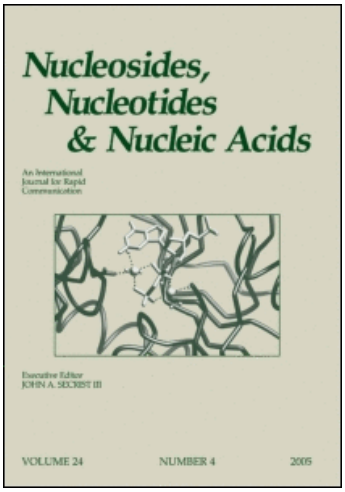

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

High Resolution NMR Studies of DNA Hairpins with Four Nucleotides in the Loop Region

C. W. Hilbers[a]; M. J.J. Blommers[a]; F. J.M. Van de Van[a]; J. H. Van Boom[b]; G. A. Van Der Marel[b]
[a] Laboratory of Biophysical Chemistry, Faculty of Science, University of Nijmegen, Nijmegen, The Netherlands [b] Laboratory of Organic Chemistry, Gorlaeus Laboratories, State University of Leiden, Leiden, The Netherlands

To cite this Article Hilbers, C. W. , Blommers, M. J.J. , Van de Van, F. J.M. , Van Boom, J. H. and Van Der Marel, G. A.(1991) 'High Resolution NMR Studies of DNA Hairpins with Four Nucleotides in the Loop Region', Nucleosides, Nucleotides and Nucleic Acids, 10: 1, 61 — 80
To link to this Article: DOI: 10.1080/07328319108046436
URL: http://dx.doi.org/10.1080/07328319108046436

NUCLEOSIDES & NUCLEOTIDES, 10(1-3), 61-80 (1991)

HIGH RESOLUTION NMR STUDIES OF DNA HAIRPINS WITH FOUR NUCLEOTIDES IN THE LOOP REGION.

C.W. Hilbers*, M.J.J. Blommers and F.J.M. van de Ven,
Laboratory of Biophysical Chemistry, Faculty of Science, University of Nijmegen, Nijmegen, The Netherlands.

J.H. van Boom and G.A. van der Marel,
Laboratory of Organic Chemistry, Gorlaeus Laboratories, State University of Leiden, Leiden, The Netherlands.

INTRODUCTION

When considering DNA structure the first picture that comes to mind is that of the celebrated double helix. Notwithstanding the overwhelming importance of this structure it is interesting however that refined structural studies that have been conducted during the last ten years have demonstrated that deviations, at the nucleotide level, from the idealized double helix and its consonant symmetry properties may lead to very interesting global structural variations.

Another result that has led to the notion that DNA has a structural richness that goes beyond the double helix has come from sequence determinations of intercistronic regions which are most likely involved in the regulation of gene expression. For instance, inverted repeats or palindrome sequences are often found in non-coding regions of the genome. It has been demonstrated for synthetic as well as naturally occurring circular DNA that these sequences can be converted into cruciform structures (1,2). Even more intricate structures have been found for DNA with complementary purine and pyrimidine stretches. Under supercoiled pressure these sequences were shown to be S1-nuclease sensitive and a number of subsequent experiments have indicated that this DNA folds into a socalled H-DNA conformation the core of which consists of a triple helix (3,4).

Guanine rich sequences such as $d(TT\text{-}GGGG\text{-}TT\text{-}GGGG)_n$ are found intelomeric ends of eukaryotic chromosomes and such nucleotide stretches are able to form anti-parallel quadruplex structures containing guanine base tetrads (5,6).

In all of these examples hairpins are formed as essential elements of the overall folding pattern. The loopfolding in such hairpins forms the subject of the present contribution.

We have been interested for some time in the question whether it is possible to describe loopfolding in nucleic acid hairpins by a set of phenemenological and physical rules and have applied NMR spectroscopy to obtain insight in the conformational properties underlying these rules. In this context one may ask to what accuracy the conformational parameters have to be known to be able to formulate such rules and to what accuracy is NMR able to provide conformational information about nucleic acid structure. At this time these questions cannot be answered exhaustively but some important features can certainly be made apparent and will be discussed in this paper.

NMR DATA AND STRUCTURE OF NUCLEIC ACIDS.

Structural constraints for the description of the conformation of nucleic acids are obtained from Nuclear Overhauser effects (NOEs) and J-coupling information, i.e. NOEs yield distance constraints and three bond J-coupling constants yield torsion angles. The amount and accuracy of this information is, however, limited and it is important to decide, given a set of data, to what detail a structure can be described. We have approached this problem by applying multiconformational analysis to a dinucleotide unit and a two basepair step which are thought to be part of a double helix. Structural information obtainable for nucleic acids from NMR spectroscopy does not normally reach beyond these structural elements except when intricate folding occurs as in loop regions or as in more complicated structures like tRNA. The multiconformational analysis was carried out with the computer program HELIX.

The conformational space of a dinucleotide is, provided bond distances and bond angles are constant, determined by the backbone torsion angles and the pseudorotation parameters of the sugar ring. Thus, proper sampling of the conformational space requires the systema-

tic variation of eleven parameters, i.e. the phase angle of pseudorotation, P, and the pucker amplitude of both sugars, the glycosidic torsion angles, χ, of the 5'- and 3'-nucleotide and the backbone torsion angles ε, ζ, α, β and γ. In the conformational analysis procedure these parameters are varied systematically and this leads to a large number of different conformations each of which can be thought of as a point of an eleven dimensional 'lattice' in conformational space. The density of these points depends on the number of steps in which the torsion angles are sampled. Using a 'search tree' algorithm the individual conformers are then tested against the input constraints and rejected if these are violated. Application of the 'search tree' algorithm prevents unnecessary calculations, e.g. if a particular value of the glycosidic torsion angle, χ, leads to a violation of the intra-residue H1'-H8 distance constraint there is no use in continuing the construction of dinucleotides with such a value for χ. Such calculations were carried out for all possible combinations of nucleotides. The results obtained for one example, the dinucleotide GC, are listed in Table 1. The helix parameters as well as the backbone torsion angles are presented for proton-proton distances with different but defined accuracies. The results show that these structural parameters are not so well-defined by the NMR distance constraints. In general it appears to be very important to have estimates of the distances between the ring protons (H6/H8 and/or H5) of adjacent bases and of imino-protons of adjacent base pairs. The differences in the spread of torsion angles and helix parameters calculated for different dinucleotides is then moderate for most parameters. An example is added to Table 1 for the dinucleotide CG.

