5740–5751 *Nucleic Acids Research, 2006, Vol. 34, No. 20* *Published online 13 October 2006*
doi:10.1093/nar/gkl617

Solution structure of a purine rich hexaloop hairpin belonging to PGY/MDR1 mRNA and targeted by antisense oligonucleotides

Flore Joli[1,2,*], Nadia Bouchemal[1], Alain Laigle[1], Brigitte Hartmann[2,*] and Edith Hantz[1]

[1]Laboratoire BioMoCeTi, CNRS UMR 7033, UFR SMBH, Université Paris 13, 74 rue Marcel Cachin, 93017 Bobigny cedex, France and [2]Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

Received May 3, 2006; Revised August 6, 2006; Accepted August 7, 2006

ABSTRACT

A preferential target of antisense oligonucleotides directed against human PGY/MDR1 mRNA is a hairpin containing a stem with a G•U wobble pair, capped by the purine-rich $^{5'}$r(GGGAUG)$^{3'}$ hexaloop. This hairpin is studied by multidimensional NMR and restrained molecular dynamics, with special emphasis on the conformation of south sugars and non-standard phosphate linkages evidenced in both the stem and the loop. The hairpin is found to be highly structured. The G•U wobble pair, a strong counterion binding site, displays structural particularities that are characteristic of this type of mismatch. The upper part of the stem undergoes distortions that optimize its interactions with the beginning of the loop. The loop adopts a new fold in which the single-stranded GGGA purine tract is structured in A-like conformation stacked in continuity of the stem and displays an extensive hydrogen bonding surface for recognition. The remarkable hairpin stability results from classical inter- and intrastrand interactions reinforced by numerous hydrogen bonds involving unusual backbone conformations and ribose 2′-hydroxyl groups. Overall, this work emphasizes numerous features that account for the well-ordered structure of the whole hairpin and highlights the loop properties that facilitate interaction with antisense oligonucleotides.

INTRODUCTION

The transmembrane P-glycoprotein (P-gp), encoded by the *MDR1* gene, acts as an energy-dependent drug-efflux pump and is involved in cellular drug excretion (1). The over-expression of P-gp is directly linked to multidrug resistance (2), a major problem in cancer therapeutics. Antisense oligonucleotides directed against the *PGY1*/*MDR1* human mRNA inhibit the expression of P-gp (3–7). The most efficient antisense oligomers are all complementary to GGGAUG RNA sequences containing the AUG initiation site and belonging to a hairpin hexaloop (5–8). The GGGA motif seems to be the preferential target for mRNA inhibition, as it is recognized by 48% of the most potent antisense oligonucleotides reported in the literature (9).

The main purpose of the present study is to determine the structure of the antisense target RNA hairpin in order to highlight the properties that allow efficient hybridization. Also, this hairpin presents some features that are of general interest from a purely structural point of view. The stem–loop of interest is constituted by a 6 bp stem comprising a G•U mismatch and the G-rich hexaloop $^{5'}$r(GGGAUG)$^{3'}$. G•U pairs are highly conserved in functional RNA and have been shown to play essential roles in a wide range of processes (10). The geometrical and conformational properties of G•U wobble pairs have been described (11), but essentially on the basis of crystal structures. Three solution studies of a single G•U wobble pair embedded within RNA helices have given detailed structural information about this mismatch (12–14). The complete structural investigation undertaken in the present study should help to evaluate the common characteristics of this type of mismatch. Concerning the loop region, in contrast to the well-studied purine-rich tetraloops, little is known about the solution properties of purine-rich hexaloops. Perusal of the recent website SCOR (http://scor.lbl.gov), revealed that among the known hexaloop solution structures only two contain five purines, i.e. UGAAAG [1R2P; (15)] and GUAAAA [1BVJ; (16)]. These loops include A-rich parts on their 3′ sides and, therefore, we anticipated that their structures would differ strongly from the GGGAUG one in which a G-rich region is located at the 5′ side. Thus, determining the structure of GGGAUG loop should provide additional insight into the folds encountered in purine-rich loops.

*To whom correspondence should be addressed. Tel: +33 158415167; Fax: +33 158415026; Email: fjoli@smbh.univ-paris13.fr
*Correspondence may also be addressed to Brigitte Hartmann. Tel: +33 158415169; Fax: +33 158415026; Email: brigitte.hartmann@ibpc.fr

In the present work, study of the solution conformation of the GGGAUG loop closed by a stem containing a G•U pair was undertaken by NMR and molecular modeling. Long trajectories of molecular dynamics in explicit solvent should further the investigation of the hairpin flexibility. We find that the whole hairpin is highly ordered thanks to numerous standard and non-standard interactions in both the stem and the loop. Finally, the hexaloop adopts a new fold in which the GGGA intrinsic properties account for the hairpin ability to form duplex with oligonucleotides.

MATERIALS AND METHODS

RNA sequence and sample preparation

The RNA sequence and the residue numbering are shown Figure 1. Unlabeled RNA strand was synthesized by RNATEC (Leuven, Belgium) and 98% uniform $^{13}C/^{15}N$-labeled RNA strand was synthesized by SILANTES (München, Germany). The two samples were purified by HPLC. The samples were dissolved in aqueous solution (10 mM sodium phosphate buffer, 0.01 mM EDTA, pH 6.9), heated to 80°C and slowly cooled to room temperature. No further salt was added to avoid the formation of duplex RNA at higher salt concentrations. Final concentrations were 1 mM for the unlabeled sample and 0.8 mM for the $^{13}C/^{15}N$ labeled sample.

UV measurements

UV thermal denaturation was carried out at 260 nm for RNA concentrations from 10^{-5} to 10^{-4} M with a KONTRON UVIKON Spectrophotometer with a heating rate of 0.2°C/min. Calculating the first derivative of the absorption versus temperature profile after correction of the background absorbance owing to the buffer revealed a single transition at 58°C irrespective of the sample concentration. This ensured that a single monomer species, i.e. the hairpin structure, exists at 25 and 35°C, the temperatures of NMR experiments.

NMR measurements

NMR experiments were performed using a 500 MHz VARIAN Unity INOVA spectrometer equipped with a 5 mm gradient indirect detection probe or a 5 mm triple (^1H, ^{13}C, ^{15}N) detection probe. The spectra were referenced to internal TSP (^1H), methanol (^{13}C), ammonium chloride (^{15}N) and phosphoric acid (^{31}P).

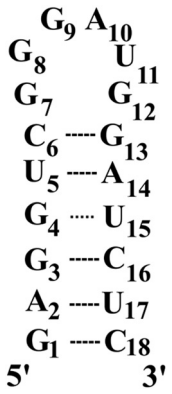

Figure 1. Sequence and secondary structure of the 18mer RNA hairpin.

NMR data were processed on a Silicon Graphics workstation using the FELIX (MSI program San Diego, CA, USA), NMRpipe (http://spin.niddk.nih.gov/bax/software/NMRPipe/) and SPARKY (http://www.cgl.ucsf.edu/home/sparky/) softwares.

A set of NOESY experiments in H_2O (90% H_2O, 10% ^2H_2O) were performed at either 1 or 25°C using a jump and return water suppression scheme (17) and 150, 250 and 300 ms mixing times. Standard 2D experiments in 99.9% ^2H_2O were used for the ^1H chemical shift assignments. NOESY (250 and 400 ms mixing times), TOCSY (60 ms mixing time) and DQF-COSY with a ^{31}P decoupling pulse experiments were recorded at temperatures of 25 and 35°C, using very low power presaturation on the residual HDO resonance. Typical two-dimensional spectra were acquired with 1024 complex points in t2 and 512 increments in t1, employing the TPPI scheme (18). The recycle delay was 3 s. Data were zero filled to 4K × 2K points in both dimensions and apodized with skewed phase-shifted sine-bell functions. In order to assign the ^{31}P chemical shifts (δP) through phosphorus coupling with the H3′ sugar proton, the following ^{31}P-^1H experiments were performed at the same temperatures with spectral widths of 5000 Hz (^1H) and 1100 Hz (^{31}P): a proton detected heteronuclear COSY (19), an HSQC and an HSQC-TOCSY (20,21). From this dataset, all the δP could be assigned.

A uniform $^{13}C/^{15}N$-labeled RNA sample in H_2O was used to achieve the assignment with the BioPack package of Lukavsky and Puglisi (22). ^1H–^{15}N HSQC spectrum was used to confirm imino proton assignments. The hydrogen bonding patterns were observed directly from the heteronuclear HNN-COSY experiment (23). Two and three-dimensional HCCH-COSY (24) and ^1H-^{13}C HSQC-NOESY experiments were acquired in order to complete the spin system assignment of sugar protons, using spectral widths of 2500 Hz (^1H), 2500 Hz (^1H) and 6900 Hz (^{13}C). The data were processed to a final size of 1 K × 256 × 128 points. These data resulted in the identification of the sugar conformations. Intranucleotide correlations of sugar and base resonances were achieved using a triple-resonance multiple-quantum TROSY-HCN experiment (25). G- or A-HNC-TOCSY-CH (26) and C- or U-HNCCCH (27,28) were performed to correlate imino or amino protons to their corresponding base H8/H6/H2 protons. For all two- and three-dimensional heteronuclear experiments, ^{13}C or ^{15}N WURST or GARP broadband decoupling was applied during the acquisition period.

Inter-proton distances between non-exchangeable protons were obtained from the intensities of their cross-peaks in NOESY spectra at 25°C with 75, 150 and 250 ms mixing times and were compared with the build-up rate of H5-H6 NOE cross-peaks (2.45 Å) of U5, U15 and U17, using the distance extrapolation method (29). A total of 129 distances were extracted from the spectra, and expressed with an uncertainty of ±10% ($d_{measured}$ ± 10%) in order to take into account experimental errors. Furthermore, 111 distances derived from visual inspection were added to this set of quantitative distances. These quantitative distances corresponded to cross-peaks that were (i) strong but overlapping (d < 5 Å);

(ii) too weak to be precisely measured (5 Å < d < 6 Å) and (iii) absent in well-resolved regions (d > 6 Å).

