Determination of Aliphatic Side-Chain Conformation Using Cross-Correlated Relaxation: Application to an Extraordinarily Stable 2′-aminoethoxy-Modified Oligonucleotide Triplex

Teresa Carlomagno,[†,||] Marcel J. J. Blommers,[*,‡] Jens Meiler,[†] Bernard Cuenoud,[§] and Christian Griesinger[*,†,⊥]

Contribution from the Institut für Organische Chemie, Universität Frankfurt, Marie-Curie Strasse 11, D-60439 Frankfurt, Germany, and Novartis Pharma AG, Core Technologies, P.O. Box, CH-4002, Basel, Switzerland, and Novartis Horsham Research Center, Wimblehurst Road, West Sussex, RH12 4AB, England

Received July 14, 2000. Revised Manuscript Received April 9, 2001

Abstract: The structural basis for the extraordinary stability of a triple-stranded oligonucleotide in which the third strand contains 2′-aminoethoxy-substituted riboses is investigated by NMR spectroscopy. The enhanced stability of the modified triplex in comparison to the unmodified DNA triplex of the same sequence can be attributed to strong interactions of the aminoethoxy groups of the third strand with the phosphate groups of the purine strand. In molecular dynamics calculations the aminoethoxy side chain was found to be rather flexible, allowing for the presence of hydrogen bonds between the aminoethoxy group of the third strand and two different phosphates of the backbone of the second strand. To investigate the conformational preference of the aminoethoxy side chain a new NMR method has been developed which relies on CH−CH dipolar−dipolar cross-correlated relaxation rates. The results indicate that the aminoethoxy side chains adopt mainly a *gauche*$^+$ conformation, for which only one of the two hydrogen bonds inferred by NMR and molecular dynamics simulations is possible. This demonstrates a highly specific interaction between the amino group of the third strand and one of the phosphate groups of the purine strand.

Introduction

The most general way to interfere with cellular processes is the inhibition of enzyme activity by small molecules. Alternatives are the binding of oligonucleotides to mRNA (antisense approach) or to DNA (antigene approach). In the antigene technology[1,2] a particular promoter region of transcription factors can be targeted, thereby transactivating or repressing the transcription of genes. Moreover, chemically modified oligonucleotides linked to psoralen can precisely place point mutations at endogeneous chromosomal loci, and they represent a new tool for gene knockout and sequence manipulation.[3] In case the DNA contains a polypurine or polypyrimidine sequence a triple helix can be formed with an antigene oligonucleotide. Recently, we have investigated various modified oligonucleotides with the aim of designing a particularly stable triple helix motif. 2′-Aminoethoxy substituents at the ribose rings were found to form extraordinarily stable triplexes,[4] showing an enhanced binding to duplex DNA of about 3.5 °C per modified nucleotide in comparison to an unmodified strand and over 1000-fold faster rate of association. It has been shown that 2′-aminoethoxy-substituted oligonucleotides bind strongly to DNA and are able to displace a DNA binding protein.[5] Furthermore, inhibition of protein transcription was observed in cells in a dose-dependent way.[5] This constitutes the first successful implementation of the antigene approach in vivo. The stability of these triplexes can be explained by ionic interactions of the positively charged amino group in the third strand and the negatively charged phosphate(s) in the second strand. The presence of specific interactions in a self-complementary triplex with a 2′-aminoethoxy-modified third strand, including hydrogen bonds between the two strands, has been investigated by standard homonuclear NMR and by molecular dynamics (MD) simulations.[4] In addition to the triple base-pair formation, the aminoethoxy side chain adopts a conformation, in which the nitrogen and oxygen atoms are *gauche*$^+$, which enables the amino group to form a hydrogen bond with the pro-*R* oxygen of the phosphate of residue $i - 1$ of the DNA. An alternative conformation of the 2′-aminoethoxy side chain, in which the nitrogen and oxygen are *trans*, allows the formation of a hydrogen bond involving the amino group and the pro-*R* oxygen of residue *i* of the DNA duplex. Evidence of the occurrence of this *trans* conformation of the aminoethoxy side chain could not be obtained experimentally from the homonuclear NMR

* Author to whom correspondence should be addressed.
† Universität Frankfurt.
‡ Novartis Pharma AG, Core Technologies.
§ Novartis Horsham Research Center.
|| Present address: Department of Molecular Biology & The Skaggs Institute of Chemical Biology, The Scripps Research Institute, MB 33, 10550 North Torrey Pines Road, La Jolla, CA 92037.
⊥ Also: Max Planck Institut für Biophysikalische Chemie, Am Fassberg, 37077 Göttingen, Germany.

(1) Neidel, S. *Anti-Cancer Drug Des.* **1997**, *12*, 433−442.
(2) Giovannangeli, C.; Hélène, C. *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 413−421. Majumdar, A.; Khorlin, A.; Dyatkina, N.; Lin, F.-L. M.; Powell, J.; Liu, J.; Fei, Z.; Khripine, Y.; Watanabe, K. A.; George, J.; Glazer, P. M.; Seidman, M. M *Nat. Genet.* **1998**, *20*, 212−214.
(3) Cuenoud, B.; Casset, F.; Hüsken, D.; Natt, F.; Wolf, R. M.; Altmann, K.-H.; Martin, P.; Moser, H. E. *Angew. Chem., Int. Ed.* **1998**, *37*, 1288−1291.
(4) Blommers, M. J. J.; Natt, F.; Jahnke, W.; Cuenoud, B. *Biochemistry* **1998**, *37*, 17714−17725.
(5) Stütz, A. M.; Höck, J.; Natt, F.; Cuenoud, B.; Woisetschläger, M. *J. Biol. Chem.* **2001**, *276*, 11759−11765.

10.1021/ja002592r CCC: $20.00
Published on Web 07/06/2001

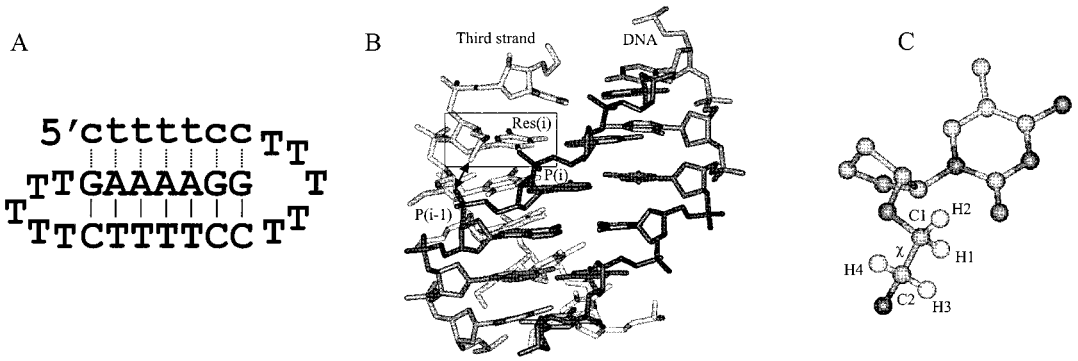

Figure 1. (A and B) Schematic and three-dimensional representation of the unimolecular triplex. (C) 2′-Aminoethoxy-substituted nucleoside drawn with balls and sticks together with the used atom numbering of the protons and carbons of the CH_2-CH_2 moiety and the torsion angle χ. The H_1 and H_2 protons correspond to the Hβ_1 and Hβ_2 protons, respectively, of protein side chains as defined by IUPAC, whereas the H_3 and H_4 protons correspond to Hγ_1 and Hγ_2 protons, respectively.

data,[4] but was detected in MD simulations[4] (R. M. Wolf and F. Casset, personal communication). Due to the large line width (approximately 25−30 Hz for the protons and 30 Hz for the carbons), the question about the presence of other conformations than *gauche*$^+$ could not be fully answered at the time and will be here thoroughly investigated by new NMR techniques and supported by additional MD simulations and calorimetric measurements.

