## Branched Nucleic Acids

# NMR Spectroscopic Determination of the Solution Structure of a Branched Nucleic Acid from Residual Dipolar Couplings by Using Isotopically Labeled Nucleotides** 

Bernd N. M. van Buuren, Jürgen Schleucher, Valentin Wittmann, Christian Griesinger, Harald Schwalbe, and Sybren S. Wijmenga*

Branched nucleic acids play important biological roles, and well-known examples such as the Hammerhead ribozyme, the all-RNA enzyme, a three-way junction ( 3 H according to IUPAC nomenclature ${ }^{[12]}$ ), and the Holliday junctions ${ }^{[16]}$ or DNA four-way junctions (J6, Figure 1a; 4H according to IUPAC nomenclature ${ }^{[12]}$ ) have recently been reviewed. ${ }^{[2 a-c]}$ The Holliday junctions form the central intermediates in both homologous and site-specific recombination. In solution, in the presence of multivalent ions and thus under physiological salt conditions, their structure consists of two quasi-continuously stacked helices in an essentially canonical conformation (Figure 1c). ${ }^{[2 b, c]}$ Four recent crystal structures of $4 \mathrm{H}^{[2 d-g]}$ confirm these structural features, but no coherence is observed in the value of the interhelix angle $\psi$ (defined in Figure 1b); three structures show right-handed interhelix
[*] Prof. Dr. S. S. Wijmenga
Laboratory of Physical Chemistry-Biophysical Chemistry
University of Nijmegen
Toernooiveld 1, 6225ED Nijmegen (The Netherlands)
Fax: $(+31) 24-3652112$
E-mail: sybrenw@sci.kun.nl
Dr. B. N. M. van Buuren
Unilever Research \& Development
Olivier van Noortlaan 120, 3133AT Vlaardingen (The Netherlands)
Dr. J. Schleucher
Department of Medical Biochemistry and Biophysics Umeå University, S-90187 Umeå (Sweden)
Dr. V. Wittmann
Johann Wolfgang Goethe-Universität Frankfurt/M. Institut für Organische Chemie und Chemische Biologie Marie-Curie-Strasse 11, 60439 Frankfurt/M. (Germany)
Prof. Dr. C. Griesinger
Max-Planck-Institut für biophysikalische Chemie
Am Fassberg 11, 37077 Göttingen (Germany)
Prof. Dr. H. Schwalbe
Johann Wolfgang Goethe-Universität Frankfurt/M.
Zentrum für Biomolekulare Magnetische Resonanz Institut für Organische Chemie und Chemische Biologie Marie-Curie-Strasse 11, 60439 Frankfurt/M. (Germany)
[ $*^{*}$ ] This work was supported by grants from the Kempe Minne Foundation (BVB), the Swedish National Research Council (SW), Bioteknik Medel Umeå University (SW), the DFG (GR1211/2-4, En111/11-3) (C.G., H.S.), the Karl Winnacker foundation (H.S.), and the European LSF for Biomolecular NMR (ER-BCT950034). We thank Janny Hof for critical reading of the manuscript.
(2) Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author. Included are the fit results, the original $J_{C H}$ values as a function of magnetic field, and the derived mRDCs.


Figure 1. a) Sequence and arm designation for J6. The ${ }^{13} \mathrm{C}$-labeled thymines in two samples are indicated as boxed residues (solid/dashed outline). b) The interhelix angle $\psi$, strand numbering, and conformation designations following 4 H nomenclature. The parallel and antiparallel orientations of the AD and BC helices are defined by the corresponding drawings. Left-handed and right-handed indicate the orientation of the AD helix relative to the BC helix. The grey region was excluded for steric reasons. c) The chosen molecular axis frame and designation of the coaxially stacked helices.
angles, ${ }^{[2 d-f]}$ while the angle in the fourth structure is lefthanded. ${ }^{[2 g]}$ In the solution structure ${ }^{[3]}$ of the J6 model for a Holliday junction, based on NMR spectroscopy using NOE and $J$ coupling constants, the quasi-continuously stacked helices AD and $\mathrm{BC}^{[33]}$ have a left-handed interhelix angle (Figure 1b). The determination of the relative helix orientations is the central question concerning the structure of branched nucleic acids.

Herein we describe how a relatively small set of longrange NMR restraints can be used to determine the global structure of branched nucleic acids. This is shown with the example of the Holliday junction, and the previous assignment of a left-handed interhelix angle is confirmed. The approach relies on the measurement of a set of residual dipolar couplings (RDCs) in a DNA sample with thymidine residues that are ${ }^{13} \mathrm{C}$ labeled in the deoxyribose moiety of the nucleotide ${ }^{[4]}$ (Figure 1a). This chemical labeling strategy allows rapid determination of the overall conformation of branched nucleic acids.

