Biophysical Chemistry, 45 (1992) 143–152
Elsevier Science Publishers B.V., Amsterdam

BIOCHE 01705

A computational and experimental study of the bending induced at a double–triple helix junction

Jacques Chomilier [a,1], Jian-Sheng Sun [a], David Andrew Collier [a,2], Thérèse Garestier [a], Claude Hélène [a] and Richard Lavery [b]

[a] *Laboratoire de Biophysique, CNRS UA 481, INSERM U 201, Muséum National d'Histoire Naturelle, 43 rue Cuvier, 75 231 Paris Cédex 05, (France)*
[b] *Laboratoire de Biochimie Théorique, CNRS UA 77, IBPC, 13 rue Pierre et Marie Curie, 75005 Paris, (France)*

(Received 6 February 1992; accepted in revised form 23 July 1992)

Abstract

We have studied the conformation of a 17 base-pair homopyrimidine · homopurine triple helix formed on a fragment of duplex DNA derived from Simian Virus SV40. Gel retardation assays indicate that an 80 base-pair fragment has an altered conformation when the triple helix is formed, which is most likely to result from an induced bend in the DNA. Investigation of the detailed conformation of the double helix–triple helix junctions has been performed by means of molecular modelling. Bending on the 5′ and 3′ sides of the third strand oligonucleotide are not located at equivalent positions with respect to the junctions, which is explained in terms of base stacking. The junction effects on DNA structure, induced by the requirement for cytosine protonation in the Hoogsteen-bonded strand to form CGC^+ base triplets, are also discussed.

Keywords: Molecular modelling; Triple helix; DNA bending; SV40

1. Introduction

Sequence-specific recognition of the major groove of double-stranded DNA at homopurine · homopyrimidine sequences can be achieved by homopyrimidine oligonucleotides [1,2]. Thymine and protonated cytosine form Hoogsteen-type hydrogen bonds with Watson–Crick A · T and G · C base pairs, respectively, and lead to triple helical structures. The homopyrimidine oligonucleotide is bound in a parallel orientation with respect to the homopurine-containing strand of duplex DNA. Oligonucleotide binding to the major groove of DNA has the potential to interfere with biological processes, e.g. by preventing binding of regulatory proteins that are involved in controlling gene expression. It has been shown recently that triple helix formation could inhibit restriction endonuclease cleavage and transcription factor binding at specific sequences [3,4]. Intramolecular triple helix (i.e. H-DNA) formation has also been reported in linear DNA or supercoiled plasmids [5].

Junctions between double and triple helix were experimentally shown to be strong binding sites for intercalating drugs [6]. Intercalation is much

Correspondence to: Dr. Jian-Sheng Sun, Laboratoire de Biophysique, CNRS UA 481, INSERM U 201, Muséum National d'Histoire Naturelle, 43 rue Cuvier, 75 231 Paris Cédex 05, France.

[1] Permanent address: Laboratoire de Minéralogie-Cristallographie, CNRS UA 09, Université Paris VI, 4 Place Jussieu, 75 252 Paris Cédex 05, France.

[2] Permanent address: Institute of Psychiatry, Section of Genetics, De Crespigny Park, London SE5 8AF, UK

stronger on the 5′ than on the 3′ end of the homopyrimidine oligonucleotide. In this study, we have examined the conformation of the duplex–triplex junctions by molecular modelling. This was aimed at understanding the asymmetry of the duplex–triplex junctions and determining the effect of the junction on the global bending of the DNA fragment.

The sequence studied was a fragment of SV40 viral DNA including a 17 base pair homopurine · homopyrimidine segment from positions 4424 to 4440. Several sequence alterations have been examined by molecular modelling in order to determine the effect of sequence variation on the bending of the DNA. Gel retardation experiments provide direct evidence for triple helix-induced DNA-bending.

In this paper, strands I and II will refer to the 17 base-long homopyrimidine and homopurine sequences, respectively, of the duplex DNA target; strand III will be the homopyrimidine oligonucleotide involved in Hoogsteen base pairing with strand II. For simplicity, nucleotides will be numbered by reference to strand II, whose sequence is as follows:

5′TCGTTT **AAAAAAGAAGAGAAAGG**
TAGAAG3′

To accommodate strand III, the major groove of the target must be sufficiently wide, so an A-like DNA is an acceptable conformation for the triple helix [7]. The double helix part has been chosen to be the B-DNA standard conformation proposed to occur under physiological conditions. Even if it is not likely to be the optimal conformation, it is a reasonable starting structure and it was verified that starting from an A-form in the double helical segment did not change our general conclusions.