The side chain conformation of molecules in solution is usually investigated by NMR scalar coupling constants. Homonuclear E.COSY techniques[6−8] can be applied to small molecules; for larger systems heteronuclear coupling constants[9−11] are most commonly used, although they do not provide the stereospecific assignment of the protons. For $CH-CH_2$ groups $^3J_{HH}$ coupling constants can be measured in HCCH-E.COSY experiments[12−14] also for molecules of large size.

Here we present an alternative methodology which relies on cross-correlated relaxation rates (CCR rates)[15,16] to derive the conformation of CH_2-CH_2 moieties in aliphatic side chains, together with the stereospecific assignment of the protons. Since CCR rates are linearly dependent on the τ_c, the measured effect increases with the size of the molecule, partially compensating the overall signal decay due to the increased autorelaxation. For this reason the method is a valid alternative to coupling constants for molecules of large size, where the increased line width and the poor signal-to-noise pose serious limitations on the applicability of E.COSY type experiments. The information obtainable from three CH−CH dipolar−dipolar CCR rates is equivalent to the information obtainable from three coupling constant values, as will be shown in the Results section. Furthermore, this information is a unique and valid alternative in those cases where coupling constants cannot be measured, as for example for bound ligands in the fast-exchange regime.[17,18]

This new approach has been applied to a 2′-aminoethoxy-modified oligonucleotide triplex (Figure 1A), to investigate the question of the specificity of hydrogen bond interactions between the aminoethoxy side chain of the third strand and the backbone phosphate(s) of the second strand.

Methods

Differential Scanning Calorimetry. Thermal denaturation studies (micro-DSC) have been performed between 20 and 105 °C with a Microcal MCS-DSC (Microcal Inc., Northhampton, MA, U.S.A.) at a scanning rate of 30 °C h^{-1}. The sample cell was filled with 0.023 mM DNA in a buffer containing 50 mM phosphate, 100 mM NaCl at pH 5.7. Prior to sample insertion, the sample was heated to 60 °C and shaken for 5 min. Two to five runs per sample were performed, and the samples showed reversible melting behavior. Data analysis was performed with the Origin software. The thermograms were fitted with several theoretical denaturation models (two state, non-two state, one melting event, two melting events, independent model, etc.) as provided by the manufacturer.

Molecular Dynamics Simulations. Computation of the molecular dynamics of the triplex has been carried out using DISCOVER 2.98 (Molecular Simulations Inc.), applying the AMBER force field and using a representative structure obtained by NMR[4] as starting structure. Charges and force field parameters of lysine have been used for the atoms of the aminoethoxy substituent. The electrostatic interactions were scaled down by 4r to compensate for the absence of explicit solvent molecules. The simulation ran for 1000 ps with a time step of 1 fs. A second molecular dynamics trajectory run in explicit solvent used a layer of 5 Å of water molecules around the triplex. The ensemble of 1 triplex molecule and 784 water molecules was energy minimized using 200 steps of conjugate gradient minimization using a dielectric constant of 1. The ensemble was subjected to a molecular dynamics simulation of 200 ps.

Design and Synthesis of the Oligonucleotide. The oligonucleotide is 5′-ctt**t**tccTTTTTCCTTTTCTTTTTGAAAAGG (Figure 1A), where c is 2′-aminoethoxy-C5-methyl-cytidine, t is 2′-aminoethoxy-thymidine, and the capitals denote standard deoxyribonucleotides. **t** is 2′-(^{15}N,^{13}C-aminoethoxy)thymidine. This oligonucleotide forms a unimolecular triplex structure. The synthesis of this oligonucleotide has been described before.[4]

NMR Methodology. A CH_2-CH_2 group is schematically depicted in Figure 1C. The CH−CH dipolar−dipolar CCR rates between vectors

(6) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1985**, *107*, 6394−6396.
(7) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Chem. Phys.* **1986**, *85*, 6837−6852.
(8) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Magn. Reson.* **1987**, *73*, 474−579.
(9) Bax, A. Vuister, G. W.; Grzesiek, S.; Delaglio, F.; Wang, A. C.; Tschudin, R.; Zhu. G. *Methods Enzymol.* **1994**, *239*, 79−105.
(10) Hennig, M.; Ott, D.; Schulte, P.; Löwe, R.; Krebs, J.; Vorherr, T.; Bermel, W.; Schwalbe, H.; Griesienger, C. *J. Am. Chem. Soc.* **1997**, *119*, 5055−5056.
(11) Schwalbe, H.; Marino, J. P.; King, C. G.; Wechselberger, R.; Bermel, W.; Griesinger, C. *J. Biomol. NMR* **1994**, *4*, 631−644.
(12) Olsen, H. B.; Ludvigsen, S.; Sørensen O. W. *J. Magn. Reson. A* **1993**, *104*, 226−230.
(13) Olsen, H. B.; Ludvigsen, S.; Sørensen O. W. *J. Magn. Reson. A* **1993**, *105*, 321−322.
(14) Carlomagno, T.; Schwalbe, H.; Rexroth, A.; Sørensen, O. W.; Griesinger, C. *J. Magn. Reson.* **1998**, *135*, 216−226.
(15) Reif, B.; Hennig, M.; Griesinger, C. *Science* 1997, *276*, 1230−1233.
(16) Wokaun, A.; Ernst, R. R. *Mol. Phys.* **1978**, *36*, 317
(17) Carlomagno, T.; Felli, I. C.; Czech, M.; Fischer, R.; Sprinzl, M.; Griesinger, C. *J. Am. Chem. Soc.* **1999**, *121*, 1945−1948.
(18) Blommers, M. J. J.; Stark, W.; Jones, C. E.; Head, D.; Owen, C. E.; Jahnke, W. *J. Am. Chem. Soc.* **1999**, *121*, 1949−1953.