In NMR spectra of liquid samples, RDCs contain information about long-range orientation and complement the local NMR restraints, NOEs, and $J$ coupling constants. The RDCs are observed for molecules for which isotropic rotational tumbling is partially restricted. ${ }^{[5-7]}$ This restriction can be induced by non-isotropic media such as liquid crystals or phages. ${ }^{[5 a, e-h]}$ However, partial alignment can also be conferred by the intrinsic properties of a molecule. For
example, the anisotropy of the magnetic susceptibility, ${ }^{[8,9]}$ a pronounced feature of double-stranded oligonucleotides, induces a partial alignment of a molecule in strong homogeneous magnetic fields and thereby gives rise to magnetic-field-induced RDCs (mRDCs). ${ }^{[5 a-d, j-\mathrm{n}, 9]}$

The anisotropy of the molecular magnetic susceptibility tensor $\left(\chi_{\mathrm{mol}}\right)$ is caused by the magnetic susceptibility of the individual nucleobases $\left(\chi_{\mathrm{b}}\right)$. The latter are dominated by base ring currents with the long axis $(z)$ of $\chi_{\mathrm{b}}$ perpendicular to the base plane. Their principal components ( $\chi_{\mathrm{b}, x x}, \chi_{\mathrm{b}, y,}, \chi_{\mathrm{b}, z z}$ ) can be written as $(0,0,-13) \times 10^{-34} \mathrm{~m}^{3}$ in SI units. The average value of $\chi_{\mathrm{b}, z z}$ per base can be calculated on the basis of the CottonMouton effect ${ }^{[8]}$ or from quantum mechanics ${ }^{[10]}$ with good accuracy (Tables 1 and 2 ). $\chi_{\text {mol }}$ can be calculated as the tensor

Table 1: Estimate of the base $\chi$ values.

|  | $\mathrm{CG}\left[-10^{-34} \mathrm{~m}^{3}\right]$ | $\mathrm{AT}\left[-10^{-34} \mathrm{~m}^{3}\right]$ |
| :--- | :--- | :--- |
| $\mathrm{ElOPt}^{[\mathrm{ab]}}$ | -25.4 | -19.9 |
| $\mathrm{JB}^{[b]}$ | -20.4 | -26.2 |
| $\mathrm{HM}^{[b]}$ | -23.8 | -30.4 |
| $\mathrm{GP}^{[b]}$ | -26 | -36 |
| Average all |  | $-26 \pm 5$ |
| Average base | $-13 \pm 2.5$ |  |

[a] Value of $\chi$ from reference [8]; [b] JB: $\chi$ from ring currents with ring currents obtained using Johnson-Bovey equations from reference [10a]; HM: $\chi$ from ring currents with ring currents obtained using HaigMallion equations from reference [10a]; GP: $\chi$ from ring currents with ring currents obtained from reference [10b]. For a CG base pair the ring currents of $C$ and $G$ were added; an analogous method was used for an AT base pair. The values for the ring currents are relative to those for benzene and proportional to the magnetic susceptibility tensor. The $\chi$ values of the CG and AT base pairs were obtained by multiplying the ring current with $\chi$ for benzene $\left(-12.7 \times 10^{-34} \mathrm{~m}^{3[9]}\right)$. The GP ring currents were multiplied with a factor 1.7 to take into account the fact that GP used lower ring currents values, because part of the base susceptibility is attributed to bond magnetic susceptibility. The $\chi$ tensor is oriented perpendicular to the base. Consequently, its principal components can be written as $\left(0,0, \chi_{z z}\right)$, where $\chi_{z z}$ indicates the $z$ principal component. The average of all different base pair estimates leads to an average value $\chi_{z z}$ per base of $13 \pm 2.5 \times 10^{-34} \mathrm{~m}^{3}$.