2. Methods

2.1 Gel retardation experiments

We have used a gel retardation assay [8–10] to test the ability of triple helix formation to induce a bend in the target duplex DNA. Bent DNA migrates more slowly in polyacrylamide gels than straight DNA, the extent of retardation being related to the extent of bending. In addition, gel retardation is least when the bend occurs close to the end of the fragment under study and is most when the bend is located at the center. Thus the movement of the locus of a bend through a fragment of DNA, by tandem digestion with restriction enzymes, can provide information on the position of curvature. We have performed this experiment using the 17-bp target for triple helix formation from SV 40, cloned into a plasmid (pBluescript-), such that it is flanked by multiple restriction sites (Fig. 1). The plasmid (pDC1) was constructed as described [11]. Complementary oligonucleotides were synthesised so that when annealed, HindIII compatible ends were generated. The resulting doule stranded oligomer was then ligated into the unique HindIII site of pBluescript KS-, and recombinant plasmid prepared by standard procedures. The oligonucleotides used had the sequence

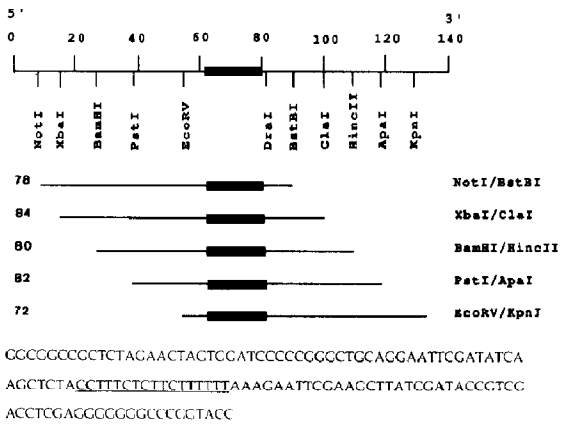


Fig. 1. Diagram showing the construct used for gel retardation experiments. The black box represents the site of triplex formation. The various restriction fragments used for the gel mobility assay are shown with their length on the left. The sequence of the DNA fragment used in the gel mobility experiments is shown below the figure, with the triple helix target underlined.

5′AGCTTCGAATTCGTTTAAAAAAGAAGA-
GAAAGGTAG3′

and

$^{5'}$AGCTTCTACCTTTCTCTTCTTTTTTAAAC-GAATTCGA$^{3'}$

For DNA bending analysis, pDC1 was digested with a pair of restiction enzymes, precipitated with ethanol and resuspended in TE buffer. Restriction fragments were then incubated in the presence or absence of triple-strand-forming oligonucleotides in TA buffer (50 mM Tris-acetate, pH 6.0, 10 μM spermine) at 30°C for 30 min. The mixtures were then applied to a pre-cooled 20% polyacrylamide, 1% bisacrylamide gel in TA buffer (pH 6.0) and electrophoresed at 4°C for 16–20 h. DNA was visualized by staining with ethidium bromide.

2.2 Molecular modelling

Conformational energy minimization was performed with the JUMNA program [12] (version IIIE) which is specifically designed for studying nucleic acids and directly employs a combination of helical and internal coordinates. The force field employed has been described in detail elsewhere [13]. It is remarked that to model solvent effects a sigmoidal distance dependent dielectric is used [14] and counterion damping is modelled by reducing phosphate net charges to -0.5 electron. New constraints have been included in JUMNA IIIE: one can now force the nucleic acid to bend around a cylinder of a chosen radius of curvature and position by including a Lennard–Jones potential representing the interaction between the cylinder and DNA in the force field.

Charges were calculated with a Hückel–Del Re procedure [15] which determined monopoles fitted to quantum mechanical results. Cytosines belonging to strand III were protonated in order to allow two hydrogen bonds with guanines in a Hoogsteen orientation.

Starting coordinates for the complexes studied were obtained from crystallographic data on triple helix fibres [16] or on B-DNA [17]. The triple helix fibre conformation which is closer to the A-form than to the B-form (typical $X_{disp} = -3.5$ Å) is termed the T-form throughout this paper.

To facilitate junction formation between the B-form and the T-form, the rise was doubled, and the twist accordingly reduced. The first step in the minimisation procedure consisted of locking the sugar and helical parameters for all nucleotides, leaving only the backbone flexible. In the second step only X and Y displacements were minimized, while at the junction the rise and twist were relaxed. The third step allowed rise and twist to change for all nucleotides and the fourth step released X and Y displacement, but locked inclination and tip. Finally, all the conformational parameters were minimized. To ensure that the minimized conformations were stable, a number of starting structures were used and the complexes were also checked by over- and under-twisting deformations.

For reasons of computational cost, the complex formed by the 17-mer oligonucleotide and its 29-mer target was split into three parts. The first is denoted as the "5′-junction" and is composed of six triplets and six base pairs, long enough to avoid end effects. The sequence of its strand II (following our convention, see introduction) is $^{5'}$TCGTTTAAAAAA$^{3'}$ and the junction is localized between T6 and A7. The "3′-junction" is also composed of six base pairs and six triplets, whose sequence in strand II is $^{5'}$GAAAGG-TAGAAG$^{3'}$, with the junction between G6 and T7. The five nucleotides in the centre of the triple helix between the 5′- and the 3′-junctions were also minimised and two nucleotide triplets were added at each end in order to avoid any end effects in the region of interest. The sequence of this triple helix segment was thus $^{5'}$AAGAAGA-GA$^{3'}$.

