Nucleic Acids Research, Vol. 20, No. 7 1777–1784

Triple helix formation by purine-rich oligonucleotides targeted to the human dihydrofolate reductase promoter

Scott W.Blume, Jay E.Gee, Kedar Shrestha and Donald M.Miller*
Comprehensive Cancer Center, Departments of Biochemistry and Internal Medicine, University of Alabama at Birmingham, VA Medical Center, Birmingham, AL, USA

Received September 17, 1991; Revised and Accepted March 11, 1992

ABSTRACT

The ability of oligodeoxynucleotides to form specific triple helical structures with critical regulatory sequences in the human dihydrofolate reductase (DHFR) promoter was investigated. A battery of purine-rich oligonucleotides targeted to the two purine·pyrimidine strand biased regions near the DHFR transcription initiation site was developed. The stable triple helical structures formed by binding of the oligonucleotides to the native promoter double helix were dominated by G*G·C triplets, with interspersed C*C·G and A*A·T alignments. Mismatches between the oligonucleotide and the purine-rich strand of the target significantly destabilized third strand binding, and a G*A·T alignment was particularly unfavorable. Formation of a pur·pur·pyr triple helical structure results in a localized limitation of access to the native double helical DNA and produces sequence dependent conformational alterations extending several nucleotides beyond the triplex-duplex boundary. Although they differ only by the insertion of two A·T base pairs, the distal and proximal purine·pyrimidine regions can be targeted individually due to the high degree of sequence specificity of triple helical alignment. Triplex formation overlapping any of three consensus transcriptional regulatory elements and collectively covering 50% of the DHFR core promoter is now possible with this set of oligonucleotides.

INTRODUCTION

The binding of an exogenous molecule to a particular sequence of DNA within the genome could be used to alter double helical conformation or limit accessibility of sequence specific DNA binding factors to that location and thereby potentially inhibit transcription or replication. Numerous low molecular weight natural and synthetic molecules which bind DNA inhibit molecular interactions by steric hindrance or by structural damage, but this interaction is largely non-sequence specific. Our lab has shown that the antibiotic mithramycin, which binds the minor groove of G-C rich DNA (1), blocks binding of transcription factor Sp1 to G/C box sequence elements in the human dihydrofolate reductase (DHFR) promoter (2) and the human c-myc promoter (3). The results of DNAse protection assays indicate that mithramycin and Sp1 recognize overlapping sequences, and gel mobility shift analyses demonstrate mithramycin interference with Sp1 binding. A corresponding inhibition by mithramycin of promoter-dependant in vitro transcription of a DHFR template was also seen. Although mithramycin treatment of DHFR gene-amplified, methotrexate resistant cells resulted in a dramatic inhibition of DHFR gene expression, the cells did not become methotrexate sensitive because of transcriptional interference for many genes beyond the intended target.

The relatively low degree of sequence specificity exhibited by the common DNA-binding drugs might be increased by utilizing a nucleic acid structure to recognize the particular base pattern in the double helix. Triple helical formation was first observed by using synthetic nucleic acid homo- or copolymers (4–7). This phenomenon is characterized by the binding of a third nucleic acid strand within the major groove of the native double helix, stabilized by non-Watson–Crick hydrogen bonding between nucleotide residues in the third strand and those of one of the strands of the native double helix. Intramolecular triplex formation (H-DNA) was later detected (8–11) in biological sequences exhibiting a bias for purines on one strand and pyrimidines on the other (a pur·pyr sequence). Intermolecular triple helical formation could also be induced by addition of a pyrimidine-rich oligonucleotide as the third strand, with base sequence identical to, but orientation opposite, the pyrimidine-rich strand of the target duplex (12–19). This pyr·pyr·pur mode of triplex formation was based upon C^{+}*G·C and T*A·T triplet sequence specificity, generally requiring an acidic pH for protonation of cytosine residues in the third strand. Very recently it was discovered that, in the presence of Mg^{++} and at neutral pH, a purine-rich third strand may bind to a pur·pyr sequence (20–26). The pur·pur·pyr model of triplex formation is based on G*G·C and A*A·T triplets as observed in the triple stranded structures poly(dG)(dG)(dC) (7,11,23) and poly (A)(dA)(dT) (23). A description of the characteristics and an understanding of the behavior of the 2purine:1pyrimidine triple helical model is currently in development.

* To whom correspondence should be addressed

We have applied this concept to the problem of identifying an agent which will selectively recognize and bind to specific sequences within the human DHFR promoter. Formation of a sequence specific intermolecular triple helix which exhibits a concentration dependent interference with regulatory protein binding has been demonstrated (27). In the present work the targets for triple helical binding within the human DHFR promoter have been significantly expanded and the nature of the 2pur·1pyr structure further detailed.

