New junction models for alternate-strand triple-helix formation

Thérèse de Bizemont*, Jian-Sheng Sun, Thérèse Garestier and Claude Hélène

Background: Oligonucleotide-directed triple-helix (triplex) formation can interfere with gene expression but only long tracts of oligopyrimidine•oligopurine sequences can be targeted. Attempts have been made to recognize short oligopurine sequences alternating on the two strands of double-stranded DNA by the covalent linkage of two triplex-forming oligonucleotides. Here we focus on the rational optimization of such an alternate-strand triplex formation on a DNA duplex containing a 5′-GpT-3′/3′-CpA-5′ or a 5′-TpG-3′/3′-ApC-5′ step by combination of (G,T)- and (G,A)-containing oligonucleotides that bind to the oligopurine strands in opposite orientations.

Results: The deletion of one nucleotide in the reverse Hoogsteen region of the oligonucleotide provides the best binding at the 5′-GpT-3′/3′-CpA-5′ step, whereas the addition of two cytosines as a linker between the two oligonucleotides is the best strategy to cross a 5′-TpG-3′/3′-ApC-5′ step. Energy minimization and experimental data suggest that these two cytosines are involved in the formation of two novel base quadruplets.

Conclusions: These data provide a rational basis for the design of oligonucleotides capable of binding to oligopurine sequences that alternate on the two strands of double-stranded DNA with a 5′-GpT-3′/3′-CpA-5′ or a 5′-TpG-3′/3′-ApC-5′ step at the junction.

Address: Laboratoire de Biophysique, INSERM U 201, CNRS URA 481, Muséum National d'Histoire Naturelle, 43 rue Cuvier, 75231 Paris Cedex 05, France.

*Present address: Imperial Cancer Research Fund, Clare Hall Laboratories, Blanche Lane, South Mimms, Herts EN6 3LD, UK.

Correspondence: Jian-Sheng Sun
E-mail: sun@cimrs1.mnhn.fr

Key words: alternate-strand triplex, oligonucleotide, quadruplet

Received: **20 May 1998**
Revisions requested: **16 June 1998**
Revisions received: **5 October 1998**
Accepted: **19 October 1998**

Published: **24 November 1998**

Chemistry & Biology December 1998, **5**:755–762
http://biomednet.com/elecref/1074552100500755

Introduction

In the antigene strategy proposed to control gene transcription, an oligopyrimidine•oligopurine sequence of double-helical DNA is recognized by a third-strand oligonucleotide that binds to the major groove and forms a local triple helix ([1,2]; for reviews see [3,4]). The orientation of the third strand depends on its base sequence. In the pyrimidine motif, the (C,T)-containing oligonucleotide binds in a parallel orientation to the target oligopurine sequence via Hoogsteen T•AxT and C•GxC$^+$ base triplets; in the purine motif, the (G,A)-containing oligonucleotides bind in an antiparallel orientation via reverse Hoogsteen C•GxG and T•AxA base triplets; and for oligonucleotides containing guanine and thymine residues, the orientation of the third strand depends on the number of 5′-GpT-3′ and 5′-TpG-3′ steps, and on the length of guanine and thymine tracts. The orientation can be parallel, via Hoogsteen T•AxT and C•GxG base triplet formation, or antiparallel involving reverse Hoogsteen T•AxT and C•GxG base triplets (for a review see [5]).