Interestingly, a correlation analysis of the backbone angles obtained in the multiconformation calculation shows that these angles are not independent from one another. Linear correlations are observed when one plots the torsion angles α, β and ε or the helix parameters twist, slide and shift against each other (cf. Figs. 1 and 2). This suggests that the structure can be defined more accurately, when, in addition to distance constraints obtained through NOE measurements, some of the backbone angles can be restricted to a certain domain. The results obtained for the dinucleotide L2-L3, i.e. the central thymidi-

TABLE 1. Spread of torsion angles and helix parameters of a GC dinucleotide above the dashed line and of a CG dinucleotide below the dashed line satisfying B-type constraints with specified inaccuracies.

Bounds*	twist	roll	tilt	rise*	slide*	shift*	α	β	γ	ε	ζ
+/-0.1	7.4	2.3	5.3	0.5	0.8	1.0	20	25	20	30	15
+/-0.2	16.1	8.3	6.7	0.7	1.2	2.0	40	35	35	40	20
+/-0.3	33.7	14.5	13.8	1.2	2.2	3.2	65	55	55	70	35
+/-0.2	29.2	13.2	7.0	1.0	1.8	2.2	45	40	50	45	20

nes, of the loop region of the hairpin formed by the hexadecanucleotide d(ATCCTA-TTTA-TAGGAT) (vide infra) may serve as an example. The relevant structural constraints derived from the NMR data are collected in Table 2. They served as input for the multiconformational analysis. Variation of the backbone angles in steps of 10° then leads to 220 structures which fulfil the input constraints. The conformations of a representative set are superimposed in Fig. 3. It is clear that this dinucleotide element of the hairpin molecule is reasonably well-defined by the NMR constraints.

ASPECTS OF FOLDING IN DNA HAIRPIN-LOOPS WITH FOUR NUCLEOTIDES

Our interest in loopfolding in nucleic acids arose when we considered the stability of a series of DNA hairpin molecules formed by the partly selfcomplementary sequences d(ATCCTA-T_n-TAGGAT) in which n varied between 0 and 8. NMR and UV-melting experiments on these compounds demonstrated that an optimal stability is obtained for loops composed of four or five nucleotides (7). These results differed from those available at that time for RNA hairpins. UV-melting experiments led to the conclusion that for the sequences A6CmU6 (m=4,5,6 or 8) (8) and A4GCmU4 (m=4,5 or 6) (9) optimal stabilities are obtained for

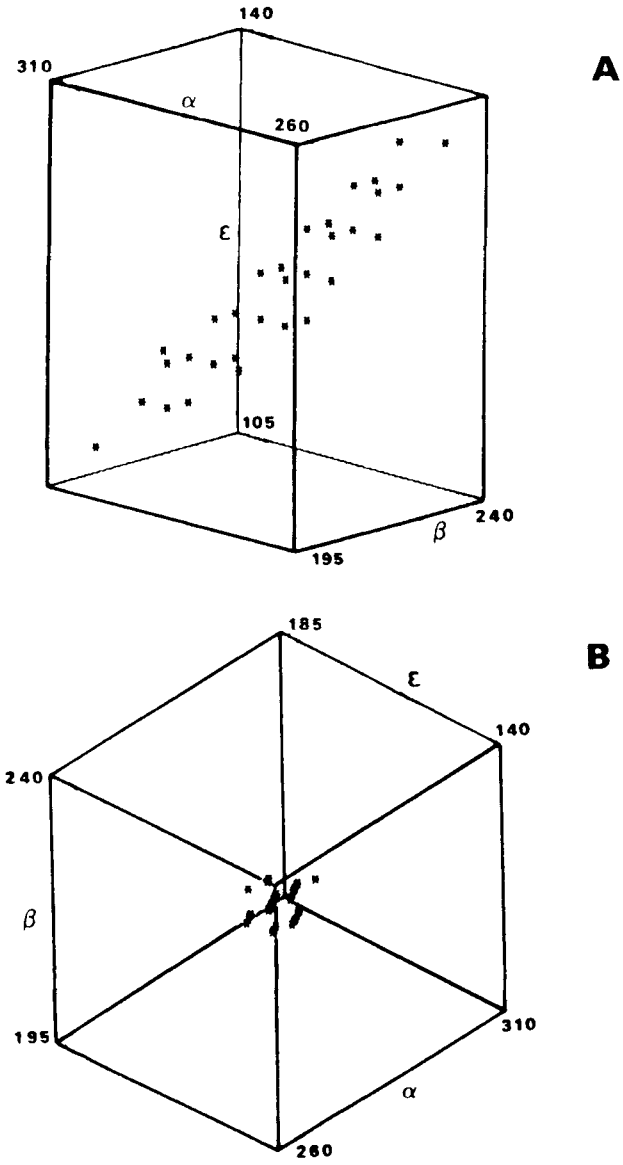

FIG. 1.
Plot showing the linear correlation between the torsion angles α, β and ε of a dinucleotide.
A) Side view,
B) View along the body diagonal.