Calculations were carried out on a Silicon 4D/120 workstation using the Insight software from Biosym, fully interfaced with JUMNA. Calculations took roughly 30 seconds per iterations and roughly 600 cycles were required for convergence.

Analysis of the minimized conformations was carried out with the CURVES algorithm [18,19] which respects the conventions laid out at Cambridge in 1988 [20]. To determine the site where the molecule bends, we paid particular attention to the roll parameter, since bending is directly related to roll. For a more global approach we

also calculated the scalar product of the normals of the first and the last base pairs of the molecules studied, which led to a rough estimate of the global angle of curvature of the DNA fragment.

3. Results and discussion

Figure 2 presents a gel and a plot of the relative mobility versus the central base-pair position of the restriction fragment, as derived from gel retardation assays. The six fragments presented here have about the same length (around 80 nucleotides) and they all contain the triple helix at various loci (see Fig. 1). As expected, triple helix formation resulting from binding of the 17 nucleotides long strand III, decreases the mobility of the DNA fragments. The K value plotted in Fig. 2 is the ratio of the mobility of the fragment with the triple helix to that of the duplex alone, and is representative of the retardation independently of the length of the fragment. This plot indicates that the mobility of the fragment is lowest when the triple helix is near the center of the restriction fragment. This is diagnostic of an induced bend in the duplex DNA upon triple helix formation, since highest retardation occurs with the bend at the center of the fragment. The induced bend is at the triple helix site, although resolution is not sufficient to determine at which junction the bend occurs, or whether both junctions are involved.

To study the stability of the complex formed by the double helix and the oligonucleotide, Fig. 3a presents the base stacking energy of the base pairs or base triplets as a function of the nucleotide number; this notation follows the 5′ to 3′ direction of strand II. Base stacking energy, between base pairs and/or base triplets, has been chosen to illustrate the stability of the complex because the main contribution to the complexation energy is due to interaction between bases of various strands. Actually, with JUMNA, the energy was decomposed for each nucleotide in three parts (base, sugar and phosphate group) and the contribution for each strand was also separated. So one is able to calculate the energy between all bases of one strand as well as the energy between the bases of two strands. Within any single strand, base–base contribution to the energy is negative, although sugar–base and sugar–phosphate interactions are more stabilizing. But when considering interstrand contributions (I and II, II and III, I and III) base–base interactions are the first contribution to the stabilizing energy.

The total 29-mer sequence of the target, too long to be computed within a single simulation, was artificially rebuilt by "manually" linking the

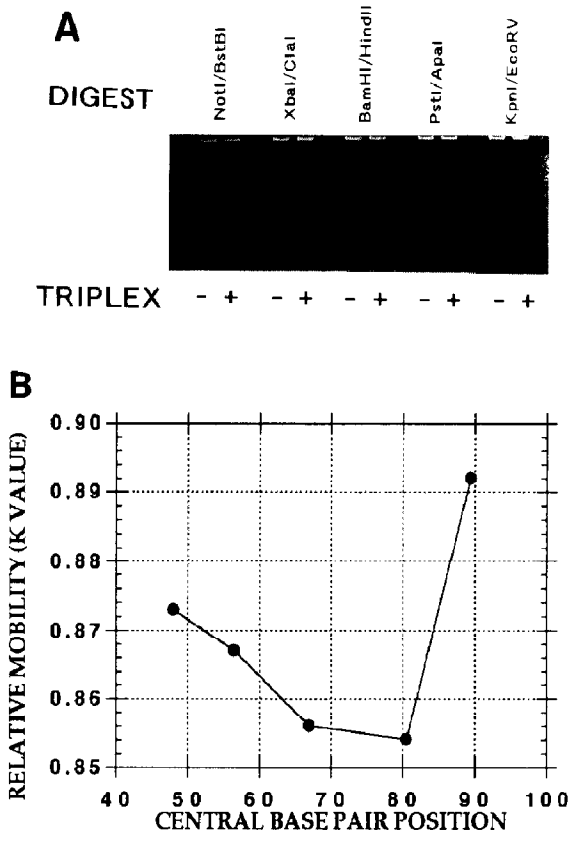


Fig. 2. Relative mobility of the restriction fragments, all containing the 17 bp target for the oligonucleotide at various loci, as a function of the position of the central base pair of the restriction fragment relative to the start of the polylinker sequence of the plasmid. (a) Typical polyacrylamide gel showing migration of the duplex DNA with (+) or without (−) the triplex formed at the homopurine·homopyrimidine segment. The direction of electrophoresis is from top to bottom. The upper band of the two in the lane marked PstI/ApaI is an artifact of the low temperature gel and should be ignored. (b) Plot of K, the ratio of the mobility of the fragment with the triple helix formed, with respect to the duplex alone. Its value is an indication of the extent of bending of the fragment.