Table 2: Ring currents calculated for the four nucleic acid bases $G, A, T$, and C. ${ }^{[a]}$

|  | GP | JB | HM |
| :--- | :--- | :--- | :--- |
| G | $0.94(0.62)$ | $1.30(0.73)$ | $1.51(0.74)$ |
| A | $1.56(1.00)$ | $1.78(1.00)$ | $2.04(1.00)$ |
| T | $0.11(0.07)$ | $0.28(0.16)$ | $0.35(0.17)$ |
| C | $0.28(0.18)$ | $0.31(0.17)$ | $0.37(0.18)$ |

[a] GP, JB, and HM are as defined in Table 1; the values in parentheses indicate the ring current values relative to that for adenine.
sum of the individual $\chi_{\mathrm{b}}$ values as long as the structure of a nucleic acid molecule is known [Eq. (1)].
$\chi_{\mathrm{mol}}=\sum_{i=1}^{N}\left(\overrightarrow{\boldsymbol{v}}_{\mathrm{b}, i}^{T} \otimes \vec{v}_{\mathrm{b}, i}\right) \chi_{\mathrm{b}, z z, i}$

Here, $\nu_{\mathrm{b}}$ is a unit row vector perpendicular to the base plane $i, N$ the number of bases in the molecule, and the symbol $\otimes$ represents the direct matrix product of the column and row vector.

When non-isotropic media ${ }^{[5 a, e-h, 6,7]}$ are used to align molecules, the size (axial and rhombic component) and orientation (three Euler angles) of the alignment tensor in a molecular frame have to be derived from the experimental RDCs; this requires a minimum of five RDCs. The relative domain orientation in an $n$ domain molecule with known domain structure is defined by three Euler angles for each pair of domains, so that in addition a minimum of three parameters (RDCs) are needed to derive the relative orientation of each domain. ${ }^{[5 a, d, 7]}$ In the procedure proposed by Dosset et al., ${ }^{[77]}$ for each domain an alignment tensor is first derived and the relative domain orientations are subsequently obtained by reorienting each domain such that all alignment tensors have the same common orientation. This procedure can be simplified in the case of magnetic-fieldinduced alignment, because the size and orientation of the alignment tensor can be calculated from each trial structure. Such reduction in complexity is also obtained if the size and shape of the alignment tensor can be predicted from the shape and/or charge distribution of the molecule. ${ }^{[7 \mathrm{~d}, \mathrm{e}]}$

Residual dipolar couplings ${ }^{1} D_{\text {CH }}$ have been measured in the sugar moiety of ten thymine residues of the J6 model. The residual dipolar coupling $D_{\mathrm{CH}, i}$ of a dipolar $\mathrm{C}-\mathrm{H}$ vector $i$ can be calculated in a chosen molecular reference frame by direct matrix multiplications of $\chi_{\text {mol }}$ and $\vec{v}_{\mathrm{CH}, i}$, where the latter is the unit row vector pointing along the $\mathrm{C}-\mathrm{H}$ vector $i$ expressed in the chosen reference frame [Eq. (2)].
$D_{\mathrm{CH}, i}=-S\left[\frac{h}{2 \pi^{2}} \gamma_{\mathrm{C}} \gamma_{\mathrm{H}} \mathrm{r}_{\mathrm{CH}}^{-3} \frac{1}{4 \pi} \frac{1}{15 k T} B_{o}^{2} \frac{3}{4}\right]\left[\left(\vec{v}_{\mathrm{CH}, i} \vec{\chi}_{\mathrm{mol}} \vec{v}_{\mathrm{CH}, i}^{T}\right)\right.$ $-\left(\frac{1}{3}\right.$ Trace $\left.\left.\overrightarrow{\vec{\chi}}_{\text {mol }}\right)\right]$

Here, $h$ is Planck's constant, $\gamma_{\mathrm{C}}$ and $\gamma_{\mathrm{H}}$ are the gyromagnetic ratios of the ${ }^{13} \mathrm{C}$ and ${ }^{1} \mathrm{H}$ nuclear spins, $r_{\mathrm{CH}}$ is the distance between the directly bonded C and H nuclei, $k T$ is the thermal energy, and $B_{\mathrm{o}}$ is the magnetic field strength in Tesla. The order parameter $S^{[5 a, d, 11]}$ takes into account the reduction in the size of $D_{\mathrm{CH}, i}$ by intramolecular dynamics and has been assumed to be 1 here.

