Journal of Biomolecular NMR, **24:** 1–14, 2002.
KLUWER/ESCOM

Refinement of d(GCGAAGC) hairpin structure using one- and two-bond residual dipolar couplings

Petr Padrta, Richard Štefl*, Lukáš Králík, Lukáš Žídek & Vladimír Sklenář**
National Centre for Biomolecular Research, Masaryk University, Kotlářská 2, CZ-611 37 Brno, Czech Republic

Received 16 January 2002; Accepted 12 July 2002

Key words: DNA hairpin, NMR, refinement, residual dipolar couplings

Abstract

The structure of the ^{13}C,^{15}N-labeled d(GCGAAGC) hairpin, as determined by NMR spectroscopy and refined using molecular dynamics with NOE-derived distances, torsion angles, and residual dipolar couplings (RDCs), is presented. Although the studied molecule is of small size, it is demonstrated that the incorporation of diminutive RDCs can significantly improve local structure determination of regions undefined by the conventional restraints. Very good correlation between the experimental and back-calculated small one- and two-bond ^{1}H-^{13}C, ^{1}H-^{15}N, ^{13}C-^{13}C and ^{13}C-^{15}N coupling constants has been attained. The final structures clearly show typical features of the miniloop architecture. The structure is discussed in context of the extraordinary stability of the d(GCGAAGC) hairpin, which originates from a complex interplay between the aromatic base stacking and hydrogen bonding interactions.

Introduction

Hairpins are nucleic acids sequences consisting of a single-stranded loop region closed by a base-paired stem. They have been shown to play a significant role in a number of biological processes. Short DNA minihairpin sequences, in particular d(GCGAAGC) and d(GCGAAAGC), occur frequently in biologically important regions (Yoshizawa et al., 1997). The d(GCGAAGC) fragment has been found in the replication origins of phage ϕX174 (Arai et al., 1981) and herpes simplex virus (Elias and Lehman, 1988), in a promoter region of an *E. coli* heat-shock gene (Cowing et al., 1985), and in rRNA genes (Hirao et al., 1994). Hairpins (or hairpin-like structures) may play a major role in expansion events of triplet repeat expansion diseases like the X syndrome, Huntington's disease, and Friedreich's ataxia. The latter disease is associated with expansion of d(GAA)·d(TTC) repeats (Suen et al., 1999).

The most interesting feature of minihairpins of the general sequence d(GCGNAGC) (N = A, G, C, T) is their extraordinary stability represented by high melting temperature, polyacrylamide gel mobility, and resistance against nucleases (Hirao et al., 1992, 1994). The thermodynamically most stable fragment is d(GCGAAGC), i.e., N = A in the general sequence, with $T_m = 76\,^{\circ}C$ (Yoshizawa et al., 1997). According to the thermodynamic studies of Yoshizawa et al., d(GCGNAGC) fragments have much smaller $\Delta G_{37\,^{\circ}C}$ value for the loop (from +0.3 to −0.4 kcal mol^{-1}) than stable RNA hairpins such as r(UNCG) (1.0 kcal mol^{-1}). The d(GAA) loop has the lowest value of −0.4 kcal mol^{-1} suggesting that the structure of d(GCGAAGC) is stable even with only two G·C base pairs in the stem. Corresponding RNA sequences are much less stable (Hirao et al., 1992, 1994) indicating that the stability of these loops depends on the presence of deoxyribose sugar on the backbone.

*Current address: Institute for Molecular Biology and Biophysics, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.
**To whom correspondence should be addressed. E-mail: sklenar@chemi.muni.cz

The extraordinary stability of the mini-hairpins could be exploited for stabilizing other oligonucleotides (Jolles et al., 1997), e.g., for stabilization of mRNA in *in vitro* protein synthesizing systems or for the isolation of full-length RNAs from cells (Hirao et al., 1993, Yoshizawa et al., 1994). Correct understanding of behavior of the mini-hairpins requires detailed knowledge of their structures. A model of the DNA hairpin d(GCGAAGC) based on NMR data was previously proposed by Hirao et al. (1994).

In general, problems in the structure determination of nucleic acids originate mostly from low proton density, a small number of NOE contacts, and a lack of the long range restraints between the elements of secondary structure due to an elongated shape of most molecules. Recently, NMR measurements in partially aligned phase (e.g., bicelles, filamentous phages) have been successfully used to extract residual dipolar couplings (RDCs). RDCs provide additional long-range and local geometrical restraints using a small degree of molecular alignment with the static magnetic field (Prestegard, 1998; Tjandra and Bax, 1997). The use of RDCs in nucleic acid structure determination is usually based on relatively large ^{1}D(CH) and ^{1}D(NH) couplings (MacDonald et al., 2001; Warren and Moore, 2001; Tsui et al., 2000; Vermeulen et al., 2000; Tjandra et al., 2000). In this paper, we present a structure of d(GCGAAGC) hairpin calculated using the current NMR methodology with structural information extracted from NOE data, scalar couplings, and both large (^{1}D(CH) and ^{1}D(NH)) and small (^{1}D(NC), ^{2}D(CH), and ^{2}D(NH)) dipolar couplings. As will be shown, the inclusion of small RDCs improves the local structure in regions that are underdetermined by the conventional restraints.

Materials and methods

NMR sample

Unlabeled sample of the d(GCGAAGC) hairpin was synthesized by the phosphoramidite method, dissolved in 200 µl of sodium phosphate buffer (pH 6.62) and divided to prepare one H_2O and one D_2O sample. To the H_2O sample, 350 µl H_2O and 50 µl D_2O were added. The D_2O sample was dissolved in 99.96% D_2O and dried twice with a stream of nitrogen gas directly in the 5-mm sample tube. Finally, the sample was dissolved in 99.998% D_2O. The samples were annealed by heating in the water bath to 90 °C followed by slow cooling to room temperature. The final concentration and pH of the heptamer were 6.0 mM and 6.7, respectively.

Samples of 1 mM uniformly ^{13}C, ^{15}N-labeled d(GCGAAGC) hairpin (purchased from Silantes GmbH, München, Germany) were prepared similarly as described above. The H_2O sample was diluted to 0.5 mM concentration and divided into two parts. One part was used as an isotropic sample while the other part was used to prepare a partially aligned sample. The partially aligned sample contained approximately 20 mg ml^{-1} Pf1 phage prepared using a slightly modified procedure of Hansen et al. (1998) and dialysed against 10 mM sodium phosphate buffer of pH 6.7.

NMR experiments

The NMR experiments were performed on a Bruker AVANCE 500 MHz spectrometer equipped with a z-gradient triple resonance ^{1}H/^{13}C/BB and ^{1}H/^{13}C/^{15}N probeheads. All measurements were carried out at 30 °C, unless indicated otherwise. The data were processed on SGI computers (Indy, O2, Octane) using Bruker NMR Suite programs.

The following spectra of the unlabeled sample were measured: 2D NOE spectra (Jeener et al., 1979) in D_2O with mixing times 100, 200 and 300 ms, 2D NOE in H_2O at 4 °C with mixing times 100 and 150 ms using WATERGATE solvent suppression (Piotto et al., 1992), 2D TOCSY spectra (Braunschweiler and Ernst, 1983) with MLEV-17 mixing (Bax and Davis, 1985) with mixing times 60, 70, 90 and 110 ms, 2D DQF-COSY spectrum (Piantini et al., 1982), 2D ^{1}H-^{13}C HSQC spectra with very high resolution (Bodenhausen and Ruben, 1980), and 2D J-resolved ^{1}H-^{31}P correlation spectra with proton detection (Sklenář et al., 1986) using the selective excitation of the H3′ or H4′, H5′, and H5″ protons. All spectra were collected with States-TPPI quadrature detection in t_1 (Marion et al., 1989). The recycle delay of 2.2 s was employed. The ^{1}H, ^{13}C, and ^{31}P signals were referenced indirectly in all 2D experiments (Wishart et al., 1995; Markley et al., 1998).