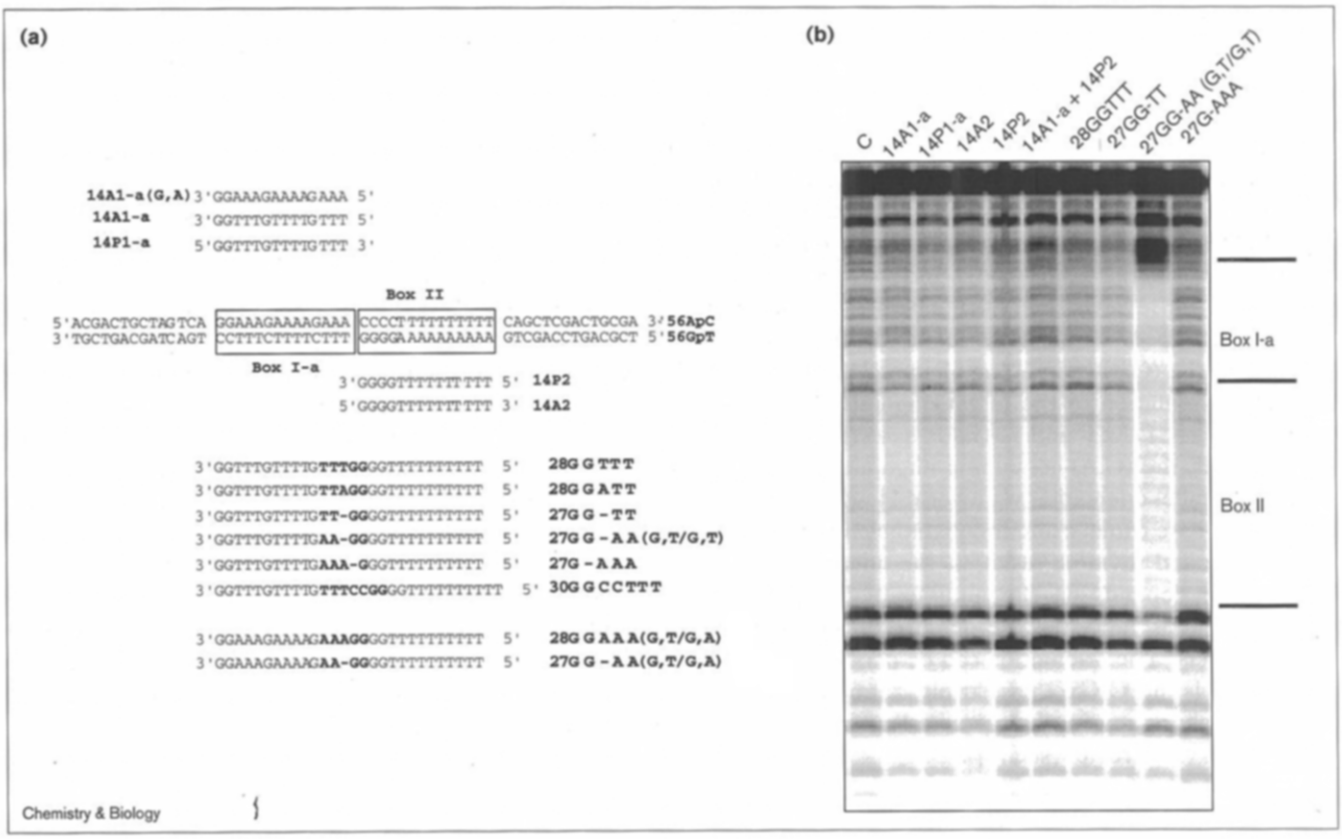
The range of DNA sequences available for triple-helix (triplex) formation can be extended by targeting two oligopurine sequences that alternate on the two strands of DNA. Two short triple helices can be formed in which the third-strand oligonucleotides are hydrogen bonded to the corresponding oligopurine sequences and these third-strand oligonucleotides must be linked together to fully cover the target site. Different strategies have been proposed to link the two triplex-forming oligonucleotides at the junction between the two target sequences: first, when the two third strands have the same hybridization orientation, the 3′ ends (or the 5′ ends) of the two oligonucleotides can be joined by a chemical linker [6–10] or a base-to-base linkage [11]; and second, when the two third-strand oligonucleotides have opposite orientations to their purine targets, it is possible to synthesize a single oligonucleotide (with only phosphodiester linkages) that can bind to the two triplex sites by switching from one oligopurine strand to the other across the major groove at the junction. 'Artificial' linkers are therefore not required at the junction of alternating oligopyrimidine•oligopurine tracts, providing the conformational constraints are properly released. Consequently, a standard oligonucleotide that can easily be synthesized without any further chemical modification can bind to such an extended target sequence by switching from one oligopurine strand to another at the 5′-purine-pyrimidine-3′, or 5′-pyrimidine-purine-3′ step (hereafter designated as 5′-RpY-3′ or 5′-YpR-3′ junction, respectively). Previous work [12–21] using an empirical approach has shown that triple helices can be formed with short third-strand oligonucleotides linked together by phosphodiester bonds. For example, Beal and Dervan [14] have shown that for a 5′-$(Y)_n(R)_m$-3′ sequence, two additional thymines are necessary at the junction to cross the major groove, whereas no additional nucleotide was introduced for the 5′-$(R)_m(Y)_n$-3′ sequence.

The present work is devoted to the development of 'switch triple-helix' or 'alternate-strand triple-helix' formation using a rational approach involving molecular modeling. We describe the optimization of the crossing over at two junctions, the 5′-GpT-3′/3′-CpA-5′ 5′-$(R)_m(Y)_n$-3′ and 5′-TpG-3′/3′-ApC-5′ 5′-$(Y)_m(R)_n$-3′ junctions, for switch triple-helix formation. We have taken advantage of the opposite orientations adopted by (T,G)-containing oligonucleotides according to their sequences in order to recognize alternating oligopurine sequences. We show by DNase I footprinting and gel-retardation experiments that the most efficient way for switching across the 5′-GpT-3′/3′-CpA-5′ junction consists of removing a nucleotide in the third strand on the reverse Hoogsteen side of the junction. Conversely, for the 5′-TpG-3′/3′-ApC-5′ junction, the addition of two cytosines at the junction gives the best switch triple-helix-forming oligonucleotide. On the basis of molecular-modeling studies and the mapping of chloroacetaldehyde-accessible sites, we propose that the stabilization of the switch triplex is a result of the formation of two hydrogen-bonded base quadruplets involving the additional cytosines and two adjacent C•GxG base triplets on the reverse Hoogsteen side of the 5′-TpG-3′/3′-ApC-5′ junction step.