The relative orientation of the two semi-continuously stacked helices of 4 H (Figure 1c) is defined by three parameters. In any axes frame, the AD helix orientation is specified by its three Euler angles, the polar angles, $\theta_{\mathrm{AD}}$ and $\phi_{\mathrm{AD}}$, and the rotation around the helix axis $\omega_{\mathrm{AD}}$ ( BC helix orientation: $\theta_{\mathrm{BC}}, \phi_{\mathrm{BC}}$, and $\omega_{\mathrm{BC}}$ ). We defined a right-handed reference frame in which the AD helix lies along the $z$ axis ( $\theta_{\mathrm{AD}}=0, \phi_{\mathrm{AD}}=0$ ) and the $x$ axis is perpendicular to the plane spanned by the AD and BC helix axes. The BC helix then rotates by definition in the $y z$ plane ( $\phi_{\mathrm{BC}}=90$ ) and $\theta_{\mathrm{BC}}$ equals the interhelix angle $\psi\left(\theta_{\mathrm{AD}}=0\right)$. In this molecular axis frame, the conformation of 4 H is defined by the interhelix angle $\psi$, and the angles $\omega_{\mathrm{AD}}$ and $\omega_{\mathrm{BC}}$ indicate the rotation about their own axes. From a given set of RDCs (minimum of three required), two sets of solutions for the conformation of 4 H
can be derived. Thus, for the first set of two solutions (set 1), the 4 H conformation is defined by the angles $\omega_{\mathrm{AD} 1}$ and $\omega_{\mathrm{BC} 1}$ with the interhelix angle equal to $\psi_{1}$ and $\psi_{1}+180^{\circ}$. The second set of two solutions (set 2 ) is defined by the angles $\omega_{\mathrm{AD} 1}$ and $\omega_{\mathrm{BC} 1}+180^{\circ}$ with the interhelix angle again equal to $\psi_{1}$ and $\psi_{1}+180^{\circ}$. Set 2 is sterically disallowed, because strands 1 and 3 (Figure 1 b ) cross from the AD to the BC helix. Furthermore, $\omega_{\mathrm{AD} 1}+180^{\circ}$ leads only to equivalent solutions involving rotation of the whole molecule. The two solutions of set 1 are reduced to one unique solution when $-60^{\circ} \leq \psi<-180^{\circ}$, because the crossing strands intermingle for $60^{\circ} \leq \psi<180^{\circ}{ }^{[12]}$ and thus exclude this region (Figure 1 b ). As shown below, we found a value close to $-90^{\circ}$ and thus one unique solution was obtained.

For each conformation $\chi_{\text {mol }}$ was calculated and subsequently $D_{\mathrm{CH}}^{\text {calcd }}$ determined. The optimal conformation was found by comparing $D_{\mathrm{CH}}^{\text {calcd }}$ with $D_{\mathrm{CH}}^{\text {exp }}$ as the minimum root mean square difference (rmsd). In total 28 mRDCs were derived from least-squares fits of ${ }^{1} J_{\mathrm{CH}}$ coupling constants and measured ${ }^{[13]}$ at four different magnetic fields ( $400-800 \mathrm{MHz}$; see Supporting Information). They are given in Table 3 as

Table 3: Experimental dipolar couplings $D$ of the model Holliday junction J6, as determined from measurements at four magnetic fields. ${ }^{[\text {a] }}$

| Arm | Residue | $D\left(\mathrm{Cl}^{\prime}, \mathrm{H1}^{\prime}\right)$ <br> $[\mathrm{Hz}]$ | $D\left(\mathrm{C}^{\prime}, \mathrm{H}^{\prime}\right)$ <br> $[\mathrm{Hz}]$ | $D\left(\mathrm{C}^{\prime}, \mathrm{H}^{\prime}\right)$ <br> $[\mathrm{Hz}]$ |
| :--- | :--- | :--- | :---: | :---: |
| A | T5 | $-1.7^{[b]}$ | 1.2 | 0.1 |
| A | T35 | 0.6 | -0.2 | 0.6 |
| A | T40 | 0.6 | 0.0 | 1.4 |
| D | T29 | -1.2 | 0.7 | 2.1 |
| D | T32 | -3.0 | -0.3 | -0.6 |
|  |  |  |  |  |
| B | T9 | 0.5 | -0.7 | 3.1 |
| B | T12 | - | 0.2 | 1.8 |
| B | T13 | 0.5 | 1.1 | 2.3 |
|  |  |  |  |  |
| C | T20 | 1.6 | - | 1.8 |
| C | T21 | 1.0 | 1.6 | 2.3 |