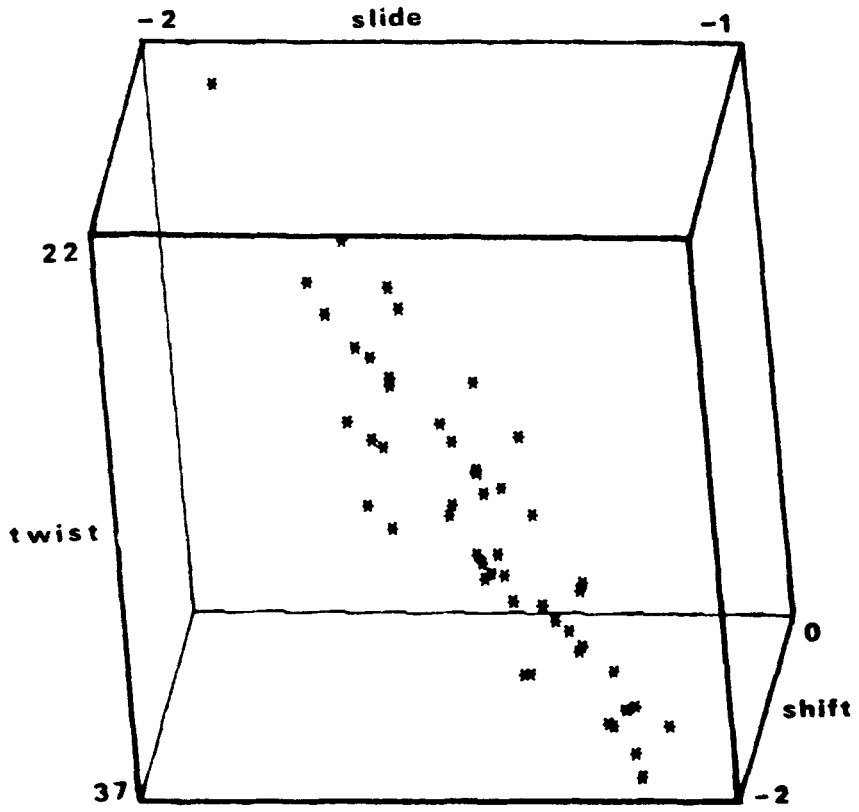

FIG. 2.
Plot showing the linear correlation between the helix parame ters twist, slide and shift.

hairpins with six or seven nucleotides in the loop region. We explained this difference by assuming a different loop architecture for RNA and DNA hairpins. In the proposed model the folding of the loop was dictated by the conformation adopted by the stem, in other words whether it belongs to the A- or B-type double helix. Thus, in RNA hairpins the loop is formed by extending the A-type conformation of one strand in the 5'-direction by five nucleotides. Then via a sharp turn the first strand is connected to the opposite strand by one or two nucleotides. In DNA hairpins a different course is followed. There the

TABLE 2. Structural constraints used in the multi conformational analysis of the dinucleotide element d(TpT), comprizing the second and third nucleotides in the loop, L2 and L3 respectively, of the hairpin formed by d(ATCCTA-TTTA-TAGGAT).
a) Values in left column are in degrees.
b) The standard angular notation is used:

$$\begin{matrix} & \alpha & & \beta & & \gamma & & \delta & & \varepsilon & & \zeta & \\ O3'- & P & - O5' & - & C5' & - & C4 & - & C3' & - & O3' & - & P \end{matrix}$$

P(5') is the phase angle of the pseudorotation of the 5'-sugar; Φm(5') is the pucker amplitude of the 5'-sugar, χ is the glycosidic torsion angle of the 5'-residue.

P(5')	=	166 - 189	L2H2' - L3H6	=	3.7 - 5.0 Å
Φm(5')	=	41 - 44	L2H2''- L3H6	=	1.8 - 30 Å
χ (5')	=	212 - 221	L2H3' - L3H6	=	3.6 - 5.0 Å
β	=	t	L2H1' - L3H5'	=	3.4 - 7.0 Å
γ	=	g^+	L2H6 - L3CH3	=	4.0 - 30 Å
P(3')	=	198	L2H1' - L3H6	=	4.0 - 30 Å
Φm(3')	=	41	L2H2''- L3H5'	=	3.1 - 7.0 Å
χ(3')	=	165 - 210			

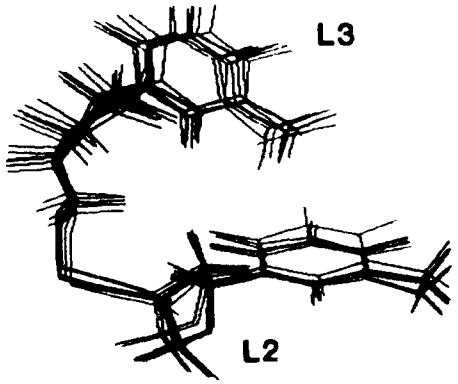

FIG. 3.
Superposition of structures obtained by means of multiconformation analysis performed for the dinucleotide step, L2-L3, of the loop region of the hairpin formed by the hexadecanucleotide d(ATCCTA-TTTA-TAGGAT). The similarity of the conformers indicates that this part of the loop structure is resonably well-defined by the NMR constraints.

normals make an angle of about 30°. This suggests (vide infra) that an interchange of the two complementary bases will not bring them into the proper position to form a base pair. Indeed, the energy minimized hairpin structures obtained for the molecules with the loop sequences-GTTC- and -ATTT- do not exhibit base pair formation between the first and the last residue in the loop (cf. Fig. 4B and 4D). The hairpin with the -ATTT- loopsequence was further examined by building a structure in which the A and the T residues were forced to form a base pair while at the same time the sugar phosphate backbone was readjusted to accommodate the new situation. This yielded a high energy structure. When the constraints were removed and the molecule was subjected to further energy minimization the A-T pair was disrupted again. Thus our molecular mechanics calculations strongly suggest that a -PyTTPu- loop (Py = pyrimidine, Pu = purine) allows base pair formation within a loop of four residues, while a -PuTTPy- loop does not. These observations were endorsed by optical melting experiments and NMR experiments.

A compilation of melting temperatures measured for a variety of hairpins is given in Table 3. For reasons of clarity we have also incorporated some results presented earlier (19); the molecules are grouped in homologous series. We first consider the replacement of thymidines by adenines. Examination of Table 3A shows that in all cases where one expects hairpin molecules with four-membered loops to be formed the melting temperature is 50° C or lower, except for the molecule with four thymidines in the loop region. As we have seen for this molecule a somewhat increased stability may be expected. Our NMR experiments have shown that the first and the last thymidine in the loop form a T-T wobble pair (11). The molecules in which it is in principle possible to form a Watson-Crick base pair beween the first and the last base of the loop also exhibit to melting temperatures below 50°C except for the loop sequence -TTTA-, but including the hairpin with the loop sequence -ATTT-. The first hairpin has a melting temperature of 57°C the latter of 50°C suggesting, in correspondence with the molecular mechanics calculations, the presence and the absence of an extra base pair respectively. NMR experiments confirmed these conclusions: for the hairpin formed by d(ATCCTA-TTTA-TAGGAT) an extra iminoproton resonance is observed between 11.5 and 14.5 ppm which could be assigned to a T-A pair, for the hairpin formed by d(ATCCTA-ATTT-