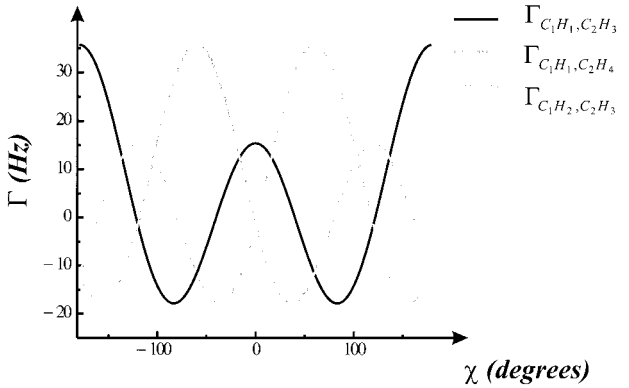

Figure 2. Dependence of the three CCR rates $\Gamma_{C_1H_1,C_2H_3}$, $\Gamma_{C_1H_2,C_2H_3}$, and $\Gamma_{C_1H_1,C_2H_4}$ on the torsional angle χ of Figure 1. The continuous black line represents $\Gamma_{C_1H_1,C_2H_3}$, the continuous gray line $\Gamma_{C_1H_2,C_2H_3}$ and the dashed gray line $\Gamma_{C_1H_1,C_2H_4}$. The calculation is for a τ_c of 4.0 ns and $S^2_{a.} = 1$. From the values of all three rates it is possible to determine the angle χ without ambiguity and to stereospecifically assign protons.

$C_1H_1-C_2H_3$, $C_1H_1-C_2H_4$, $C_1H_2-C_2H_3$, and $C_1H_2-C_2H_4$ depend on the dihedral angle χ. The rates are:[19]

$$\Gamma_{C_1H_i,C_2H_j} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\hbar^2 \gamma_H^2 \gamma_C^2}{r_{CH}^6} J_{ij}(0)$$

$$J_{ij}(\omega) = \frac{2}{5}\left[S_{ij}^2 \frac{\tau_c}{1+\omega^2\tau_c^2} + (1-S_{ij}^2)\frac{\tau_{eff}}{1+\omega^2\tau_{eff}^2}\right]\frac{3\cos^2\theta_{ij}-1}{2} \quad (1)$$

where γ_H and γ_C are the gyromagnetic ratios of proton and carbon, respectively, r_{CH} is the length of the CH bond, $i = 1, 2$, $j = 3, 4$, S_{ij} is the order parameter relevant to the correlated motions of vectors C_1H_i and C_2H_j, θ_{ij} is the projection angle between the vectors C_1H_i and C_2H_j, and $\tau_{eff}^{-1} = \tau_c^{-1} + \tau_e^{-1}$, with τ_e as correlation time for the internal motions. In the expression for $J_{ij}(\omega)$ isotropic diffusion of the oligonucleotide is assumed. Defining $\chi = 0$ for eclipsed vector pairs $C_1H_1-C_2H_3$ and $C_1H_2-C_2H_4$, the four angles θ_{ij} depend on the dihedral angle χ:

$$\begin{aligned}
\cos\theta_{13} &= -\cos^2(109°) + \sin^2(109°)\cos(\chi) \\
\cos\theta_{14} &= -\cos^2(109°) + \sin^2(109°)\cos(\chi + 120°) \\
\cos\theta_{23} &= -\cos^2(109°) + \sin^2(109°)\cos(\chi - 120°) \\
\cos\theta_{24} &= -\cos^2(109°) + \sin^2(109°)\cos(\chi)
\end{aligned} \quad (2)$$

The two $C_1H_1-C_2H_3$ and $C_1H_2-C_2H_4$ dipolar−dipolar CCR rates are always equal, independent of the value of χ. The variation of the three independent CCR rates $\Gamma_{C_1H_1,C_2H_3}$, $\Gamma_{C_1H_1,C_2H_4}$, and $\Gamma_{C_1H_2,C_2H_3}$ between vectors $C_1H_1-C_2H_3$, $C_1H_1-C_2H_4$, and $C_1H_2-C_2H_3$ with the dihedral angle χ is shown in Figure 2. The curve for rate $\Gamma_{C_1H_1,C_2H_3}$ is symmetric around $\chi = 0$, while the two rates $\Gamma_{C_1H_1,C_2H_4}$ and $\Gamma_{C_1H_2,C_2H_3}$ assume different values for $\pm\chi$. The combination of the three CCR rates and NOEs[4] allows the unambiguous determination of the dihedral angle χ and of its sign, which means that the stereospecific assignment of the protons is given.

NMR Experiments. The pulse sequence used to measure CH−CH dipolar−dipolar CCR rates (Figure 3) is a slightly modified version of the quantitative Γ-HCCH experiment published recently.[20] The CCR rates are extracted from the intensity ratio of the signals obtained in a *cross* and in a *reference* experiment.[20] The delay Δ' is equal to zero in the *cross* experiment and to $1/(4J_{CH})$ in the *reference* experiment. A 3D correlation is necessary to distinguish between the contribution of the two protons of the methylene groups, which appear in t_2 at the same carbon frequency. For a CH_2-CH_2 group four peaks are obtained at frequency $(\omega_{H_1},\omega_{C_1},\omega_{H_3})$, $(\omega_{H_2},\omega_{C_1},\omega_{H_3})$, $(\omega_{H_1},\omega_{C_1},\omega_{H_4})$ and $(\omega_{H_2},\omega_{C_1},\omega_{H_4})$. In the *cross* experiment the transfer between the following operators is obtained in the deviation between point *a* and point *b* through the respective CCR rates:

$$\begin{aligned}
&1.\ \hat{\rho}_{\text{Cart.},1} = 4C_{1x}C_{2y}H_{1z} \xrightarrow{\Gamma_{C_1H_1,C_2H_3}} 4C_{1y}C_{2x}H_{3z} = \hat{D}_1 \quad \omega_{H_1},\omega_{C_1},\omega_{H_3} \\
&2.\ \hat{\rho}_{\text{Cart.},2} = 4C_{1x}C_{2y}H_{2z} \xrightarrow{\Gamma_{C_1H_2,C_2H_3}} 4C_{1y}C_{2x}H_{3z} = \hat{D}_1 \quad \omega_{H_2},\omega_{C_1},\omega_{H_3} \\
&3.\ \hat{\rho}_{\text{Cart.},1} = 4C_{1x}C_{2y}H_{1z} \xrightarrow{\Gamma_{C_1H_1,C_2H_4}} 4C_{1y}C_{2x}H_{4z} = \hat{D}_2 \quad \omega_{H_2},\omega_{C_1},\omega_{H_4} \\
&4.\ \rho_{\text{Cart.},2} = 4C_{1x}C_{2y}H_{2z} \xrightarrow{\Gamma_{C_1H_2,C_2H_4}} 4C_{1y}C_{2x}H_{4z} = \hat{D}_2 \quad \omega_{H_2},\omega_{C_1},\omega_{H_4}
\end{aligned} \quad (3)$$