MATERIALS AND METHODS

Human dihydrofolate reductase promoter fragment

The plasmid pDHFR1.8 contains the first exon and first intron of the human dihydrofolate reductase gene along with 1.25 kb of 5′ flanking sequences (28). The DHFR core promoter was obtained by subcloning a Not I- Ava II fragment (−112 to +56 relative to the transcription start site) to produce pDHFR.19. For DNAse protection and Hpa II interference assays, the insert of pDHFR.19 was isolated by Pst1-EcoR1 or EcoR1-HindIII digestion and 3′ end-labelled on the coding (purine-rich) or non-coding (pyrimidine-rich) strands respectively.

Oligonucleotides

Oligonucleotides were synthesized by the automated phosphoramidite method, eluted using reverse phase chromatography, and analyzed by polyacrylamide electrophoresis for relative purity. Experiments described in this manuscript utilized the following sequences:

dist-19 =	5′ - C GGGG C GGGGGGG C GGGG C - 3′
dist-14 =	5′ - C GGGG C GGGGGGG C - 3′
alt-dist-14 =	5′ - C GGGGGGG C GGGG C - 3′
dist-C =	5′ - G CCCC G CCCCCCC G CCCC G - 3′
cmn-12 =	5′ - C GGGG C GGGGGG - 3′
prox-F =	5′ - C GGGG C GGGGGGAG C - 3′
prox-G =	5′ - C GGGGGGAG C AGGGG T - 3′
prox-H =	5′ - GGAG C AGGGG T AAA C - 3′
prox-C =	5′ - G CCCC G CCCCCCTC G TCCCC A TTT G - 3′
prox-B =	5′ - GGGGA C GAGGGGGG C GGGG - 3′

Triple helical binding

The labelled DHFR promoter fragment (100,000 cpm, ~40nM) or a pre-hybridized double stranded oligonucleotide target (0.2μM) with sequence of one of the pur·pyr regions was incubated with the appropriate triplex forming oligonucleotide (10−40 μM) with 10mM Tris-Cl pH7.2 and 5mM $MgCl_2$ at room temperature for 45 minutes.

Gel mobility shift

The products of triple helical binding reactions (or hybridization reactions under identical conditions) were analyzed by electrophoresis on a 5% native polyacrylamide gel run at 110V for 100−150 minutes. The gel and buffer contained 5mM $MgCl_2$ and no EDTA (22).

DNAse 1 protection

Following incubation to allow triple helical binding, limited DNAse digestion (12.5u/ml for control samples, 75u/ml for samples containing oligonucleotide, 30 seconds on ice) was carried out. After extraction and precipitation, products were analyzed on an 8% denaturing polyacrylamide gel.

Hpa II interference

Incubation for triple helical binding was followed by addition of Hpa II (1u/μl) without altering buffer conditions, and allowing digestion to proceed at room temperature for another 75 minutes. After extraction and precipitation, reaction products were analyzed on an 8% denaturing polyacrylamide gel.

RESULTS

Purine·pyrimidine targets for triple helical binding

The critical sequences for function of the human DHFR promoter appear to be contained within a 75 base pair region surrounding the transcription initiation site (−55 to +20), probably involving transcription factor binding to at least four specific sequence elements (28−34). (The initiation site for the divergent transcript Rep-3 (−89) is positioned so near that of DHFR that regulatory promoter sequences for the two genes almost certainly are shared or overlap). Within this region are two purine·pyrimidine biased sequences (distal, −57 to −41; proximal, −24 to −6). The relationship of these pur·pyr sequences to the DHFR promoter architecture is shown in Figure 1. Each of the pur·pyr sequences contains a consensus G/C box element which binds transcription factor Sp1. The proximal pur·pyr sequence also overlaps a 'CAA box/ element 3' homologous in sequence and position among the three mammalian DHFR genes which have been examined. Together the two pur·pyr sequences cover approximately 50% of this critical promoter region. We have developed a set of purine-rich oligodeoxynucleotides targeted to these two pur·pyr sequences, determined the sequence requirements and specificity of this mode of triple helical formation at these biologically relevant sites, and investigated the consequential effects of their binding to the native promoter double helix.

Triplex formation at the distal pur·pyr sequence of the human DHFR promoter

Most studies of the pur·pur·pyr model of triplex formation, including the present work, have utilized a target sequence of oligo dG·dC, or a close derivative with interspersed individual A, T, or C residues in the purine-rich strand. The presence of single A's or pyrimidines within the target appears not to greatly disturb the potential for 2pur:1pyr triplex formation. The difficulty in producing triple helical structures involving multiple A*A·T alignments with apparent steric or size constraints create

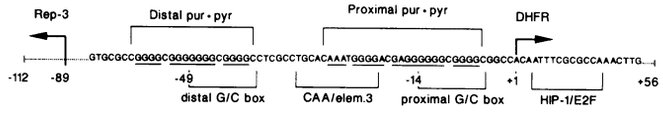

Fig. 1. Purine·pyrimidine strand biased sequences within the architecture of the human DHFR promoter. The restriction fragment utilized for DNAse I protection and Hpa II interference assays was obtained from a subclone of pDHFR1.8 (28) and contains sequences −112 to +56 relative to the DHFR transcription initiation site. The divergent transcription start site for the Rep-3 gene is positioned at −89 (33). The sequence shown is that of the purine-rich (DHFR-coding) strand. The consecutive purine residues constituting the two pur·pyr regions are underlined. The four consensus transcriptional regulatory elements were recognized by homology among the murine, CHO, and human DHFR promoters (28−34). Protein binding has been detected for each element (30,32,34, and our unpublished results). The two pur·pyr sequences contain or overlap three of these important elements and together cover 50% of the DHFR core promoter.

concern over the general applicability of this triplet to formation of three stranded nucleic acid structures (20).