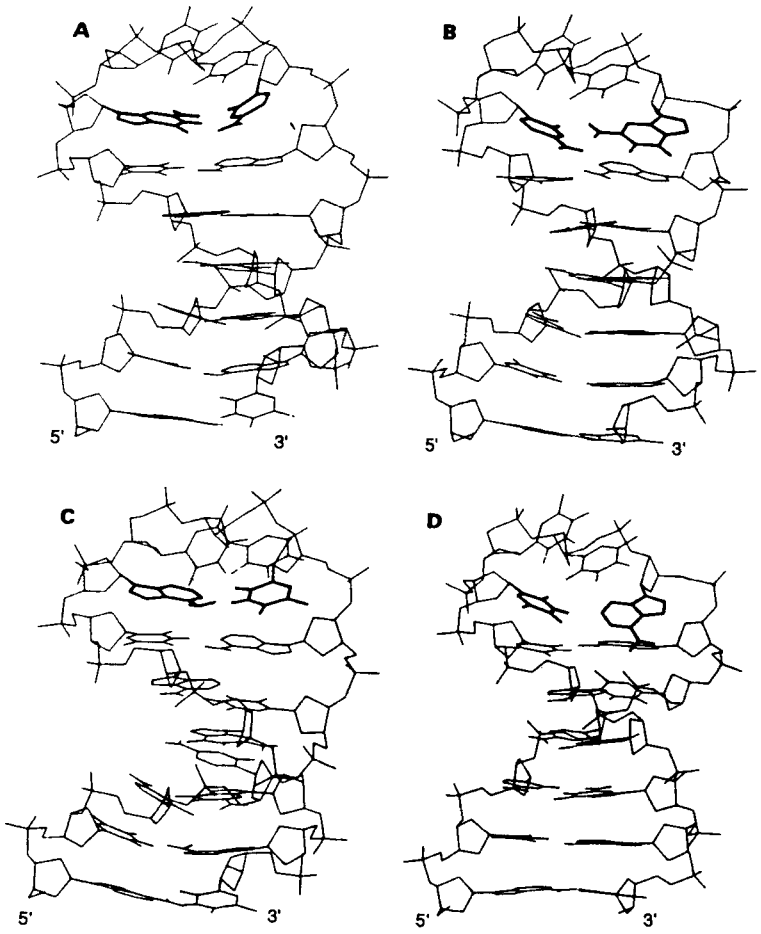

FIG. 4. Energy minimized structures of the hairpin formed by d(ATCCTA-CTTG-TAGGAT) (A), d(ATCCTA-GTTC-TAGGAT) (B), d(ATCCTA-TTTA-TAGGAT) (C) and d(ATCCTA-ATTT-TAGGAT) (D). Note that in case of the loop sequence -CTTG- and -TTTA- an additional base pair can be formed (A,C) while this does not occur for the -GTTC- and ATTT- loop sequence (B,D).

TABLE 3. Melting temperatures and enthalpies of the hairpin to coil transitions of DNA hairpins.

	oligonucleotide	Tm(°C)	ΔH (kcal/mol)
A	d(ATCCTA-TTTT-TAGGAT)	53	41
	d(ATCCTA-ATTT-TAGGAT)	50	39
	d(ATCCTA-TTTA-TAGGAT)	57	45
	d(ATCCTA-TAAA-TAGGAT)	49	36
	d(ATCCTA-AAAA-TAGGAT)	43	27
	d(ATCCTA-AAAC-TAGGAT)	47	35
	d(ATCCTA-AACA-TAGGAT)	48	34
B	d(ATCCTA-GTTC-TAGGAT)	53	38
	d(ATCCTA-CTTG-TAGGAT)	62	46
	d(ATCCTA-CTGG-TAGGAT)	57	45
	d(ATCCTA-CGTG-TAGGAT)	61	43
	d(ATCCTA-CAAG-TAGGAT)	55	41
C	d(ATCCTA-GAAA-TAGGAT)	53	35
	d(ATCCTA-GTTA-TAGGAT)	55	42

TAGGAT) such a resonance could not be found. Structural information so far available for hairpins with four nucleotides in the loop suggest that these loops are tightly packed. Introduction of the more bulky purine residues in the loop region, e.g. in our case two adenines instead of the central thymidines, is therefore expected to decrease hairpin stability. This has indeed been observed in optical melting experiments (cf. Table 3A).

A similar picture develops when one considers molecules in which the first and the last base in the loop sequence is a C and a G respectively or vice versa. Although the melting temperatures for this series of molecules are generally somewhat higher than those discussed above (cf. Table 3A and 3B). In this series we have also introduced a more bulky residue in place of one of the central thymidines. When the second base in the loop becomes a G the melting temperature hardly changes, while a G in the third position in the loop has a more important effect (cf. Table 3B). NMR experiments show that a CG base pair is formed in the loop sequences -CGTG-, -CTGG- and -CAAG-.

Sofar, the results fit into a rather simple scheme that explains the stability of DNA hairpins with four nucleotides in the loop region

and which at this point can be summarized as follows. A four-membered loop in a DNA hairpin can be formed by extending the stacking in stem in the 3'-direction by three nucleotides. Then after a sharp turn of the backbone the remaining gap can be closed by one nucleotide of which the base turns inwards into the loop. This fourth base may pair with the first in the loop . This has been observed for the -PyTTPu- but not for the -PuTTPy- loop sequence. Moreover, the possible pairing may be influenced by the character of the central bases.

We shall see, however, that this scheme has to be amended. Recently, reports have appeared in the literature about DNA as well as RNA hairpins with loop sequences which render those hairpins unusually stable. An example is provided by the sequence d(CG-GAAA-CG) which forms a hairpin with a loop sequence -GAAA- (14). This is a rather supprizing result. Within the scheme discussed above we would expect that this loop sequence with the bulky purine residues leads to a diminished hairpin stability as for instance examplified by the hairpin with four A's connecting the strands of the stem (cf. Table 3A). Subsequent melting experiments forced us, however, to revise this point of view. The melting temperature of the hairpin formed by d(ATCCTA-GAAA-TAGGAT) is 10°C higher than that of d(ATCCTA-AAAA-TAGGAT). In addition, extension of these experiments showed that the hairpin formed by d(ATCCTA-GTTA-TAGGAT) is even more stable, approaching that of the d(ATCCTA-TTTA-TAGGAT) molecule (Table 3A and 3C). Apparently, it is not the -GAAA- sequence, but the positioning of a G in the first and an A in the last position of the loop that confers extra stability to the hairpin.