Chemical shift data are deposited in the BioMagResBank (BMRB-7090).

Simulation methodologies

Molecular mechanisms. The JUMNA, Junction Minimization of Nucleic Acids, (30) approach models nucleic acids using helical coordinates to position each nucleotide and internal coordinates (restricted to torsion angles and a limited number of valence angles) to treat changes in nucleotide conformation. This approach is very useful in the present study to rapidly generate and test numerous hairpin conformations. Calculations were performed using the Parm99 force field (31) modified for AMBER (32). Solvent damping of electrostatic interactions was included using a sigmoidal distance-dependent dielectric function defined by a slope of 0.16 and a plateau value of 80 (33). Counterion effects were taken into account by reducing the net charge on each phosphate group to -0.5 e^-. The use of an implicit solvent is not as accurate as an explicit one, however, previous studies (34–36) have shown that valuable results are obtained by such methodology.

Molecular dynamics. The simulations were performed using AMBER 7.0 program (32) and the Parm99 force field (31). The molecular dynamics were carried out on the hairpin neutralized with 17 Na^+ counterions (one Na^+ per phosphate group) and explicitly solvated by a 12 Å water shell in all directions (5641 TIP3P water molecules). After 2250 cycles of energy minimization, the minimized system was heated to 300°K, rescaling the velocities as necessary, and coupling to a heat bath using the Berendsen algorithm (37). During each of these phases, harmonic restraints were imposed on the atomic positions of the oligomers and then slowly relaxed over several periods of 50 ps until an unrestrained system was achieved. The simulations were then performed at constant temperature and pressure (NTP) using the Berendsen algorithm. Bond lengths involving hydrogen atoms were constrained using the SHAKE algorithm (38), which enabled an integration time step of 2 fs. Long-range electrostatic interactions were treated using the Particle Mesh Ewald (PME) approach (39,40) with a 9 Å direct space cutoff, a direct sum tolerance criteria of 10^{-5}, and a reciprocal space charge grid spacing of roughly 1 Å. During the production phase, translations and rotations of the stem–loop oligomer were removed every 100 steps.

Two MD were performed under different sets of constraints extracted from the NMR experiments and described in the 'Results' section. The distances restraints, comprising those of hydrogen bonds, were applied via parabolic potentials with a force constant of 10 Kcal/Mol/Å^2, around a central flat-bottomed well. The flat-bottomed well covered the experimental range of the distances, including experimental errors (d_{measured} ± 10%). The torsion angles, when they were taken into account, were constrained by a force constant of 50 Kcal/mol in south (C2′-endo; P = 100 to 240°) or north (C3′-endo; P = −60 to 100°) for the sugars puckers and within the g^+ (60° ± 40°), *trans* (180° ± 40°) and g^- (300° ± 40°) regions for the backbone. Note that torsion angles were defined as follows ε: $C_{4'}$–$C_{3'}$–$O_{3'}$–P; ζ: $C_{3'}$–$O_{3'}$–P–$O_{5'}$; α: $O_{3'}$–P–$O_{5'}$–$C_{5'}$; β: P–$O_{5'}$–$C_{5'}$–$C_{4'}$; γ: $O_{5'}$–$C_{5'}$–$C_{4'}$–$C_{3'}$.

Generation of initial structures

The initial model was built and minimized with the internal coordinate program JUMNA (30). We started from a first strand composed of the $^{5'}$r(GAGGUCGGGAUG)$^{3'}$ sequence and a second strand containing the $^{5'}$r(GAUCUC)$^{3'}$ sequence. The six bases of the two strands were paired in an A-form RNA double helix while the loop part was in an extended conformation. Constrained distances were applied progressively by steps of 0.2 Å between 3′-G12 of the first strand and G13-5′ of the second strand (initially $d \sim 20$ Å) until the single strand part is folded into a loop ($d \sim 2$ Å). Starting from this folded hairpin, a total of 52 loop structures was then generated by randomization of the base positions, in order to reproduce different stacking features. The cross-RMSD (root mean square deviation) between these loop structures were from 2 to 5 Å, ensuring that they covered a large range of conformations.

Coordinates

Structures have been submitted to the Protein Data Bank (accession number 2GVO).

RESULTS

NMR spectroscopy

Hydrogen bonds. The NOE cross-peaks related to hydrogen bonds in the stem are clearly observed from 1°C (Figure 2) to 35°C, demonstrating marked stability for the corresponding 6 bp of the duplex region. These data indicate that all the bases adopt standard Watson–Crick (WC) pairs, except for the G•U mismatch whose hydrogen binding gives rise to the characteristic signature of a wobble conformation. The imino protons belonging to G4 (11.53 p.p.m.) and U15 (11.91 p.p.m.) are upfield shifted with respect to those of the WC base pairs (between 12 and 15 p.p.m.). These assignments were confirmed by the imino cross-peaks of G3 and G4 and by the ^{15}N-HSQC spectrum optimized for imino resonances, in which the guanine and uracil imino atoms are well separated, appearing at 146 p.p.m. for N1 of G3 and at 158 p.p.m. for N3 of U15 (41). In addition to these characteristic chemical shifts, strong NOE cross-peaks are observed between H1 of G4 and H3 of U15 (Figure 2). The absence of cross-peak between H1 of G4 (11.53 p.p.m.) and H5 of U15 (5.45 p.p.m.) excludes a G•U bifurcated conformation. Finally, NOE cross-peaks related to hydrogen bonds were not detected for the spins of the hexaloop region.

Intra and internucleotide distances. No unusual features were observed for the intranucleotide cross-peaks. The weak intensity of intranucleotide H6/H8-H1′ NOE cross-peaks indicated that the χ torsion angle of all the bases adopted the anti conformation, including those of the loop nucleotides.

The stem region, excluding the terminal bases that exhibited fraying, was defined by 26 quantitative and 21 qualitative sequential distances. The H6/H8$_{i+1}$-H1′$_i$ distances (>4 Å) were very close to the A-form typical ones (∼4.5 Å) (42). Nevertheless, the remarkably short H2′$_i$-H6/H8$_{i+1}$ (2.5 Å)

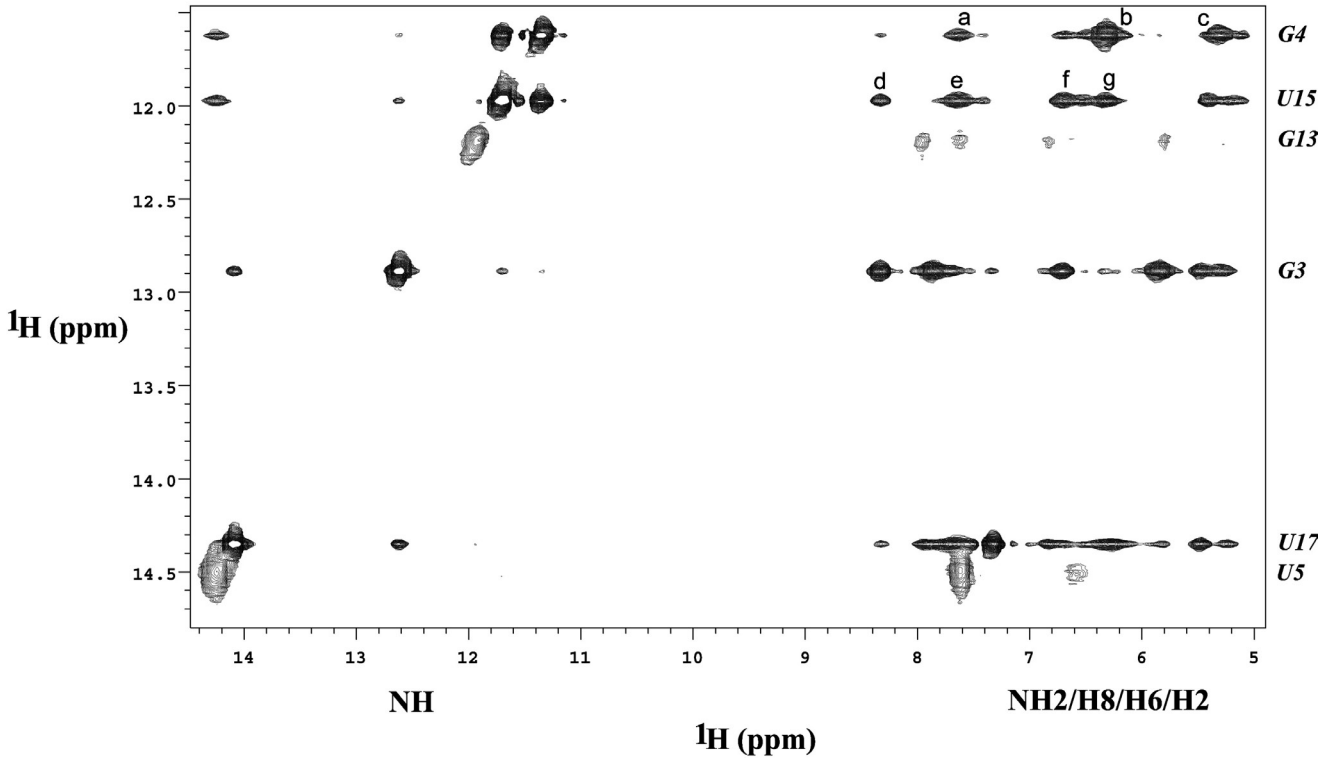

Figure 2. Section of the 2D-watergate-NOESY spectrum in H_2O at 1°C (mixing time 200 ms) containing NOE connectivities involving G4•U15 wobble resonances. Assignments are shown on the right of the spectrum and letters correspond to a, H2 (A14); b, NH_2 (G4); c, NH_2 (G4); d, NH_2 (C16); e, H2 (A14); f, NH_2 (C16); g, NH_2 (G4).

and H6/H8$_i$-H6/H8$_{i+1}$ (4.2 Å) distances of the G3pG4• U15pC16 nucleotides suggested that distortions were occurring in the vicinity of the G4•U15 wobble pair.