The operators before the arrow represent the initial density operator in the Cartesian basis $\hat{\rho}_{\text{Cart.},i}$ at point *a*, and the operators after the arrow represent the finally detected magnetization terms $\hat{D}_i$ at point *b*. In the *reference* experiment the transfer between the pairs of operators of eq 3 takes place through both CCR rates and coupling constants. In fact, since the delay Δ' is equal to $1/(4J_{CH})$ only $^1/_4$ of the initial operator is transferred through coupling constants to the final one. For the first pair of operators of eq 3 (line 1), for example, the evolution of coupling constants for a delay $1/(4J_{CH})$ generates 16 terms from the initial operator $4C_{1x}C_{2y}H_{1z}$, eight of which are listed in the first column of Table 1. These terms can be transferred to the final operator $4C_{1y}C_{2x}H_{3z}$ through the mechanisms listed in the second column of Table 1, which comprise the four CCR rates $\Gamma_{C_1H_1,C_2H_3}$, $\Gamma_{C_1H_1,C_2H_4}$, $\Gamma_{C_1H_2,C_2H_3}$, and $\Gamma_{C_1H_2,C_2H_4}$, the CCR rates $\Gamma_{C_1H_1,C_1H_2}$ and $\Gamma_{C_2H_3,C_2H_4}$ and the NOE between the geminal protons. Moreover, some of the eight terms of Table 1 can be transferred to $\hat{D}_2$ through the same mechanisms listed above. As several pathways of coherence transfer through relaxation rates are possible, in addition to the transfer brought about by scalar couplings, the resulting intensities of the peaks in the reference experiment are either larger or smaller than expected from pure *J*-coupling transfer. Unlike the case of a CH−CH moiety, for the CH_2-CH_2 group the desired CCR rate cannot be extracted from the ratio of the intensity of a peak in the *cross* and in the *reference* experiment, I_{cross} and I_{ref}, according to:

$$R = I_{cross}/(I_{ref}\cos\pi TJ_{C_1H_j}\cos\pi TJ_{C_2H_i}\sin\pi TJ_{C_1H_i}\sin\pi TJ_{C_2H_k}) = \tanh(\Gamma_{C_1H_i,C_2H_k}T) \quad (4)$$

where T is the length of the constant time. Using eq 4, the corresponding CCR rate would be either overestimated or underestimated, depending on the deviation of the reference peak intensity from that expected in case of an ideal transfer through coupling constants alone. The four CH−CH dipolar−dipolar CCR rates in the CH_2-CH_2 moiety are obtained by fitting the experimental data to simulated intensity ratios, calculated taking full auto- and cross-correlated relaxation into account, as described in the next section.

Simulations. The pulse sequence is simulated between point *a* and point *b* for the *cross* and *reference* experiments. An initial set of 64 Cartesian operators is used, comprising the 16 operators which develop from the initial term of eq 3, line 1, through J_{CH} evolution during the time $\Delta' = 1/(4J_{CH})$, and the corresponding operators deriving from the initial terms of eq 3, lines 2, ,3 and 4. The Cartesian operators are transformed to the single-element operator basis through the transformation:

$$\hat{\rho}_{\text{s.e.}} = U^{-1}\hat{\rho}_{\text{Cart.}}\,U \quad (5)$$

(19) Brüschweiler, R.; Case, D. A. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 27−58; Griesinger, C.; Hennig, M.; Marino, J. P.; Reif, B.; Schwalbe, H. In *Biomolecular NMR*; Berliner, L., Krishna, R., Eds.; Biological Magnetic Resonance, Vol. 16; Kluwer Academic/Plenum Press: 1999; pp 259−367; Reif, B.; Diener, A.; Hennig, M.; Maurer, M.; Griesinger, C. *J. Magn. Reson.* **2000**, *143*, 45−68.

(20) Felli, I. C.; Richter, C. Griesinger, C.; Schwalbe, H. *J. Am. Chem. Soc.* **1999**, *121*, 1956−1957.

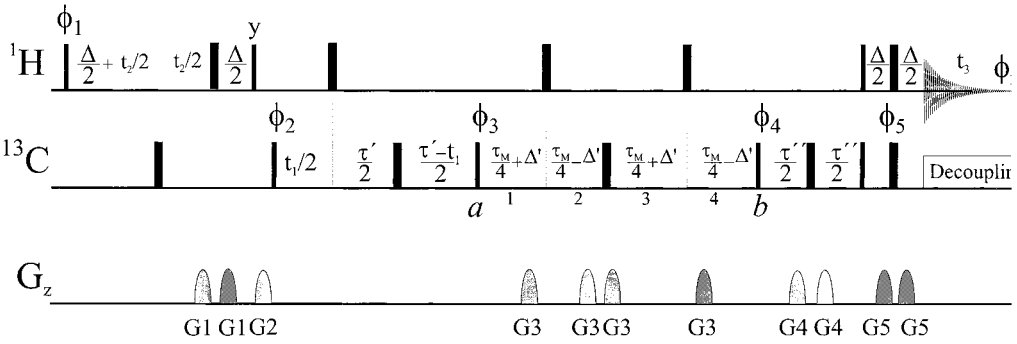

Figure 3. Pulse sequence for the quantitative Γ-HCCH experiment for the measurement of CH−CH dipolar−dipolar CCR rates in $H_nCCH_{n'}$ groups. The sequence is a three-dimensional $H(t_2)-C(t_1)-C-H(t_3)$ correlation with the following parameters values: $\tau' = \tau'' = 7.0$ ms $(1/(4J_{CC}))$ for C−C−C chains); $\Delta = 3.7$ ms; $\Delta' = 0.87$ ms in the *reference* experiment and $\Delta' = 0$ in the *cross* experiment; $\tau_M = n/J_{CC} = 28$ ms for $n = 1$. $\phi_1 = x, -x$; $\phi_2 = 2(x), 2(-x)$; $\phi_3 = 4(x), 4(-x)$; $\phi_4 = 8x, 8(-x)$; $\phi_5 = 16x, 16(-x)$; $\phi_{rec.} = \phi_1 + \phi_2 + \phi_3 + \phi_4$. The default phase is *x*. Quadrature detection is achieved in t_1 and in t_2 by States-TPPI on ϕ_2 and ϕ_1, respectively. ^{13}C spins are irradiated with hard pulses. The power for carbon decoupling during acquisition is 2.5 kHz. The experiments were acquired on a Bruker DRX-800 MHz spectrometer.

Table 1. In the First Column Eight of the Sixteen Operators Which Are Generated in the Reference Experiment by Coupling Constant Evolution, Starting from the Initial Operator of Line 1, Eq 3, Are Reported[a]