We first designed third strand oligonucleotides with sequences identical to the purine-rich strand of the target double helix. Thus triple helical alignments for the DHFR promoter are dominated by G*G·C triplets, and interruptions in the oligo dG·dC pattern in the target are matched with corresponding interruptions in the triplex forming oligonucleotide.

The distal purine·pyrimidine sequence of the human DHFR promoter presents a symmetrical target for triple helix formation. The first purine-rich oligonucleotide targeted to this sequence, designated dist-19, was designed to bind the entire distal pur·pyr region (−58 to −40):

5′ - C GGGG C GGGGGGG C GGGG C - 3′

A second, smaller oligonucleotide, dist-14, was designed to bind to only a portion of the distal pur·pyr target sequence:

5′ - C GGGG C GGGGGGG C - 3′

Triple helical binding of these oligonucleotides to a restriction fragment of the human DHFR promoter containing the two pur·pyr sequences could be detected by DNAse protection assays (Figure 2). The areas of protection produced on both the coding and non-coding strands approximated the size and position of the expected triple helical structures. The downstream (relative to the direction of DHFR transcription) boundaries of protection produced by the two oligonucleotides were identical, but for dist-14, bands were seen in the upstream portion of the target sequence which were not present with the larger oligonucleotide. This pattern is consistent with binding of dist-14 to the downstream portion of the target (−53 to −40), and orientation parallel to the pyrimidine-rich strand of the native duplex.

A third purine-rich oligonucleotide, alt-dist-14, was targeted to the upstream portion of the distal pur·pyr sequence (−58 to −45):

5′ - C GGGGGGG C GGGG C - 3′

Single, double, and triple stranded structures can be distinguished by differential electrophoretic mobility on a native polyacrylamide gel (14,22,35) (Figure 3). When incubated with a labelled single stranded complementary oligonucleotide (dist-C), dist-19, dist-14, and alt-dist-14 each hybridized to form an antiparallel duplex. When incubated with a labelled double stranded oligonucleotide (dist-C·dist-19) with the sequence of the distal pur·pyr region, each of the distal targeted oligonucleotides bound antiparallel to the purine-rich strand of the duplex to form a pur·pur·pyr triplex.

A recognition sequence for restriction endonuclease Hpa II (5′- CCGG −3′) is located at the upstream boundary of the distal pur·pyr region. In order to examine the effect of triplex formation on the accessibility of the double helix (16,18,36), and to confirm the orientation of the oligonucleotide third strand (25,37,38), triple helical structures formed by each of the distal targeted oligonucleotides with the DHFR promoter were tested for interference with Hpa II digestion (Figure 4). If no exogenous oligonucleotide was added, the full length 204 bp DHFR promoter fragment was digested by Hpa II to a labelled 132 bp fragment and an unlabelled 72 bp fragment. However, if the promoter fragment was pre-incubated with dist-19 under standard triplex forming conditions, >80% of promoter molecules were protected from Hpa II restriction. Thus, formation of a triple

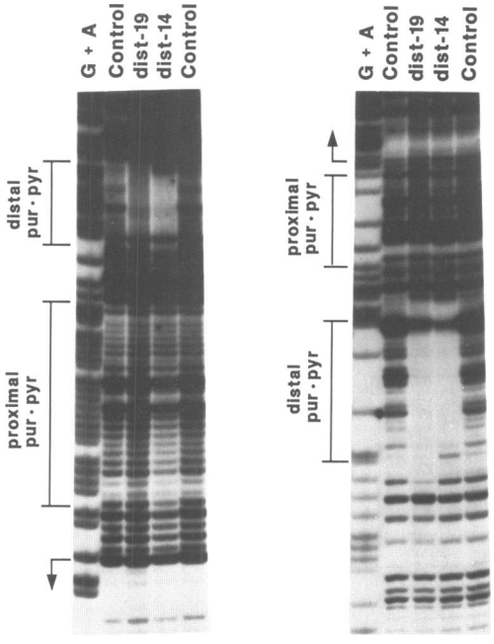

Fig. 2. DNAse I protection assays of triple helix formation at the distal pur·pyr region of the human DHFR promoter. Digestion products for the coding (purine-rich) and non-coding (pyrimidine-rich) strands of the native promoter double helix are shown in the left and right panels respectively. Maxam−Gilbert G+A sequencing reaction lanes and control DNAse 1 digests of the labelled 204 bp promoter fragment are included for position reference. The positions of the two pur·pyr regions and the DHFR transcription start site (arrow) are indicated. Areas of endonuclease protection and enhancements resulting from conformational alterations induced by oligonucleotide third strand binding are exhibited, demonstrating specific binding sites for the purine-rich oligonucleotides dist-19 and dist-14 within the distal pur·pyr region.