To see whether this observation can still be understood in terms of the hairpin stability scheme discussed above, the hexadecanucleotide d(ATCCTA-GTTA-TAGGAT) was subjected to 2D-NMR studies. Part of its 600 MHz ^{1}H-NOESY spectrum ($\omega 1$ = 7.0-8.5 ppm; $\omega 2$ = 5.0-6.5 ppm) recorded at ambient temperatures with a mixing time of 300 ms is presented in Fig. 5. In this spectral region cross peaks are observed between the H8/H6 base- and the sugar H1'-resonances. Using the standard sequential analysis procedure (15-17) this spectrum was analyzed. When one starts at the cross peak generated by the 5'-terminal A1 residue the sequential assignment can be continued to the T8H1'-resonance; when one

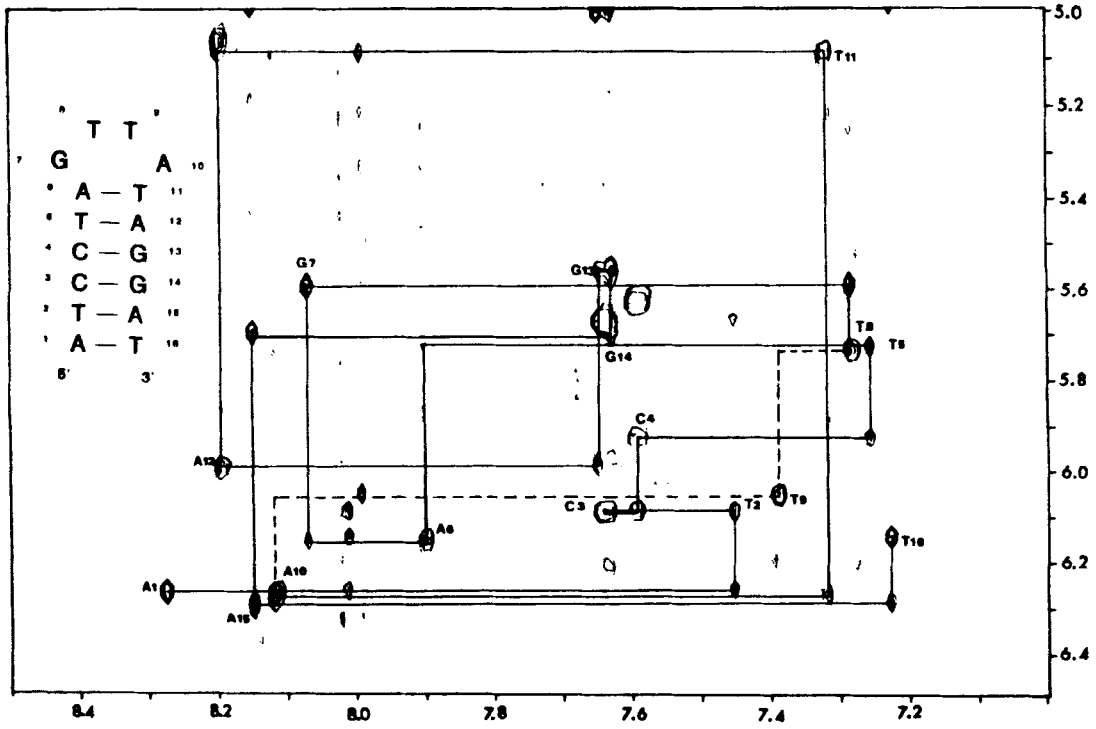

FIG. 5. Part of the 600 MHz phase sensitive 1H-NOESY spectrum (cross peaks between H6/H8 and H1'-resonances) of d(ATCCTA-GTTA-TAGGAT) Dissolved in D_2O. The mixing time was 0.3 s. The sequential analysis of the cross peaks is indicated. The drawn lines represent the connectivity diagram for cross peaks between neighbouring residues. The dashed lines show that the isolated intra-residue cross peak of T9 cannot be connected to its neighbours.

commences at the 3'-end of the sequence with the cross peak from T16, the assignments can be carried through to the intra-residue cross peak of A10. The same conclusion follows from the analysis of the cross peaks between the H8/H6 and the H2'/H2" resonances (not shown). A more detailed analysis of the H8/H6-H1'cross peak intensities shows that the intensity of the intra-residue cross peak of A10 is abnormally high. This can be seen more clearly by considering the cross sections taken along the vertical axis of the NOESY spectrum through the diagonal resonance positions of A10H8, C4H6 and A6H8 (Fig. 6). It is obvious that the intensity of the H8-H1' cross peak of A10 is much higher than