$4C_{1x}C_{2y}H_{1z}\cos(\pi J_{C_1H_1}\Delta')\cos(\pi J_{C_1H_2}\Delta')\cos(\pi J_{C_2H_3}\Delta')\cos(\pi J_{C_2H_4}\Delta')$	$\Gamma_{C_1H_1,C_2H_3} \rightarrow \hat{D}_1$ or $\Gamma_{C_1H_1,C_2H_4} \rightarrow \hat{D}_2$
$\underline{4C_{1y}C_{2x}H_{3z}\sin(\pi J_{C_1H_1}\Delta')\cos(\pi J_{C_1H_2}\Delta')\sin(\pi J_{C_2H_3}\Delta')\cos(\pi J_{C_2H_4}\Delta') = \hat{D}_1}$	$\Gamma_{C_2H_3,C_2H_4} \rightarrow \hat{D}_2$ or NOE $\rightarrow \hat{D}_2$
$4C_{1x}C_{2y}H_{2z}\sin(\pi J_{C_1H_1}\Delta')\sin(\pi J_{C_1H_2}\Delta')\cos(\pi J_{C_1H_3}\Delta')\cos(\pi J_{C_2H_4}\Delta')$	$\Gamma_{C_1H_2,C_2H_3} \rightarrow \hat{D}_1$ or $\Gamma_{C_1H_2,C_2H_4} \rightarrow \hat{D}_2$
$4C_{1y}C_{2x}H_{4z}\sin(\pi J_{C_1H_1}\Delta')\cos(\pi J_{C_1H_2}\Delta')\cos(\pi J_{C_2H_3}\Delta')\sin(\pi J_{C_2H_4}\Delta') = \hat{D}_2$	NOE $\rightarrow \hat{D}_1$ or $\Gamma_{C_2H_3,C_2H_4}$
$16C_{1y}C_{2x}H_{1z}H_{2z}H_{4z}\cos(\pi J_{C_1H_1}\Delta')\sin(\pi J_{C_1H_2}\Delta')\cos(\pi J_{C_2H_3}\Delta')\sin(\pi J_{C_2H_4}\Delta')$	$\Gamma_{C_1H_1,C_1H_2} \rightarrow \hat{D}_2$
$16C_{1y}C_{2x}H_{1z}H_{2z}H_{3z}\cos(\pi J_{C_1H_1}\Delta')\sin(\pi J_{C_1H_2}\Delta')\sin(\pi J_{C_2H_3}\Delta')\cos(\pi J_{C_2H_4}\Delta')$	$\Gamma_{C_1H_1,C_1H_2} \rightarrow \hat{D}_1$
$16C_{1x}C_{2y}H_{1z}H_{3z}H_{4z}\cos(\pi J_{C_1H_1}\Delta')\cos(\pi J_{C_1H_2}\Delta')\sin(\pi J_{C_2H_3}\Delta')\sin(\pi J_{C_2H_4}\Delta')$	$\Gamma_{C_1H_1,C_2H_4} \rightarrow \hat{D}_1$ or $\Gamma_{C_1H_2,C_2H_3} \rightarrow \hat{D}_2$
$16C_{1x}C_{2y}H_{2z}H_{3z}H_{4z}\sin(\pi J_{C_1H_1}\Delta')\sin(\pi J_{C_1H_2}\Delta')\sin(\pi J_{C_2H_3}\Delta')\sin(\pi J_{C_2H_4}\Delta')$	$\Gamma_{C_1H_2,C_2H_4} \rightarrow \hat{D}_1$ or $\Gamma_{C_1H_2,C_2H_3} \rightarrow \hat{D}_2$

[a] The operator corresponding to D1 in eq 3 is underlined. Operators of the first column can be turned into either D1 or D2 by the relaxation mechanisms listed in the second column.

where U is the unitary transformation from the single element operator basis to the Cartesian operators and $\hat{\rho}_{s.e.}$ and $\hat{\rho}_{Cart.}$ are the two density matrices in the respective single-element and Cartesian operators basis. The $\hat{\rho}_{s.e.}$ evolve during the constant time between points *a* and *b* according to the Liouville−von Neumann differential equation for the density operator:

$$\frac{d\hat{\rho}_{s.e.}}{dt} = -i\hat{\hat{H}}_0\hat{\rho}_{s.e.} - \hat{\hat{\Gamma}}(\hat{\rho}_{s.e.} - \hat{\rho}^0_{s.e.}) \tag{6}$$

where $\hat{\hat{H}}$ is the Hamiltonian superoperator containing the frequency evolution and the *J*-coupling terms and $\hat{\hat{\Gamma}}$ is the relaxation superoperator. In the calculations, the inhomogeneous part of the Liouville−von Neumann equation $\hat{\hat{\Gamma}}(\hat{\rho}^0_{s.e.})$ has been neglected, since it is experimentally averaged to zero by phase cycling. If we define $\hat{\hat{L}}_i$ as the matrix representing $-i\hat{\hat{H}}_0 - \hat{\hat{\Gamma}}$ during the time t_i, $\hat{\hat{P}}_H$, and $\hat{\hat{P}}_C$ as the matrices representing a π pulse on the proton and on the carbon, respectively, the density operator at point *b* can be expressed as follows:

$$\hat{\rho}_{s.e.}(b) = e^{\hat{\hat{L}}_4 t_4}\hat{\hat{P}}_H e^{\hat{\hat{L}}_3 t_3}\hat{\hat{P}}_C e^{\hat{\hat{L}}_2 t_2}\hat{\hat{P}}_H e^{\hat{\hat{L}}_1 t_1}\hat{\rho}_{s.e.}(a) \tag{7}$$

where t_i for $i = 1, 2, 3, 4$ are the time intervals reported in Figure 3. The $\hat{\rho}_{s.e.}(b)$ at point *b* is transformed back into the Cartesian operators basis ($\hat{\rho}_{Cart.}(b)$). The intensity I_{ij} of the peaks at the four frequency combinations of eq 3, is obtained from the efficiency of the transfer to the detectable operator $\hat{D}_i$ at point *b* for an initial density operator $\hat{\rho}_{Cart.,j}$:

$$I_{ij} = \langle \hat{D}_i | \hat{\rho}_{Cart.}(b) \rangle \tag{8}$$

where $i,j = 1,2$. In the calculation the effect of the NOE between geminal protons and of the relaxation rates $\Gamma_{C_1H_1,C_1H_2}$ and $\Gamma_{C_2H_3,C_2H_4}$ during the times $\tau'$and τ'' before *a* and after *b*, respectively, is taken into account, allowing for mixing of the product operators terms containing H_{mz} magnetization ($m = 1, 3$) with the terms containing H_{nz} magnetization ($n = 2, 4$). The CCR rates $\Gamma_{C_1H_1,C_1H_2}$ and $\Gamma_{C_1H_3,C_1H_4}$ and the NOE between geminal protons, used as fixed parameters in the fitting, were calculated assuming a correlation time $\tau_c = 4.0$ ns and an autocorrelated order parameter S^2_a of 0.8 ($NOE_{H_1,H_2} = -6.7$ Hz; $\Gamma_{C_1H_1,C_1H_2} = \Gamma_{C_2H_3,C_2H_4} = -9.8$ Hz), as derived from $^{13}C-T_1$ and heteronuclear NOE relaxation measurements for the CH_2-CH_2 moiety under consideration ($^{13}C-R_1 = 3.65$ Hz for C_1 and 3.45 Hz for C_2 and heteronuclear NOE = 1.5 for C_1 and 1.4 for C_2).