5′- and 3′-junctions, each one being composed of 12 mers, to a triple stranded 9-mer whose two last triplets at each end were removed when analyzed by CURVES. This splicing was intended to avoid end effects that might occur because each part of the molecule was separately minimized, although under the same conditions. The mean values of the stacking energies were −10 kcal/mol for the double helices, regardless of the sequence, and between −13 and −25 kcal/mol for the triple helix, showing strong sequence dependence.

At the 5′ end of the oligonucleotide the transition from a stacking energy typical of a double helix to a value typical of a triple helix occurs exactly at the 5′-junction between the last base pair of the double helix (number 6) and the first base triplet of the triple helix (number 7, Fig. 3a), as if there were two triplets following each other. This indicates that the 5′ end of the oligonucleotide places itself in a position that maximizes the stacking with the last base pair of the double helix. The 5′-junction is represented as a stereo view on Fig. 4.

Inside the triple helix, the stacking energy is the strongest at the position where there is a protonated cytosine on strand III. We checked, by studying various sequences of the triple helix, that this systematic stacking enhancement occurs at each step where a protonated cytosine is preceded by or followed by a thymine along strand III. The stacking energies for successive triplets TAT/CGC$^+$, CGC$^+$/TAT, TAT/TAT and CGC$^+$/CGC$^+$ are respectively around −24, −20, −16 and −13 kcal/mol.

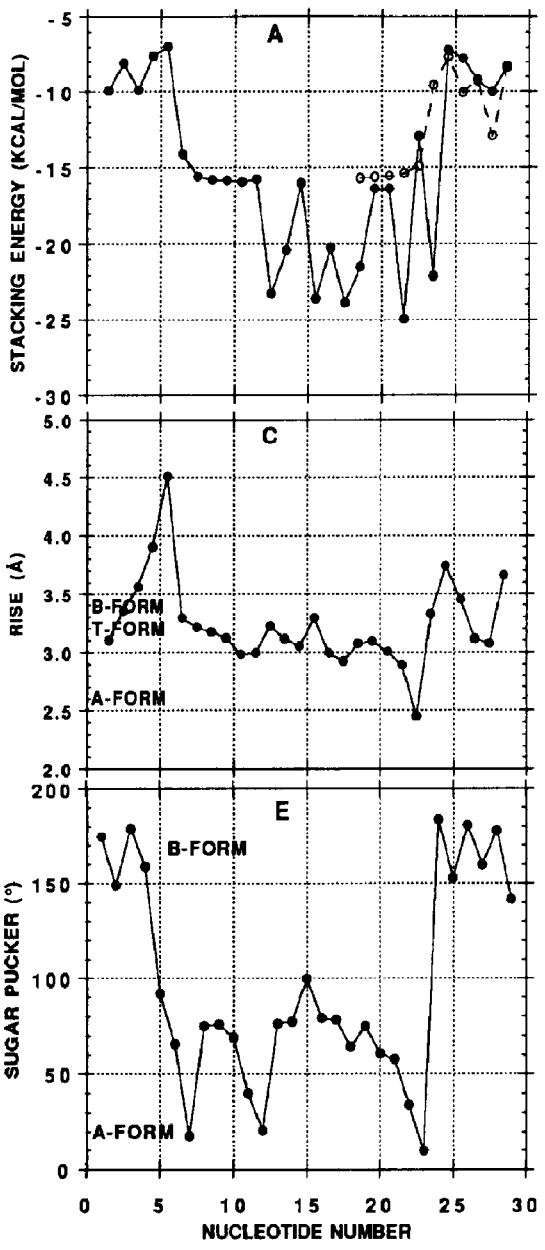
For the 3′ end of the oligonucleotide, there is an enhancement of the stacking energy at the junction (CGC$^+$/TA) to −22 kcal/mol, more negative than the value for an equivalent sequence within a triple helix (CGC$^+$/TAT). To separate the effect of sequence from the effect of junction upon the stacking energy, a 3′-junction whose sequence is $^{5'}$AAAAAATTTGCT$^{3'}$ (instead of $^{5'}$GAAAGGTAGAAG$^{3'}$), i.e. a reversed sequence as compared to the 5′-junction, was minimized; its stacking energy is also reported on Fig. 3a. One observes the same stacking energy between all TAT base triplets, as it is for the 5′ junction. At this new 3′-junction, the stacking energy for TAT/TA is now −10 kcal/mol, typical of a double helix, and is weaker than the stacking at the 5′-junction; thus junctions are not equivalent from the point of view of base stacking. As the first nucleotide of strand II after the 3′-junction is the same for both sequences, one can deduce that the reduction of stacking for TAT/TA compared to CGC$^+$/TA, is not due to the junction, but to the sequence.