Two-dimensional^{1}H-^{13}C and ^{1}H-^{15}N HSQC spectra (Bodenhausen and Ruben, 1980) and a set of spectra recorded using spin-state-selective excitation IS[T] experiments (Žídek et al., 2001) were obtained for the samples of the ^{13}C,^{15}N-labeled DNA hairpin in isotropic and anisotropic media, respectively. A slight modification of HC[C] version of the spin-state-selective excitation experiment, utilizing 3 ms

double-excitation Q3 Gaussian cascade for simultaneous selection of C4′ and C5′, was used for the measurement of two-bond ^{1}H-^{13}C coupling constants related to torsion angle γ.

Extraction of restraints

Inter-proton distance restraints

Distance restraints for non-exchangeable protons were extracted from the 2D NOE spectrum (mixing time 200 ms) of the unlabeled DNA measured in D_2O at 30 °C. Distance restraints for exchangeable protons were obtained from the 2D NOE (mixing time 150 ms) spectrum of the unlabeled DNA in H_2O at 4 °C. The inter-proton distances were determined from 2D NOE cross-peak volumes using the isolated spin-pair approximation (ISPA) approach using average H2′-H2″ cross-peak intensity as a standard reference for the distance of 1.75 Å. The cross-peak volumes were determined by the 'summation over ellipse' method as implemented in program SPARKY 3.76 (T.D. Goddard and D.G. Kneller, SPARKY 3, University of California, San Francisco). The lower and upper bounds were set to 0.8 and 1.4 multiples of calculated distances, respectively.

Torsion angle restraints – backbone

The ^{1}H-^{31}P coupling constants were determined by the line shape simulation of the H3′-P and H4′-P crosspeaks in the 2D J-resolved ^{1}H-^{31}P correlation spectra using program CHEOPS (Macaya et al., 1992; Schultze and Feigon, unpublished program). The ^{3}J(C4′P5),^{3}J(C2′P3), ^{3}J(C4′P3) couplings were estimated from the E.COSY patterns of H4′-C4′ and H2′-C2′ crosspeaks in the 2D HSQC spectra (Schmieder et al., 1992). The ^{2}J(H4′C5′), ^{2}J(H5′C4′) and ^{2}J(H5″C4′) coupling constants were determined as the distance between peak maxima in spin-state-selective HC[C] spectra.

Unambiguous values of the torsion angles β and ε were determined by finding common solutions of the appropriately parameterized (Wijmenga and Buuren, 1998) Karplus equations for the three-bond couplings using program MULDER (Padrta, manuscript in preparation). The torsion angles β and ε were established from ^{3}J(H5′P5), ^{3}J(H5″P5), ^{3}J(C4′P5), and from ^{3}J(H3′P3), ^{3}J(C2′P3), ^{3}J(C4′P3), respectively (Wijmenga and Buuren, 1998). Torsion angle γ was semi-quantitatively determined from the ratio of ^{3}J(H4′H5′) and ^{3}J(H4′H5″) (Wijmenga and Buuren, 1998) and qualitatively corroborated from ^{4}J(H4′P). For torsion angle γ in residues A4 and A5, the results were confirmed by the values of ^{2}J(H4′C5′), ^{2}J(H5′C4′), and ^{2}J(H5″C4′).

Accurate values of the backbone torsion angles α and ζ are difficult to obtain with the current methodology. These torsion angles, as well as the glycosidic torsion angle χ, were not determined in this study and were left unrestrained in the structure calculations (see below).

Torsion angle restraints – sugars

The endocyclic torsion angle restraints were established using ^{3}J(H1′H2′), ^{3}J(H1′H2″), ^{3}J(H2′H3′), and ^{3}J(H2″H3′). These coupling constants were extracted by line shape simulation of the crosspeaks in DQF-COSY spectra using program CHEOPS (Macaya et al., 1992, Schultze and Feigon, unpublished program). In order to describe the conformational behavior of the deoxyribose rings, the concept of pseudorotation (Altona and Sundaralingam, 1972) in the two-state model approximation, as implemented in program PSEUROT, version 6.3 (de Leeuw and Altona, 1983), was used. The ranges of the endocyclic torsion angles for two sugar conformers (conformer pair) considered in the two-state model were calculated from the output of PSEUROT procedure MANY using home-written program MULDER (P. Padrta, in preparation). In short, program MULDER assumes that output of PSEUROT procedure MANY, due to the employed Newton–Raphson optimization, contains not only converged conformer pairs but also pairs that fell to the nearest local minimum. MULDER provides a graphical display of all characteristics of conformer pairs resulting from the PSEUROT procedure MANY. Using preselected criteria, the 'unacceptable' conformer pairs can be easily sorted out. The following criteria were used in our study: sugar pucker amplitude (Φ_m) was required to be within the range 30° to 45°, inclusively, and differences between the calculated and experimental couplings had to be less than ±0.5 Hz. The results of PSEUROT and MULDER analysis were incorporated in the molecular dynamic protocol as follows. In cases, where the population of a major conformer was larger than 85%, the sugar endocyclic torsions were restrained to the range of a major conformer obtained by the procedure described above. In other cases, the endocyclic torsions were not restrained.

Residual dipolar couplings

Splittings used for the determination of the one-bond ^{1}H-^{13}C and ^{1}H-^{15}N residual dipolar coupling constants, ^{1}D(CH), were measured as the distance between the maxima of the individual peaks of the doublets in the HSQC spectra. Couplings used for the determination of small RDCs were evaluated as the difference between peak maxima in pairs of spectra recorded by the IS[T] experiments. An automatic peak-picking procedure as implemented in program SPARKY was used when analyzing both HSQC and IS[T] spectra. In the case of small couplings, undesired cross-talk peaks were eliminated by the linear combination of individual IS[T] spectra described in the original report (Žídek et al., 2001). The residual dipolar couplings were calculated as the difference between values measured in isotropic and partially aligned sample. All RDC were normalized relatively to one-bond ^{1}H–^{13}C couplings (Clore et al., 1998) according to equation $D_{AB}(\text{normalized}) = D_{AB}(\gamma_C\gamma_H\langle r_{CH}^{-3}\rangle / \gamma_A\gamma_B\langle r_{AB}^{-3}\rangle)$.

Refinement protocol

The restrained MD calculations were performed with the SANDER module of program AMBER 6.0 (Pearlman et al., 1995) and the Cornell et al. (1995) force field. This force field has been extensively used in studies of nucleic acids and has shown excellent performance (Cheatham and Kollman, 2000) by providing a balanced description of base stacking and H-bonding interactions of purine and pyrimidine bases (Hobza et al., 1997). The calculations were done using *in vacuo* simulations.

Refinement protocol implemented in our laboratory consists of two steps. In the first step, the restrained MD is used to calculate the ensemble of structures from completely randomized starting conformations with the use of NOE and torsion angle restraints. In the second step, the ensemble of converged structures is subjected to a new round of restrained MD calculations with inclusion of RDCs (optimizing all five independent components of the alignment tensor) in addition to NOEs and torsion angles. These two steps are necessary since the use of RDCs in the MD calculations causes a serious problem in convergence if starting from completely random structures (Meiler et al., 2000). Each restrained MD run includes simple simulated annealing as depicted in detail in Figure 1. No cutoff was applied (by setting cutoff of 100 Å). Using the target temperature of 0 K and tight time constant of heat bath coupling, it was possible to bring the system close to 0 K. After simulated annealing, the system is relaxed and free of thermal motions, making averaging over a portion of the trajectory and/or final minimization of the structure unnecessary.

The model of the oligonucleotide of sequence d(GCGAAGC) was built with the use of the NUCGEN module of AMBER (Pearlman et al., 1995). Starting structures were prepared using 3–8 ps unrestrained MD at elevated temperatures ranging from 600 K to 1500 K, with tight time constant of heat bath coupling (0.01 ps). Variation of temperatures and lengths of simulations ensured that each of the starting structures were in different regions of conformational space.