Results

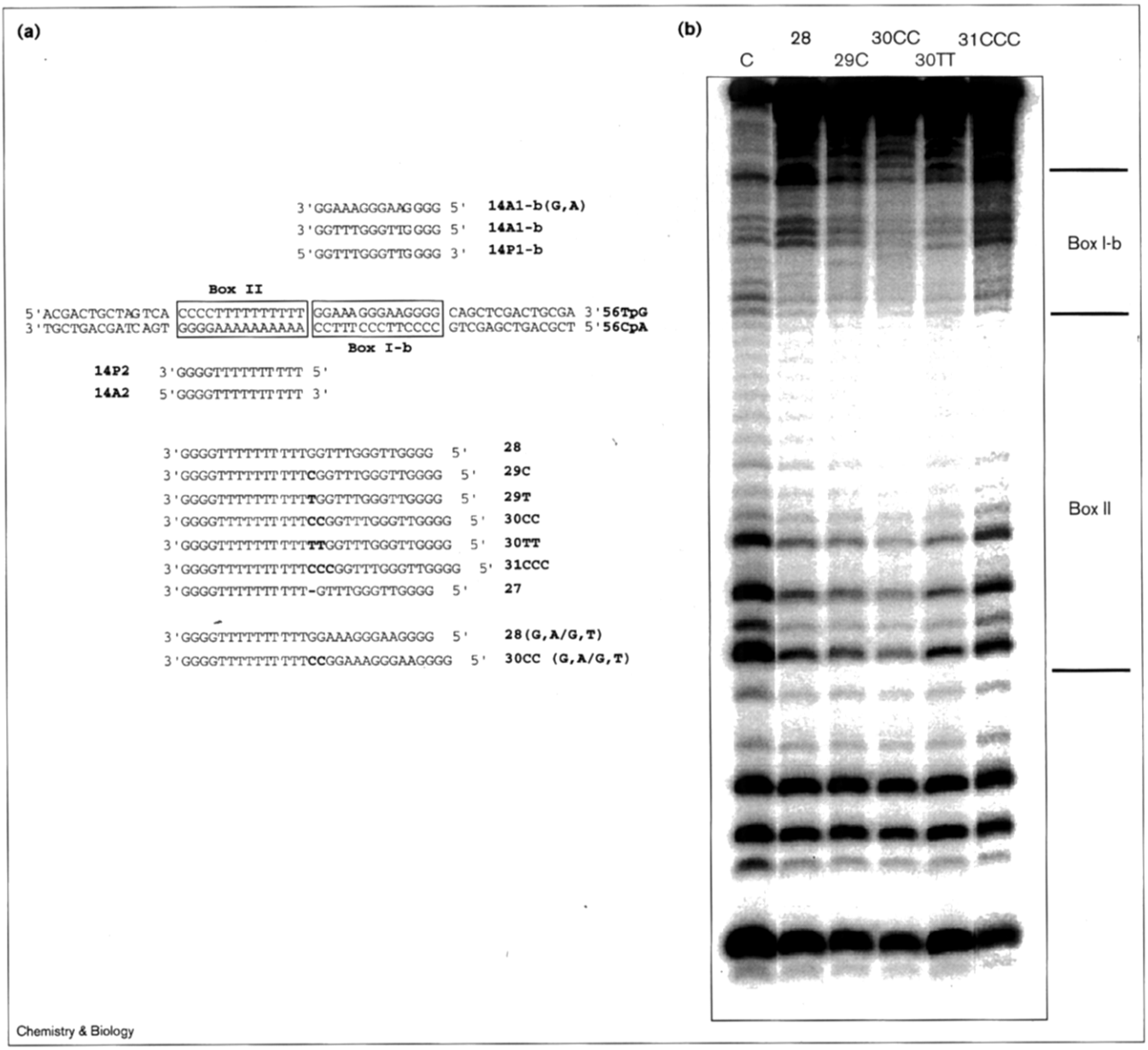
In this study, we have investigated the formation of triple-helical complexes on two 56 base pair DNA fragments shown in Figures 1a and 2a, each of which contains two 14-mer oligopurine sequences alternating on the two strands of the DNA double helix. The 14 base pair oligopyrimidine•oligopurine boxes are characterized by a 5′-GpT-3′/3′-CpA-5′ junction (Figure 1a) and a 5′-TpG-3′/3′-ApC-5′ junction (Figure 2a). One box (box II) is shared by the two 56 base pair fragments; in contrast, box I-a and box I-b have slightly different sequences. Box II contains a single 5′-ApG-3′ step, whereas box I-a and box I-b have five and four (5′-ApG-3′ + 5′-GpA-3′) steps, respectively. For each of the two oligopyrimidine•oligopurine sequences in the two 56 base pair fragments, it is possible to design

Figure 1



Alternate-strand recognition at a 5′-GpT-3′/3′-CpA-5′ junction step. **(a)** Sequence of the 56 base pair DNA fragment and the oligonucleotides used for footprinting experiments and gel-retardation assays. The target contains a 5′-GpT-3′/3′-CpA-5′ junction site. Oligopyrimidine•oligopurine sequences are boxed. The 5′-GpT-3′/3′-CpA-5′ junction step in the duplex and the nucleotides in the third strand oligonucleotides involved in the junction are indicated in bold. **(b)** DNase I footprinting experiments at 37°C of the 56 base pair fragment (56GpT•56ApC) in which the 56GpT strand was 5′-end labeled with the following oligonucleotides: C (nonspecific oligonucleotide), 14A1-a, 14P1-a, 14A2, 14P2, 14A1-a+14P2, 28GGTTT, 27GG-TT, 27GG-AA or 27G-AAA. The concentration of all oligonucleotides was 20μM. Only the 27GG-AA(G,T/G,T) gives a clear footprint on both oligopyrimidine•oligopurine boxes.