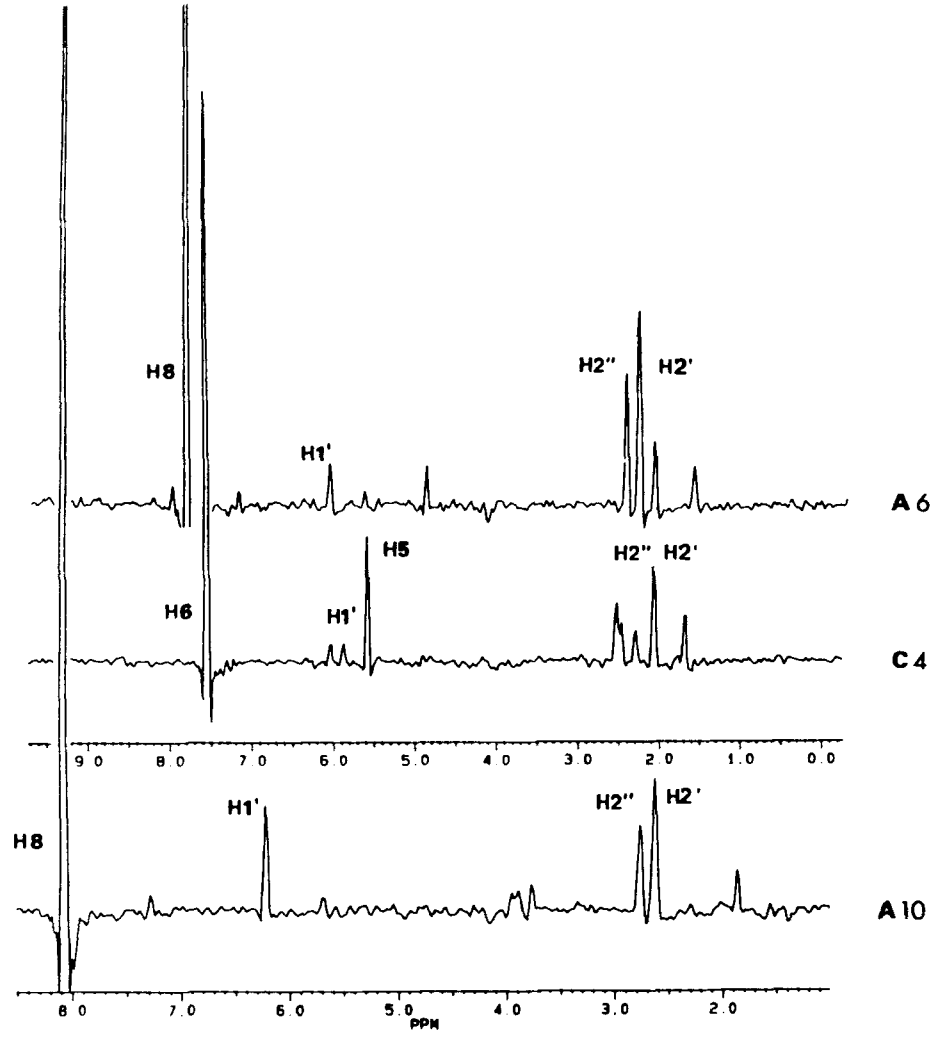

FIG. 6. Cross sections through the NOESY spectrum of Fig. 5 along F1 taken for three different F2 values sothat cross peaks from the H6/H8 ring puston resonances to other proton resonances are visible for residues A6, C4 and A10. The assignment of various cross peaks are indicated.

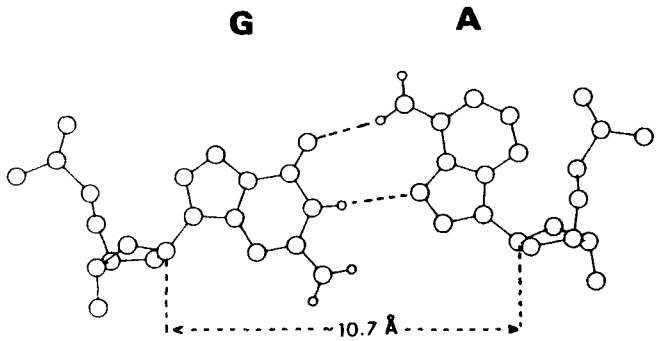

FIG. 7. Scheme of the configuration of a G(anti)-A(syn) base pair.

that of the C4 and A6 H6/H8-H1' cross peaks, but comparable to that of the H5-H6 cross peak of C4. This suggests that the distance between H8 and H1' of A10 is of the same magnitude, i.e. ≈ 2.5Å, which is possible when A10 adopts a syn-conformation. The situation is similar to that encountered in the hairpin formed by d(ATCCTA-TTTA-TAGGAT), where indeed A10 is in a syn-conformation and as a result the adenine base is in the right position to form a Hoogsteen base pair with the first thymidine in the loop (19). A similar configuration can be adopted by the -GTTA- loop sequence. With the adenine in a syn-conformation a G(anti)-A(syn) wobble pair can be formed with G7N1H connected to A10-N7 and G7-O6 connected to A10-N6H (cf. Fig. 7). Such a base pair has been observed before in the crystal structure of the double helix formed by d(CGCGAATTAGCG) (18). The sugar C1' atoms are at a normal 10.6 Å in this base pair sothat it fits in a double helix with minimal disturbances.

If we summarize the information available from our NMR experiments it can be conluded that both G7 and A10 are stacked upon the double helical stem. This together with the syn-conformation adopted by A10 makes a strong case for the existence of the G(anti)-A(syn) base pair. The formation of this base pair could explain the enhanced stability of our hairpin and would fit in the hairpin stability scheme outlined in the above. Observation of the G7 iminoproton resonance between 11.5 and 14.5 ppm would endorse this conclusion (19), but in the iminoproton spectrum of d(ATCCTA-GTTA-TAGGAT) we do

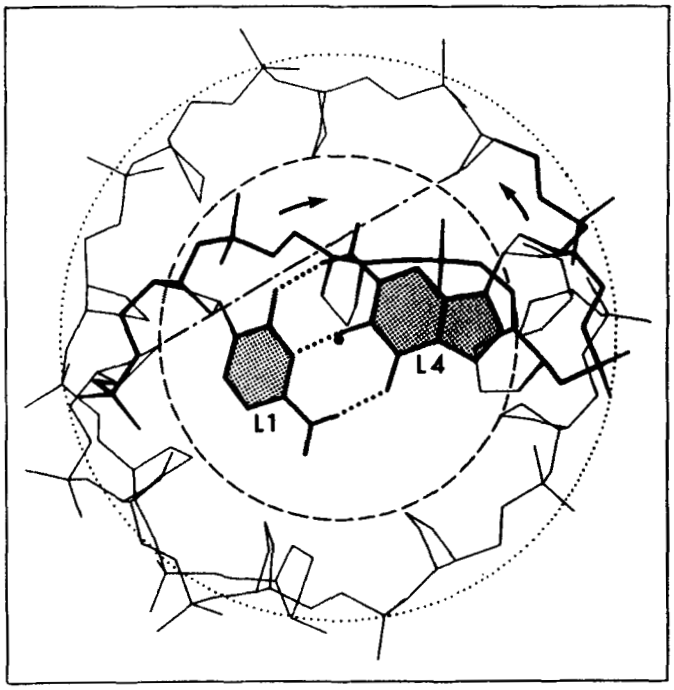

FIG. 8. View along the helix axis of the hairpin structure in Fig. 4A. For reasons of clarity all bases have been omitted except the C-G base pair which is formed in the loop and which is viewed from the side of the stem. The sugar phosphate backbone in the loop is accentuated by thick lines and the 5'-3' direction is indicated by arrows. The dotted circle is the projection of the cylinder on which in an ideal helix the phosphorous atoms are located; the dashed circle is the projection of the cylinder on which in an ideal helix the C1' atoms are located. The C1' atoms of base pair A-T (6), which are at a normal distance are connected by a dashed-dotted line. Note that the C1' of the first residue in the loop, L1, has moved away from the C1'-circle into the direction of the helix axis, indicated by the central dot.