The calculated intensity ratios of the peaks in the *cross* ($\Delta' = 0$) and in the *reference* experiment ($\Delta' = 1/(4J_{CH})$) are fitted to the experimental values, targeting the minimum of the function:

$$f = \sqrt{\sum_{i,j=1}^{2}\left(\frac{(R^{exp}_{ij} - R^{calc}_{ij})}{\delta_{ij}}\right)^2}/4 \tag{9}$$

where R^{exp}_{ij} and R^{calc}_{ij} are the experimental and calculated ratios of the intensities of peak *ij* in the *cross* and in the *reference* experiment and δ_{ij} is the experimental error on R^{exp}_{ij}. $\Gamma_{C_1H_1,C_2H_3}$, $\Gamma_{C_1H_1,C_2H_4}$, and $\Gamma_{C_1H_2,C_2H_3}$ have been optimized starting from their apparent values derived from eq 4. $\Gamma_{C_1H_2,C_2H_4}$ was set equal to $\Gamma_{C_1H_1,C_2H_3}$. The optimization procedure was a grid search repeated several times with decreasing step size until 0.2 Hz. In the first iteration the step size was equal to 20.0 Hz and a range of ±20 Hz around the starting values of Γ was explored. The point with lowest *f* value was used as the new starting point for the next iteration of the grid search, and the step size was reduced to $^3/_4$ of the initial step size. A total of 16 iterations were used until the step size was equal to the smallest δ_{ij} (1). (The program is available on request.)

Results

Thermodynamic Analysis. The thermodynamic properties of the intramolecular triplex with a third strand containing seven 2′-aminoethoxy-modified residues have been compared with those of the unmodified triplex (Figure 4). For the unmodified oligonucleotide, a melting transition is observed at 57 °C (ΔH = 40 kcal/mol). The melting curve could not be fitted with a two-state model. The hairpin with the Watson−Crick duplex stem of seven base pairs showed a duplex-to-single strand transition at 57 °C, associated with the same melting enthalpy. Combination of these data suggests that the measured transition

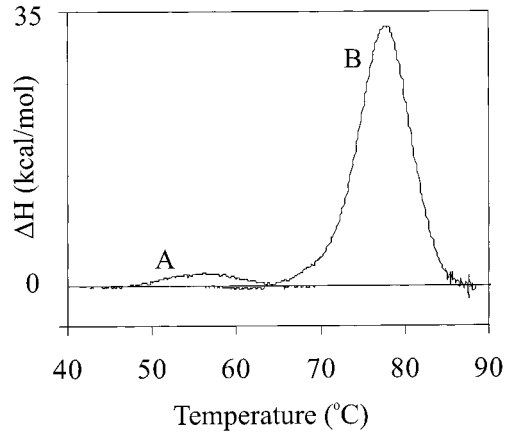

Figure 4. Differential scanning calorimetry curves of the 31-mer oligonucleotides at pH 5.7. The secondary structure is drawn in Figure 1A. (A) The reference oligonucleotide with standard DNA nucleotides. The three cytosines in the third strand are C5-methylated. (B) The 2′-aminoethoxy-modified oligonucleotide with the same nucleotide sequence. The three cytosines in the third strand are C5-methylated. All seven residues contain a 2′-aminoethoxy-substituted ribose.

represents the duplex-to-single strand transition. A clear triplex-to-duplex transition expected at lower temperatures is not visible. 1D NMR spectra of the imino protons have been measured as a function of the temperature between 0 and 45 °C. Much more than the expected 14 imino proton resonances are observed downfield from 12 ppm, suggesting the presence of the triplex in equilibrium with other stable conformations at low temperature.

The modified triplex with seven 2′-aminoethoxy-substituted riboses melts at 78 °C associated with $\Delta H = 266$ kcal/mol. The melting temperature and enthalpy are significantly higher than those of the unmodified oligonucleotides. Also in this case, the curve could only be fitted using a non-two-state model, suggesting a complex multistate melting transition triplex-to-duplex-to-single strand oligonucleotide. 1D NMR experiments are in agreement with the presence of a stable single species between 0 and 55 °C. All imino protons experience line broadening at elevated temperatures. Therefore, it can be concluded that the transition around 78 °C corresponds to a triplex-to-single strand transition.

Molecular Dynamics Simulations. The structure of the unimolecular triplex was previously determined by NMR spectroscopy[4] and MD simulations. One of the NMR structures of the ensemble has been subjected to further investigation by MD. The simulations have been initially carried out in vacuo. The triplex structure remained stable during the simulation. The deoxyribose moieties of the DNA strands undergo expected transitions between the N-type and S-type conformations, whereas the 2′-substituted riboses of the third strand remain in an N-type conformation. The amino group of the aminoethoxy substituent forms hydrogen bonds alternatively with the pro-*R* oxygen of the phosphate groups in position *i* and $i - 1$ of the polypurine strand of the DNA. Hydrogen-bond formation is typically correlated with the torsion angle χ in the aminoethoxy moiety (Figure 5). When the torsion angle χ is *gauche*$^+$, the amino group of residue *i* is within hydrogen-bonding distance with the pro-*R* oxygen of residue $i - 1$ in the second strand (Figure 5A). The aminoethoxy group has significant flexibility, and conformational transitions are observed for χ between the *gauche*$^+$ and the *trans* state. In the *trans* conformation, the amino group is within hydrogen-bonding distance with the pro-*R* oxygen of residue *i* in the second strand (Figure 5B). The third hypothetically stable conformer, that is, *gauche*$^-$, in which the amino group points away from the Hoogsteen−Crick groove, is practically unobserved.

In the MD simulations using a layer of explicit solvent molecules, the dynamical features of the aminoethoxy moiety described above could not be observed. On the contrary, the aminoethoxy moiety behaves like a rigid anchor connecting the two strands, and only a short transition to the *gauche*$^-$ conformation occurs.

Cross-Correlated Relaxation Experiments on the 2′-Aminoethoxy-Modified Oligonucleotide Triplex. For cross-correlated relaxation experiments, ^{13}C isotopically enriched molecules are needed. The triplex under investigation contains one selectively labeled aminoethoxy substituent in the fourth residue of the third strand. In Figure 6 the quantitative Γ-HCCH *cross* (B) and *reference* (A) experiments for the labeled 2′-aminoethoxy side chain are shown. Two peaks are missing in the *cross* experiment, meaning that the corresponding apparent CCR rate is close to zero. The four protons corresponding to the CH_2-CH_2 group of the 2′-aminoethoxy are named H_1, H_2, H_3, and H_4, for chemical shifts going from low to high field. The $C-CH_2-N$ protons resonate around 3.6 ppm ($^{13}C = 42.3$ ppm) and the $O-CH_2-C$ protons around 4.5 ppm ($^{13}C = 60$ ppm). The apparent CCR rates, measured according to eq 4, are $\Gamma_{C_1H_1,C_2H_3} = \Gamma_{C_1H_2,C_2H_4} = 0.0 \pm 0.4$, $\Gamma_{C_1H_1,C_2H_4} = 11.3 \pm 0.2$, and $\Gamma_{C_1H_2,C_2H_3} = -3.4 \pm 0.3$. The effective CCR rates, obtained from the fitting procedure described above, are $\Gamma_{C_1H_1,C_2H_3} = \Gamma_{C_1H_2,C_2H_4} = -6.3 \pm 0.4$, $\Gamma_{C_1H_1,C_2H_4} = 27.3 \pm 0.2$, and $\Gamma_{C_1H_2,C_2H_3} = -4.7 \pm 0.3$. The larger the rate the more its apparent value deviates from the effective one, due to the large contribution of the transfer through cross-correlated relaxation to the intensity of the reference peak. The experimental relative intensities of the four reference and cross peaks could be well reproduced in the simulations as well.