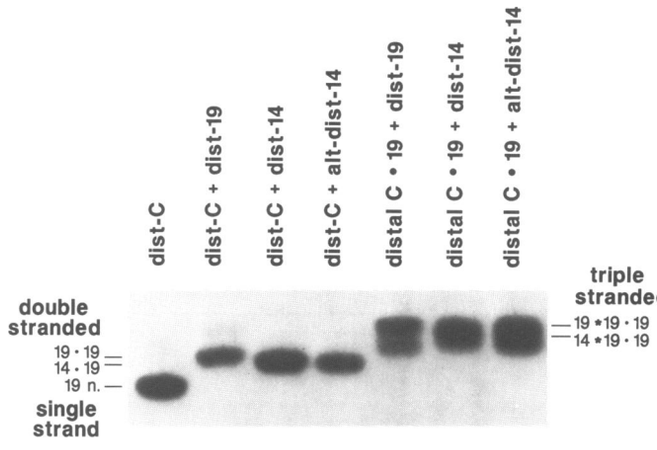

Fig. 3. Gel mobility shift assay of duplex and triplex formation by distal targeted oligonucleotides. A labelled 19 base C-rich oligonucleotide (dist-C; lane 1) with the sequence of the pyrimidine-rich strand of the distal pur·pyr region was used for duplex formation (hybridization) with the purine-rich oligonucleotides (lanes 2−4). A labelled 19 bp double stranded oligonucleotide (dist-C·dist-19, lane 2) with the sequence of the distal pur·pyr region was used as the target for triple helical binding by the distal targeted oligonucleotides (lanes 5−7). The relative mobilities of the single stranded, double stranded, and triple stranded structures in the native polyacrylamide gel are indicated. Since the distal pur·pyr sequence is symmetrical, the purine-rich distal targeted oligonucleotides can orient antiparallel to a single stranded pyrimidine-rich complement to form a duplex, or bind antiparallel to the purine-rich strand of the distal duplex to form a triplex.

helical structure overlapping the recognition sequence of the restriction endonuclease interfered with the activity of the enzyme at that site. On the other hand, pre-binding of dist-14 to the DHFR promoter did not prevent Hpa II digestion in > 90% of molecules. This result is consistent with the footprinting data which indicated dist-14 bound only to the downstream portion of the pur·pyr target; the triple helical structure formed would not overlap with the recognition sequence of the restriction enzyme at the upstream boundary. Protection of the Hpa II site by alt-dist-14, the mirror image of dist-14, indicates that it bound to the upstream portion of the pur·pyr region. These findings indicate a highly localized effect of pur·pur·pyr triplex formation on accessibility of the double helix, and specify an orientation of the purine-rich oligonucleotide third strand parallel to the pyrimidine-rich strand of the native double helix.

Triplex formation by the proximal pur·pyr sequence of the DHFR promoter

The symmetry of the distal pur·pyr region provided identical third strand to target triplet alignments regardless of orientation of the purine-rich oligonucleotide. In contrast, the proximal pur·pyr sequence is not symmetrical, and orientation of the oligonucleotide third strand becomes a critical factor. The first proximal targeted oligonucleotides were designed to bind in a parallel fashion to the purine-rich strand of the target duplex, as exemplified by prox-B:

5′ - GGGGA C GAGGGGGG C GGGG - 3′ (target = −24 to −6)

No evidence of triple helix formation was detected for these oligonucleotides. In spite of the similarity in sequence of the distal