not see an extra resonance in this spectral region. This does not, however, disprove the presence of a G-A(syn) pair. In yeast $tRNA^{Phe}$, m^2_2G26 forms a base pair with A44 on top of the anticodon stem (20). The resonance position of the $m^2_2G26N1H$ iminoproton is, however, at 10.4 ppm (21,22). A similar situation is encountered in our hairpin, where the G7 iminoproton resonance is situated at 10.8 ppm (not shown). Patel and coworkers (23) as well as Kan and coworkers (24) have observed iminoproton resonances of G-A pairs in the 11.5 - 14.5 ppm

region but these investigators showed that in their molecules both the G and the A are in an anti-conformation. Traub and Sussman have analysed the secondary structure of ribosomal RNA (25). They conclude that the G(anti)-A(syn) base pair may be a better model than the G(anti)-A(anti) base pair to fit in the end of an undistorted RNA double helical stem of a hairpin. The present experiments suggest that this is true for DNA haipins as well.

At this point we return to the question how the terminal phosphates of the base pair formed in the four-membered loop are brought sufficiently close to one another that the distance between them can be spanned by two nucleotides and why, if a base pair is formed between complementary bases (vide supra), this takes place for -PyTTPu- and is unlikely for -PuTTPy- loop sequences. We first consider a solution suggested by the energy minimization experiments performed for the hairpin formed by d(ATCCTA-CTTG-TAGGAT). As has been mentioned already, the C-G base pair formed in the loop of the calculated model is strongly buckled. The reason for this may be gleaned from Fig. 8, where the hairpin is presented in such a manner that the double helical stem is viewed along the helix axis. In a standard B-type double helix the sugar C1' atoms fall on a circle (the dashed circle in Fig. 8) and the distance between the C1' atoms of a base pair amounts to 10.7 Å. Examination of Fig. 8 shows that this is not so for the base pair formed in the loop. Although the C1' atom of the fourth residue, L4, remains on the circle, the C1' atom of the first residue, L1, does not as a result of which the two carbons approach each other to a distance of 9.7Å. This impairs the formation of a flat Watson-Crick base pair. The change in the position of the C1' atom of L1 is brought about by gradual changes in the backbone angles (with respect to a B-type backbone) between the sugars of residues A6 and L1. In turn the 3'-phosphate of L1 has moved rather close to the helix axis. In this way the distance between this phosphate and the 5'-phosphate of L4, which like its C1' atom remains at the B-type position, has diminished to such an extent that it can be spanned by two nucleotides. The bend in the buckled base pair is displaced from the dyad axis which can be defined with respect to the two C1'atoms. This introduces an asymmetry so that interchange of the bases no longer allows formation of the standard hydrogen bonds and therefore of the base pair. This is demonstrated in Fig. 9.

FIG. 9. Schematic orientation of the first (left) and the fourth base (right) in the loop to bring out that for a -CTTG- sequence these bases are in a position to form a buckled base pair (A) while for the -GTTC- sequence this is not possible.

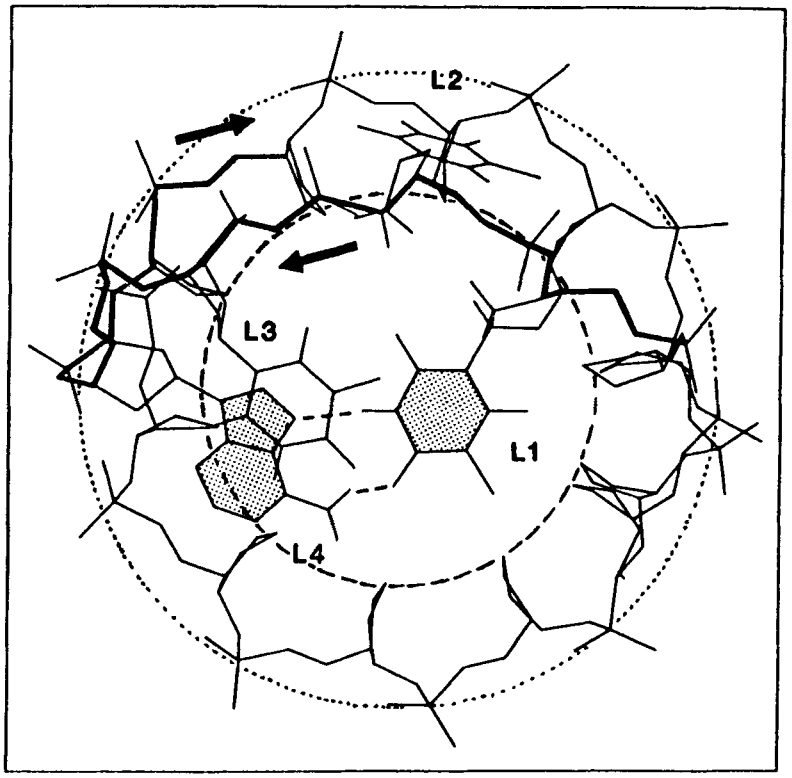

FIG. 10. View along the helix axis of the hairpin formed by d(ATCCTA-TTTA-TAGGAT). For reasons of clarity all bases have been omitted except for the bases of the loop. The first and the last base, L1 and L4, in the loop form a Hoogsteen base pair. The loop of the hairpin is pointing towards the reader. The sugar phosphate backbone in the loop is accentuated by thick lines and the 5'-3' direction is indicated by arrows. The dotted circle is a projection of the cylinder on which in an ideal helix the phosphorous atoms are located; the dashed circle is a projection of the cylinder in which in an ideal helix the C1' atoms are located.