Results

Resonance assignment

The assignment (BMRB entry 5282, see also Supplementary material) of all nonexchangeable and imino ^{1}H, ^{31}P, ^{15}N, and of all ^{13}C nuclei directly bonded to assigned protons was done using conventional methods. The assignment was carried out independently of the assignment of Hirao et al. (1994). The H5′/H5″ protons were not stereospecifically assigned due to the lack of experimental data. Interestingly, H4′ of A4 resonates as far up-field as 2 ppm, i.e., in the H2′/H2″ range, which indicates that the proton is located directly above or below an aromatic ring. A broad signal of A5 H2′/H2″ (Figure 2, H2′ is overlapped with H2″) indicates a conformational exchange of these protons. The sequential connectivities in H2′/H2″-H6/H8 region of the 200 ms NOESY can be traced from G1 to A4 and from A5 to C7. Only the weak crosspeaks of A4 H4′-A4 H8 and A4 H4′-A5 H8 connect residues A4 and A5, supplying the missing link in the H2′/H2″-H6/H8 region. This suggests that the sharp turn in the loop occurs between residues A4 and A5.

The connectivities of the G1 and G6 imino to amino protons, as well as the chemical shifts of the imino protons, are typical for the Watson–Crick pairing between residues G1-C7 and G6-C2. On the other hand, the imino proton of G3 resonates at 10.4 ppm indicating that this proton is not involved in intramolecular hydrogen bonding.

Torsion angles and sugar pucker analysis

Analysis of relevant scalar couplings revealed that torsion angles β and ε for all residues fell into the *trans* region, with exception of the ε torsion in G3 which

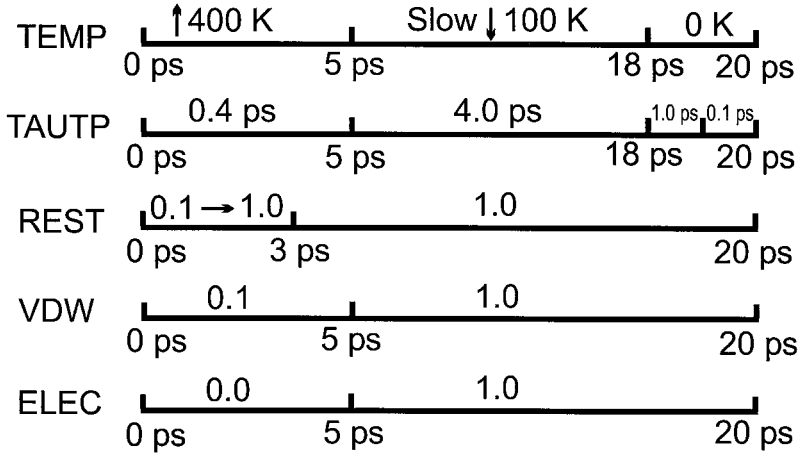

Figure 1. Scheme of simulated annealing protocol shows variation of the temperature, restraint weights, and force field terms in the course of simulation. The temperature (TEMP) of the system is rapidly increased to 400 K and maintained until the fifth picosecond. Subsequently, the system is gradually cooled down from 400 to 100 K (for 13 ps), and finally the temperature is set to 0 K for the last two picoseconds. The time constant for heat bath coupling (TAUTP) is the parameter that helps the system to follow the temperature profile. Initially, the TAUTP is adjusted to low value of 0.4 ps in order to closely follow the system heating. During cooling down, the TAUTP is increased to 4.0 ps to obtain a more 'realistic' sampling of the conformational space. At the end, the TAUTP is decreased to 1.0 ps (for the penultimate picosecond) and to 0.1 ps (for the last picosecond). The relative weight of NMR restraints (REST) increases from 0.1 to 1.0 during the first three picoseconds (in absolute numbers, the force constants for distance restraints were ramped from $k_{djst} = 3.2$ kcal mol^{-1} Å^{-2} to 32 kcal mol^{-1} Å^{-2}, force constants for torsion angle restraints were ramped from $k_{angle} = 50$ kcal mol^{-1} rad^{-2} to 500 kcal mol^{-1} rad^{-2}, and force constants for RDC restraints were ramped from $k_{RDCs} = 0.1$ kcal mol^{-1} Hz^{-2} to 1 kcal mol^{-1} Hz^{-2}). Note that for RDCs, unlike NOEs and torsion angles, only harmonic potential is currently implemented as penalty function in program AMBER. The weight of van der Waals (VDW) and eletrostatics (ELEC) terms are decreased ten-times and turned off during the first five picoseconds, respectively. This is done only when starting from a random structure in order to obtain fast convergence. The VDW and ELEC terms are turned on during the whole simulation when RDCs are included. In case of RDCs employment, a relatively short integration time-step of Newton's equations is used (0.2 fs) to prevent atom overlap, which could be enforced by RDC restraints. Otherwise, 1 fs step is used.

could not be monitored due to spectral overlap. All γ torsion angles were found in *gauche+* region, with exception of residue A5 where torsion angle γ fell into *trans* region. The deviation of γ torsion in A5 was also verified by values of two-bond scalar coupling constants (Wijmenga and Buuren, 1998): ^{2}J(H4′C5′) $= -1.2$ Hz, ^{2}J(H5′C4′) $= +1.8$ Hz, and ^{2}J(H5″C4′) $= -2.0$ Hz.

The analysis of the measured ^{3}J(HH) couplings by programs PSEUROT and MULDER showed that in all residues, where the couplings were measurable (G1, C2, G3 and A4), the sugar conformational equilibrium was shifted towards the S-conformer. The exception was the terminal C7 residue, where the PSEUROT analysis suggested a balanced S/N equilibrium with the approximate ratio 60/40. Results of the analysis are summarized in Table 1. Spectral overlap prevented such a conformational analysis of the A5 and G6 deoxyribose rings.

Residual dipolar couplings

Comparison of spectra measured in isotropic and partially aligned phases allowed to extract 13 one-bond ^{1}H–^{13}C couplings for sugars and 30 one- and two-bond couplings of the types ^{1}D(NC), ^{1}D(NH), ^{1}D(CH), ^{1}D(CC), ^{2}D(CH), ^{2}D(NH) for purine and pyrimidine bases. In the case of the bases, 4–7 couplings were obtained for residues C2, G3, A4, G6 and C7 (see Žídek et al., 2001, for the data). Low signal intensity did not allow to obtain precise couplings in G1 and A5. For the five bases, internal consistency of the data was tested before the MD calculations by exploiting the fact that only three couplings are independent in planar bodies, allowing (re)calculation of the remaining couplings from a chosen set of three RDCs (L. Žídek et al., in preparation). Out of 29 tested RDCs, 5 couplings violated the internal consistency and were discarded as unreliable. Two of the discarded RDCs were cytosine ^{1}D(C5C6) couplings determined from spin-state-selective HC[C] spectra with line-shape distortion in the indirect dimension. The reason why the other three discarded couplings