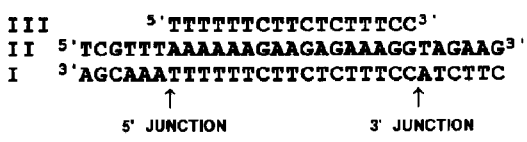
There is clearly an asymmetry in the stacking interactions at the two junctions, the base pair neighbor to the triple helix behaving as a triplet at the 5′-junction and as a doublet at the 3′-junction. This difference could find its origin in the right-handed twist of DNA. For the 5′-junction, because of the direction of twist, the first base of strand III tends to stack with the upstream base of strand II. At the 3′-junction, the opposite situation occurs and the last base of strand III is not stacked with the next downstream Watson–Crick base pair but rather located without neighbor in the major groove. Nevertheless, the presence of a net positive charge on the last cytosine of strand III, producing strong electrostatic interactions included in the base stacking energy, finally results in an enhancement of the stacking energy for the native sequence.

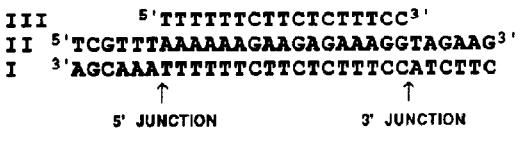
This asymmetry is intimately connected to the deformation of the molecule. Actually the kinks occuring at the 5′- and 3′-junctions, estimated by the angle of curvature deduced from the scalar product of the normals of the first and last base pairs or triplets, are 28° and 26°, respectively. They are identical within the accuracy of our calculations. The details of the structural conformations were further refined, by using some pertinent parameters (Rise, Twist, X_{disp} and Roll as well as sugar pucker), also reported on Fig. 3 as a function of the nucleotide number of strand II. Note that this analysis, performed with CURVES, determines a global axis by a minimization procedure which includes strand III when present. X_{disp}, Rise, Twist and Roll were calculated for base pairs of strands I and II, and sugar pucker is given for bases of strand II. Therefore, these parameters show the modifications induced by strand III on the configuration of the target sequence. As for the stacking energy, 5′- and 3′-

junctions (six base pairs and six base triplets each) have been artificially connected, i.e. a triple stranded DNA was inserted between them. Typical values of these parameters for canonical A-, B- and T-form are also given for comparison.

When considering base pair X displacement,







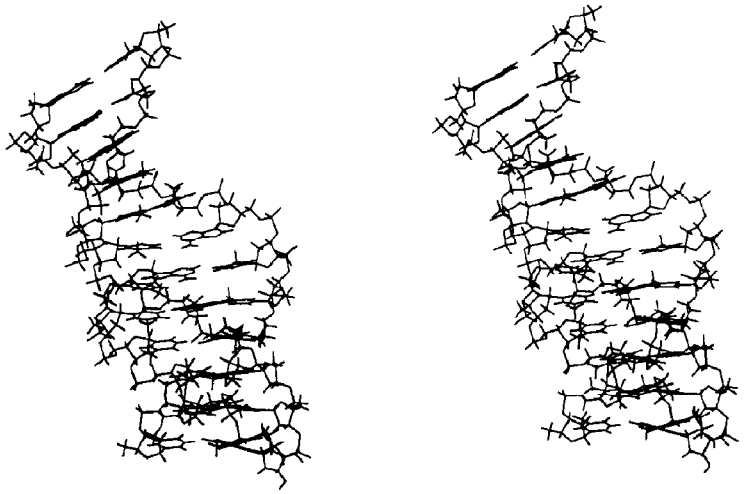


Fig. 4. Stereo view of the 5′-junction. The sequence on strand II is: $^{5'}$**GAAAGG** TAGAAG$^{3'}$.

the range of values is fairly large, extending from −1 Å, typical of a B-form of DNA, up to −3.5 Å, typical value for a T-form. Whichever junction is considered, there is a sharp transition in the values of X_{disp}, but its location is not the same according to the junction. As the presence of strand III tends to favor base stacking at the 5′-junction, X_{disp} at position 6 has the value of a typical triple helix. This means that the base pair before the junction has adjusted to enhance stacking with the first base triplet. But at the 3′-junction, as there is no change in stacking due to the presence of strand III, base triplet 23 has returned to the value of X_{disp} typical of a double helix.

The inter base pair rise is fairly constant (between 2.9 and 3.4 Å) for most of the sequence, lying in between typical values for A (2.6 Å) and B-DNA (3.4 Å). At position 5.5, i.e. in between base pairs number 5 and 6 (the two last pairs of the double helix upstream of the 5′-junction), the rise strongly escapes from canonical values. This can be explained by the presence of the third strand starting at position 7 which gives rise to an enhanced stacking of base pair 6 on the first triplet; combined with an inversion of rolls between positions 6 and 7 (see Fig. 3f), this tends to increase the distance between the penultimate and the last base pairs of the double helix. At the 3′-junction, we already discussed that the en-