Interpretation of the rates in terms of conformational equilibrium of the three staggered conformations. We assume an equilibrium of the three staggered conformations that interconvert on a time scale not relevant for relaxation. Invoking the Occam's razor principle, which suggests to prefer a simple model over a more complicated one, we refrained from using models in which the interconversion between the conformations is relevant for relaxation. The three values for $\Gamma_{C_1H_1,C_2H_3}$, $\Gamma_{C_1H_2,C_2H_3}$, and $\Gamma_{C_1H_1,C_2H_4}$ are interpreted assuming the presence in equilibrium of the three lowest energy staggered conformations with populations p_{g+}, p_{g-}, and p_t. The experimental data are best fitted to calculated CCR rates according to $\Gamma = p_{g+}\Gamma^{g+} + p_{g-}\Gamma^{g-} + p_t\Gamma^t$ with $\Gamma^{g+} = -11.6$ Hz and $\Gamma^t = 35.7$ Hz.

The populations of the three staggered conformations and the order parameters S_1^2 and S_2^2 are optimized through a grid search using eq 9 as target function. S_1^2 is applied to the CCR rate involving pairs of CH vectors which define a dihedral angle of 180° in absence of internal motion (for example, C_1H_1, C_2H_3 in a *trans* conformation); S_2^2 is applied to those CH vectors for which the dihedral angle is ±60° in absence of internal motion (for example, C_1H_1, C_2H_3 in a *gauche*$^\pm$ conformation). Rigorously, different order parameters should be used for each pair of CH vectors in the three staggered conformations. Our choice assumes that anisotropic internal motions of the CH_2-CH_2 group are axially symmetric around the bond axis. Since the number of fitted parameters (two order parameters and two populations) is higher than the number of experimental data (three CCR rates), multiple solutions are found in the following ranges: $p_{g+} = 81-97\%$; $p_{g-} = 3-12\%$; $p_t = 0-7\%$; $S_1^2 = 0.80-1.00$ and $S_2^2 = 0.50-0.85$. It should be noted that the

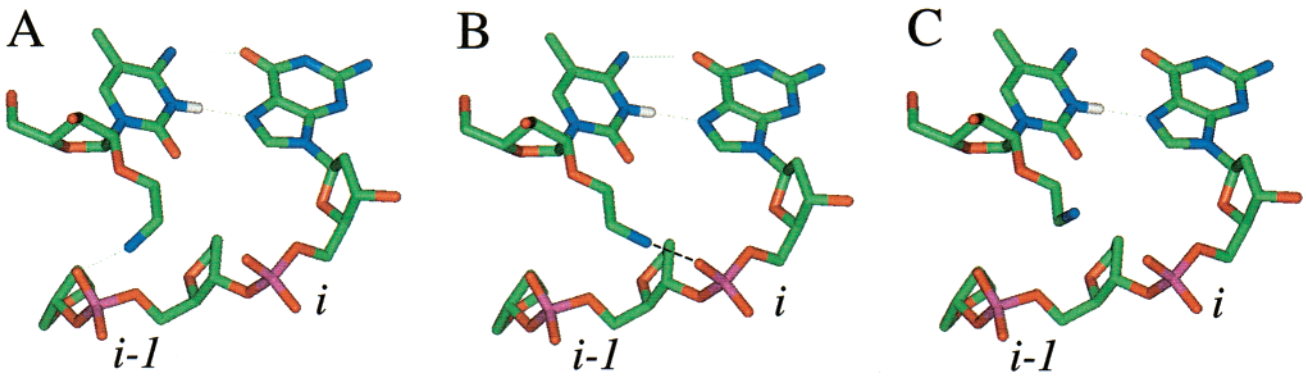

Figure 5. Stick models which represent the dual recognition between a 2′-aminoethoxy-substituted residue in the third strand with the polypurine strand of the DNA: Hoogsteen base-pair formation as well as interactions of the positively charged amino group of the aminoethoxy side chain with the phosphate backbone of the polypurine strand. Molecular dynamics simulations indicate the presence of three conformations of the aminoethoxy side chain, i.e., *gauche*$^+$ (A), *trans* (B), and *gauche*$^-$ (C) (see text). The first two conformations allow a hydrogen-bond interaction of the amino group with the pro-*R* oxygens of the phosphates in position $i-1$ and i, respectively.

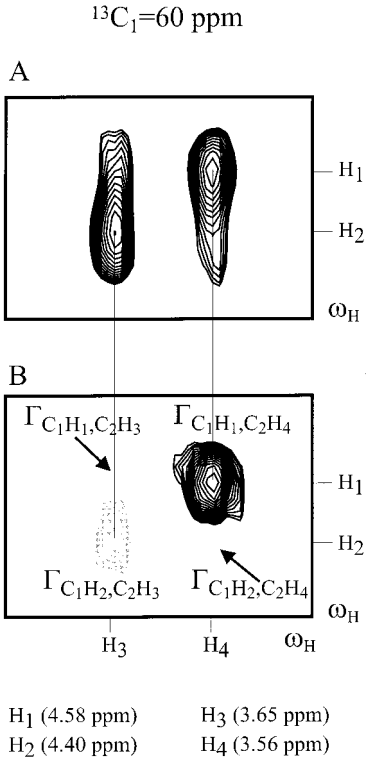

Figure 6. Two-dimensional plane ($\omega_2-\omega_3$) at the ^{13}C frequency ω_1 = 60 ppm of the three-dimensional spectrum, acquired with the sequence of Figure 3, for 0.5 mM of the 2′-aminoethoxy-modified rY-dR-dY triplex sample in D_2O at pH 5.7 and 25 °C. *Reference* experiment ($\Delta' = 0.87$ ms) in panel A, *cross* experiment ($\Delta' = 0$) in panel B. The peak in the *cross* experiment corresponding to the rate $\Gamma_{C_1H_1,C_2H_4}$ is strong and positive; the one corresponding to the rate $\Gamma_{C_1H_2,C_2H_3}$ is weak and negative. The two peaks corresponding to the equal rates $\Gamma_{C_1H_1,C_2H_3}$ and $\Gamma_{C_1H_2,C_2H_4}$ are missing, suggesting an apparent value of zero for these rates.

order parameter $S_1^2 = 0.80-1.00$ and the prevalence of a single conformation reflect the large cross-correlated relaxation rate of the protons *trans* to each other. A model that includes jumps between the staggered conformations on a time scale relevant for relaxation would necessarily reduce these cross-correlated relaxation rates requiring even higher populations of the *gauche*$^+$ conformation. Thus, although the motional model assumed is a simplification, it provides a lower limit for the population of the main conformation. We also investigated the effect of anisotropy of the oligonucleotide. We find that the orientation dependence of the cross-correlated relaxation rates do not exceed ±5% of the maximum cross-correlated relaxation rate and translate linearly into the order parameters. The population distribution found is essentially unchanged.