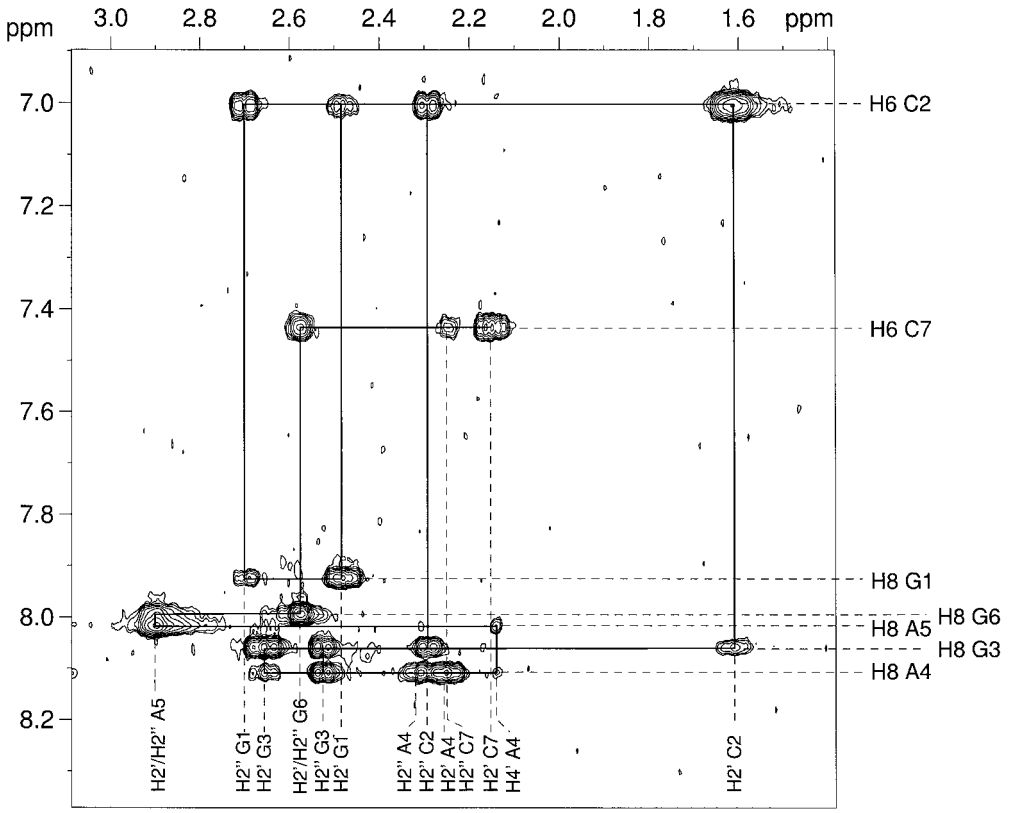

Figure 2. Region of 2D NOE spectrum showing sequential connectivities between H2′/H2″ and aromatic protons H8/H6. The spectrum was measured at 303 K with mixing time of 250 ms.

(^{1}D(N7H8) in A4, ^{1}D(N9H8) in A4, and ^{1}D(N1C1′) in G6) were not consistent with the remaining data was not obvious from an inspection of the spectra. The set of RDCs used for structure determination was thus reduced to 13 sugar and 25 base couplings which either passed the test of internal consistency or were considered reliable based on precision of their determination (i.e., large ^{1}D(CH) couplings).

Restrained molecular dynamics

In the first step of the refinement protocol, 14 prefolded structures were calculated from 20 completely randomized conformations. The converged ensemble was used in further r-MD calculations using the sets of restraints listed in Table 2. Most of these r-MD calculations were preliminary tests of influence of individual types of restraints and of their combinations (refinements employing restraint sets **1** to **7**). These refinements also provided preliminary structures that allowed assignment of the H2 protons in the bases A4 and A5. The final structures were thus obtained using all restraints including the additional H2 NOEs (set **9** in Table 2). The statistical description of the refined structures is presented in Table 3, measured torsion angles are shown in Table 4. The final structures were deposited with Protein Data Bank (PDB ID 1KR8). The representative structure (structure No. 7 in the deposited PDB entry) is presented in Figure 3.

Discussion

Sugar pucker torsion angle restraints

Structural information contained in the values of the sugar endocyclic torsion angles is often compromised by conformational flexibility of the sugar pucker. This problem was addressed by programs PSEUROT and MULDER, investigating the conformational space of the sugar rings. The first preliminary runs of r-MD tested the applicability of PSEUROT calculated data to structure refinement of the studied molecule. The experimental structure determination was first performed

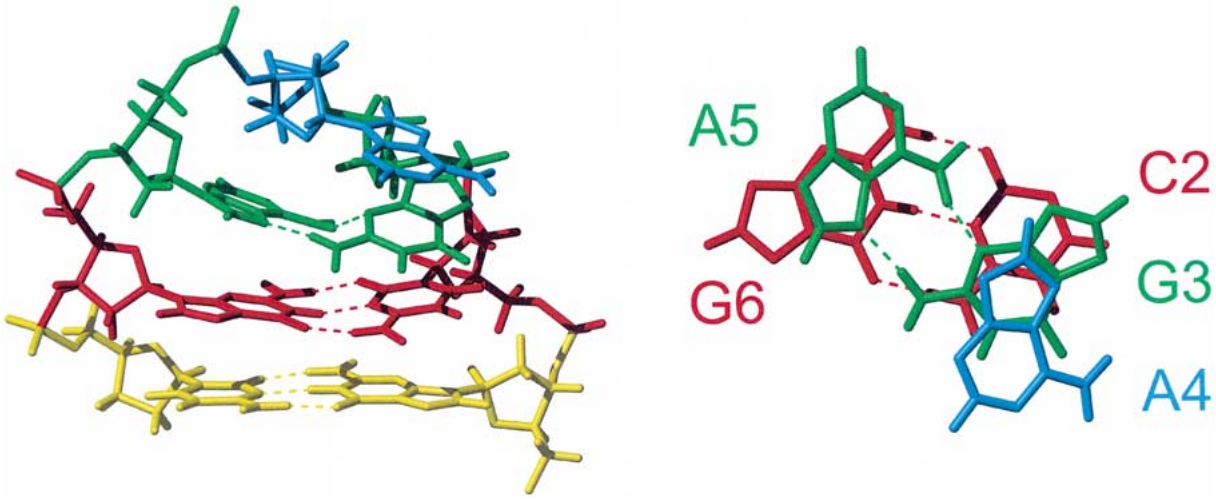

Figure 3. The representative structure (PDB ID 1KR8, model 7) calculated with final set of restraints (set **9** in Table 2) including NOEs, backbone torsions angles, and RDCs (left). Top view of residues C2 to G6 is shown on the right. Individual base pairs are color coded and inter-base hydrogen bonds are depicted by dashed lines.

Table 1. Pseudorotation phase angle P and sugar pucker amplitude Φ_m of sugar conformers 1 and 2 and the average fraction of the second conformer (X_2) of the deoxyribose rings in d(GCGAAGC). All these parameters were calculated from three-bond J_{HH} coupling constants in the two-state approximation by procedure MANY in program PSEUROT 6.3. The final intervals of pseudoroation parameters were calculated from PSEUROT output by program MULDER after applying the following filtering criteria: $|\Delta J_i| \leqslant 0.5$ Hz, $30° \leqslant \Phi_m \leqslant 45°$, major case assumed ($X_2 \geqslant 0.85$). The intervals of pseudorotation parameters are presented in the range from 0 to 360 degrees, the X_2 fraction is unitless. In major case (residues G1, C2, G3, and A4), the endocyclic torsion angle restraints were calculated from intervals of P_2 and $\Phi_{m,2}$ while neglecting the minor conformer (P_1, $\Phi_{m,1}$)

Residue	P_1	$\Phi_{m,1}$	P_2	$\Phi_{m,2}$	X_2
G1	8–95	30–42	163–197	30–42	0.85 ± 0.07
C2[a]	288–229	30–42	124–145	39–45	0.99 ± 0.05
G3[b]	130–75	30–42	157–184	36–42	0.94 ± 0.05
A4	298–236	30–42	127–155	37–42	0.97 ± 0.07
A5[c]	–	–	–	–	–
G6[c]	–	–	–	–	–
C7[d,e]	359–36	33–42	119–137	30–42	0.62 ± 0.03

[a] $|\Delta J_{H1'H2'}| \leqslant 1$ Hz.
[b] All $|\Delta J_i| \leqslant 1$Hz.
[c] Not measurable.
[d] Balanced case ($X_2 < 0.85$).
[e] All $|\Delta J_i| \leqslant 0.2$ Hz.

using restrained molecular dynamics employing only NOEs and set of torsions, which included relatively reliable values of the backbone angles β, γ, and ε (set **1** in Table 2). Starting from 14 structures prefolded by the first step of the refinement protocol (see Materials and methods), a family of 14 structures exhibiting good convergence (average pairwise RMSD = 0.39 Å) was obtained. In the second refinement, the endocyclic torsion angles of residues with the sugar conformational equilibrium shifted towards a single conformer (residues G1, C2, G3, and A4, see Table 1) were added to set **1**, creating set **2** (Table 2). For 13 converged structures, the RMSD was also 0.39 Å. The comparison of the structures resulting from sets **1** and **2** showed that inclusion of the semiquantitative sugar torsion angle restraints did not introduce any