Concerning the stem–loop junctions and the loop (bases 6–13), the presence of 37 detectable sequential distances involving base protons strongly suggests the absence of extruded bases. All the H1′, H2′ and H3′ sugar protons are correlated to the H6/H8 of the next base, the only exception being for H2′ of G12 that does not display cross-peaks with H8 of G13 in the NOESY spectra. Similarly, the (H6/H8)$_i$ and (H6/H8)$_{i+1}$ protons give measurable cross-peaks except for A10pU11 and G12pG13. Thus, two stacking segments seem to be indicated by the sequential NOE cross-peaks, the first one composed of nucleotides C6–A10 and the second one of U11 and G12. All the G12 protons are located at too great a distance (>5 Å) from the H8 of the G13 stem base for efficient stacking.

Sugars and backbone angles. Owing to numerous overlaps in the 2D spectra, 3D HCCH-COSY experiments were necessary to obtain the complete assignment of all protons resonances of the sugar moieties. The assignment of the A14 sugar protons is illustrated Figure 3. Ribose conformations were determined from $^3J_{H1'-H2'}$ coupling constants in the DQF-COSY spectrum (Figure 4). The stem sugars mainly populate the north (C3′-endo) conformation ($^3J_{H1'-H2'}$ < 3 Hz), with the exception of this of U5 at the 3′ side of the G•U wobble pair that exhibits a strong preference for the south (C2′-endo) domain ($^3J_{H1'-H2'}$ > 7 Hz). The inverse situation is observed in the loop where five sugars predominantly adopt a south conformation, but the G7 sugar clearly prefers the north conformation.

The ^{31}P resonances cover a large chemical shift range, from −0.9 to 0.28 p.p.m. (Table 2, Figure 5). In the lower part of the stem, G1-G4•U15-C18, all the phosphorus chemical shifts were characteristic of a standard A double helix. However, in the stem region adjacent to the loop and in the loop itself, numerous chemical shifts of phosphate groups appeared to lie outside the canonical range, shifted towards either upfield (G4pU5, G7pG8, G13pA14) or downfield (U5pC6 and all the phosphates between G8 and G13) regions.

Exploration of the hairpin conformational space by constrained mechanical mechanics

The determination of the hairpin structure is well-defined from the numerous proton–proton distances and the sugar puckers that were accessible from the NMR spectra. Extracting the backbone conformation from NMR data is not straightforward as unusual δP cannot be immediately interpreted in terms of torsion angle conformations. To overcome this problem, we investigated the stability of a series of backbone conformations of the phosphate linkages affected by unusual δP with JUMNA (30). Owing to the use of internal coordinates and implicit solvent, this program allows rapid and systematic minimization of hairpin structures. All the minimizations were made under the following basic constraint set: (i) the stem hydrogen bonds, (ii) the totality of the proton–proton distances, (iii) the 18 sugar puckers and (iv) the backbone angles of phosphate junctions that displayed the

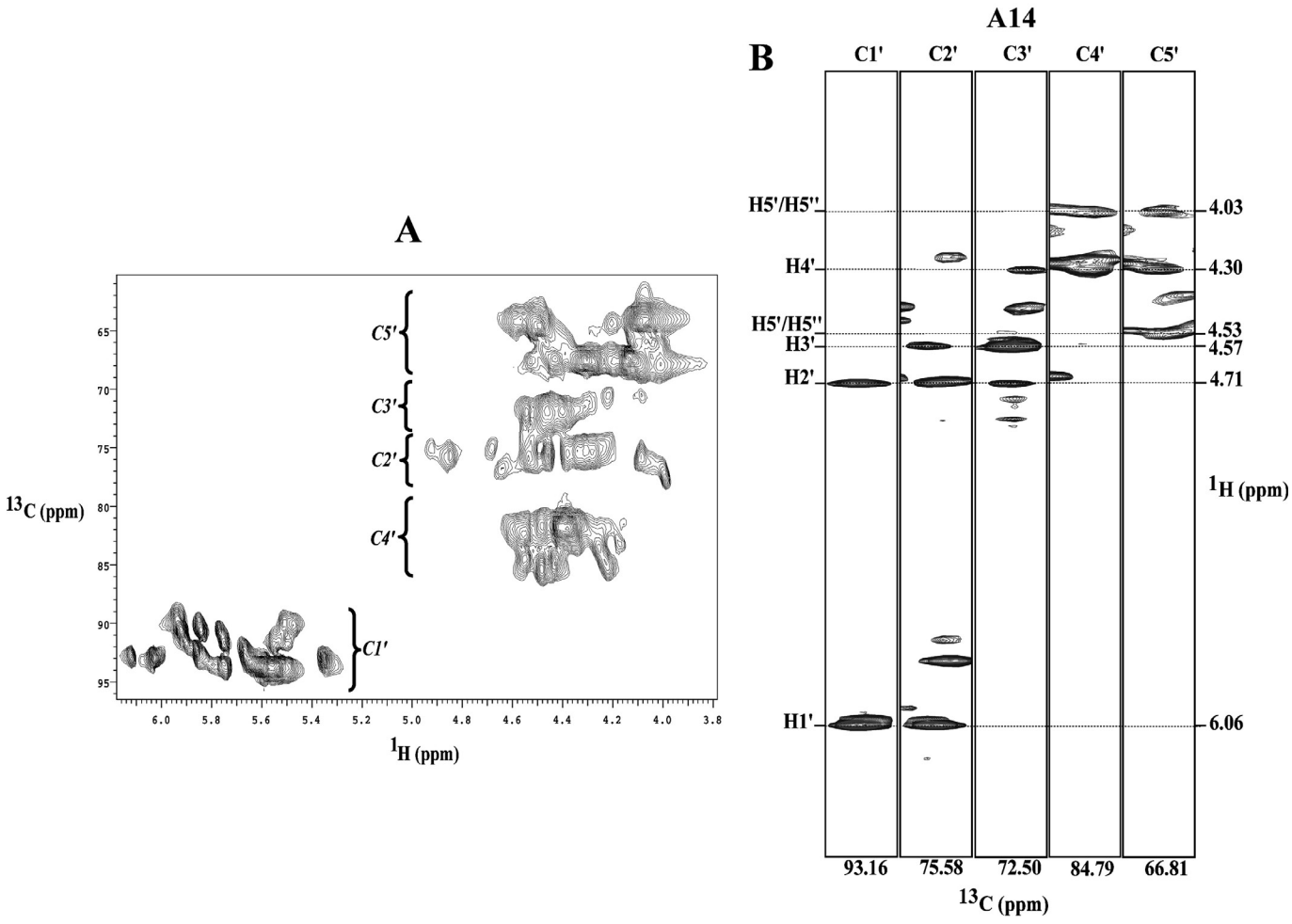

Figure 3. Assignment of ribose resonances. (**A**) The ribose region of a 2D ^{1}H–^{13}C HSQC experiment at 25°C. (**B**) ^{1}H–^{1}H tiles from slices through the different ribose carbon regions of A14 from a HCCH-COSY at 25°C.

standard δP (∼−0.45 p.p.m.) of canonical A-RNA (ε/ζ/α/β/γ:*t*/g^{-}/g^{-}/*t*/g^{+}). The simulations differ with respect to the additional restraints imposed on unusual phosphate groups. However, only certain ε/ζ/α/β/γ combinations have been demonstrated. The constraints applied on the stem phosphate groups that exhibit either upfield (G4pU5, G13pA14, δP ∼ −0.8 ppm) or downfield (U5pC6, δP ∼ −0.03 p.p.m.) phosphorus chemical shifts were derived from a previous analysis of X-ray structures (43). This study reported that only four stable non-standard ε/ζ/α/β/γ combinations existed within nucleic acid double helices, these combinations being moreover correlated with 5′ and 3′-sugar puckers. For instance, 5′-north sugars (the cases of G4pU5 and G13pA14) were always associated with ε/ζ:t/g^{-} canonical conformation. In the hexaloop, the G7pG8 δP was strongly shifted upfield and the next five δP, from G8 to G13, were shifted downfield. The latter δP cover so narrow a range (from 0 to 0.3 p.p.m.) that they can be considered to correspond to identical conformations. Thus, only two types of δP needed to be interpreted in term of conformation in the loop. Here, we take advantage of the previous abasic hexaloop modeling (44) that showed that the hexaloop fold necessitated a backbone extension, obtained by the simultaneous presence of south sugars and unusual ε/ζ/α/β/γ phosphate linkages, which were classified into only five families.

Thus, following these relationships, we systematically explored the acceptable combinations for the dinucleotide moieties of interest. In the converged structures, constraints for the hydrogen bonds, the intranucleotide distances and the sugars conformations were well reproduced, as well as the stem internucleotide distances. Thus, these structures are chosen according to (i) the number of sequential distance violations in the loop and (ii) the lowest total potential energy that is obviously related to the stability of the imposed backbone torsion angles.