According to the energy minimization calculations a similar picture arises for the hairpin in which a T-A base pair has replaced the C-G base pair. Detailed NMR studies show, however, that nature prefers a somewhat different solution. To permit a direct comparison of the results of these studies with the hairpin discussed in the above, the structure of the hairpin formed by d(ATCCTA-TTTA-TAGGAT) is presented in Fig. 10 in the same manner (compare Fig. 8 and 10). Examination of Fig. 8 shows that in the loop a Hoogsteen base pair is formed as was mentioned already. Moreover, residue L2 is turned into the minor groove (26) and this makes stacking pattern different from that in the hairpin d(ATCCTA-CTTG-TAGGAT) (cf. Fig. 4A). It is noted in passing, that the base stacking in this latter molecule is the same as that found for the hairpin with four thymidines in the loop by means of NMR (11). Moreover, in contrast to these molecules, in d(ATCCTA-TTTA-TAGGAT) L3(T9) is stacked upon L4(A10). This has become possible because L2 has turned into the minor groove. Yet the course of the phosphate backbone is essentially the same in both type of molecules as

In summary we find that, for the DNA hairpins for which reasonably accurate conformational data are available the backbone of the loop region follows a course which can be derived by applying the rule that the stacking of the bases in the stem is continued in the 3'-direction by 3 nucleotides after which a sharp turn in the backbone occurs sothat the loop can be closed by one nucleotide. The actual stacking of the bases does not need to follow this prescription, but depends on the base sequence in the loop.

ACKNOWLEDGEMENTS

NMR spectra were recorded at the Dutch National Hf NMR Facility at Nijmegen, which is supported by SON. Dr. S.S. Wijmenga and J.J.M. Joordens are thanked for expert technical assistance.

REFERENCES

1. Lilley, D.M.J. (1980) Proc. Natl. Acad. Sci. USA 77, 6468-6472.

2. Naylor, L.H., Lilley, D.M.J. and van de Sande, J.H. (1986) EMBO J. 5, 2407-2413.

3. Mirkin, S.M., Lyamichev, V.I., Drushlyak, K.N., Dobrynin, V.N., Fillipov, S.A. and Frank-Kamenetskii, M.D. (1987) Nature 330, 495-497.

4. Htun, H. and Dahlberg, J.E. (1988) Science 241, 1791-1796.

5. Williamson, J.R., Raghuranan, M.K. and Cech, T.R. (1989) Cell, 59, 871-880.

6. Sundquist, W.I. and Klug, A. (1989) Nature 342, 825-829.

7. Hilbers, C.W., Haasnoot, C.A.G., de Bruin, S.H., Joordens, J.J.M., van der Marel, G.A. and van Boom, J.H. (1985) Biochimie, 67 685-695.

8. Uhlenbeck, O.C., Borer, P.N. Dengler, B. and Tinoco, I. (1973) J. Mol. Biol. 73, 483-496.

9. Gralla, J. and Crothers, D. (1973) J. Mol. Biol. 73, 497-511.

10. Haasnoot, C.A.G., Hilbers, c.W., van der Marel, G.A. van Boom, J.H., Singh, U.C., Pattabiraman, N. and Kollman, P.A. (1986) J. Biomol. Struct. Dyns. 3, 843-857.

11. Blommers, M.J.J., Haasnoot, C.A.G., Hilbers, C.W., van Boom, J.H. and van der Marel, G.A. (1987) Structure and Dynamics of Biopolymers, NATO ASIS Series E: Applied Sciences 133, 78-91.

12. Orbons, L.P.M., van Beuzekom, A.A. and Altona, C. (1987) J. Biomol. struct. Dyns. 4, 965-987.

13. Haasnoot, C.A.G., Blommers, M.J.J., Hilbers, C.W. (1987) Springer Series in Biophysics 1: Structure, Dynamics and Function of Biomolecules (Eds. A. Ehrenberg, R. Rigler, A. Gräslund, L. Nilsson) 212-216.

14. Hiroa, I, Nishimura, Y. Naraoka, T., Watanabe, K., Arata, Y. and Miura, K. (1989) Nucl. Acids Res. 17, 2223-2231.

15. Haasnoot, C.A.G., Westerink, H.P., van der Marel, G.A. and van Boom, J.H. (1983) Biomol. Struct. Dyn. 1, 131-149.

16. Hare, D.R., Wemmer, D.E., Chou, S.-H., Drobny, G. and Reid, B.R. (1983) J. Mol. Biol. 171, 319-336.

17. Scheek, R.M., Russo, N., Boelens, R., Kaptein, R. and van Boom, J.H., (1983) J. Am. Chem. Soc. 105, 2914-2915.

18. Brown, T., Hunter, W.N., Kneale, G. and Kennard, O. (1986) Proc. Natl. Acad. Sci. U.S. 83, 2402-2406.

19. Blommers, M.J.J., Walters, J.A.L.I., Haasnoot, C.A.G., Aelen, J.M.A., van der Marel, G.A., van Boom, J.H. and Hilbers, C.W. (1989) Biochemistry, 28, 7491-7498.

20. Rich, A. and Raj Bhandary, U.L. (1976) Ann. Rev. Biochem. 45, 819-837.

21. Heerschap, A. Haasnoot, C.A.G. and Hilbers, C.W. (1983) Nucl. Acids. Res. 11, 4501-4520.

22. Johnston, P.D., and Redfield, A.G. (1981) Biochemistry 20, 1147-1156.

23. Patel, D.J. Kozlowski, S.A., Ikuta, S. and Itakura, K (1984) Fedn. Proc. 43, 2662-2670.

24. Kan, L.S., Chandrasegaran, S., Pulford, S.M., Miller, P.S. (1983) Proc. Natl. Acad. Sci. USA, 80, 4263-4265.

25. Traub, W. and Sussman, J.L. (1982) Nucleic Acids Res. 19, 2701-2708.

26. Blommers, M.J.J., van de Ven, van der Marel, G.A., van Boom, J.H. and Hilbers, C.W. (1990) Eur. J. Biochem., submitted.