Fig. 3. Effect of triple helix formation on some helical parameters or energy contribution as a function of the nucleotide number along strand II of the 29 mer double-stranded target. Third strand runs from positions 7 to 23. The whole molecule has been split into three parts in order to be able to minimize it with NUMNA: 5′-junction (12-mer), triple helix (5-mer) and 3′-junction (12-mer). (a) Stacking energy: filled circles correspond to the sequence shown below the figure. The open circles represent the calculated stacking energy for a triple helix that would have the following sequence of strand II around the 3′-junction: $^{5'}$AAAAAATTTGCT$^{3'}$ (instead of $^{5'}$GAAAGGTAGAAG$^{3'}$), i.e. a reversed sequence as compared to the 5′-junction. (b) X displacement of base pairs of strands I and II; negative values mean a displacement of the bases towards the minor groove. (c) Rise between two successive base pairs of strands I and II. (d) Twist between two successive base pairs of strands I and II. (e) Sugar conformation expressed in terms of the pseudo rotation angle of the bases of strand II. (f) Roll angle between successive base pairs of strands I and II. A positive value corresponds to a bending directed towards the major groove. Filled circles correspond to the 29-mer native sequence reported below the figure, opened circles to a 3′-junction whose sequence is reverse of the 5′-junction (same as for Fig. 3a). Two single mutations at position 5 of the 5′-junction are also plotted: cytosine (open triangles) and adenine (open squares). For stacking energy, rise, twist and roll, half integer values of the nucleotide number are given, as these parameters are measured between two successive base pairs. Canonical A-, B- and T-forms are indicated in the plots of the helical parameters, except for the X displacement of the A-form, which is −5.3 Å.

hanced stacking of the first base pair on the last triplet (position 23.5 of Fig. 3a) is mostly due to the presence of the protonated cytosine at the last position of strand III. In addition with the absence of large change in rolls at the 3′-junction (see Fig. 3f), successive base planes remain rather parallel and this does not lead to large deviation of rise from canonical values.

Twist also presents a rather constant value (around 31°) all along the triple stranded helix, in between T- and A-forms. In the double helix, twist is larger, as it is for B-DNA, but not as regular as in the triple helix, indicating a more flexible behavior of double-stranded DNA. Contrary to what one might expect, an increase in rise (for instance at positions 5.5 or 24.5, see Fig. 3c) is not systematically compensated by a decrease in twist. The only case when one can find such a correlation is at 22.5, where the presence of one positive charge on each of the two last cytosines of strand III could compensate the small rise by a rather large twist because of electrostatic repulsions. The asymmetry is still present for twist when considering both junctions. The 5′-junction, except for a very localized increase at position 2.5, is globally between A- and B-forms, while twist in the double helix at the 3′-junction is larger than the canonical B-form value. This indicates that the presence of strand III, leading to small values of twist in the T-form, affects more the twist in the double helix of the 5′-junction than of the 3′-junction. The large change in twist occurring at the 3′-junction between positions 23.5 and 24.5 in Fig. 3d, is correlated to the stacking energies at these two positions, not compensated by a particular roll. So, the enhancement of base stacking at the junction is accompanied by a decrease in twist.

Sugar conformations, presented only for nucleotides of strand II on Fig. 3e because this strand is the most sensitive to the presence of strand III, present transitions between double and triple-stranded regions. The double helix at the 5′-junction has a puckering typical of a triple helix (C3′-endo) two bases before the triple helix starts, but, at the 3′ junction, the first base of the double helix is already in the conformation typical of a double helix (C2′-endo). Once again, the presence of strand III has a larger effect upon the double helix at the 5′ junction than at the 3′-junction. Comparison with NMR results is difficult because experimental data are not in agreement. According to some authors [21,22] sugars in the purine strand adopt a C2′-endo conformation, while other results [23,24] favour a C3′-endo conformation, like our prediction. Fourier Transform Infrared experiments [25] performed on triple stranded dT · dA · dT, give a new insight in this controversy. Sugar conformation seems to be correlated with hydration, starting from a C2′-endo type at high humidity, and moving towards C3′-endo conformation when it is decreased.

Our results confirm an asymmetry between the 5′- and 3′-junctions and one must now examine the flexibility of the molecule. As already mentioned, the roll plays an important role, closely related to the curvature of nucleic acids. Figure 3f shows the roll for successive base pairs of strands I and II of the native sequence. Fluctuations are important in the double helix part, but almost negligible inside the triple helix. An important variation between successive values of roll occurs between base pairs 6 and 7, i.e. at the 5′-junction; this change of roll is around 25°. As the two successive rolls are of opposite sign, the induced bending is almost compensated ($-10°$ towards the minor groove and $+15°$ towards the major groove) so the overall curvature, measured at the ends of the 12-mer 5′-junction, is not very large (28°) and directed towards the minor groove. There is no important change of roll between successive base pairs at the 3′-junction, but a succession of decreasing rolls finally gives rise to a global angle of curvature bending towards the major groove, which is about the same order as that of the 5′-junction.

Deducing the direction of bending from the examination of the rolls is difficult with irregular nucleic acids. All one can argue is the following: as the global changes of roll at the 5′- and 3′-junctions are of opposite sign, the subsequent bending at these junctions occurs towards different grooves. The fact that the oligonucleotide is 17 nucleotides long, corresponding to one and a half turns around the helical axis, favors constructive overall bending. Actually, the minor groove

of the double helix at the 5′ end of the oligonucleotide projects itself exactly on its major groove at the 3′-junction, because of the helical symmetry. So, bending towards opposite grooves at the two ends of the oligonucleotide means, in three dimensional space, an additive bending, thus enhancing the global curvature of the molecule.