Table 2. Sets of restraints

Set	NOEs	Torsion angles		RDCs		Total
		Backbone	Sugar	Base	Sugar	
1	99[a]	17	0	0	0	116
2	99[a]	17	20	0	0	136
3	99[a]	17	0	25[c]	0	141
4	99[a]	17	0	0	13	129
5	99[a]	17	0	25[c]	13	154
6	99[a]	17	20	9[d]	13	158
7	99[a]	17	20	25[c]	13	174
8	109[b]	17	20	0	0	146
9	109[b]	17	20	25[c]	13	184

[a] Sixty-four intraresidual NOEs + 25 interresidual NOEs + 10 distances for fixing the first two GC pairs; interresidual NOEs were counted only once for each proton pair.
[b] Ten (6 interresidual + 4 intraresidual) ambiguously assigned NOEs added based on the knowledge of the preliminary structure.
[c] All base RDCs.
[d] Only ^{1}D(CH) and ^{1}D(NH) base RDCs (=large RDCs).

artifacts. This test indicated that torsions evaluated by the semiquantitative PSEUROT-MULDER analysis can be used as restraints if the sugar dynamics favors one conformation. Violations of the experimental restraints observed for the resulting structures are summarized in Table 3.

Sugar and base RDC restraints

Further tests addressed the use of RDCs in r-MD calculations. Special attention was paid to the small RDCs, which were included as restraints in order to increase a number of restraints. Incorporation of RDCs measured in the purine and pyrimidine bases introduces several problems. Presuming that the bases are planar, which is somewhat simplified but widely accepted assumption, only three of the obtained couplings are independent. However, the three independent values are not sufficient to determine spatial orientation of a base. For that reason, the RDCs from individual bases must be combined with other RDC restraints (from sugars or other bases) in the course of structure calculation. Such calculations provide both the optimized structure and the parameters describing alignment of the whole molecule. These parameters are five independent components of the dipolar coupling tensor **D**, represented by a trace of a diagonalized matrix and by three Euler angles defining orientation of the molecular frame with respect to the principal order frame of the tensor (Losonczi et al., 1999). The trace of diagonalized matrix **D** is directly related to the distribution of measured RDCs, with D_{xx} representing the mode and D_{yy} and D_{zz} representing the extrema (Clore et al., 1998). When a large number of evenly distributed RDCs is available, the distribution histogram closely resembles the continuous powder pattern. In such cases, D_{ii} (i = x, y, z) can be estimated from the histogram. The number of RDCs utilized in this study, and in structure determination of DNA and RNA in general, is too small for reliable determination of D_{ii}. Nevertheless, the histograms were used qualitatively in this study as an independent check of results of the preliminary r-MD runs involving RDCs.

When the set of 25 base RDCs was used together with NOEs and backbone torsion angles in r-MD calculations (set **3**), a dipolar coupling tensor suggesting unrealistic spread of normalized RDCs was obtained. The calculated trace of matrix **D**, (D_{xx}, D_{yy}, D_{zz}) = (-1 ± 4, $+111 \pm 27$, -112 ± 25) Hz contrasted with the experimental data, which covered the range from -19 Hz to $+13$ Hz (see histogram in Figure 4). This may be attributed to a nearly coplanar arrangement of bases, largely stabilized by stacking interactions (see representative structure in Figure 3 and distribution pattern of base RDC vectors denoted by open circles in Figure 5). On the other hand, the replacement of 25 base RDCs with 13 sugar RDCs (set **4**) resulted in a dipolar tensor consistent with the extrema of the histogram, only slightly deviating from the mode of the distribution (D_{xx}, D_{yy}, D_{zz}) = ($+5 \pm 1$, $+15 \pm 3$, -20 ± 3) Hz. Finally, when the 13 sugar and 25 base RDCs were used together with NOEs and backbone torsion angles in r-MD calculations (set **5**), the resulting dipolar coupling tensor (D_{xx}, D_{yy}, D_{zz}) = (0 ± 2, $+29 \pm 7$, -29 ± 6) Hz agreed reasonably well with the distribution of the measured RDCs (see histogram in Figure 4). This can be attributed to the improved spatial distribution of RDCs (Figure 5). Therefore all 38 RDCs (25 from bases + 13 from sugars) were used in the final structure refinement.

Impact of small RDCs on refined structures

Impact of RDCs (especially of small RDCs, typically not used in DNA structure determination) was tested in additional preliminary r-MD runs. First, 22 large ^{1}D(CH) and ^{1}D(NH) couplings were added to set **2**, forming set **6**. In the next step, 16 small RDCs of the types ^{1}D(NC), ^{2}D(CH), and ^{2}D(NH) were added to set **6**, forming set **7**. Number of restraints used in r-MD refinements based on sets **2, 6** and **7** is summarized in Table 2. The r-MD calculations with restraint sets **6** and **7** resulted in two new families of structures, referred to as Family 6 and 7, respectively. These families were compared with Family 2, obtained by the refinement using NOEs and all torsion angles (set **2**). The comparison is presented in Figure 6.

When only large RDCs were incorporated (set **6**), the AMBER protocol with a default RDC force constant of 1 kcal mol^{-1}Hz^{-2} produced the Family 6 containing 13 structures with inferior convergence (RMSD = 0.70 Å). In order to obtain optimal convergence, the force constant had to be decreased by one order of magnitude (data not shown). However, such reduction of the force constant resulted in unrealistically high violations of the large RDC restraints, as compared to the estimated experimental error of 0.8 Hz.

The addition of the small RDCs (set **7**) allowed to increase the force constant back to its default value while obtaining good convergence of 12 refined structures (RMSD = 0.32 Å, Family 7). Violations of the

Table 3. Statistics for final ensembles of structures calculated with restraint sets **2**, **6**, **7** and **9**

Restraint set	**2**	**6**	**7**	**9**
Number of selected structures	13	13	12	14
Violations of experimental restraints				
Distance restraints				
Number of violations > 0.1 Å	0	2	1	1
Maximum violation (Å)	0.07	0.15	0.13	0.10
Torsion angle restraints				
Number of violations > 5 °	0	0	0	0
Maximum violation (°)	0.52	0.71	0.96	2.46
Dipolar restraints				
Average violation (Hz)	–	-0.05 ± 0.29	-0.01 ± 0.34	-0.01 ± 0.31
Maximum violation (Hz)	–	1.71	1.61	1.38
R.M.S. deviation from ideal geometry				
Bonds (Å)	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00
Angles (°)	2.29 ± 0.01	2.52 ± 0.06	2.78 ± 0.06	2.86 ± 0.09
Average atomic RMSD of heavy atoms (Å)				
Pairwise	0.39 ± 0.15	0.70 ± 0.25	0.32 ± 0.11	0.30 ± 0.11
From mean structure	0.27	0.49	0.21	0.21

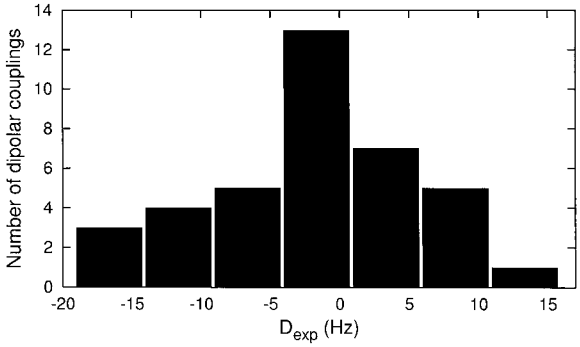

Figure 4. Histogram showing the distribution of measured normalized RDCs.