Concerning the stem, in the best structures, the upfield G4pU5 and G13pA14 phosphate linkages are associated with α/γ:t/t (ε/ζ/α/β/γ:t/g^{-}/**t**/t/**t,** non-standard conformations in bold). The α/γ:t/t of G4pU5 is totally in line with X-ray data (45–48), and NMR (14) structures for which such phosphate linkages were observed adjacent to G•U wobble pairs. The unusual U5pC6 phosphate is stabilized only when ε/ζ adopt the g^{-}/t conformation (ε/ζ/α/β/γ:***g***$^{-}$/***t***/g^{-}/*t*/g^{+}), in full agreement with numerous NMR studies of B-DNA oligomers (49,50), showing that such downfield δP are related to correlated ε/ζ crankshaft motion. In line with these findings,

the most stable loop is constituted by a ε/ζ/α/β/γ:*t*/g^-/***t***/*t*/***t*** G7pG8 junction, followed by five ε/ζ/α/β/γ:***g***$^-$/***t***/g^-/*t*/g^+ phosphate linkages. Despite 10 residual sequential distance violations (differing by 0.4–0.9 Å of $d_{exp} = d_{measured} \pm 10\%$, in order to take into account the experimental error) the fold suggested by the NMR data is reproduced by the JUMNA calculations. So, G7, G8, G9 and A10 are stacked in continuity with the stem C6 base, while the dinucleotide segment containing U11 and G12 is also stacked, but isolated from both A10 and G13. Finally, the total potential energy of the whole hairpin is largely negative (−140 Kcal/Mol). Thus, this structure is chosen to be the starting point of the molecular dynamics study.

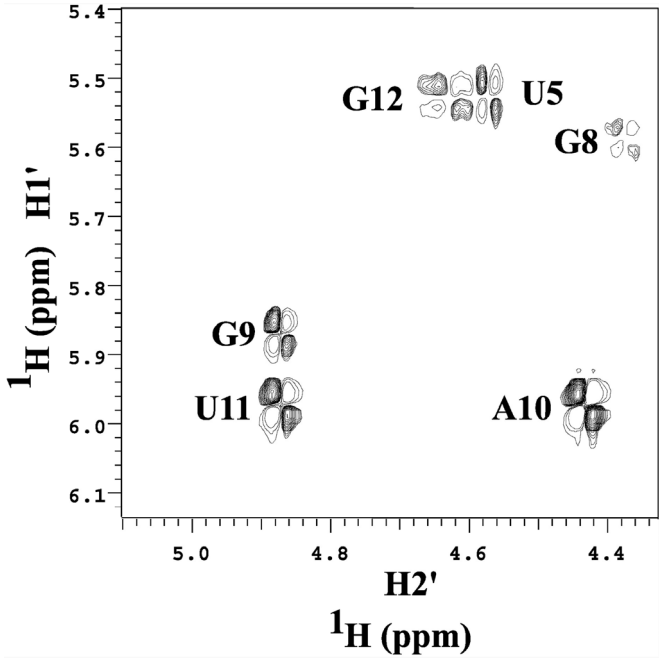

Figure 4. Expanded phase-sensitive DQF-COSY spectrum in ^2H_2O at 25°C corresponding to the H1′ to H2′ region.

The hairpin structure and dynamics

Two trajectories of 6 ns were performed on the structure selected in the previous section. The first trajectory (MD1) is made under the sequential distance and angle constraints reported in Table 1. After 2 ns of simulation, the RMSD calculated on heavy atoms between snapshots and either the starting or the average structures stabilize ∼2.5 ± 0.5 Å and 1.1 ± 0.2 Å, respectively. A second trajectory (MD2) of 6 ns is performed, removing both the sugar and the backbone angle constraints but maintaining restrained the 84 sequential distances. In MD2, the RMSD between the snapshots and the average structure is 1.8 ± 0.2 Å, reflecting an increased flexibility due to removal of certain constraints. In the two MD, all the intranucleotide distances, not constrained, spontaneously respect the experimental data and the internucleotide

Table 1. Structure statistics of NMR restraints used in MD refinements

MD1 and 2	Sequential and long-range internucleotide distances	84
MD1	Dihedral angles	85
MD1	Sugar puckers	18
MD1 and 2	Number of NOE violations	11
MD1	Dihedral backbone angle or Sugar pucker violations	0

The MD names are indicated on the left. The number of NOE violations corresponds to MD distances is not in the $d_{measured} \pm 10\%$ range. However, the worst violation attains only 0.6 Å.

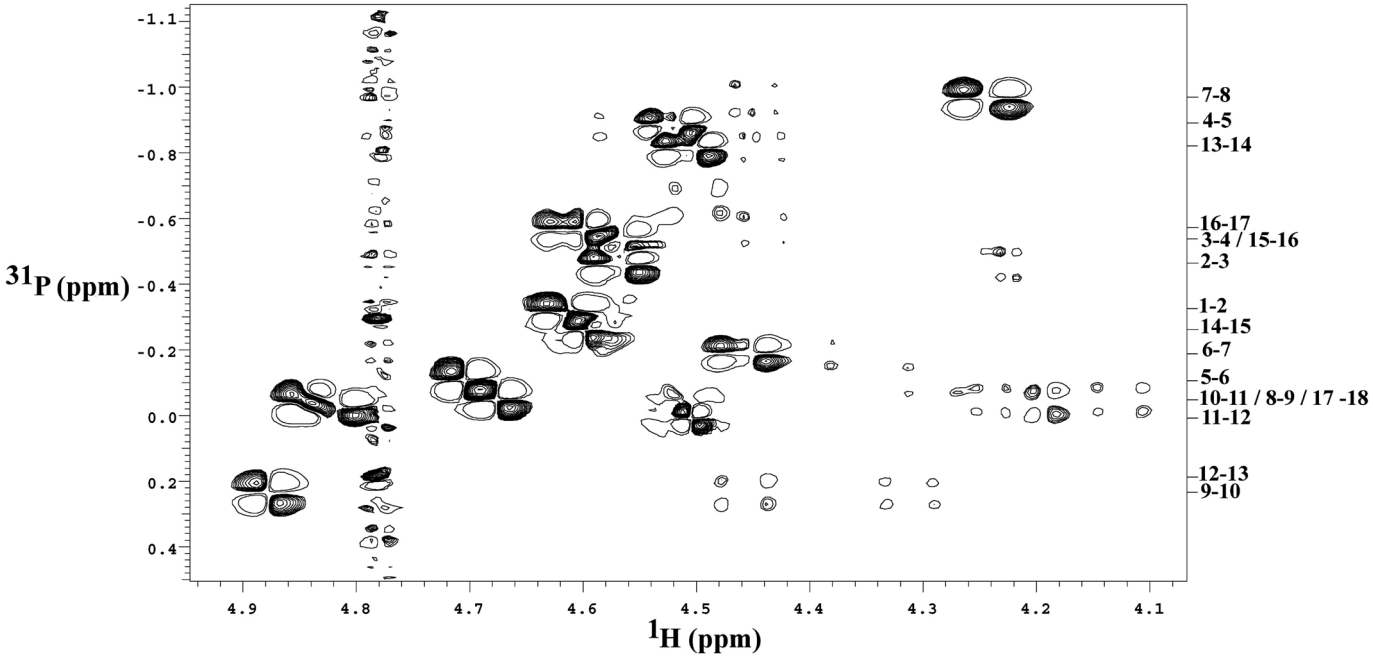

Figure 5. 2D-HCOSY experiment at 25°C showing the correlation between $H3'_i$ and $H4'_{i+1}$ sugar protons and $^{31}P_i$ nuclei. Assignments are shown on the right of the spectrum.

Table 2. ^{31}P chemical shifts (δP, p.p.m. with respect to H_3PO_4 external) and the corresponding backbone angle conformations

	Step	δP	ε	ζ	α	β	γ
Stem Strand 1	G1pA2	−0.19	t	g^-	g^-	t	g^+
	A2pG3	−0.45	t	g^-	g^-	t	g^+
	G3pG4*	−0.45	t	g^-	g^-	t	g^+
	G4*pU5	−0.83	t	g^-	t	t	t
	U5pC6	−0.03	g^-	t	g^-	t	g^+
Junction	C6pG7	−0.16	t	g^-	g^-	t	g^+
Loop	G7pG8	−0.90	t	g^-	t	t	t
	G8pG9	−0.00	**g^- (t)**	**t (g^-)**	g^-	t	g^+
	G9pA10	0.30	g^-	t	g^-	t	g^+
	A10pU11	−0.03	**g^- (g^-↔t)**	t	g^-	t	g^+
	U11pG12	0.03	g^-	**t (g^+)**	**g^- (g^+)**	t	**g^+ (t)**
Junction	G12pG13	0.17	**g^- (g^-↔t)**	t	g^-	t	g^+
Stem stand 2	G13pA14	−0.77	t	g^-	t	t	t
	A14pU15*	−0.19	t	g^-	g^-	t	g^+
	U15*pC16	−0.54	t	g^-	g^-	t	g^+
	C16pU17	−0.54	t	g^-	g^-	t	g^+
	U16pC18	−0.02	t	g^-	g^-	t	g^+

The conformations found in the unrestrained torsion MD2 dynamics are indicated under brackets when they differ from MD1. Non-standard phosphorus chemical shifts and conformations are shaded. G•U wobble bases are labeled by stars.

distances are in full agreement with d_{exp} (Table 1). Thus, the global hairpin structures are not altered when the sugar and angle restraints are removed. With or without restraints, the experimental sugar puckers are maintained. The sugars remain confined in one conformation, comprising the south sugars, and no south↔north transitions occur. Similarly, the backbone conformations in the stem are stable in both MD1 and MD2. Nevertheless, they evolve slightly in the loop in MD2. On the two steps A10pU11 and G12pG13, ε angles undergo fast $g^-\leftrightarrow t$ transitions, as suggested by X-ray data analysis (51,52) and as predicted by a previous modeling study (44). The g^-/t orientation of G8pG9 ε/ζ relaxes towards the canonical ε/ζ:t/g^- conformation. Finally, the most radical transition is observed along the U11pG12 backbone torsion angles, in which ζ/α/γ, initially in $t/g^-/g^+$, switch towards $g^+/g^+/t$. All these changes occur as soon as the angle constraints are relaxed. Then, throughout the whole MD2 trajectory, the backbone combinations remain very stable, apart from the chronic instability noticed for ε. Table 2 summarizes the backbone angle conformations in MD1 and MD2 structures.