[a] The dipolar couplings were determined from a least-squares fit of the magnetic field dependence of the ' ${ }_{\mathrm{CH}}$ coupling constants of $\mathrm{Hl}^{\prime}-\mathrm{Cl}^{\prime}, \mathrm{H}^{\prime}$ '$\mathrm{C} 3^{\prime}$, and $\mathrm{H} 4^{\prime}-\mathrm{C} 4^{\prime}$ at $400,500,600$, and 800 MHz . The measured ${ }^{1} J_{\mathrm{CH}}$ value can be written as ${ }^{1} J_{C H}=a B_{o}^{2}+{ }^{1} J_{C H}{ }^{\circ}$, where ${ }^{1} J_{C H}{ }^{\circ}$ represents the ${ }^{1} J_{C H}$ value at a magnetic field of zero and the slope $a$ contains the $D_{C H}$ information ( $a=D_{C H} / B_{o}^{2}$ ). For the least-squares fit the above equation was used. The $D_{C H}$ at a given field is then equal to ( $\mathrm{a} B_{\mathrm{o}}^{2}$ ). The $D_{\mathrm{CH}}$ values were calculated from the slope as the difference $\Delta D_{C H}$ between values at 800 and 400 MHz . Thus, they correspond to $D_{\mathrm{CH}}$ measured at an effective magnetic field of $(64-16)^{0.5} \times 100 \mathrm{MHz}=693 \mathrm{MHz}$. Two samples were used with ${ }^{13} \mathrm{C}$ labeling in the ribose; sample 1 was labeled at $\mathrm{T} 5, \mathrm{~T} 9, \mathrm{~T} 12$, T20, T32, and T35; and sample 2 at T13, T12, T29, and T40. [b] The uncertainty in the $D_{C H}$ values is estimated to be $\pm 0.5 \mathrm{~Hz}$.
$\Delta^{1} J_{\mathrm{CH}}$ between $400-800 \mathrm{MHz}$ and denoted as $D_{\mathrm{CH}}$. A variety of optimizations were performed with model helices or helices derived from two NOE-based experimental structures (Figure 2 and Supporting Information). Most striking is that $\psi$ always has a well-defined minimum $\left(-88^{\circ}<\psi_{\text {opt }}<-98^{\circ}\right.$, depending on the optimization, Figure 2 a). This is more negative than the value of approximately $-70^{\circ}$ in the NOE-


Figure 2. a) Plot of the rmsd of $D_{\mathrm{CH}}^{\text {exp }}$ and $D_{\mathrm{CH}}^{\text {calcd }}$ versus $\psi$. The AD and BC helices were modeled as B-DNA helices capped with a known stable CTTG loop ${ }^{[44]}$ (red, r) or taken from the NOE-based structure set (blue, black, z: two structures close to average). Model helices ( $r$ ): fit with 3 (r1) or $3+7$ parameters (r2) or as r 2 with base-specific $\chi_{\mathrm{b}, \mathrm{zz}}(\mathrm{r} 3$; see Supporting Information). Experimental helices: fit with 3 (bl, z1) or $3+8$ parameters (b2). b) Monte Carlo estimate of the spread in $\psi$ of fit $r 1$ (see the text). The fitting was repeated for 100 samples of the 28 mRDCs with random variation in their values according to normal distribution with $\sigma=0.5 \mathrm{~Hz}$. c) Stereo views of the mRDC-optimized 4 H (corresponding to fits r1-r3 and b2) and the NOE-based 4 H (yellow, optimized is b2) with chosen axes frame (see the text). $\chi_{\text {mol }}$ essentially coincides with this chosen axes frame (see Supporting Information).
based structure. The Monte Carlo (MC) spread in $\psi$ due to the $0.5-\mathrm{Hz}$ experimental error on $D_{\mathrm{CH}}$ is small (rmsd 3-6 ${ }^{\circ}$, Figure 2b and Supporting Information). The angles $\omega_{\text {AD }}$ and $\omega_{\mathrm{BC}}$ also have well-defined minima, but their MC errors are somewhat larger (rmsd 6-15 ${ }^{\circ}$; see Supporting Information). The use of base-specific values instead of one average value for $\chi_{\mathrm{b}, z z}$ in the calculation of $\chi_{\text {mol }}$ hardly affects the fit results (Figure 2a). Similarly, increasing or decreasing the average $\chi_{\mathrm{b}, z z}$ value by one standard deviation does not change the fit results (see Supporting Information). To verify these results, cross validation was carried out on the three-parameter fit $\mathrm{M}(3)$ _r1 (Figure 2) by randomly removing six of the 28 mRDCs in 100 different samples. For each sample the fitting was carried out and the MC spread determined to yield average values and deviations (rmsd) for $\psi_{\mathrm{opt}}, \omega_{\mathrm{AD}, \mathrm{opt}}$, and $\omega_{\mathrm{BC}, \text { opt }}$ of $-93(3)^{\circ},-159^{\circ}\left(6^{\circ}\right)$ and $-182^{\circ}\left(6^{\circ}\right)$, respectively. The similarity of these values with those found for the other fit results (see Supporting Information) further confirms the conclusions.