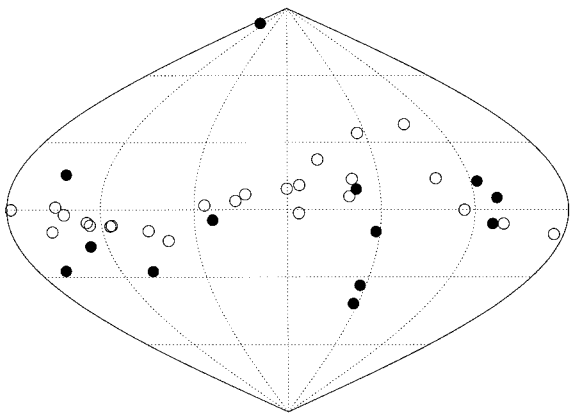

Figure 5. The spatial distribution of internuclear vectors corresponding to measured RDCs shown in Sanson–Flamsteed (sinusoidal) projection. The vectors connecting nuclei in bases and sugars are marked by open and filled circles, respectively. Note that internuclear vectors are shown as having distinct (arbitrarily chosen) orientation although they are RDC degenerate.

Table 4. Backbone torsion angles measured in structures resulting from final set of restraints. All values are presented in the range from −180° to +180°, with exception of pseudorotation phase P, which is presented in the range from 0° to 360°. The three loop-forming torsions are denoted by bold

Residue	α	β	γ	ε	ζ	χ	P	Φ_m
G1	–	–	177 ± 1	−171 ± 1	−88 ± 1	−97 ± 2	190 ± 1	34 ± 1
C2	−70 ± 1	170 ± 1	54 ± 1	175 ± 1	−88 ± 2	−103 ± 2	125 ± 1	45 ± 1
G3	−62 ± 1	177 ± 2	62 ± 2	175 ± 1	−101 ± 2	−90 ± 5	162 ± 7	38 ± 2
A4	−68 ± 1	−175 ± 3	52 ± 1	−149 ± 1	**104 ± 3**	−113 ± 4	150 ± 1	40 ± 1
A5	**80 ± 1**	−149 ± 1	**−177 ± 1**	177 ± 1	−92 ± 1	−81 ± 2	146 ± 3	44 ± 1
G6	−63 ± 1	166 ± 2	69 ± 2	171 ± 5	−93 ± 3	−113 ± 4	87 ± 11	36 ± 2
C7	−61 ± 4	172 ± 3	54 ± 3	–	–	−126 ± 3	83 ± 16	27 ± 4

RDC restraints remained in the range of 0.8 Hz. The largest difference between Families 2, 6 and 7 was observed for the base of the residue A4. Calculations using set **2**, as well as set **6**, left the orientation of A4 purine relatively undetermined (RMSD(A4) = 0.72 Å and 1.26 Å, respectively). This result can be attributed to the lack of restraints defining orientation of A4 purine. Significant improvement in the definition of A4 base was achieved by incorporating the complete set of RDCs (set **7**) into calculations (RMSD(A4) = 0.40 Å). The experimental base RDCs thus helped to define the position of the A4 base with the precision close to that of other bases. Another effect of RDC restraints (particularly those measured in sugars) was observed for sugar conformations in residues G6 and C7. While the pseudorotation phase remained in typical S-region in Family 2, O4′-endo pucker was observed in Families 6 and 7.

The comparison of Families 2, 6 and 7 showed that the small RDCs may significantly influence local conformation of the refined molecule. The effect of RDCs on (local) precision can be evaluated easily (see comparison of RMSD(A4) in the above discussion). However, accuracy of the structures obtained using RDCs is difficult to address. The fact that certain NOE restraints (those pertaining to H2 protons of bases A4 and A5) were not used in the structure refinements due to ambiguous assignment was utilized in this study. These so far ambiguously assigned crosspeaks in NOESY spectra were compared with inter-proton distances measured in structures refined with set **7**. This comparison ruled out some of the hypothetical assignments and resulted in 10 newly assigned NOEs. These NOE restraints were added to set **2**, creating set **8**, and the structure refinement was repeated. The added distance restraints improved precision of definition of base A4 (RMSD(A4) = 0.42 Å), while the overall structure converged to a structure very similar to that found using the small RDCs (set **7**). It demonstrated that RDCs did not introduce any significant artifacts into the refined structures.

Final structure refinement

The discussed tests of compatibility of available experimental restraints showed that all restraints can be combined without danger of introducing internal inconsistency. The final r-MD structure refinement was thus performed using all available restraints, i.e., set **7** extended by 10 NOE restraints assigned based on the knowledge of the preliminary structure. The final set of restraints (set **9**) defined the studied molecule relatively well, including residues A4 (underdetermined by set **2**) and A5 (underdetermined by sets **2** and **7**). All 14 starting structures converged in the r-MD calculation into a structural family exhibiting pairwise RMSD of 0.30 Å. Correlation of the experimental and back-calculated RDCs is presented in Figure 7. The final structures (PDB ID 1KR8) were similar to structures obtained in preliminary r-MD runs (RMSDs between average structures from final Family 9 and Families 2 and 7 were found to be 0.60 Å and 0.13 Å, respectively). Structural details are described in the following sections.

Sugar pucker

The analysis of the sugar pucker in final structures revealed that residues G1, C2, G3, A4, and A5 fell into S-conformation range (Table 4). Interestingly, the pseudorotation phase of residues G6 and C7 moved towards the N-region in the final structure. This is in contrast to the model by Hirao et al. (1994), who re-

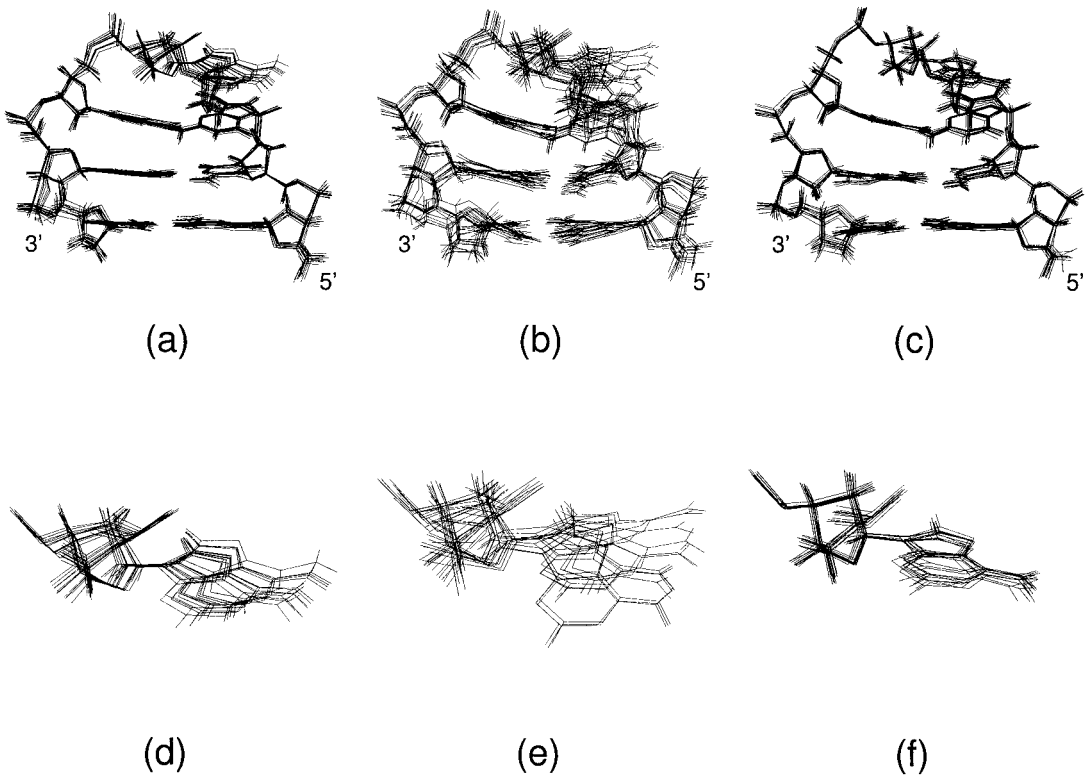

Figure 6. Preliminary hairpin structures calculated by the r-MD protocol described in Materials and Methods. (a) Family 2 consisting of 13 converged structures obtained using NOE, backbone and sugar torsion angle restraints (set **2**). (b) Family 6 consisting of 13 converged structures obtained using NOE, backbone torsion angle, and large RDC restraints (set **6**). (c) Family 7 consisting of 12 converged structures obtained using NOE, backbone torsion angle, and all RDC restraints (set **7**). The improvement in definition of the base A4 in Family 6 compared to Family 2 is demonstrated in panels (d), (e) and (f), respectively.

