* **1998**, *39*, 2793. (p) Huang, X.; Yu, P.; LeProust, E.; Gao, X. *Nucleic Acids Res.* **1997**, *25*, 4758. (q) Tolbert, T. J.; Williamson, J. R. *J. Am. Chem. Soc.* **1996**, *118*, 7929. (p) De Voss, J. J.; Hangeland, J. J.; Townsend, C. A. *J. Org. Chem.* **1994**, *59*, 2715. (r) Nikonowicz, E.P.; Michnicka, M.; DeJong, E. *J. Am. Chem. Soc.* **1998**, *120*, 3813. (s) Louis, J. M.; Martin, R. G.; Clore, G. M.; Gronenborn, A. M. *J. Chem. Biol.* **1998**, *273*, 2374. (t) Tolbert, T. J.; Williamson, J. R. *J. Am. Chem. Soc.* **1997**, *119*, 12100. (u) Quant, S.; Wechselberger, R. W.; Wolter, M. A.; Wörner, K.-H.; Schell, P.; Engels, J. W.; Griesinger, C.; Schwalbe, H. *Tetrahedron Lett.* **1994**, *35*, 6649. (v) Agrofoglio, L. A.; Jacquinet, J.-C.; Lancelot, G. *Tetrahedron Lett.* **1997**, *38*, 1411. (z) Serianni, A. S.; Bondo, P. B. *J. Biomol. Struct. & Dinam.* **1994**, *11*, 1133.
3. Dickerson, R. E.; Drew, H. Y. R. *J. Mol. Biol.* **1981**, *149*, 761.
4. Földesi, A.; Maltseva, T. V., Chattopadhyaya, J. *Tetrahedron* **1998**, *in press*
5. Kim, Y.; Prestegard, J. H. *J. Magn. Reson.* **1989**, *84*, 9.
6. Eberstadt, M.; Gemmecker, G.; Mierke, D. F.; Kessler, H. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1671.
7. Schwalbe, H.; Marino, J. P.; King, G. C.; Wechselberger, R.; Bermel, W.; Griesinger, C. *J. Biomol. NMR* **1994**, *4*, 631.
8. Schwalbe, H.; Marino, J. P.; Glaser, S. J.; Griesinger, C. *J. Am. Chem. Soc.* **1995**, *117*, 7251.
9. Zimmer, D. P.; Marino, J. P.; Griesinger, C. *Magn. Reson. Chem.* **1996**, *34*, S177.
10. Ono, A.; Makita, T.; Tate, S.-i.; Kawashima, E.; Ishido, Y.; Kainosho, M. *Magn. Reson. Chem.* **1996**, *34*, S40.
11. Varani, G.; Aboul-ela, F.; Allain, F. H.-T. *Progr. NMR Spectr.* **1996**, *29*, 51.
12. Schmieder, P., Ippel, J. H., van den Elst, H., van der Marel, G. A., van Boom, J. H., Altona, C., Kessler, H. *Nucleic Acids Res.*. **1992**, *20*, 4747.
13. Szyperski, T., Ono, A., Fernandez, C., Iwai, H., Tate, S.-I., Wüthrich, K., Kainosho, M. *J. Am. Chem. Soc,* **119**, 9901 (1997).
14. Sattler, M.; Fesik, S. W. *Structure.* **1996**, *4*, 1245.
15. James, T. L. *Current Op. Struct. Biol.* **1991**, *1*, 1042.
16. Allain, F. H.; Varani, G. *J. Mol. Biol.* **1995**, *250*, 333.
17. (a) Shaka, A. J.; Barker, P. B.; Freeman, R. *J. Magn. Reson.* **1985**, *64*, 547. (b) Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1983**, *52*, 335,
18. Maltseva, T. V.; Földesi, A.; Chattopadhyaya, J. *Perkin Trans II.* **1998**, *in press*
19. Bodenhausen, G.; Kogler, H.; Ernst, R. R. *J. Magn. Reson.* **1984**, *580*, 370.
20. Neuhaus, D.; Wagner, G.; Vasak, M.; Kagi, J. H. R.; Wuthrich, K. *Eur. J. Biochem.* **1985**, *151*, 257.
21. Maltseva, T. V.; Földesi, A.; Chattopadhyaya, J. *Magn. Reson. Chem.* **1998**, *36*, 227.
22. A. Abragam, *The pronciples of Nuclear Magnetism.* Clarendon Press, Oxford (1961)