Despite minor discrepancies, the global similarity of the MD1 and MD2 structures is very clear. This is reflected by the low values obtained for the RMSD between the two average structures that are 0.9 Å for the stem and 1.3 Å for the loop. The superposition of a randomly chosen set of 10 structures is shown in Figure 6, illustrating the overall structure and dynamics of the hairpin.

The stem and the G•U wobble. During the MD1 and MD2 simulations, the 6 bp stem forms a well-defined structural unit, the average RMSD with respect to the average structures being 0.7 ± 0.2 Å. Analysing the stem helicoidal parameters with the program CURVES (53,54), we find that, overall, the double helix resembles a canonical A-form RNA, with bases displaced towards the minor groove (X-$disp_{av}$ = −3.1 ± 0.9 Å), low twists (29.0 ± 6°) and rises

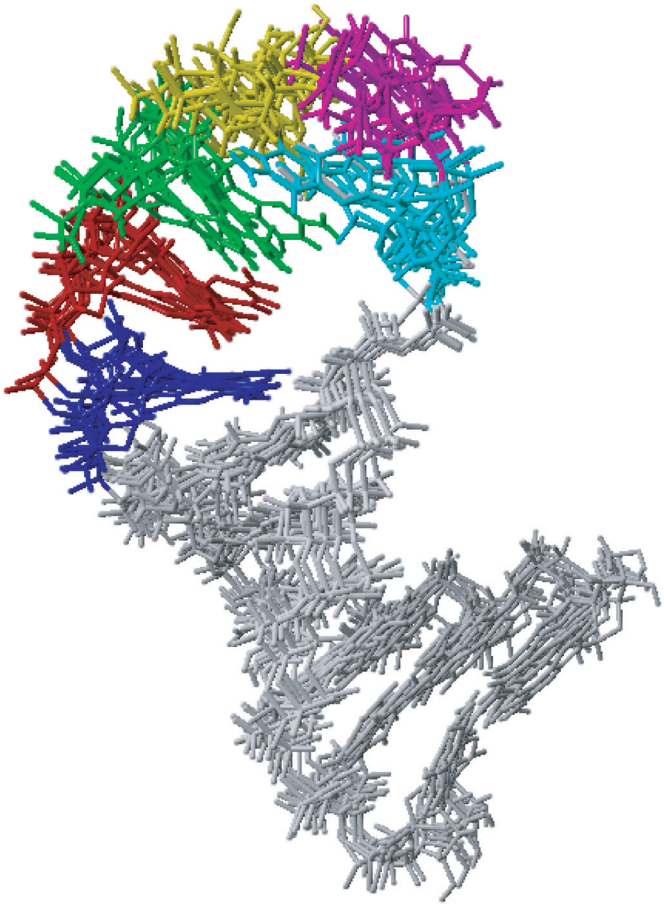

Figure 6. A total of 10 hairpin structures extracted from MD1. The superimpositions were made with respect to the stem part of the hairpin. The stem in light grey and the loop is in color.

Table 3. Stem interbase helicoidal parameters

Base step	Rise	Tilt	Roll	Twist
G1pA2•C18pU17	2.9 ± 0.4	−9 ± 4	4 ± 5	29 ± 6
A2pG3•U17pC16	3.0 ± 0.4	−3 ± 4	9 ± 5	29 ± 5
G3pG4*•C16pU15*	3.2 ± 0.4	−6 ± 5	19 ± 7	18 ± 5
G4*pU5•U15*pA14	3.1 ± 0.4	6 ± 8	21 ± 10	31 ± 6
U5pC6•A14pG13	3.4 ± 0.4	7 ± 5	−16 ± 6	30 ± 7

G•U wobble bases are labeled by stars. Rise, Tilt, Roll and Twist values are in °.

(3.1 ± 0.4 Å). Nevertheless, the helicoidal inter-base parameter analysis made step by step (Table 3) reveals that the G4•U15 wobble induces some noticeable perturbations. The base pair itself shows a shear value of −2.2 ± 0.3 Å, G3pG4•U15pC16 is under-twisted and G4pU5•U15pA14 slightly over-twisted, and the 5′ and 3′ side rolls are widely positive. Despite the fact that G•U wobble distortions can be influenced by the identity of the base pairs immediately adjacent to the mismatch (10,11), the under-/over-twisting observed here around G4•U15 is reproduced in three solution structures containing a single G•U wobble pair in CGG•CUG (12), UGU•AUA (13) and GGG•CUC (14) fragments. The G4•U15 effect is also reflected in the G4pU5•A14pU15 peculiar stacking composed by extensive intra-strand overlaps

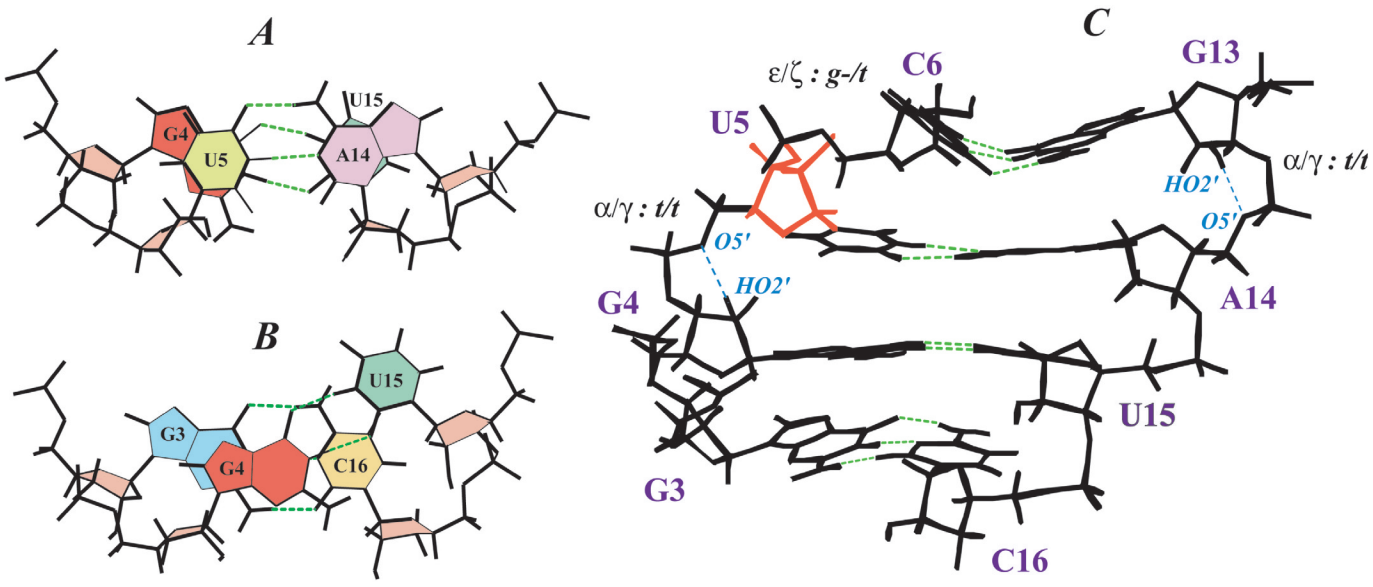

Figure 7. On the left, stacking patterns around the G4•U15 wobble pair: intrastrand stacking observed between (**A**) G4•U15 and U5•A14 and intra and interstrand stacking found between (**B**) G3•C16 and G4•U15. On the right, the view of the average structure of upper part of the stem and loop–stem junction (**C**). The U5 south sugar is in red. Non-standard backbones are explicitly mentioned. Hydrogen bonds are indicated by green and blue dashed lines.

(Figure 7A), a feature originally described by Mizuno and Sundaralingam (55) and regularly found around G•U wobble pair (11–14). Nevertheless, the value of G4pU5•A14pU15 stacking energy (−13 Kcal/Mol) is comparable with what would be expected in a regular A-form helix (56), in which both intra- and inter-strand interactions are observed, as found for G3pG4•U15pC16 (Figure 7B). Another factor that contributes to stabilize the G•U wobble is provided by the α/β/γ:*t*/*t*/*t* G4pU5 backbone, that could be also characteristic of this type of mismatch (14). Indeed, the α/γ:*t*/*t* angles hold O5′ and HO2′ of G4 in a conformation suitable for hydrogen bonding (Figure 7B), 85% of the time in MD1 and MD2. Finally, we detect a Na^+ ion located in the major groove around the G4•U15 wobble pair throughout the entire simulations. This ion remains close (≤2.5 Å) to either O6 of G4 and O4 of U5 (75% of time) or O6 of G3 (25% of time). This observation is in full agreement with previous studies (10,57) highlighting that, in the G•U wobble pair, the presence of co-planar uracil O4 and guanine N7 and O6 defines a region of deep negative electrostatic potential field able to attract mono or multivalent cations.

Also, U5pC6•G13pA14 contains many structural features that diverge from a canonical A-form (Figure 7C). The last base pair C6•G13 exhibits a marked positive buckle (37 ± 9°, see Figure 7C), a parameter known to optimize the stacking (58). Thus, G7, the first base in the loop interacts with both C6 (stacking energy of −5.5 Kcal/Mol) and C13 (stacking energy of −3 Kcal/). This buckle results in a striking negative roll between U5 and C6, a feature known to be favored in double helices by ε/ζ:g^-/*t* phosphate group (43) that is precisely found in the U5pC6 linkage. In turn, this ε/ζ:g^-/*t* phosphate imposes a south conformation on the 5′-sugar (here, the U5 sugar) for simple steric reasons (43). The helical distortions of the G13pA14 step of the opposite strand are less pronounced, but may need to be stabilized by the corresponding α/γ:*t*/*t* backbone that, similar to G4pU5, allows an O5′-HO2′ hydrogen bond 85% of the time (Figure 7C).