Although the fits with three adjustable parameters gave well-defined minima, the correlation between $D_{\mathrm{CH}}^{\text {exp }}$ and $D_{\mathrm{CH}}^{\text {calcd }}$ showed a systematic offset for the $\mathrm{H}^{\prime}-\mathrm{C3}^{\prime}$ and $\mathrm{H}^{\prime}{ }^{\prime}-\mathrm{C4}^{\prime}$
mRDCs (see Supporting Information). Variation in sugar puckering and orientation may affect the orientation of these vectors (but not $\mathrm{H1}^{\prime}-\mathrm{Cl}^{\prime}$ ). Note that in DNA the energy well for the S-puckered deoxyribose is relatively wide, and sugars are often not found to be $100 \%$ S-puckered even in regular helices. Furthermore, the ribose rings are oriented approximately parallel to the helix axis, so that the sugar pucker motion, repuckering, and reorientation will mainly affect the $\omega$ angles of these vectors. Therefore, we operationally included small additional $\omega$ rotations, where the $\mathrm{H} 3^{\prime}-\mathrm{C} 3^{\prime}$ and $\mathrm{H} 4^{\prime}-\mathrm{C} 4^{\prime} \mathrm{mRDCs}$ were treated as separate groups in the A , B, C, and D helices, and found indeed improved fit results (see Supporting Information). However, the number of adjustable parameters $(\leq 11)$ remains much smaller than the number of experimental mRDCs (28).

Model helices gave similar results and the best fits (Figure 2a, red curve), and the mRDC-optimized 4H structures nearly overlap (Figure 2c, red structure). Compared with the NOE-based structure, the main change, apart from the small change in $\psi$, is that the BC helix has rotated by about $40^{\circ}$ around its own axis (Figure 2c, yellow versus red structure), leading to a net opening of the cavity formed by the facing major grooves. In the mRDC-optimized NOEbased structure (Figure 2c, blue structure), the BC helix has rotated approximately $40^{\circ}$ further. Optimization of the other experimental structure leads to a smaller cavity opening, and the resulting structure is close to the yellow structure shown in Figure 2c. Generally, the fits using experimental helices are not as good (Figure 2a, red versus blue and black structures), thus leading to the larger uncertainty in the value of $\omega_{\mathrm{BC}}$. This can be attributed to the difficulty of using NOEs to define global features even within the regular helix domains.

We have assumed a single rigid structure for the calculation of $\chi_{\mathrm{mol}}$ and $D_{\mathrm{CH}}$ [Eqs. (1) and (2)]. It is of interest to consider potential hinge motion, that is, motion of the BC helix with respect to the AD helix. Internal motion is usually taken into account by an order parameter $S$ [Eq. (2)], which simply scales the mRDCs, and the optimized angles should be viewed as effective angles. ${ }^{[11]}$ Assuming that the potential hinge motion can be accounted for in this way a rough estimate of its amplitude can be obtained. The errors in $\chi_{\text {mol }}$ and $D_{\mathrm{CH}}$ are about $10 \%$ and $15 \%$, respectively, which correspond to an overall error of around $18 \%$ in the tensor and allows for an overall $S$ value of about 0.82 . If we further assume that the hinge motion can be described by axial wobbling in a cone according to $\frac{1}{2} \cos \varepsilon(1+\cos \varepsilon)=S$, the amplitude of this motion would imply an opening half-angle of about $28^{\circ}$. Thus, this would allow replacement of the model with a fixed orientation of the two helices with an interhelical angle of about $-90^{\circ}$ by a model in which the structure has an effective angle of $-90^{\circ}$ with a spread of around $\pm 28^{\circ}$. It should be noted, however, that the effect of hinge motion is more complex than a simple scaling of the dipolar couplings, because the hinge motion differently affects the axial and rhombic components of $\chi_{\text {mol }}$ as well as its orientation with respect to the AD and BC helices (for certain $\mathrm{C}-\mathrm{H}$ vectors hinge motion may even lead to an increase in $D_{\mathrm{CH}}$ ). An accurate estimation of the effect of hinge motion would require simulation of these motions.

In summary, we have shown with the example of the Holliday junction that the global orientation of branched nucleic acids can be derived from a small set of mRDCs by taking advantage of the reduction in required parameters, thus extending approaches based on phage-induced RDCs. ${ }^{[15]}$ The same type of analysis can be extended to other branched nucleic acids in a relatively straightforward fashion. For instance, under physiological salt conditions, RNA and DNA 3H show coaxial stacking of two of the three arms. The global conformation is then defined by three angles that describe the relative helix orientations and would require the same minimum number of mRDCs as for the 4 H discussed here. Similarly, for more strongly branched nucleic acids, such as 5 H and 6 H sometimes found in RNAs, the global conformation is defined by the relative orientation of three helices when coaxial stacking is present and would require a minimum of six mRDCs (six Euler angles define the relative helix orientations). The mRDC size, although sufficiently large here, can easily be increased by using larger magnetic fields or larger nucleic acids (e.g. extension of each of the 4 H helices by four base pairs leads to $\Delta D_{\mathrm{CH}}^{\max }$ of 16 Hz for 900 vs. 400 MHz ). Also, with present cryoprobe technology, magnetic-field-induced dipolar couplings can be determined at natural abundance.