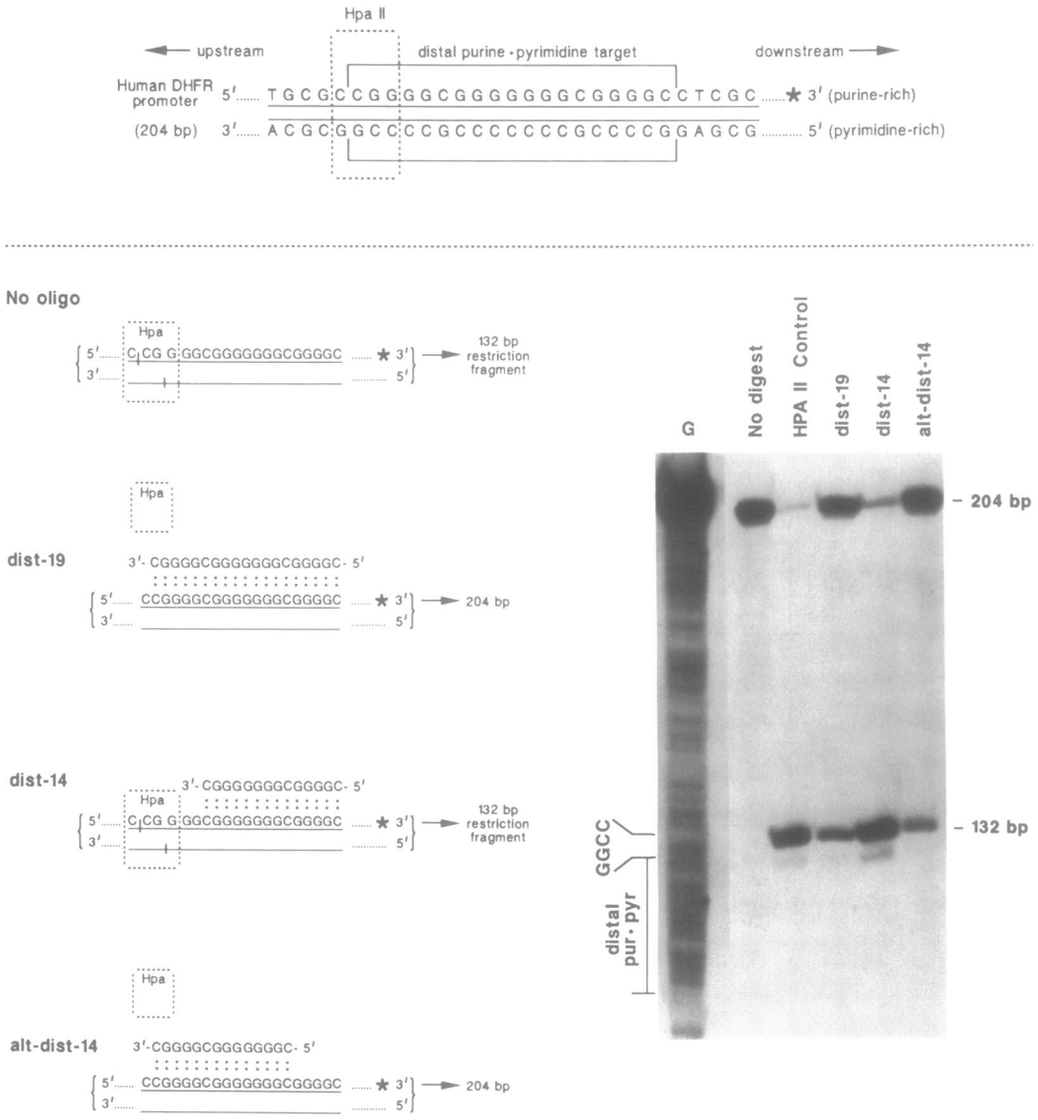

Fig. 4. Interference with restriction endonuclease Hpa II activity by triple helical binding of oligonucleotides targeted to the distal pur·pyr region of the DHFR promoter. Top Panel: Relationship of the Hpa II recognition site to the distal pur·pyr region within the human DHFR promoter. Left panel: Illustration of the effect of sequence specific triplex formation on restriction endonuclease digestion. Right panel: Denaturing polyacrylamide gel analysis of reaction products. The 204 bp DHFR promoter fragment is 3′-end labelled on the purine-rich (coding) strand. Hpa II digestion produces a 132 bp labelled restriction fragment. A Maxam−Gilbert G sequence lane is included for reference. Binding of dist-19 or alt-dist-14 but not dist-14 blocks Hpa II digestion of the DHFR promoter fragment, indicating that pur·pur·pyr triplex formation results in a localized limitation of access to the native double helix.

and proximal pur·pyr regions, oligonucleotides targeted to the distal sequence did not participate in triple helix formation with the proximal pur·pyr sequence either. We considered the possibilities that either the inclusion of A residues within the proximal pur·pyr sequence, or a secondary structure dependent on the flanking sequences of the proximal target might make triplex formation at this site impossible. Following definitive determination of the antiparallel orientation of binding of the purine-rich third strand at the distal pur·pyr region, a new group of oligonucleotides directed at the proximal target sequence was synthesized. For these, the purine-rich third strand was made identical in sequence to the purine-rich strand of the native duplex, but parallel in orientation to the pyrimidine-rich strand. Prototypical of these is prox-F:

3′ - C GAGGGGGG C GGGG C - 5′ (target = −19 to −5)

In addition, in order to compare the inherent ability of the proximal and distal pur·pyr sequences to accommodate a purine-rich third strand, use was made of the similarity between the two targets to design a 12 base oligonucleotide with sequence common to both regions, cmn-12:

3′ - GGGGGG C GGGG C - 5′ (targets = −16 to −5 and −51 to −40)

The binding of each of these oligonucleotides to a target duplex was assayed initially by gel mobility shift analysis (Figure 5). Both cmn-12 and prox-F exhibit stable triple helical binding to the labelled double stranded oligonucleotide with the sequence of the proximal pur·pyr target. Triplex formation between cmn-12 and a labelled double stranded oligonucleotide with the sequence of the distal pur·pyr target is also demonstrated, but only a minor degree of triple helical binding of prox-F (a three base terminal mismatch) to the distal pur·pyr target is seen, and the triple helical structure formed is unstable, dissociating during electrophoresis (note the smear between the duplex and triplex positions). Both cmn-12 and prox-F form stable hybrids with the labelled C-rich single stranded oligonucleotide, even though prox-F is a three base terminal mismatch with the sequence. The electrophoretic stability of the mismatch duplex but not the mismatch triplex suggests that the requirement for compatible sequence alignment may be more stringent for stable triple helix formation than for hybridization.