Figure 2

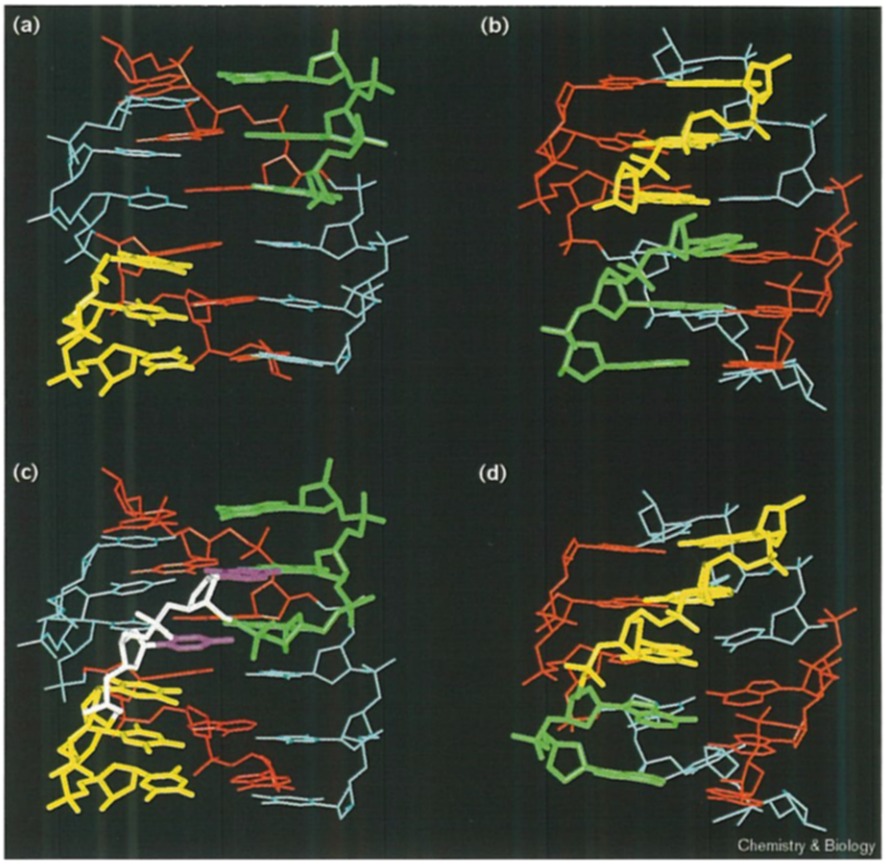


Alternate-strand recognition at a 5′-TpG-3′/3′-ApC-5′ junction step. **(a)** Sequence of the 56 base pair DNA fragment and the oligonucleotides used for footprinting experiments and gel-retardation assays. The target contains a 5′-TpG-3′/3′-ApC-5′ junction site. Oligopyrimidine•oligopurine sequences are boxed. The 5′-TpG-3′/3′-ApC-5′ junction step in the duplex and the nucleotides in the third-strand oligonucleotides involved at the junction are indicated in bold. **(b)** DNase I footprinting at 25°C of the 56 base pair fragment 56TpG•56CpA in which the 56TpG strand was 5′-end labeled with the following oligonucleotides: C (nonspecific oligonucleotide), 28, 29C, 30CC, 30TT or 31CCC. All oligonucleotides were at a concentration of 20 μM.

(G,T)-containing oligonucleotides (14-mers) that bind either in a parallel orientation (Hoogsteen hydrogen-bonding interactions) to the oligopurine sequence of box II or in an antiparallel orientation (reverse-Hoogsteen hydrogen bonding) to the oligopurine sequence of box I-a or I-b. These different orientations are due to the difference in the number of 5′-ApG-3′ + 5′-GpA-3′ steps and in the length of G and A tracts in the two oligopurine sequences, in agreement with the conclusions of previous studies [13,20]. The antiparallel (G,T)-oligonucleotide can also be replaced by a (G,A)-containing oligomer.

The two 14-mer (G,T)-containing oligonucleotides that bind to each of the two boxes occupy different positions in the major groove at the junction, as described previously [14,22]. The distance between the 3′-OH and the

Figure 3



Energy-minimized models of the adjacent (unlinked) two triplexes **(a)** at the 5′-TpG-3′/3′-ApC-5′ step and **(b)** at the 5′-GpT-3′/3′-CpA-5′ step, as well as the optimized alternate-strand triplexes **(c)** at the 5′-TpG-3′/3′-ApC-5′ step and **(d)** at the 5′-GpT-3′/3′-CpA-5′ step. The oligopurine and oligopyrimidine strands are red and blue, respectively. The third strand oligonucleotides are yellow and green, which indicate Hoogsteen and reverse Hoogsteen binding parts, respectively. Hydrogen atoms are omitted for clarity. In the model shown at bottom left, the additional cytosine bases are shown in purple, whereas their sugar–phosphate backbone is in white to illustrate the linkage at the junction.

5′-phosphate groups of the two oligonucleotides in the vicinity of the junction is considerably shorter at the 5′-GpT-3′/3′-CpA-5′ junction than at the 5′-TpG-3′/3′-ApC-5′ junction (Figure 3a,b, respectively).

We have used molecular modeling by energy minimization and experimental techniques such as DNase I footprinting, gel-retardation assays and chemical probing to design a single oligonucleotide that recognizes simultaneously, with the highest possible affinity, the two alternating oligopyrimidine•oligopurine sequences on each of the 56 base pair fragments.

Alternate-strand recognition at the 5′-GpT-3′/3′-CpA-5′ junction

Inspection of the model of two adjacent unlinked triplexes at the 5′-GpT-3′/3′-CpA-5′ step reveals that the 5′- and 3′-terminal OH groups of the two triplex-forming oligonucleotides are close at the junction (Figure 3b). Molecular modeling suggested that the deletion of a nucleotide at the 5′-GpT-3′/3′-CpA-5′ junction should be sufficient to allow the third strand to cross the major groove without too much distortion [22] (Figure 3d). The nucleotide can be removed either from the 3′ end of the Hoogsteen portion (as in 27G-AAA) or from the 5′ end of the reverse Hoogsteen portion (as in 27GG-TT or 27GG-AA(G,T/G,T) (Figure 1a). From molecular-modeling studies [22], a better stacking interaction at the junction was obtained when the last nucleotides in the reverse Hoogsteen portion were adenines instead of thymines (27GG-AA(G,T/G,T) compared with 27GG-TT (Figure 1a).

At 37°C, DNase I footprinting experiments (Figure 1b) clearly showed that only 27GG-AA gave a clear footprint. Under these conditions neither 14A1-a nor 14P2 alone exhibited binding to either of the target oligopyrimidine•oligopurine boxes. Oligonucleotides with the same length as the target did not bind to it (28GGTTT and 28GGATT).