## Experimental Section

$\left[{ }^{13} \mathrm{C}_{5}\right]$-Thymidine phosphoramidite (6): Two J6 samples (Figure 1a) were synthesized according to Scheme 1 with thymidine residues ${ }^{13} \mathrm{C}$ labeled in the deoxyribose moiety of the nucleotide ${ }^{[4]}$ (see Figure 1a). $\left[{ }^{13} \mathrm{C}_{6}\right]$ Glucose was converted into 1,2 -di- $O$-acetyl-3,5-di- $O$-benzoyl ribofuranose (1) and glycosylated to give nucleoside $\mathbf{2}$ according to a published procedure. ${ }^{[4]} 2^{\prime}$-Deoxygenation was carried out following a procedure described by Robins et al. for the synthesis of $2^{\prime}$ deoxyuridine. ${ }^{[16]}$ Briefly, 2 was deprotected and treated with the Markiewicz reagent to give $3^{\prime}, 5^{\prime}$-protected nucleoside 3, which was further converted into its $2^{\prime}-O$-phenoxythiocarbonyl derivative. Reductive deoxygenation with tri- $n$-butyltin hydride and the freeradical initiator AIBN in warm toluene provided the thymidine derivative 4. Fluoride-induced removal of the silyl protecting group and treatment with dimethoxytrityl chloride gave 5 , which was phosphitylated to yield phosphoramidite 6 . The product was ready for use in automated DNA synthesis.

Received: April 10, 2003
Revised: September 30, 2003 [Z51632]

Keywords: Holliday junctions • isotopic labeling • NMR spectroscopy • nucleic acids • structure elucidation





Scheme 1. Synthesis of 6 . The asterisks indicate the ${ }^{13} \mathrm{C}$-labeled positions. AIBN $=$ azobisisobutyronitrile, $\mathrm{Bz}=$ benzyl, DIEA $=N, N$-diisopropylethylamine, $\mathrm{DMAP}=4$-dimethylaminopyridine, $\mathrm{DMTr}=4,4^{\prime}$-dimethoxytriphenylmethyl, OTf = trifluoromethanesulfonate, TBAF = tetrabutylammonium fluoride, TIPDS = tetraisopropyldisilyl, TMS = trimethylsilyl.
[1] a) D. M. J. Lilley, R. M. Clegg, S. Diekmann, N. C. Seeman, E. von Kitzing, P. J. Hagerman, Nucleic Acids Res. 1995, 23, $3363-$ 3364; b) R. Holliday, Genet. Res. 1964, 5, 282-304.
[2] a) A. Kuzminov, Microbiol. Mol. Biol. Rev. 1999, 63, 751-813; b) S. C. West, Annu. Rev. Genet. 1997, 31, 213-244; c) D. M. J. Lilley, Q. Rev. Biophys. 2000, 33, 109-159; d) J. Nowakowski, P. J. Shim, G. S. Prasas, C. D. Stout, G. F. Joyce, Nat. Struct. Biol. 1999, 6, 151-156; e) B. F. Eichman, J. M. Vargason, B. H. M. Mooers, P. S. Ho, Proc. Natl. Acad. Sci. USA 2000, 97, 3971 3976; f) M. Ortiz-Lombardía, A. González, R. Eritja, J. Aymamí, F. Azorín, M. Coll, Nat. Struct. Biol. 2000, 6, 913-917; g) J. Nowakowski, P. J. Shim, C. D. Stout, G. F. Joyce, J. Mol. Biol. 2000, 300, 93-102.
[3] a) B. N. M. van Buuren, J. Schleucher, S. S. Wijmenga, J. Biomol. Struct. Dyn. 2000, 11, 237-243; b) B. N. M. van Buuren, T. Herman, S. S. Wijmenga, E. Westhof, Nucleic Acids Res. 2002, 30, 507-514.
[4] S. Quant, R. W. Wechselberger, M. A. Wolter, K. Wörner, P. Schell, J. W. Engels, C. Griesinger, H. Schwalbe, Tetrahedron Lett. 1994, 35, 6649-6652.
[5] a) J. H. Prestegard, H. M. Al-Hashimi, J. R. Tolman, Q. Rev. Biophys. 2000, 33, $371-424$; b) A. Bax, N. Tjandra, Nat. Struct. Biol. 1997, 4, 254-256; c) N. Tjandra, J. G. Omichinski, A. M. Gronenborn, G. M. Clore, A. Bax, Nat. Struct. Biol. 1997, 4, 732 738; d) J. H. Prestegard, Nat. Struct. Biol. 1998, 5, 517-522; e) A. Bax, N. Tjandra, J. Biomol. NMR 1997, 10, 289-292; f) M. R. Hansen, L. Mueller, A. Pardi, Nat. Struct. Biol. 1998, 5, 1065-1074; g) G. M. Clore, R. M. Starich, A. M. Gronenborn, J. Am. Chem. Soc. 1998, 120, 10571-10572; h) G. M. Clore, A. M.; Gronenborn, N. Tjandra, J. Magn. Reson. 1998, 131, 159-162; i) G. Cornilescu, L. J. Marquardt, M. Ottiger, A. Bax, J. Am. Chem. Soc. 1998, 120, 6836-6837; j) H. C. Kung, K. Y. Wang, I. Goljer, P. H. Bolton, J. Magn. Reson. Ser. B 1995, 109, 323-325; k) H. M. Al-Hashimi, A. Majundar, A. Gorin, A. Kettani, E.