Finally, our 6 ns MD trajectories allow investigation of the stem flexibility. Whatever the set of constraints, i.e. with or without the sugar and backbone restraints, the average value of atomic fluctuations calculated on heavy atoms for each unit (1.2 Å in average, excluding the ends) is similar to those found in our own studies of double helices (59). The G4 and U15 units exhibit the smallest atomic fluctuation values (0.9 Å) whereas the U5, A14, C6 and G13 ones are all slightly above the average. This increased U5•A14 mobility is reflected by the G4pU5•U15pA14 roll and tilt standard deviations that are 3° greater than those of the 5′ stem part (Table 3).

The fairly rigid conformation of G•U wobble pair deviates from the canonical A-double helix structural characteristics, in terms of both sugar-backbone and helical parameters. Also, the local conformation of the last stem base pair (C6:G13) is clearly and strongly influenced by the loop. However, despite the large number of distortions, the stem structure exhibits numerous features that account for the stability of this 6 bp part of the hairpin.

The hexaloop. As stated previously, all the constraints applied in the loop region are reproduced in the model structures. The average RMSD between loop part structures extracted from the trajectory and the average one is 1.0 ± 0.2 Å in MD1 and 1.4 ± 0.2 Å in MD2, a low value that reflects a well ordered loop.

The average loop structure is shown Figure 8. The four purines G7, G8, G9 and A10 are stacked together (−5 Kcal/Mol per dinucleotide step) in continuity with C6, the last stem base. The clear overlap of G7 with C6 and, to a lesser extent, with G13, should be an element that controls the relative orientation of the loop with respect to the stem (Figure 6). So, owing to this stability, we observe a regular

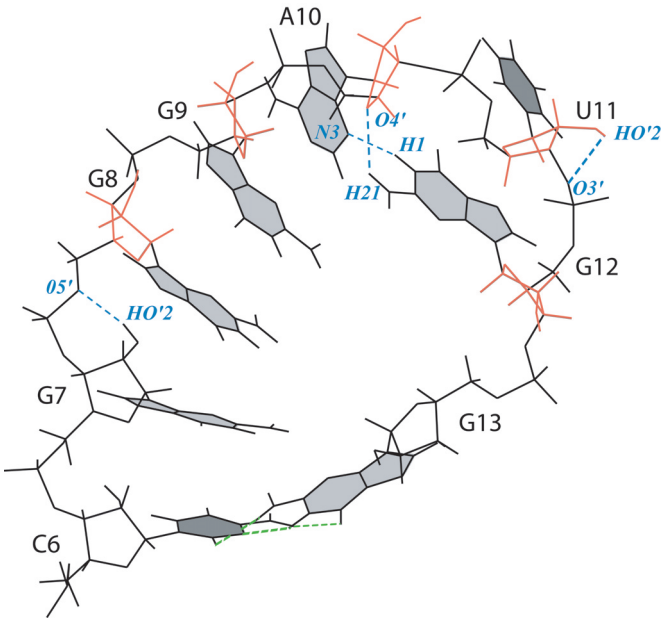

Figure 8. A view of the average structure of the last stem base pair (C6•G13) and the loop. The hydrogen bonds are indicated by dashed lines. South sugars are in red. Hydrogen bonds are indicated by green and blue dashed lines.

helical course of the nucleotide chain from the beginning of the stem to A10. The GGGA block behaves as A-like single chain, with rise and twist average values of 3.3 ± 0.3 Å and 30 ± 9°, respectively. In this loop region, the G7pG8 α/γ:t/t angles, as for the same backbone combinations encountered in the stem, allow the formation of a hydrogen bond between O5′ and HO2′ of G7 60% of the time in both the MD trajectories. The near inversion of the chain direction occurs between A10 and U11 and, consequently, disrupts the stacking. No sharp backbone turn is observed and the chain inversion results in combined changes of U11 χ angle and inclination. The χ of U11 (220° ± 16°) disrupts the regularity observed in G8, G9, A10 and G12 units in which χ (277 ± 15°) are closer to B-form (262°) than to A-form (202°; 209 ± 9° in the stem), as expected for bases attached to south sugars. In addition, the inclination of U11 is inverted compared with A10, generating a strong negative tilt (−40° ± 6°) between these bases. The loop ends with U11 and G12 that interact but remain far from the stem. Within the U11 unit, one hydrogen bond is formed between OH2′ of the sugar and the O3′ belonging to U11pG12 (50% of the time in MD1 and MD2). In addition, 50% of the time, G12 is in a position to promote two supplementary hydrogen bonds with A10 and its sugar, between (i) N1-H (G12) and N3 (A10) and (ii) N2-H21 (G12) and O4′ (A10). To guarantee good sampling of the possible relations between the loop bases, we performed supplementary dynamics simulations, constraining G12 to be close to G9 using a simple distance restraint. The sequential distances remain in agreement with the NMR data, and one hydrogen bond is formed between N2-H of G9 and 06 of G12. However, when the constraint is removed, G12 immediately reverts to its initial position in order to interact with A10, ensuring the stability of this hydrogen bond pattern.

The hexaloop requires backbone extension and curvature. The extension is manifest through $P_i - P_{i+1}$ distances that are on average 6.6 Å in the loop, against 6 Å in the stem or in an A-RNA. Actually, distending the loop is the main function of the numerous south sugars found in this part of the hairpin. In line with previous studies, (44,60) the distance between the O5′ and O3′ atoms increases by 0.5 Å when the enclosed sugar changes from north to south conformation. From an energetic point of view, the south sugar cost should be limited since it was shown (61,62) that C3′- and C2′-endo conformations were equally probable in rA and rG ribonucleosides, i.e. in systems devoid of any double strand specific constraints. The fold of the backbone is not completely independent of the abundance of south sugars as it seems to be achieved by the G8–G13 tract of ε/ζ:g^-/t phosphate linkages that need the presence of C2′-endo sugars (43). Substituting for instance the five ε/ζ:g^-/t conformers by the standard ε/ζ:t/g^- combination increases by 4 Å the distance between the C1′ of G8 and the C1′ of G13 and greatly distorts the loop–stem junction. In contrast, five α/γ:t/t confer the correct curvature to close the loop but destroy U11–G12 interactions. So, the five south sugars contribute to stretching the loop and help in selecting ε/ζ/α/β/γ combinations that do not destroy base stacking and loop closure. Thus, the sugar puckering preferences are important for the overall loop stability.

The well-ordered structure of the loop is reflected by the atomic fluctuations that attain an average of 1.9 Å per unit, a moderate value not so far from the stem one (1.2 Å). Upon calculating separately the fluctuations of sugar-backbone and bases atoms, we find that, in the angle constraint free MD2 trajectory, the base fluctuations slightly decrease with respect to MD1. In parallel, the MD2 backbone flexibility obviously increases, as it is not restrained. This observation suggests that the backbone motions, in particular the fast $t \leftrightarrow g^-$ transitions of ε, optimize base interactions. The amplitude of the base fluctuations increases progressively from G7 (1.3 Å) to U11 (2.2 Å) but strongly diminishes for the last base G12 (1.5 Å). So, the bases that benefit from extensive interactions with other loop components have the best defined positions, while the two bases that are located at the point of backbone inversion appear to be the most flexible. However, their fluctuations are not sufficient to break the stacking or to significantly alter the global conformation of the loop.

Finally, as the loop is a preferential target for DNA antisense, we evaluated the atom accessible surfaces, calculated by rolling a sphere of 1.4 Å radius over the van der Waals surface of atoms as previously described (35). Apart from the U11 apex base, in which almost all atoms are accessible, the major groove atoms (in reference to the major groove atoms in a double helix) of the bases are more accessible (∼7.5 Å^2 in average) than the minor groove ones (∼4 Å^2). Focusing now on the atoms engaged in Watson–Crick hydrogen-bonding, obviously buried in the stem part, it becomes clear that, apart from H–N1 and H–N3 of G12, all the loop donor and acceptor atoms present extensive surfaces to the solvent, as shown in Figure 9.

So, the stem is capped by a well-structured hexaloop, stabilized by both stacking interactions and hydrogen bonds. The backbone course is regular, even at the point where the inversion occurs. The elongation of the sugar-backbone

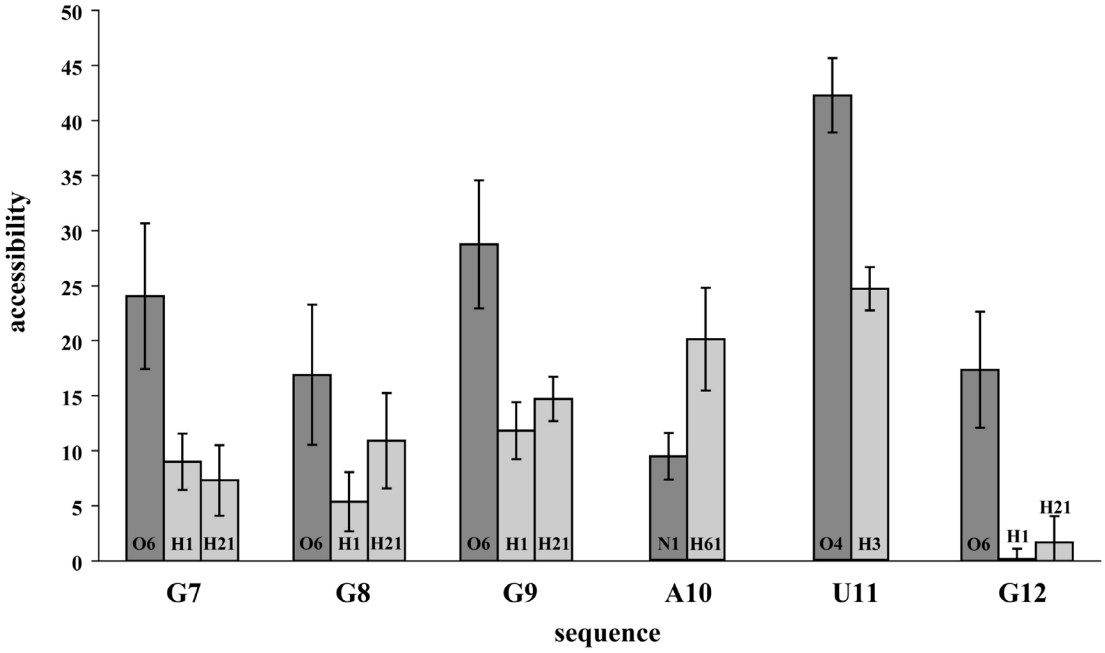

Figure 9. Accessibility (Å^2) calculated along the loop for the W–C acceptor (dark gray) and donor (light gray) atoms.

required by the hexaloop fold is mainly procured by the C2′-endo sugars. The most striking feature is the uniform stacking of the four successive purines GGGA that, moreover, expose to solvent their Watson–Crick faces, ready to base pair with a complementary strand of RNA or DNA.