By using a 204 bp DNA restriction fragment (containing DHFR sequences −112 to +56) as the target, the effects of flanking sequences could be incorporated for a more physiologic assay of intermolecular triplex formation. Figure 6 illustrates the DNAse protection patterns produced by the binding of oligonucleotides targeted to one or both of the purine·pyrimidine regions of the DHFR promoter fragment. Alt-dist-14 protected only the distal pur·pyr region; cmn-12 afforded DNAse protection within both the distal and proximal targets; and prox-F footprinted primarily over the proximal pur·pyr region.

The upstream boundary of the alt-dist-14 footprint is identical to that of dist-19 (Fig. 2), while the downstream boundary is altered, indicating that the shorter molecule binds to the upstream portion of the distal pur·pyr region (−58 to −45) as expected from the results of the Hpa II interference experiment. The

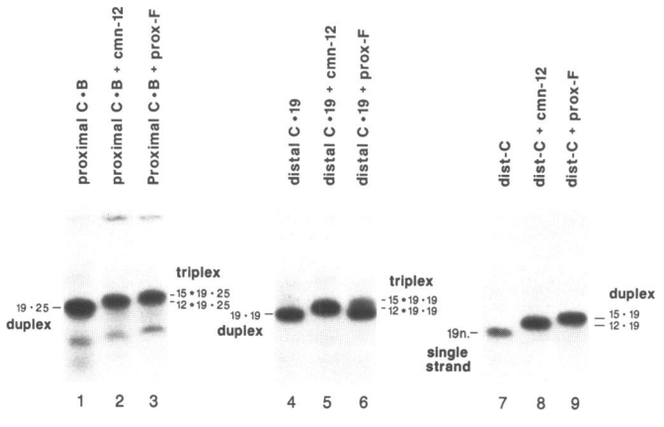

Fig. 5. Gel mobility shift assay of triple helix formation by matched and mismatched oligonucleotides on the distal and proximal pur·pyr sequences of the DHFR promoter. Labelled, double stranded oligonucleotides (prox-C·prox-B, lane 1; dist-C·dist-19, lane 4) with the sequence of the proximal or distal pur·pyr regions respectively were used as targets for triplex formation by purine-rich oligonucleotides. Both cmn-12 and prox-F bind the proximal target with which they align with sequence identical to (but orientation opposite of) the purine-rich strand (lanes 2 & 3). Cmn-12 also forms a stable triplex with the distal target (lane 5), but limited binding of prox-F to the distal target, for which it contains a 3 base terminal mismatch, is seen; the triple helical structure produced is unstable, dissociating during electrophoresis (lane 6). Both cmn-12 and prox-F exhibit stable duplex formation (hybridization) (lanes 8 & 9) with dist-C (lane 7) despite a 3 base terminal non-complementary alignment between prox-F and the C-rich sequence.

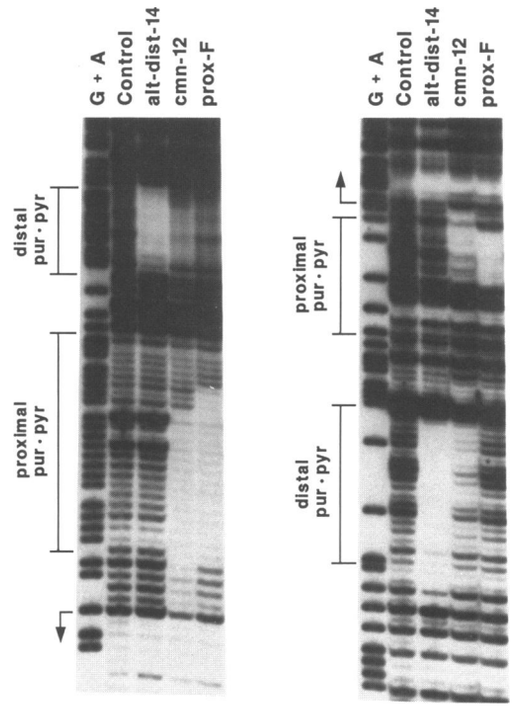

Fig. 6. DNAse I protection assays of triplex forming oligonucleotides targeted specifically for one pur·pyr sequence (alt-dist-14; prox-F) or both (cmn-12). Digestion products for the coding (purine-rich) and non-coding (pyrimidine-rich) strands of the DHFR promoter fragment are shown in the left and right panels respectively. A Maxam−Gilbert G+A sequence lane and control DNAse digest are included in each panel. The positions of the two pur·pyr regions and the DHFR transcription start site (arrow) are indicated. Triple helical binding of the purine-rich oligonucleotides to the human DHFR promoter is demonstrated by inhibition of endonuclease activity secondary to conformational alterations induced by the presence of a third nucleotide strand in the major groove of the native duplex. Although the distal and proximal pur·pyr sequences differ only by insertion of two A·T base pairs, they can be individually targeted for triplex formation due to the high degree of specificity of sequence alignment of the third strand.