ported that all sugars of the d(GCGAAGC) hairpin are in C2′-endo conformation. The observed pseudorotation phase (approximately 80°, see Table 4) can be interpreted as an average of C2′-endo and C3′-endo conformers, i.e., a conformation close to O4′-endo. This interpretation was supported by MD-independent PSEUROT analysis (see Results and Table 1) for residue C7 (no data were available for residue G6 due to spectral overlap). Note that the change of the pseudorotation phase in G6 and C7 was not observed in Family 2, calculated without RDCs. Similar effect of RDCs was reported for DNA dodecamer d(CGCGAATTCGCG) by Tjandra et al. (2000) where the sugar RDCs caused a shift in sugar pucker towards the N-region in several residues. Tjandra et al. explained this change of the pseudorotation phase by dynamic equilibrium between C2′-endo and C3′-endo conformations, which was also supported by independent analyses of homo- and hetero-nuclear J-couplings (Sanderson et al., 1983; Rinkel and Altona, 1987).

Backbone torsion angles

Analysis of backbone torsion angles in the final structures (Table 4) showed that all torsion angles are in their regular regions ($\alpha(g^-)$, $\beta(t)$, $\gamma(g^+)$, $\varepsilon(t)$, $\zeta(g^-)$, $\chi(anti)$), with exception of torsion angles $\zeta4(g^+)$, $\alpha5(g^+)$ and $\gamma5(t)$. Irregular backbone conformation between the residues A4 and A5 was also revealed by missing crosspeaks in the connectivity network of 2D NOE spectra. Furthermore, the torsion angles ε4 and β5 are shifted slightly towards the *gauche*− region. This is only in partial agreement with the findings of Hirao et al. (1994) who also reported γ5 in the *trans* region but proposed different conformations for ζ4, α5, and even β5: $\zeta4(t)$, $\alpha5(g^-)$, $\beta5(g^+)$, $\gamma5(t)$. The *trans* conformation of β5 presented here is supported by the high value of ^{3}J(C4′P5) $\approx$ 10 Hz (data not shown). The unusual values of torsion angles ζ4, α5 and γ5 are consistent with the assumption that formation of a miniloop is conditioned by a deformation of at least three torsion angles from their regular domains (Hi-

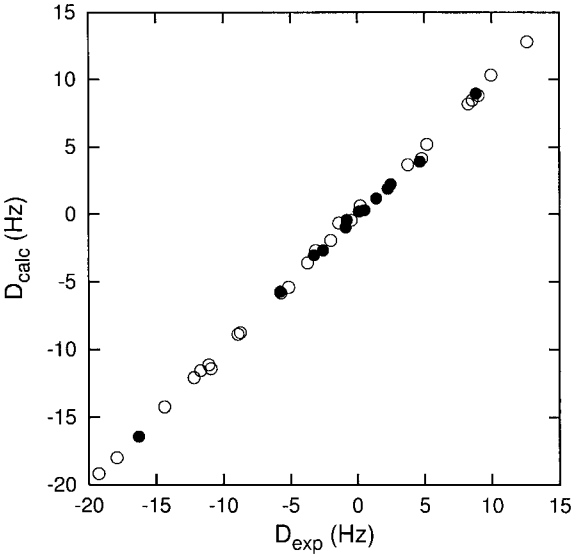

Figure 7. The correlation of calculated versus experimental values of normalized RDCs. The open circles denote the base RDCs, the full circles denote the sugar RDCs. The calculated values are averaged over all structures from the final family deposited with PDB (ID 1KR8). Average violation was -0.01 ± 0.31 Hz, experimental error of the normalized RDCs was estimated to be less than 0.8 Hz for large one-bond ^{13}C-^{1}H RDCs and less than 6 Hz for small RDCs, with the exception of N1-C1′ (7.5 Hz).

rao et al., 1994; Yoshizawa et al., 1997; Lescrinier et al., 1999; Chou et al., 2000). For similar loops, various authors reported the following unusual values: Lescrinier et al., the GTA loop in d(GCGTAGC) (Lescrinier et al., 1999): $\zeta4(t)$, $\alpha5(t)$, $\beta5(t)$, $\gamma5(t)$, Chou et al., the AAA loop in d(GTACAAAGTAC) (Chou et al., 1996, 1997): $\zeta6(g^+)$, $\alpha7(g^+)$, $\beta7(t)$, $\gamma7(t)$.

Bases

Chemical shifts of the imino protons of residues G1 and G6 (Table S1) indicated the presence of hydrogen bonds between G1-C7 and C2-G6. Assuming that these residues form two canonical Watson-Crick base pairs, the relevant hydrogen bonds were artificially restrained in the structure calculation. In addition to the backbone conformational changes discussed above, formation of a miniloop (residues G3, A4, and A5) requires unusual orientation of the loop bases. Bases G3 and A5 are arranged in the sheared G-A mismatch pair and the unpaired base A4 is stacked above G3. Despite the fact that the G3-A5 pair was not fixed by any artificial restraints during r-MD, whereas the first two G-C pairs were fixed by setting 10 distances corresponding to three conventional G-C pair H-bonds, the geometry corresponding to the sheared G-A pair was established. The sheared G-A mismatch is stabilized by two hydrogen bonds NH_2(G3)-N7(A5) and NH_2(A5)-N3(G3).

The presence of the sheared G-A mismatch was also indicated by the following evidence. The chemical shift of the H1 imino proton of G3 (10.4 ppm) fell in the typical range for protons not involved in hydrogen bonding – this excluded the presence of a classic G-A mismatch. A cross peak between H1′ of G3 and hydrogen of amino group from A5 was observed in the 2D NOE spectrum with the intensity corresponding to the distance of 2.5 Å measured in the final structure. In addition, there was a small cross peak between the imino proton of G3 and hydrogen of amino group of A5 with the intensity indicating the distance above 4 Å. This definitively excluded classic G-A mismatch architecture.

Hairpin stability

The structural details described for the GAA loop also offer explanation of the extraordinary stability of the d(GCGAAGC) hairpin. Although sequences of general formula GCGNAGC are known to form the duplex structure at certain conditions (Yoshizawa et al., 1997; Hirao et al., 1994; Lescrinier et al., 1999; Zhu et al., 1995), only negligible amount of the duplex was observed in this study. The exceptional stability of the d(GCGAAGC) hairpin (as represented by high melting temperature and high resistance against nucleases) is expected to originate from a complex interplay between the aromatic base stacking and hydrogen bonding interactions.

The observed thermodynamic stability reported by Yoshizawa et al. (1997), however, does not imply structural rigidity of the studied hairpin. On the contrary, sugar pucker wobbling in C7, line broadening in A5, and unrestrained MD simulations showing correlated conformational changes of backbone torsions $\gamma5$, $\zeta4$ and $\alpha5$ (unpublished results) suggest certain flexibility of the molecule. The presence of flexibility in the folded form may increase the stability as this is more favourable entropically.

Conclusions

The structure of the d(GCGAAGC) hairpin was determined using NOE-derived distances, torsion angles, and RDCs as experimental restraints. Although the

studied molecule of small size and compact fold does not represent a typical case requiring RDC restraints for precise global structure determination, it was demonstrated that the incorporation of small RDCs significantly improved local structure determination of regions undefined by the conventional restraints. It should be emphasized that the RDC restraints in individual bases had to be used in combination with other restraints because only three spatially independent RDCs exist if the nucleic acid bases are assumed to be planar. The final structures clearly show typical features of the miniloop architecture such as extensive hydrogen bonding and stacking of the bases. The experimental data also revealed flexibility in certain regions of the molecule (base A5 and sugars of G6 and C7). The structure has been discussed in context of the extraordinary stability of the d(GCGAAGC) hairpin.