Skripkin, D. J. Patel, J. Am. Chem. Soc. 2001, 123, 633-640; 1) H. M. Al-Hashimi, A. Gorin, A. Majumdar, D. J. Patel, J. Am. Chem. Soc. 2001, 123, 3179-3180; m) H. M. Al-Hashimi, Y. Gosser, A. Gorin, W. Hu, A. Majumdar, D. J. Patel, J. Mol. Biol. 2002, 315, 95-102; n) H. M. Al-Hashimi, J. R. Tolman, A. Majumdar, A. Gorin, D. J. Patel, J. Am. Chem. Soc. 2001, 123, $5806-5807$; p) N. Sibille, A. Pardi, J. P. Simorre, M. Blackledge, J. Am. Chem. Soc. 2001, 123, 12135-12146.
[6] a) N. Tjandra, S. I. Tate, A. Ono, M. Kainosho, A. Bax, J. Am. Chem. Soc. 2000, 122, 6190-6200; b) A. Vermeulen, H. Zhou, A. Pardi, J. Am. Chem. Soc. 2000, 122, 96389647.
[7] a) M. W. F. Fischer, J. A. Losonczi, J. L. Weaver, J. H. Prestegard, Biochemistry 1999, 38, 9013-9022; b) N. R. Skrynnikov, N. K. Goto, D. Yang, W. Choy, J. R. Tolman, G. A. Mueller, L. E. Kay, J. Mol. Biol. 2000, 295, 1265-1273; c) E. T. Mollova, R. Hansen, A. Pardi, J. Am. Chem. Soc. 2000, 122, 11561-11562; d) C. A. Bewley, G. M. Clore, J. Am. Chem. Soc. 2000, 122, 6009-6016; e) M. Zweckstetter, A. Bax, J. Am. Chem. Soc. 2000, 122, 3791-3792; f) P. Dosset, J. C. Hus, D. Marion, M. Blackledge, J. Biomol. NMR 2001, 20, 223-231.
[8] G. Weill in Molecular Electro-Optics (Ed.: S. Krause), Plenum, New York, 1981, pp. 473-483.
[9] A. A. Bothner-By in Encyclopedia of Nuclear Magnetic Resonance (Eds.: D. M. Grant, R. K. Harris), Wiley, Chicester, 1995, pp. 2932-2938.
[10] a) D. A. Case, J. Biomol. NMR 1995, 6, 341-346; b) C. Giessner Prettre, B. Pullman, Q. Rev. Biophys. 1987, 20, 113-172.
[11] J. Meiler, J. J. Prompers, W. Peti, C. Griesinger, R. Brüschweiler, J. Am. Chem. Soc. 2001, 123, 6098-6107.
[12] E. von Kitzing, D. M. Lilley, S. Diekmann, Nucleic Acids Res. 1990, 18, 2671-2683.
[13] N. Tjandra, A. Bax, J. Magn. Reson. 1997, 124, 512-515.
[14] J. H. Ippel, H. Van den Elst, G. A. Van der Marel, J. H. Van Boom, C. Altona, Biopolymers 1998, 46, 375-393.
[15] K. Bondensgaard, E. T. Mollova, A. Pardi, Biochemistry 2002, 41, 11532-11542.
[16] M. J. Robins, J. S. Wilson, F. Hansske, J. Am. Chem. Soc. 1983, 105, 4059-4065.