To determine whether the substitution of adenines for thymines in the reverse Hoogsteen portion of the 27/28-mers played any role in binding, the (G,T)-containing sequence was replaced by a (G,A)-containing sequence that could form reverse-Hoogsteen hydrogen bonds as well. The 28-mer 28GGAAA (G,T/G,A) did not bind to its target at 37°C and therefore behaved as 28GGTTT or 28GGATT. In contrast, 27GG-AA (G,T/G,A) with one nucleotide deleted at the 5′ end of the reverse-Hoogsteen portion exhibited a clear footprint over both boxes I-a and

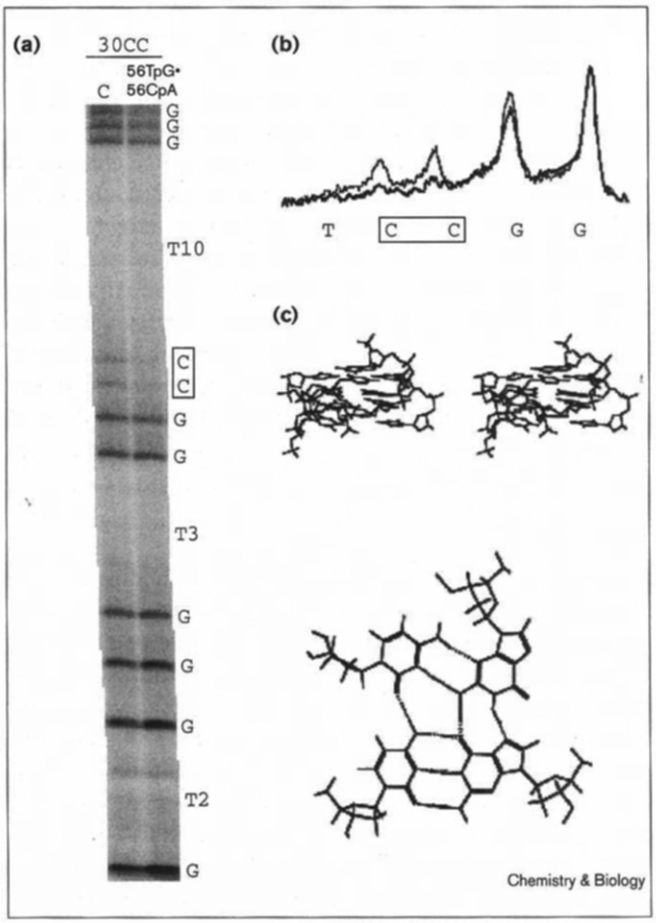
II (data not shown) as did 27GG-AA(G,T/G,T). Similar results were obtained by gel-shift experiments. Gel-retardation assays indicated that a 30-mer with two additional cytosines (30CC) that gave the highest binding to the sequence with the 5′-TpG-3′/3′-ApC-5′ junction step (see below) was not able to bind to the 56 base pair fragment (data not shown). So far, the most efficient oligonucleotides at inducing triplex formation were 27GG-AA (G,T/G,T) and 27GG-AA (G,T/G,A).

Alternate-strand recognition at the 5′-TpG-3′/3′-ApC-5′ junction

In contrast to the 5′-GpT-3′/3′-CpA-5′ junction, the 5′- and 3′-terminal OH groups of the two triplex-forming oligonucleotides are much further apart at the 5′-TpG-3′/3′-ApC-5′ junction (Figure 3a). Additional nucleotides have to be introduced to cross the major groove without too much distortion. As shown in Figure 2a, the thymine and the guanine at the junction are involved in T•AxT Hoogsteen and C•GxG reverse-Hoogsteen base triplets, respectively. Molecular modeling by energy minimization suggested that the more stable solution to cross the major groove was to incorporate two cytosines in the third strand [22] (Figure 3c). The backbone of the third strand adopts an S shape that allows the two cytosines to form hydrogen bonds with the two C•GxG base triplets on the reverse Hoogsteen portion of the triplex forming two base quadruplets (see Figures 3c and 4c). Introducing bases other than cytosine would not allow these interactions to occur; purines would introduce steric problems at the junction and thymines could be accommodated but would form only one hydrogen bond with C•GxG base triplets (instead of three for cytosines).

To test this hypothesis, several experiments were carried out. Figure 2a shows the sequence of the oligonucleotides that were investigated for their binding to the 56 base pair duplex with a 5′-TpG-3′/3′-ApC-5′ junction between the two 14 base pair oligopyrimidine•oligopurine boxes. Footprinting experiments shown in Figure 2b indicated that the third-strand oligonucleotide with the highest affinity was a 30-mer (30CC) with two cytosines added at the junction between the two (G,T)-containing oligonucleotides bound in a parallel and antiparallel orientations with respect to the two oligopurine sequences. Three cytosines (31CCC) did not stabilize the triplex structure as compared with the 28-mer. Replacement of both cytosines by a single cytosine (29C) or by two thymines (30TT) allowed the third strand to bind but with a lower affinity than 30CC. The reverse Hoogsteen (G,T)-containing part of 30CC can be replaced by a (G,A)-containing sequence. Gel-retardation assays performed at 25°C indicated that the triplex containing the third strand 30CC(G,A/T,G) was more stable than the triple helix with the corresponding oligonucleotide 28(G,A/G,T) containing no additional cytosines (data not shown).

Figure 4



Alternate-strand recognition at a 5′-TpG-3′/3′-ApC-5′ junction step: formation of a base quadruplet between additional cytosines and reverse Hoogsteen C•GxG bases triplets near the junction. **(a)** Mapping of chloroacetaldehyde-modified sites of the oligonucleotide 30CC (5′-end labeled) at 4°C with different duplexes: C (nonspecific duplex) and 56TpG•56CpA (specific duplex). Duplexes were at 20 μM. To enhance the reactivity of chloroacetaldehyde-modified sites, formic acid was added prior to hot piperidine treatment; this explained why guanine residues were also revealed on the electrophoresis gel. **(b)** Quantification of chloroacetaldehyde-modified sites of the oligonucleotide 30CC (5′-end labeled) at 4°C with different duplexes (see Figure 4a): C (thin line), 56TpG•56CpA (thick line). **(c)** Top: stereoview of the model shown on Figure 3c involving two additional cytosines at the 5′-TpG-3′/3′-ApC-5′ step. Hydrogen atoms are omitted for clarity; bottom: suggested base quadruplet formed by the additional cytosines and the reverse Hoogsteen C•GxG base triplets.