Note added in proof

Quite recently, a flat-bottom potential for RDC restraints has been implemented in AMBER 7 (courtesy by David Case). We have recalculated the structure of d(GCGAAGC) hairpin with the new implementation, using restraint set **9** and including the estimated experimental errors of residual dipolar couplings (see caption to Figure 7). The resulting structures are very similar to those obtained with the original AMBER implementation (Family 9): RMSD between mean structures of the original Family 9 and the new ensemble is 0.31 Å. The new structures show better precision (average pairwise RMSD $= 0.24 \pm 0.10$ Å), improved covalent geometry of purine and pyrimidine bases, and more realistic correlation of calculated *versus* experimental values of normalized RDCs (Figure S1, available in Supplementary material). The maximum difference between the experimental and back-calculated RDC is 6.76 Hz, the average RDC violation is -0.28 ± 2.15 Hz.

Supplementary material

Tables S1 and S2 with ^{1}H and ^{31}P, and ^{13}C chemical shifts of d(GCGAAGC) hairpin, respectively as well as Figure S1 showing the correlation of calculated *versus* experimental values of normalized RDCs using the flat-bottom potential incorporated in AMBER 7 are available as Supplementary material on request from the authors.

Acknowledgements

This work was supported by grant LN00A016 from the Ministry of Education of the Czech Republic and an EMBO fellowship to R.Š. The authors also thank the Academic Supercomputer Center in Brno and Prague for providing with access to computer facilities, J. Feigon for providing the unlabeled sample, and K. Kubíček and M. Wimmerová for preparation of Pf1 phage.

References

Altona, C. and Sundaralingam, M. (1972) *J. Am. Chem. Soc.*, **15**, 8205–8212.

Arai, K., Low, R., Kobori, J., Shlomai, J. and Kornberg, A. (1981) *J. Biol. Chem.*, **256**, 5273–5280.

Bax, A. and Davis, D.G. (1985) *J. Magn. Reson.*, **65**, 355–360.

Bodenhausen, G. and Ruben, D. (1980) *Chem. Phys. Lett.*, **69**, 185–189.

Braunschweiler, L. and Ernst, R.R. (1983) *J. Magn. Reson.*, **53**, 521–528.

Clore, G.M., Gronenborn, A.M. and Bax, A. (1998) *J. Magn. Reson.*, **133**, 216–221.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz Jr., K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.A. (1995) *J. Am. Chem. Soc.*, **117**, 5179–5197.

Cowing, D.W., Bardwell, J.C.A., Craig, E.A., Woolford, C., Hendrix, R.W. and Gross, C.A. (1985) *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 2679–2683.

de Leeuw, F.A. and Altona, C. (1983) *J. Comput. Chem.*, **4**, 428–438.

Elias, P. and Lehman, I.R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2959–2963.

Hansen, M.R., Mueller, L. and Pardi, A. (1998) *Nat. Struct. Biol.*, **5**, 1065–1074.

Hirao, I., Kawai, G., Yoshizawa, S., Nishimura, Y., Ishido, Y., Watanabe, K. and Miura, K.-i. (1994) *Nucl. Acids Res.*, **22**, 576–582.

Hirao, I., Nishimura, Y., Tagawa, Y.-i., Watanabe, K. and Miura, K.-i. (1992) *Nucl. Acids Res.*, **20**, 3891–3896.

Hirao, I., Yoshizawa, S. and Miura, K. (1993) *FEBS Lett.*, **321**, 169–172.

Hobza, P., Kabeláč, M., Šponer, J., Mejzlík, P. and Vondrášek, J. (1997) *J. Comput. Chem.*, **18**, 1136–1150.

Cheatham, T.E. and Kollman, P.A. (2000) *Annu. Rev. Phys. Chem.*, **51**, 435–471.

Chou, S.H., Tseng, Y.Y. and Chu, B.Y. (2000) *J. Biomol. NMR*, **17**, 1–16.

Chou, S.H., Zhu, L.M., Gao, Z.M., Cheng, J.W. and Reid, B.R. (1996) *J. Mol. Biol.*, **264**, 981–1001.

Chou, S.-H., Zhu, L. and Reid, B.R. (1997) *J. Mol. Biol.*, **267**, 1055–1067.

Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.

Jolles, B., Refregiers, M. and Laigle, A. (1997) *Nucl. Acids Res.*, **25**, 4608–4613.

Lescrinier, E., Sheng, S., Schraml, J., Busson, R. and Herdewijn, P. (1999) *Nucleos. Nucleot.*, **18**, 2721–2744.

Losonczi, J.A., Andrec, M., Fischer, M.W.F. and Prestegard, J.H. (1999) *J. Magn. Reson.*, **138**, 334–342.

Macaya, R.F., Schultze, P. and Feigon, J. (1992) *J. Am. Chem. Soc.*, **114**, 781–783.

MacDonald, D., Herbert, K., Zhang, X., Polgruto, T. and Lu, P. (2001) *J. Mol. Biol.*, **306**, 1081–1098.

Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.

Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright, P.E. and Wuthrich, K. (1998) *Eur. J. Biochem.*, **256**, 1–15.

Meiler, J., Blomberg, N., Nilges, M. and Griesinger, C. (2000) *J. Biomol. NMR*, **16**, 245–252.

Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham III, T.E., DeBolt, S., Ferguson, D.M., Seibel, G.L. and Kollman, P.A. (1995) *Comp. Phys. Commun.*, **91**, 1–41.

Piantini, U., Sørensen, O.W. and Ernst, R.R. (1982) *J. Am. Chem. Soc.*, **104**, 6800–6801.

Piotto, M., Saudek, V. and Sklenář, V. (1992) *J. Biomol. NMR*, **2**, 661–665.

Prestegard, J.H. (1998) *Nat. Struct. Biol.*, **NMR Supplement**, 517–522.

Rinkel, L.J. and Altona, C.J. (1987) *J. Biomol. Struct. Dyn.*, **4**, 621–649.

Sanderson, M.R., Mellema, J.-R., van der Marel, G.A., Wille, G., van Boom, J.H. and Altona, C.J. (1983) *Nucl. Acids Res.*, **11**, 3333–3362.

Schmieder, P., Ippel, J.H., van den Else, H., van der Marel, G.A., van Boom, J.H., Altona, C.J. and Kessler, H. (1992) *Nucl. Acids Res.*, **20**, 4747–4751.

Sklenář, V., Miyashiro, H., Zon, G., Miles, H.T. and Bax, A. (1986) *FEBS Lett.*, **208**, 94–98.

Suen, I.-S., Rhodes, J.N., Christy, M., McEwen, B., Gray, D.M. and Mitas, M. (1999) *Biochim. Biophys. Acta*, **1444**, 14–24.

Tjandra, N. and Bax, A. (1997) *Science*, **278**, 1111–1114.

Tjandra, N., Tate, S., Ono, A., Kainosho, M. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 6190–6200.

Tsui, V., Zhu, L.M., Huang, T.H., Wright, P.E. and Case, D.A. (2000) *J. Biomol. NMR*, **16**, 9–21.

Vermeulen, A., Zhou, H.J. and Pardi, A. (2000) *J. Am. Chem. Soc.*, **122**, 9638–9647.

Warren, J.J. and Moore, P.B. (2001) *J. Biomol. NMR*, **20**, 311–323.

Wijmenga, S. and Buuren, B. (1998) *Prog. NMR Spectrosc.*, **32**, 287–387.

Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) *J. Biomol. NMR*, **6**, 135–140.

Yoshizawa, S., Kawai, G., Watanabe, K., Miura, K.-i. and Hirao, I. (1997) *Biochemistry*, **36**, 4761–4767.

Yoshizawa, S., Ueda, T., Ishido, Y., Miura, K., Watanabe, K. and Hirao, I. (1994) *Nucl. Acids Res.*, **22**, 2217–2221.

Zhu, L.M., Chou, S.H., Xu, J.D. and Reid, B.R. (1995) *Nat. Struct. Biol.*, **2**, 1012–1017.

Žídek, L., Wu, H., Feigon, J. and Sklenář, V. (2001) *J. Biomol. NMR*, **21**, 153–160.