DNAse I protection pattern produced by cmn-12 confirms that the proximal and distal target sequences can participate in triplex formation equally, and demonstrates stable triple helical binding by a third strand as small as 12 nucleotides. The footprint induced by prox-F indicates that the inclusion of an A residue in the oligonucleotide and its target sequence is compatible with triple helical binding. The minor degree of DNAse protection produced by prox-F at the distal pur·pyr region is a result of unstable mismatch triple helical binding.

An attempt was made to extend the sequence of the human DHFR promoter which could be targeted by an exogenously added third strand oligonucleotide. Prox-G represents an oligonucleotide targeted further upstream within the proximal pur·pyr region:

3′ - T GGGGA C GAGGGGGG C - 5′ (target = −25 to −10)

The successful triplex formation by this oligonucleotide as determined by both gel mobility shift and DNAse protection, indicates that multiple individual C and A residues can be accommodated in the purine-rich strand of the target, and that a T residue is also allowable, at least as a terminal nucleotide. An attempt to progress even further upstream was made with the 15mer, prox-H:

3′ - C AAA T GGGGA C GAGG - 5′ (target = −29 to −15)

This oligonucleotide did not participate in any detectable triplex formation. The multiple near terminal consecutive A residues may be responsible for this negative result; the internal T represents another possibility. Decreased stabilization of triplex formation attributable to a lower number of G*G·C triplets may contribute to the failure of prox-H to bind its target duplex (37).

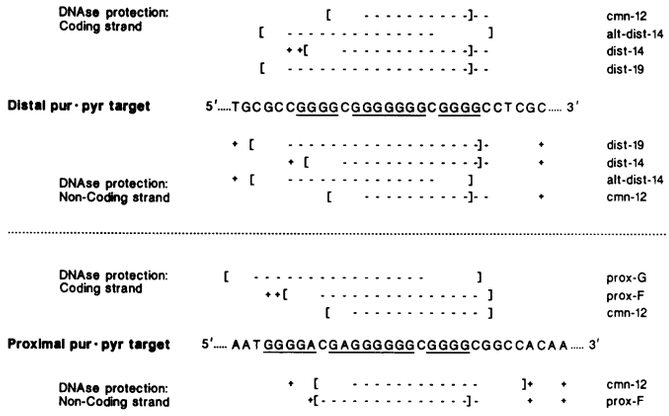

Fig. 7. Summary of data from DNAse I protection assays of triple helix formation by purine-rich oligonucleotides targeted to the human DHFR promoter. The sequences (purine-rich strand) of the distal and proximal pur·pyr regions are given. Results of DNAse I digests for the coding (purine-rich) strand of the native double helix are shown above the target sequence; results for the non-coding (pyrimidine-rich) strand are shown below the target sequence. The relationship of sites of endonuclease protection (brackets) and enhancements (plus signs) to the specific target sequences of the individual triplex forming oligonucleotides (hyphenated lines) are depicted. Footprints correspond roughly in size and position to the expected triple helical structures. Inhibition of DNAse 1 digestion from the minor groove by third strand binding in the major groove is dependent on conformational alterations, which are influenced by sequence accounting for the variability of footprint boundaries. DNAse I enhancements are indicative of distortions of secondary structure near the triplex-duplex boundaries. --- = specific oligonucleotide target, [ ] = DNAse protection, + = DNAse enhancement.

Summary of DNAse protection data

Figure 7 summarizes the data obtained from DNAse I protection assays of intermolecular pur·pur·pyr triplex formation at the distal and proximal pur·pyr target sequences of the human DHFR promoter. The region of the double helix protected from DNAse I digestion by oligonucleotide binding approximates the size and position of the expected triple helical structure. Footprints generally extended 2−3 base pairs upstream of the specific target on both the coding and non-coding strands. The extent of protection at the downstream end of each triplex was more variable, commonly not covering the entire target. The variability of the downstream protection border suggests that the endonuclease may be able to approach the triple helical structure better from the downstream side (5′ end of the third strand) than from the upstream side (3′ end of the third strand). Undoubtedly the native secondary structure and the inherent sensitivity of the target sequence to DNAse contribute to the patterns of DNAse protection.

Hypersensitive sites (DNAse enhancements) produced by oligonucleotide binding were detected, primarily on the non-coding pyrimidine-rich strand, and usually 4−6 base pairs on either side of the intended target sequence. These are indicative of an alteration or distortion in secondary structure near the triplex-duplex boundaries.