To obtain information on the potential interactions involving the two cytosines at the junction in 30CC, chloroacetaldehyde was used as a probe of cytosine accessibility. Chloroacetaldehyde reacts with both the exocyclic amino group N4 of cytosine residues and the N3 position of cytosine residues when the base is not involved in hydrogen-bonding interactions. The two cytosines in 30CC reacted with chloroacetaldehyde (Figure 4a,b). In the

presence of the specific 56 base pair target sequence, the chloroacetaldehyde reaction of the two cytosines in 30CC was inhibited (Figure 4a,b). The sample was treated with formic acid before piperidine treatment to enhance the detection of chloroacetaldehyde-modified cytosines. These reactions also cleaved the oligonucleotide at guanine nucleotides. Figure 4a and b shows that guanine cleavage is not altered, whereas cytosines are protected from the chloroacetaldehyde reaction by the duplex target. These results indicate that the two cytosines are protected in the triple-helical complex. In the alternate-strand triplex, the N3 and/or N4 positions of both additional cytosines therefore seem to be involved in hydrogen-bonding interactions. Although not proving it, this result is in agreement with the model presented in Figure 4c, in which the two additional cytosines form two quadruplets with the C•GxG base triplets in the reverse Hoogsteen configuration.

Discussion

An oligonucleotide with only natural phosphodiester linkages can be synthesized to recognize two oligopurine sequences that alternate on the two strands of the DNA double helix, providing the two triplex-forming domains have opposite orientations to their target oligopurine sequences. For example, a (G,A)-containing oligonucleotide can be combined with a (C,T)-containing oligonucleotide. The former forms reverse-Hoogsteen hydrogen bonds and binds antiparallel to its oligopurine target; the latter forms Hoogsteen hydrogen bonds and bind in parallel to its oligopurine target sequence. Alternatively two (G,T)-containing oligonucleotides can be used, taking advantage of the opposite orientations that they adopt depending on the sequence: parallel when the sequence contains long G and T tracts and a small number of 5′-GpT-3′ + 5′-TpG-3′ steps, and antiparallel when G and T tracts are short and the number of ′-GpT-3′ + 5′-TpG-3′ steps is high.

In the present study, we have adopted the second strategy. We have used molecular modeling and DNase I footprinting to determine the best way to link together the two (G,T)-containing oligonucleotides at the 5′-GpT-3′/3′-CpA-5′ junction, and we found that one nucleotide should be removed in order to increase binding. Comparison of the hybridization of 27G-AAA and 27GG-AA to the duplex indicates that it is better to remove one nucleotide in the reverse Hoogsteen than in the Hoogsteen part of the third strand. This preference is a result of the higher stability of the C•GxG base triplet in Hoogsteen configuration than that of the T•AxA base triplet in reverse Hoogsteen configuration. Additional stabilization of the switch triplex is obtained when the two adjacent bases on the reverse Hoogsteen and Hoogsteen sides are adenine and guanine, respectively. The replacement of thymines by adenines in the reverse Hoogsteen part of the oligonucleotide allows a greater stabilization of the switched triplex. As 14A1a(G,A) does not bind to box I-a more strongly than 14A1-a (data not shown), this phenomenon can be explained by the stacking of base triplets occurring at the 5′-GpT-3′/3′-CpA-5′ junction step: the stacking interaction energy between reverse-Hoogsteen T•AxA and Hoogsteen C•GxG base triplets is larger than that between reverse-Hoogsteen T•AxT and Hoogsteen C•GxG base triplets. The importance of stacking interactions at the switch position as a stabilizing factor for alternate-strand recognition has been reported previously by Balatskaya *et al.* [12].

At the 5′-TpG-3′/3′-ApC-5′ junction, two nucleotides have to be added at the junction to cross the major groove. The third strand adopts an S shape at the crossing site. Additional stabilization is observed when these two nucleotides are cytosines. Molecular modeling suggested that this additional stabilization was due to the formation of base quadruplets in which the cytosines are hydrogen bonded to adjacent C•GxG base triplets on the reverse Hoogsteen side of the complex. Chemical-protection experiments using chloroacetaldehyde as a reagent provided evidence for protection of the two cytosines at the junction. The formation of base quadruplets involving a cytosine to recognize a C•GxG base triplet presents a new possible method for recognizing base sequences in DNA. The structure adopted by the third-strand oligonucleotide involves S-shape folding and requires the presence of two guanines on the reverse Hoogsteen side. This stabilization could occur at both 5′-CpG-3′ and 5′-TpG-3′ junction in 5′-$(Y)_n(R)_m$-3′ sequences.

For alternate-strand triple-helix formation, there are two types of junctions — the 5′-RpY-3′ and the 5′-YpR-3′ junction. There are three different junctions of each type: 5′-ApT-3′, 5′-GpC-3′ and 5′-GpT-3′ (which is equivalent to 5′-ApC-3′) for the 5′-RpY-3′ junction; 5′-TpA-3′, 5′-CpG-3′ and 5′-TpG-3′ (which is equivalent to 5′-CpA-3′) for the 5′-YpR-3′ junction. A comprehensive knowledge of how to deal correctly with all six junctions would theoretically increase the range of DNA sequences that can be recognized by triple-helix-forming oligonucleotides. Further studies are currently underway to establish a full switch code for all six junctions. The results presented above show that a combination of molecular modeling by energy minimization together with experimental studies (such as DNase I footprinting and gel-shift assays) can provide a solution for each of the six possible junctions. The choice of the most appropriate switch oligonucleotide (G,A/G,T or G,T/G,T) depends on the particular sequences on each side of the junction. Together with another approach consisting of linking the 3′ or the 5′ ends of two oligonucleotides that are bound within the same motif (Hoogsteen or reverse Hoogsteen) by means of a non-phosphodiester linker [8-11], the design of switch oligonucleotides presents interesting possibilities for extending the range of double-helical sequences that can be selectively recognized by oligonucleotides.