To test if there were any effect of local sequence upon curvature, several sequence changes were analyzed by molecular modelling. The most pertinent one, for the 5′-junction, occurs at the fifth position and Fig. 3f presents the rolls for the native sequence TTT, and two single mutations TCT and TAT at positions 4 to 6. The global angle of curvature for these structures is 28° for TTT, and 24° and 18° for TCT and TAT, respectively. An extra check of this curvature was done; starting form the fully minimized native sequence, the last step of minimization, when releasing all the remaining constraints, was redone but a single mutation (C or A instead of T at position 5) was introduced. Although the distribution of roll is rather similar (changes of less than 5°, see Fig. 3f), the effect of a single mutation within the double-stranded part of the 5′-junction gives rise to a significantly different global curvature. This tends to demonstrate that, unless the global curvature is due to a very localized kink, the knowledge of roll is not sufficient to quantitatively predict the bending because from one base to the next one, the direction of bending is rotated by the value of the twist.

Within the triple helix, mutation from T to C^+ at the 5′-junction has almost no effect on conformation (i.e. rolls are similar) or curvature (angles range from 28° to 31°). At the 3′-junction, bending is enhanced by the presence of the charge on the protonated cytosine. This has been studied by changing the sequence $^{5'}$**GAAAGGTAGAAG**$^{3'}$ to $^{5'}$**AAAAAA**TTTGCT$^{3'}$, i.e. the reverse (5′–3′ instead of 3′–5′) of the sequence at the 5′-junction.

This new 3′-junction gives a global angle of curvature of 17°. Mutation of the penultimate or last thymine at the 3′-junction to a protonated cytosine does not significantly increase the curvature. But two successive C^+ at the 3′ end of strand III lead to a curvature of 24°. In contrast, two neutral cytosines in strand III (so with only one hydrogen bond ensuring the Hoogsteen base pairing) lead to the smallest curvature, 13°. If one looks at the largest change of roll, i.e. between positions 24 and 25 of the 3′-junction (Fig. 3f), there is no correlation between the value of this change and the global curvature.

Besides the effect of sequence alteration on curvature, we also studied the flexibility of the molecule, i.e. the ability to bend at the lowest cost in energy. For this purpose, a cylinder is placed perpendicularly to the global axis of the fragment. The Lennard–Jones potential of this cylinder forces the DNA to adopt the curvature of the radius of the cylinder. This radius was varied from 500 Å to 50 Å and we studied the variation of the final minimized energy as a function of this radius of curvature. Our aim was to analyze the effect of a single mutation on the energy cost associated with a given curvature. To avoid protonated cytosines, only the 5′-junction calculations were performed, for the two mutated double helices ($^{5'}$TCGTATAAAAAA$^{3'}$ and $^{5'}$TCGTCTAAAAAA$^{3'}$) and the native ($^{5'}$TCGTTTAAAAAA$^{3'}$) sequence.

In order to compare the flexibility of the various sequences, Fig. 5 presents the overall angle of curvature as a function of ΔE, the difference

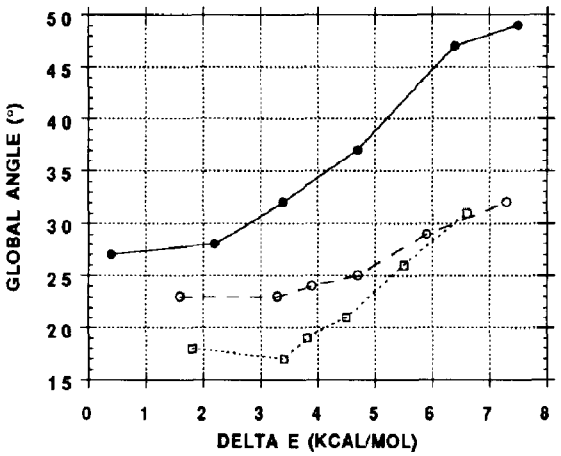


Figure 5: Angle of curvature of the 5′-junction as a function of ΔE, the energy change between the final minimized energy when curvature is imposed and the equivalent final energy for the unconstrained junction. Native sequence (— and filled circles) and two mutations in the double helix part corresponding to TCT and TAT at positions from 4 to 6 (--- and open circles) and (··· and open squares).

between the final minimized energy when curvature is imposed and the equivalent final energy for the unconstrained junction.

The first observation is that for small values of the deformation energy there is a plateau for each molecule, typically, less than 2 kcal/mol for the native and 3 kcal/mol for the mutated sequences. Beyond this value, the larger the angle of curvature, the larger the deformation energy, so bending becomes more and more costly. The second point is that, for a given deformation energy, the native sequence always produces a larger angle of curvature, indicating a higher flexibility.