Discussion

Conformational Aspects. The molecular dynamics simulations suggest two possibilities to explain the observed thermodynamic stability of the triplex with a 2′-aminoethoxy-modified third strand: the amino group-containing side chain is flexible (model A, vacuum simulations) or rigid, exclusively in the *gauche*$^+$ conformation (model B, simulations in the presence of explicit solvent).

In model A, the aminoethoxy group is considered as a flexible positively charged side chain. Time-shared hydrogen-bond interactions take place with two phosphates, thereby reducing the partial negative charge of the phosphate backbone in the DNA. This is a relatively weak and nonspecific interaction, comparable with the binding of polyanions such as spermine and spermidine. Different physical interactions influence the stability of the three staggered conformations: *gauche*$^+$ and *trans* states are favored because of interstrand hydrogen bonds, while the *gauche*$^+$ and *gauche*$^-$ rotamers are stabilized by the so-called gauche effect.[21] 1D NMR spectra of a dinucleotide containing an aminoethoxy side chain show an almost complete degeneration of chemical shifts of the CH_2 protons, indicating flexibility of the aminoethoxy moiety in absence of the DNA. The flexibility of the side chain within the triplex architecture assumed in model A would indicate that the triplex is considerably stabilized by nonspecific interactions between the positively charged amino group and the negatively charged phosphate backbone and entropic contributions from the conformational disorder.

In model B, the amino group makes an exclusive and therefore very specific interaction with only one phosphate of the DNA. In this case, the interaction between the amino group of the third strand and the phosphate of the second strand is specific, meaning that the gauche effect and the hydrogen-bond formation stabilize the triplex despite the decreased conformational entropy.

A choice between the two models is now possible with the detailed population analysis allowed by the new NMR approach. According to the NMR data the following populations ranges are possible: $p_{g+} = 81-97\%$; $p_{g-} = 3-12\%$; $p_t = 0-7\%$ with

(21) Wiberg, K. B. *Acc. Chem. Res.* **1996**, *29*, 229−234.

motional parameters $S_1^2 = 0.80-1.00$ and $S_2^2 = 0.50-0.85$. Interestingly, the order parameter S_1^2 is as large as the autocorrelated order parameter $S_{a.}^2$ (0.8) or even larger. Since the effect of internal motions on vectors with initial projection angle $\theta = 0°$ or $\theta = 180°$ are the same, if there was no internal motion within the CH_2-CH_2 group, S_1^2 would be equal to $S_{a.}^2$. In case the motions which take place inside the CH_2-CH_2 group are not correlated to the internal motions of the CH_2-CH_2 group as a whole S_1^2 is $\leq S_{a.}^2$. The fact that $S_1^2 \geq S_{a.}^2$ indicates that the two internal motions are correlated. An order parameter of 1.0 would imply that the internal motion within the CH_2-CH_2 completely compensates for the effect of the internal motions of the CH_2-CH_2 group as a whole. As the probability of this event is rather low, we impose $S_1^2 \leq 0.9$. Under this assumption the solution space is reduced to the following: $p_{g+} = 87-97\%$; $p_{g-} = 3-8\%$; $p_t = 0-5\%$; $S_1^2 = 0.8-0.9$ and $S_2^2 = 0.50-0.7$. This result shows unambiguously that the CH_2-CH_2 moiety in the 2′-aminoethoxy side chain adopts almost exclusively a *gauche*$^+$ conformation. Both the *trans* and the *gauche*$^-$ conformations are only transiently populated with a slight tendency of $p_{g-} > p_t$. This is in contrast to what is suggested by molecular dynamics simulations in vacuum, where the conformational space allows for the presence of the stable alternative *trans* conformation. In simulations with explicit solvent molecules, the aminoethoxy moiety is less flexible: the *trans* conformation is not observed, and only a small fraction of *gauche*$^-$ conformation is present. Although the simulations have been run only for a short period of 200 ps vs 1000 ps for the in vacuo simulations, they provide a better representation of the experimental data.

The experimental results allow the characterization of the central aminoethoxy side chain in the third strand. Apart from the aminoethoxy substituents at C1 and C7, the conformational properties are likely to be the same for the other aminoethoxy moieties. In case the interaction between amino group *i* and phosphate $i - 1$ is specific, it is not likely that the aminoethoxy at C1 is hydrogen-bonded. Indeed, the chemical shift and NOE data indicate disorder of this side chain.

In conclusion, the enhanced binding properties can be explained by the specific interaction of the aminoethoxy side chain with a unique phosphate group of the DNA (model B). The aminoethoxy side chains can be represented as stable spacers that are linked with covalent bonds to the third strand and with hydrogen bonds to the pro-*R* oxygen atoms of the phosphate of the polypurine strand. This interaction is able to hold the triplex structure together, even at high temperature (78 °C). The aminopropoxy-modified DNA forms less stable triplexes.[4] Unspecific ionic interactions are possible for this substituent, but formation of specific and stable hydrogen bonds is hindered by the longer side chain. No very stable hydrogen bond was observed for a model of the aminopropoxy-modified DNA by MD simulations (R. M.Wolf and F. Casset, unpublished results). This finding substantiates the importance of specific interactions in the stabilization of the triplex motif. The thermodynamic and kinetic stability of the aminoethoxy-modified triplexes is unprecedented and seems thus far to be the best candidate for targeting DNA in vivo.

NMR Methodology. The CH−CH dipolar−dipolar CCR rates allow conformational investigation of biomolecular side chains without any need for parametrization. They are easy to measure also for molecules that exhibit a broad line width and provide at the same time conformational and dynamic information. The method presented here constitutes a valid alternative to coupling constants, when no satisfactory parametrization is available or when coupling constants are difficult or impossible to measure. This last case applies to the conformational investigation of small ligands weakly bound to a macromolecule for which only "relaxation-related" physical observables (transferred NOE, transferred CCR rates) can give information about the bound conformation.

Concluding Remarks

For the system investigated here the interaction between the positively charged side chain and the DNA phosphate is highly specific. The question of the specificity of such contacts can be asked more generally for protein−DNA recognition mechanisms which include a lysine side chain−phosphate contact. While the interactions which drive protein−DNA recognition are often considered to be unspecific, the results presented here show that specific interactions are possible and may have an impact on the stability and specificity of the complex.

A good understanding of molecular recognition requires the investigation of the three-dimensional structure. The newly presented NMR approach allows population analysis of side chain conformations even for large systems. We expect this to have a notable impact on the description of the details of structure−function relationship.

Acknowledgment. This work was supported by the Fonds der Chemischen Industrie, the DFG, and the MPG. All spectra were recorded at the large-scale facility for Biomolecular NMR at the University of Frankfurt (ERBCT96034). T.C. is supported by the E.U. through a Marie Curie stipend, J.M. by the Fonds der Chemischen Industrie through a Kekulé stipend. We thank Professor Dr. H. Schwalbe, MIT, for useful discussions.

JA002592R