DISCUSSION AND CONCLUSION

Through NMR and molecular dynamics in explicit solvent, we have determined the main structural characteristics of an RNA hairpin that is a part of the mRNA *PGY/MDR1* gene coding for the P-glycoprotein. This hairpin contains a G•U wobble pair in its stem and the $^{5'}$r(GGGAUG)$^{3'}$ purine-rich hexaloop, is a preferential target of antisense oligonucleotides. Complete assignment of the ^{1}H, ^{13}C and ^{31}P chemical shifts through ^{13}C/^{15}N labeling allowed unambiguous interpretation of the COSY data.

The stem forms a stable double helical structure although its upper part, from the G•U wobble pair to the beginning of the loop, is strongly distorted with respect to a canonical A-form. The 5′-under/3′-over-twisting and the 3′-intrastrand stacking, together with the presence of a counterion in the major groove, appear to be characteristic of the G•U wobble pair as they are reproduced in the three solution structures of single G•U pairs surrounded by various neighbors (10,11,13). The accentuated buckle of the last stem base pair may optimize an adequate overlap with the first loop base. Along this upper stem part, the helical distortions are associated with non-standard phosphate linkages that actively strengthen the structure by several hydrogen bonds involving O5′ atoms and ribose 2′-hydroxyl groups.

The apical loop structure is well ordered, as shown by the atomic fluctuations that do not considerably exceed those of the stem components. The GGGA purines continuously stacked upon the stem so that their arrangement mimics a regular A-like single strand in which all the Watson–Crick atoms are largely exposed to the solvent. The strand inversion occurs between A10 and U11, which is stacked with G12, the last base of the loop. This U11G12 segment is disconnected from the stem. In addition to these stacking interactions, the loop is stabilized by various hydrogen bonds involving ribose, backbone and base atoms.

The fold described here is a new one as, to the best of our knowledge, it does not resemble any available hexaloop structures in solution. Two other hexaloops contain five purines [1R2P (15): UGAAAG and 1BVJ (16): GUAAAA], but their purine tracts are located at the 3′ side of the loops, i.e. inverted compared with our GGGAUG loop. The 1R2P and 1BVJ backbones turn between the second and the third bases, and, in the case of 1BVJ, this inversion resembles U-turns, a very common RNA feature stabilized by one or two hydrogen bonds involving the uracil unit. In our hexaloop, the 3′ location of U11 does not allow such a feature. The reversal of the backbone precedes U11, which is not implicated in hydrogen bonds with any distant base. Here, the rupture in stacking logically occurs between A10 and U11, which is less costly than a break between two purines. In addition, it is the only manner to preserve the G7–A10 interactions. This trend for purine stacking is common to 1R2P and 1BVJ hexaloops, and can occur even in the case where the purine tract is interrupted by a cytosine as shown by the GUAACA loop solution structure [1U2A, (63)]. So, in the loops of 1R2P, 1BVJ and 1U2A sequences, in addition to the one studied here, the purines seem to prevent any extrusion of bases and lead to compact folds.

In addition to the stacking pattern and the regular course of the GGGA region, the hexaloop is characterized by the enhanced accessibility of the WC atoms, defining a distinctive hydrogen bonding surface for recognition. Largely,

owing to the hydrogen bonds existing between the first and the sixth bases, no other hexaloop exhibits continuously accessible WC atoms comparable with the ones observed in our structure. A previous study (64) has shown that oligonucleotides preferentially targeted single-stranded, stacked, helically ordered segments of folded RNA in solution. Furthermore, for duplex formation to occur bases must be sterically accessible and the initial pairing has to be stable enough to allow propagation, thus favoring G rich sequences. All these conditions are met in the GGGAUG hexaloop, which is thus a very attractive structure for accommodating antisense oligonucleotides.

Overall, the results of the present work provide new insight into the features that lead to a well-defined hairpin structure. They highlight a new hexaloop fold that contains all the elements for oligonucleotide recognition.

ACKNOWLEDGEMENTS

C. Simenel and M. Delepierre (Pasteur Institut) are gratefully acknowledged for providing NMR facilities to perform the TROSY-HCN experiment used in this work. The authors thank R. Lavery for a generous allocation of computer time and C. H. du Penhoat for helpful advice and comments on the manuscript. We also thank the University Paris 13 for financial support. Funding to pay the Open Access publication charges for this article was provided by CNRS.

Conflict of interest statement. None declared.

REFERENCES

1. Juranka,P.F., Zastawny,R.L. and Ling,V. (1989) P-glycoprotein: multidrug-resistance and a superfamily of membrane-associated transport proteins. *Faseb J.*, **3**, 2583–2592.
2. Ling,V., Gerlach,J. and Kartner,N. (1984) Multidrug resistance. *Breast Cancer Res. Treat.*, **4**, 89–94.
3. Alahari,S.K., Dean,N.M., Fisher,M.H., Delong,R., Manoharan,M., Tivel,K.L. and Juliano,R.L. (1996) Inhibition of expression of the multidrug resistance-associated P-glycoprotein of by phosphorothioate and 5′ cholesterol-conjugated phosphorothioate antisense oligonucleotides. *Mol. Pharmacol.*, **50**, 808–819.
4. Brigui,I., Djavanbakht-Samani,T., Jolles,B., Pigaglio,S. and Laigle,A. (2003) Minimally modified phosphodiester antisense oligodeoxyribonucleotide directed against the multidrug resistance gene mdr1. *Biochem. Pharmacol.*, **65**, 747–754.
5. Ho,S.P., Britton,D.H., Stone,B.A., Behrens,D.L., Leffet,L.M., Hobbs,F.W., Miller,J.A. and Trainor,G.L. (1996) Potent antisense oligonucleotides to the human multidrug resistance-1 mRNA are rationally selected by mapping RNA-accessible sites with oligonucleotide libraries. *Nucleic Acids Res.*, **24**, 1901–1907.
6. Motomura,S., Motoji,T., Takanashi,M., Wang,Y.H., Shiozaki,H., Sugawara,I., Aikawa,E., Tomida,A., Tsuruo,T., Kanda,N. *et al.* (1998) Inhibition of P-glycoprotein and recovery of drug sensitivity of human acute leukemic blast cells by multidrug resistance gene (mdr1) antisense oligonucleotides. *Blood*, **91**, 3163–3171.
7. Wei,H.L., Wu,Y.J., Jing,T., Bai,D.C. and Ma,L.F. (2003) Sensitization and apoptosis augmentation of K562/ADM cells by anti-multidrug resistance gene peptide nucleic acid and antisense oligodeoxyribonucleotide. *Acta. Pharmacol. Sin.*, **24**, 805–811.
8. Kostenko,E.V., Beabealashvilly,R.S., Vlassov,V.V. and Zenkova,M.A. (2000) Secondary structure of the 5′-region of PGY1/MDR1 mRNA. *FEBS Lett.*, **475**, 181–186.
9. Tu,G.C., Cao,Q.N., Zhou,F. and Israel,Y. (1998) Tetranucleotide GGGA motif in primary RNA transcripts. Novel target site for antisense design. *J. Biol. Chem.*, **273**, 25125–25131.
10. Varani,G. and McClain,W.H. (2000) The G × U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep.*, **1**, 18–23.
11. Masquida,B. and Westhof,E. (2000) On the wobble GoU and related pairs. *Rna.*, **6**, 9–15.
12. White,S.A., Nilges,M., Huang,A., Brunger,A.T. and Moore,P.B. (1992) NMR analysis of helix I from the 5S RNA of *Escherichia coli*. *Biochemistry*, **31**, 1610–1621.
13. Allain,F.H. and Varani,G. (1995) Divalent metal ion binding to a conserved wobble pair defining the upstream site of cleavage of group I self-splicing introns. *Nucleic Acids Res.*, **23**, 341–350.
14. Ramos,A. and Varani,G. (1997) Structure of the acceptor stem of *Escherichia coli* tRNA Ala: role of the G3.U70 base pair in synthetase recognition. *Nucleic Acids Res.*, **25**, 2083–2090.
15. Sigel,R.K., Sashital,D.G., Abramovitz,D.L., Palmer,A.G., Butcher,S.E. and Pyle,A.M. (2004) Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif. *Nature Struct. Mol. Biol.*, **11**, 187–192.
16. Puglisi,E.V. and Puglisi,J.D. (1998) HIV-1 A-rich RNA loop mimics the tRNA anticodon structure. *Nature Struct. Biol.*, **5**, 1033–1036.
17. Plateau,P. and Gueron,M. (1982) Exchangeable proton NMR without base-line distorsion, using new strong-pulse sequences. *J. Am Chem. Soc.*, **104**, 7310–7311.
18. Marion,D. and Wuthrich,K. (1983) Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of 1H-1H spin–spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.*, **113**, 967–974.
19. Sklenar,V., Miyashiro,H., Zon,G., Miles,H.T. and Bax,A. (1986) Assignment of the 31P and 1H resonances in oligonucleotides by two-dimensional NMR spectroscopy. *FEBS Lett.*, **208**, 94–98.
20. Bodenhausen,G. and Ruben,D. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.*, **69**, 185–189.
21. Norwood,T.J., Boyd,J., Heritage,J.E., Soffe,N. and Campbell,I.D. (1990) Comparison of techniques for 1H detected heteronuclear 1H–15N spectroscopy. *J. Magn. Reson.*, **87**, 488–501.
22. Lukavsky,P.J. and Puglisi,J.D. (2001) RNAPack: an integrated NMR approach to RNA structure determination. *Methods.*, **25**, 316–332.
23. Wohnert,J., Dingley,A.J., Stoldt,M., Gorlach,M., Grzesiek,S. and Brown,L.R. (1999) Direct identification of NH. N hydrogen bonds in non-canonical base pairs of RNA by NMR spectroscopy. *Nucleic Acids Res.*, **27**, 3104–3110.
24. Pardi,A. and Nikonowicz,E.P. (1992) Simple procedure for resonance assignment of the sugar protons in carbon-13 labeled RNAs. *J. Am Chem. Soc.*, **114**, 9202–9203.
25. Brutscher,B. and Simorre,J.P. (2001) Transverse relaxation optimized HCN experiment for nucleic acids: combining the advantages of TROSY and MQ spin evolution. *J. Biomol. NMR.*, **21**, 367–372.
26. Simorre,J.P., Zimmermann,G.R., Mueller,L. and Pardi,A. (1996) Correlation of the guanosine exchangeable and nonexchangeable base protons in 13C-/15N-labeled RNA with an HNC-TOCSY-CH experiment. *J. Biomol. NMR.*, **7**, 153–156.
27. Simorre,J.P., Zimmermann,G.R., Pardi,A., Farmer,B.T., 2nd and Mueller,L. (1995) Triple resonance HNCCCH experiments for correlating exchangeable and nonexchangeable cytidine and uridine base protons in RNA. *J. Biomol. NMR.*, **6**, 427–432.
28. Simorre,J.P., Zimmermann,G.R., Mueller,L. and Pardi,A. (1996) Triple-resonance experiments for assignment of adenine base resonances in 13C/15N-labeled RNA. *J. Am. Chem. Soc.*, **118**, 5316–5317.
29. Baleja,J.D. and Sykes,B.D. (1991) Correlation time adjustment factors for NOE-based structure refinement. *J. Magn. Reson.*, **91**, 624–629.
30. Lavery,R., Zakrzewska,K. and Sklenar,H. (1995) JUMNA (junction minimisation of nucleic acids). *Comput. Phys. Commun.*, **91**, 135–158.
31. Cheatham,T.E., III, , Cieplak,P. and Kollman,P.A. (1999) A modified version of the Cornell *et al.* force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.*, **16**, 845–862.
32. Case,D.A., Pearlman,D.A., Caldwell,J.W., Cheatham,T.E.I., Wang,J., Ross,W.S., Simmerling,C.L., Darden,T.A., Merz,K.M., Stanton,R.V. *et al.* (2002) *AMBER 7*. University of California, San Francisco, CA.
33. Lavery,R. and Hartmann,B. (1994) Modelling DNA conformational mechanics. *Biophys. Chem.*, **50**, 33–45.