Significance

We have shown that oligonucleotides containing natural nucleotides joined together through a natural phosphodiester backbone can recognize adjacent oligopurine sequences that alternate on the two strands of the duplex. The goal was to optimize the binding of the third strand oligonucleotide that crosses the major groove either at a 5′-GpT-3′/3′-CpA-5′ or at a 5′-TpG-3′/3′-ApC-5′ junction step. In agreement with the prediction of molecular-modeling studies, the experimental results presented here showed that the backbone distortion at the 5′-GpT-3′/3′-CpA-5′ junction can be reduced by deleting one nucleotide at the 3′ end of the antiparallel part, and by choosing appropriate nucleotides near the junction in order to increase base stacking in the third strand. At the 5′-TpG-3′/3′-ApC-5′ junction, two additional cytosines are required to facilitate the switching over of the third strand oligonucleotide across the major groove. The formation of two novel base quadruplets between reverse Hoogsteen C·GxG base triplets near the junction and the additional cytosines provides a molecular basis for recognition and enhanced binding of the switch oligonucleotide.

Further work is under way in our laboratory to study the other four junction steps in alternating oligopyrimidine·oligopurine duplex DNA sequences. These studies should allow us to design stable triple helices with extended duplex targets. Such switched triple helices could provide a new basis for the development of the 'antigene' strategy aimed at controlling specifically gene expression.

Materials and methods

DNA

Oligodeoxyribonucleotides were synthesized on an automatic synthesizer by Genosys (Cambridge, England) using the phosphoramidite method. Full-length oligonucleotides were purified by denaturating polyacrylamide gel electrophoresis. Oligonucleotide concentrations were obtained by UV absorption at 260 nm on a Kontron Uvikon 860 spectrophotometer at 20°C. The 56-mers used in this study were labeled at the 5′ end with [γ^{32}P] ATP (ICN) and polynucleotide kinase (BioLabs).

DNase I footprinting experiments

DNase I footprinting studies were carried out as described by Duval-Valentin and Ehrlich [23] with the following modifications. 20nM 56 base pair duplex DNA fragment (56GpT•56ApC or 56TpG•56CpA, see Figure 1a and Figure 2a for sequences, respectively) labeled on one strand (56GpT or 56TpG) at the 5′ end was incubated overnight with 20 μM switch oligonucleotides (at different temperatures in a 10 μl solution containing 20 mM Tris-HCl pH 7.2, 50 mM NaCl, 5 mM $MgCl_2$, 0.5 mM spermine and 5mM $CaCl_2$. DNaseI (0.2 μg/ml final) was added and digestion was stopped after different periods of time (depending on the temperature) by adding 2 μl of the loading buffer (80% deionized formamide, 0.1% xylene cyanol and 0.1% bromophenol blue) and chilled at −80°C. After denaturation at 90°C, the DNaseI digestion products were loaded on a 20% denaturating polyacrylamide gel and analyzed by PhosphorImager (Molecular Dynamics).

Mapping of chloroacetaldehyde-accessible sites

Oligonucleotide radiolabeled at the 5′ end (20 nM) was incubated overnight with the duplex DNA fragment (20 μM) in a 10 μl solution containing 20 mM Tris-HCl pH 7.2, 50 mM NaCl, 5 mM $MgCl_2$ and 0.5 mM spermine. Chloroacetaldehyde was added and the reaction was stopped with ether followed by a precipitation with ethanol. After a treatment with formic acid, the pellets were dissolved in 1 M piperidine, incubated 20 min at 90°C and then lyophilized. The samples were loaded on a 20% acrylamide/7 M urea gel. The gel was exposed to a PhosphorImager screen (Molecular Dynamics).

Molecular modeling by conformational energy minimization

The calculation of molecular mechanics was carried out using the JUMNA (version 7) program package [24]. Neither water nor positively charged counter ions were explicitly included in the energy minimization. Their effects were simulated by a sigmoidal, distance-dependent, dielectric function [25], however, and by assignment of a half-negative charge for each phosphate group. Computations were carried out on a Silicon Graphics 4D/420GTXB dual processor workstation.

The coordinates of triple helices were derived from the previously published B-like triple helix [26,27] which is now widely supported by many nuclear magnetic resonance (NMR) and vibrational spectroscopic studies. Typically, the triple helices were ten base triplets in length with the junction step occurring at the center. The three last base triplets at both ends were restrained to a mononucleotide symmetry in order to decrease the end effects and to focus on the effect at the junction (the central four base triplets). A manual docking was sometimes required around the junction in order to establish an appropriate interaction and to avoid strong steric clashes during initial steps of minimization. The total complexation energy (E_{TOT}) was decomposed in terms of intermolecular interactions ($E_{DH\text{-}III}$) between the third strand (III) and the target double-helical DNA (DH), as well as the conformational deformation energy of the double helix (ΔE_{DH}) and the third strand (ΔE_{III}). Conformational deformation energies were evaluated as the difference between the corresponding energetic components before and after complexation of the third strand. It should be noted that this evaluation is approximate, especially for the third strand because the conformation of a free single-stranded DNA is generally less well defined than a free double-stranded DNA. These approximations were necessary to take into account the effect of base composition, however.

Acknowledgements

T. de Bizemont was supported by grants from the Ministère de la Recherche et de l'Enseignement Supérieur and the Association pour la Recherche contre le Cancer. The authors thank Guy Duval-Valentin for helpful discussions.