4. Conclusion

In this paper we have studied a fragment of Simian Virus SV40 containing a 17 base-pair homopurine · homopyrimidine sequence which is a target for a 17-mer homopyrimidine oligonucleotide, forming a triple helix. Gel retardation experiments indicate that triple helix formation may induce a bend in the DNA. Molecular modelling calculations are consistent with gel retardation assay and further characterize the bending induced by triple helix formation. Because of the direction of the helical twist, 5′-junction and 3′-junction are not equivalent. At the 5′ end of the oligonucleotide an enhanced stacking of the last base pair occurs, which is not the case at the 3′ end, except when a protonated cytosine is placed at the last position of the triple helix. Sequence alterations performed in the calculation showed that the native sequence was the most favorable for bending. This may be a point of interest for the control of biological processes. Actually one can expect that, according to the sequence of the oligonucleotide, triple helix formation can induce a bend or abolish an existing curvature on a DNA fragment.

Acknowledgments

This work was supported in part by Rhône-Poulenc. One of us (D. A. C.) wishes to thank the EMBO for a long term fellowship (ALTF89–119).

References

1 T. Le Doan, L. Perrouault, D. Praseuth, N. Habhoub, J.-L. Decout, N.T. Thuong, J. Lhomme and C. Hélène, Nucleic Acids Res. 15 (1987) 7749–7760.
2 H.E. Moser and P.B. Dervan, Science 238 (1987) 645–650.
3 J.-C. François, T. Saison-Behmoaras, N.T. Thuong and C. Hélène, Biochemistry 28 (1989) 9617–9619.
4 L.J. Maher III, B. Wold and P.B. Dervan, Science 245 (1989) 725–730.
5 R.D. Wells, D.A. Collier, J.C. Hanvey, M. Shimizu and F. Wohlrab, FASEB J. 2 (1988) 2939–2949.
6 L. Perrouault, U. Asseline, C. Rivalle, N.T. Thuong, E. Bisagni, C. Giovannangeli, T. Le Doan and C. Hélène, Nature 344 (1990) 358–360.
7 H.W.T. Van Vlijmen, G.L. Ramé and B.M. Pettit, Biopolymers 30 (1990) 517–532.
8 J.C. Marini, S.D. Levene, D.M. Crothers and P.T. Englund, Proc. Natl. Acad. Sci. USA, 79 (1982) 7664–7668.
9 H.M. Wu and D.M. Crothers, Nature 308 (1984) 509–513.
10 K. Zahn and F.R. Blattner, Nature 317 (1985) 451–453.
11 D.A. Collier, J.-L. Mergny, N.T. Thuong and C. Hélène, Nucleic Acids Res. 19 (1991) 4219–4224.
12 R. Lavery, Junctions and bends in nucleic acids: a new theoretical modeling approach, in: Structure and Expression vol 3: DNA Bending and Curvature, eds. W.K. Olson, M.H. Sarma, R.H. Sarma and M. Sundaralingam, (Adenine Press, Guilderland, NY, 1988) pp. 191–211.
13 R. Lavery, H. Sklenar, K. Zakrzewska and B. Pullman, J. Biomol. Struct. Dynam. 3 (1986) 989–1014.
14 B. Hingerty, R.H. Richie, T.L. Ferrel and J.E. Turner, Biopolymers 24 (1985) 427–439.
15 R. Lavery, K. Zakrzewska and A. Pullman, J. Comput. Chem. 5 (1984) 363–373.
16 S. Arnott, P.J. Bond, E. Selsing and P.J.C. Smith, Nucleic Acids Res. 3 (1976) 2459–2470.
17 S. Arnott, R. Chandrasekaran, D.L. Birdsall, A.G.W. Leslie and R. L. Ratliff, Nature 283 (1980) 743–745.
18 R. Lavery and H. Sklenar, J. Biomol. Struct. Dynam. 6 (1988) 63–91.
19 R. Lavery and H. Sklenar, J. Biomol. Struct. Dynam. 6 (1989) 655–667.
20 R.E. Dickerson, M. Bansal, C.R. Calladine, S. Diekmann, W.N. Hunter, O. Kennard, R. Lavery, H.C.M. Nelson, W.K. Olson, W. Saenger, Z. Shakked, H. Sklenar, D.M. Soumpasis, C.-S. Tung, E. von Kitzing, A. H.-J. Wang and V.B. Zhurkin, J. Mol. Biol. 205 (1989) 787–791.
21 P. Rajagopal and J. Feigon, Biochemistry 28 (1989) 7859–7870.
22 P. Rajagopal and J. Feigon, Nature 339 (1989) 637–640.
23 M.M.W. Mooren, D.E. Pulleyblank, S.S. Wijmenga, M.J.J. Blommers and C.W. Hibers, Nucleic Acids Res. 22 (1991) 437–445.
24 C. de los Santos, M. Rosen and D. Patel, Biochemistry 28 (1989) 7282–7289.
25 J. Liquier, P. Coffinier, M. Firon and E. Taillandier, J. Biomol. Struct. Dynam. 9 (1991) 437–445.