34. Varnai,P., Djuranovic,D., Lavery,R. and Hartmann,B. (2002) Alpha/gamma transitions in the B-DNA backbone. *Nucleic Acids Res.*, **30**, 5398–5406.
35. Tisne,C., Hantz,E., Hartmann,B. and Delepierre,M. (1998) Solution structure of a non-palindromic 16 base-pair DNA related to the HIV-1 kappa B site: evidence for BI-BII equilibrium inducing a global dynamic curvature of the duplex. *J. Mol. Biol.*, **279**, 127–142.
36. Tisne,C., Hartmann,B. and Delepierre,M. (1999) NF-kappa B binding mechanism: a nuclear magnetic resonance and modeling study of a GGG → CTC mutation. *Biochemistry*, **38**, 3883–3894.
37. Berendsen,H.J.C., Postma,J.P.M., Van Gunsteren,W.F., A,D.N. and Haak,J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **81**, 3684–3690.
38. Ryckaert,J.P., Ciccotti,G. and Berendsen,H.J.C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, **23**, 327–341.
39. Darden,T., York,D. and Pedersen,L.J. (1993) Particle mesh Ewald: an $N\cdot\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.
40. Cheatham,T.E., III, Miller,J.L., Fox,T., Darden,T. and Kollman,P.A. (1995) molecular dynamics simulations on solvated biomolecular systems: the particle Mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J. Am Chem. Soc.*, **117**, 4193–4194.
41. Nikonowicz,E.P. and Pardi,A. (1993) An efficient procedure for assignment of the proton, carbon and nitrogen resonances in 13C/15N labeled nucleic acids. *J. Mol. Biol.*, **232**, 1141–1156.
42. Varani,G. and Tinoco,I., Jr (1991) RNA structure and NMR spectroscopy. *Q. Rev. Biophys.*, **24**, 479–532.
43. Djuranovic,D. and Hartmann,B. (2003) Conformational characteristics and correlations in crystal structures of nucleic acid oligonucleotides: evidence for sub-states. *J. Biomol. Struct. Dyn.*, **20**, 771–788.
44. Joli,F., Hartmann,B. and Hantz,E. (2006) Structure and dynamics of phosphate linkages and sugars in an abasic hexaloop RNA hairpin. *Biophys. J.*, **90**.
45. Holbrook,S.R., Cheong,C., Tinoco,I., Jr and Kim,S.H. (1991) Crystal structure of an RNA double helix incorporating a track of non-Watson–Crick base pairs. *Nature*, **353**, 579–581.
46. Lima,S., Hildenbrand,J., Korostelev,A., Hattman,S. and Li,H. (2002) Crystal structure of an RNA helix recognized by a zinc-finger protein: an 18-bp duplex at 1.6 A resolution. *Rna*, **8**, 924–932.
47. Tanaka,Y., Fujii,S., Hiroaki,H., Sakata,T., Tanaka,T., Uesugi,S., Tomita,K. and Kyogoku,Y. (1999) A′-form RNA double helix in the single crystal structure of r(UGAGCUUCGGCUC). *Nucleic Acids Res.*, **27**, 949–955.
48. Znosko,B.M., Kennedy,S.D., Wille,P.C., Krugh,T.R. and Turner,D.H. (2004) Structural features and thermodynamics of the J4/5 loop from the *Candida albicans* and Candida dubliniensis group I introns. *Biochemistry*, **43**, 15822–15837.
49. Lefebvre,A., Mauffret,O., Lescot,E., Hartmann,B. and Fermandjian,S. (1996) Solution structure of the CpG containing d(CTTCGAAG)2 oligonucleotide: NMR data and energy calculations are compatible with a BI/BII equilibrium at CpG. *Biochemistry*, **35**, 12560–12569.
50. Djuranovic,D. and Hartmann,B. (2004) DNA fine structure and dynamics in crystals and in solution: the impact of BI/BII backbone conformations. *Biopolymers*, **73**, 356–368.
51. Murray,L.J., Arendall,W.B., 3rd, Richardson,D.C. and Richardson,J.S. (2003) RNA backbone is rotameric. *Proc. Natl Acad. Sci. USA*, **100**, 13904–13909.
52. Schneider,B., Moravek,Z. and Berman,H.M. (2004) RNA conformational classes. *Nucleic Acids Res.*, **32**, 1666–1677.
53. Lavery,R. and Sklenar,H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.*, **6**, 63–91.
54. Lavery,R. (1988) Junctions and bends in nucleic acids: a new theoretical approach. *J. Biomol. Struct. Dyn.*, **3**, 191–211.
55. Mizuno,H. and Sundaralingam,M. (1978) Stacking of Crick Wobble pair and Watson–Crick pair: stability rules of G-U pairs at ends of helical stems in tRNAs and the relation to codon-anticodon Wobble interaction. *Nucleic Acids Res.*, **5**, 4451–4461.
56. Perez,A., Noy,A., Lankas,F., Luque,F.J. and Orozco,M. (2004) The relative flexibility of B-DNA and A-RNA duplexes: database analysis. *Nucleic Acids Res.*, **32**, 6144–6151.
57. Klein,D.J., Moore,P.B. and Steitz,T.A. (2004) The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA*, **10**, 1366–1379.
58. Hunter,C.A. (1993) Sequence-dependent DNA structure. The role of base stacking interactions. *J. Mol. Biol.*, **230**, 1025–1054.
59. Djuranovic,D. and Hartmann,B. (2005) Molecular dynamics studies on free and bound targets of the bovine papillomavirus type I e2 protein: the protein binding effect on DNA and the recognition mechanism. *Biophys. J.*, **89**, 2542–2551.
60. Sorin,E.J., Engelhardt,M.A., Herschlag,D. and Pande,V.S. (2002) RNA simulations: probing hairpin unfolding and the dynamics of a GNRA tetraloop. *J. Mol. Biol.*, **317**, 493–506.
61. Foloppe,N. and MacKerell,J.A.D. (1998) Conformational properties of the deoxyribose and ribos moieties of nucleic acids: a quantum mechanical study. *J. Phy. Chem. B.*, **102**, 6669–6678.
62. Foloppe,N. and Nilsson,L. (2005) Toward a full characterization of nucleic acid components in aqueous solution: similations of nucleosides. *J. Phys. Chem. B.*, **109**, 9119–9131.
63. Stallings,S.C. and Moore,P.B. (1997) The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA. *Structure*, **5**, 1173–1185.
64. Mir,K.U. and Southern,E.M. (1999) Determining the influence of structure on hybridization using oligonucleotide arrays. *Nat. Biotechnol.*, **17**, 788–792.