References

1. Le Doan, T., *et al.*, & Hélène, C. (1987). Sequence specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo[α]thymidylate covalently linked to an azidoproflavine derivative. *Nucleic Acids Res.* **15**, 7749-7760.
2. Moser, H.E. & Dervan, P.B. (1987). Sequence specific cleavage of double helical DNA by triple helix formation. *Science* **238**, 645-650.
3. Plum, G.E., Pilch, D.S., Singleton, S.F. & Breslauer, K.J. (1995). Nucleic acid hybridization: triplex stability and energetics. *Annu. Rev. Biophys. Biomol. Struc.* **24**, 319-350.
4. Thuong, N.T. & Hélène, C. (1993). Sequence-specific recognition and modification of double-helical DNA by oligonucleotides. *Angew. Chem. Int. Ed. Engl.* **32**, 666-690.
5. Sun, J.S. & Hélène, C. (1993). Oligonucleotide-directed triple-helix formation. *Curr. Opin. Struct. Biol.* **3**, 345-356.
6. Froehler, B.C., Terhorst, T., Shaw, J.P. & McCurdy, S.N. (1992). Triple-helix formation and cooperative binding by oligodeoxynucleotides with a 3′-3′ internucleotide junction. *Biochemistry* **31**, 1603-1609.
7. Horne, D.A. & Dervan, P.B. (1990). Recognition of mixed sequence duplex DNA by alternate-strand triple helix formation. *J. Am. Chem. Soc.* **112**, 2435-2437.
8. McCurdy, S., Moulds, C. & Froehler, B. (1991). Deoxyoligonucleotides with inverted polarity: synthesis and use in triple-helix formation. *Nucleosides Nucleotides* **10**, 287-290.
9. Ono, A., Chen, C.N. & Kan, L.S. (1991). DNA triplex formation of oligonucleotide analogues consisting of linker groups and octamer segments that have opposite sugar-phosphate backbone polarities. *Biochemistry* **30**, 9914-9921.

10. Ouali, M., Bouziane, M., Ketterle, C., Gabarroarpa, J., Auclair, C. & Lebret, M. (1996). A molecular mechanics and dynamics study of alternate triple-helices involving the integrase binding site of the HIV-1 virus and oligonucleotides having a 3′-3′ internucleotide junction. *J. Biomol. Struct. Dyn.* **13**, 835-853.
11. Zhou, B.-W., *et al.*, & Hélène, C. (1995). Recognition of alternating oligopurine•oligopyrimidine tracts of DNA by oligonucleotides with base-to-base linkages. *Bioconj. Chem.* **6**, 516-523.
12. Balatskaya, S.V., Belotserkovskii, B.P. & Johnston, B.H. (1996). Alternate-strand triplex formation: Modulation of binding to matched and mismatched duplexes by sequence choice in the Pu•PuxPy block. *Biochemistry* **35**, 13328-13337.
13. de Bizemont, T., Duval-Valentin, G., Sun, J.S., Bisagni, E., Garestier, T. & Helene, C. (1996). Alternate strand recognition of double-helical DNA by (T,G)-containing oligonucleotides in the presence of a triple helix-specific ligand. *Nucleic Acids Res.* **24**, 1136-1143.
14. Beal, P.A. & Dervan, P.B. (1992). Recognition of double helical DNA by alternate strand triple helix formation. *J. Am. Chem. Soc.* **114**, 4976-4982.
15. Bouziane, M., Cherny, D.I., Mouscadet, J.F. & Auclair, C. (1996). Alternate strand DNA triple helix-mediated inhibition of HIV-1 U5 long terminal repeat integration *in vitro*. *J. Biol. Chem.* **271**, 10359-10364.
16. Jayasena, S.D. & Johnston, B.H. (1992). Oligonucleotide-directed triple helix formation at adjacent oligopurine and oligopyrimidine DNA tracts by alternate strand recognition. *Nucleic Acids Res.* **20**, 5279-5288.
17. Jayasena, S.D. & Johnston, B.H. (1993). Sequence limitations of triple helix formation by alternate-strand recognition. *Biochemistry* **32**, 2800-2807.
18. Olivas, W.M. & Maher III, L.J. (1994). DNA recognition by alternate strand triple helix formation: affinities of oligonucleotides for a site in the human p53 gene. *Biochemistry* **33**, 983-991.
19. Olivas, W.M. & Maher, L.J. (1996). Binding of DNA oligonucleotides to sequences in the promoter of the human bcl-2 gene. *Nucleic Acids Res.* **24**, 1758-1764.
20. Sun, J.S., de Bizemont, T., Duval-Valentin, G., Montenay-Garestier, T. & Hélène, C. (1991). Extension of the range of recognition sequence for triple helix formation by oligonucleotides containing guanines and thymines. *C. R. Acad. Sci. Paris, série III* **313**, 585-590.
21. Washbrook, E., Fox, K.R. (1994). Alternate-strand DNA triple-helix formation using short acridine-linked oligonucleotides. *Biochem. J.* **301**, 569-575.
22. Sun, J.S. (1995). Rational design of switched triple helix-forming oligonucleotides: extension of sequences for triple helix formation. In *Modelling of Biomolecular Structures and Mechanisms*. (Pullman, A., Pullman, B. & Jortner, J., eds), pp. 267-288, Kluwer Academic Publishers, Amsterdam.
23. Duval-Valentin, G. & Ehrlich, R. (1987). Far upstream sequences of the *bla* promoter from *Tn3* are involved in complexation with *E. coli* RNA polymerase. *Nucleic Acids Res.* **16**, 2031-2044.
24. Lavery, R. & Sklenar, H. (1988). The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* **6**, 63-91.
25. Lavery, R., Sklenar, H., Zakrzewska, K. & Pullman, B. (1986). The flexibility of the nucleic acids (II): the calculation of internal energy and applications to mononucleotide repeat DNA. *J. Biomol. Struct. Dyn.* **3**, 929-1014.
26. Ouali, M., *et al.*, & Taillandier, E. (1993). A possible family of B-like triple helix structures: comparison with the Arnott A-like triple helix. *Biochemistry* **32**, 2098-2103.
27. Ouali, M., *et al.*, & Taillandier, E. (1993). Determination of G*G•C triple helix structure by molecular modeling and vibrational spectroscopy. *J. Am. Chem. Soc.* **115**, 4264-4270.

Because *Chemistry & Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cmb – for further information, see the explanation on the contents pages.