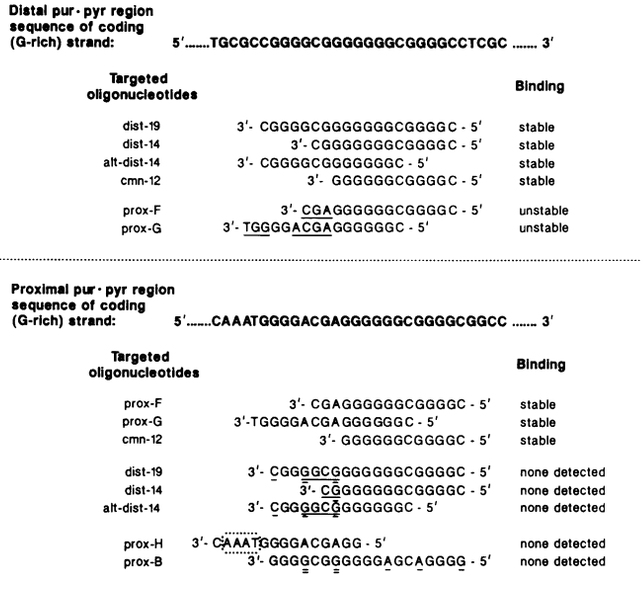

Fig. 8. Alignment of triplex forming oligonucleotides with target pur·pyr sequences. For purine-rich oligonucleotides exhibiting stable triple helical binding, sequences corresponded exactly with the purine-rich strand of the native duplex. Mismatches between proximal targeted oligonucleotides and the distal pur·pyr region include A*G·C, G*C·G, and C*G·C triplets; triple helical formation is unstable, dissociating upon electrophoresis and producing limited DNAse protection. Mismatches between distal targeted oligonucleotides and the proximal pur·pyr region include G*A·T triplets; no evidence of such triple helical structures is detected. The inability to detect binding of prox-H to the proximal target may reflect the requirement for non-Mg^{++} divalent cations for triplex formation when frequent or alternating A·T base pairs are present in the target. Prox-B had been (incorrectly) designed to bind parallel to the purine-rich strand of the native duplex; binding in the reverse of the orientation intended could not take place without misalignment of sequences. Third strand to target sequence mismatches are underlined. G*A·T alignments are double underlined. The contiguous A residues and internal T of prox-H are encircled.

Triple helical structures formed by different oligonucleotides which share one target boundary exhibit consistent alterations in the DNAse protection pattern at that common boundary (i.e. the downstream boundary of dist-19, dist-14, and cmn-12; the upstream boundary of dist-19 and alt-dist-14). The distal and proximal target sequences are very similar, differing only by the insertion of two A·T base pairs in the proximal region. Pairs of triplex forming oligonucleotides which bind analogous portions of the two target regions exhibit a similarity in DNAse protection pattern (i.e. dist-14 and prox-F; alt-dist-14 and prox-G).

Specificity of triple helical binding

The near identity of the two pur·pyr sequences has allowed us to measure the effects of some triplex third strand to target sequence mismatches. Figure 8 illustrates the alignment of the potential triple helix forming oligonucleotides with their target sequences in the human DHFR promoter. Mismatches of sequence between the oligonucleotide and the purine-rich strand of the target duplex are underlined, and the relative degree of triple helical stability for each alignment is indicated. The proximal targeted oligonucleotides exhibit detectable but unstable triple helical binding to the distal target sequence. Examination of the resulting alignments indicates that a combination of A*G·C, G*C·G, and C*G·C mismatches within the 2pur:1pyr conformation creates a significant destabilizing effect on the triplex structure. No interaction of the distal targeted oligonucleotides with the proximal pur·pyr sequence was found. It appears that the distinguishing feature of these alignments, the G*A·T triplet, is extremely unfavorable for pur·pur·pyr triple helical binding. Allowable triplet associations for the 2pur:1pyr model are undoubtedly a function of alignment of the phosphodiester backbone of the third strand in relation to the target double helix, and likely to differ from those of the 2pyr:1pur model (17,25,38).

DISCUSSION

Target sequence composition and third strand sequence alignment

In defining the sequence requirements and specificity of the 2pur:1pyr model of triple helix, two parameters must be considered: a) sequence composition of the target double helix, and b) sequence alignment of the nucleotides of the third strand with those of the native double helix.

The sequence requirements for triple helix formation on the proximal pur·pyr target (i.e. prox-F vs. prox-B) and the specific binding sites defined for the distal targeted triplex-forming oligonucleotides indicate that the third strand of the pur·pur·pyr structure must be oriented antiparallel to the purine-rich strand of the native double helix. These data are consistent with the original description by Kohwi and Kohwi-Shigematsu of an intramolecular 2pur:1pyr triplex (21) and the recent determination of orientation by Beal and Dervan for an intermolecular 2pur:1pyr structure (25).

The distal pur·pyr sequence of the human DHFR promoter consists of multiple consecutive G·C base pairs with occasional C·G interruptions. The third strand oligonucleotides targeted to this sequence have utilized C residues opposite the C interruptions in the purine-rich strand of the native duplex. The results of gel mobility shift and DNAse protection assays indicate that this C*C·G alignment is compatible with stable pur·pur·pyr triplex formation. This alignment is apparently quite specific, as the distal targeted oligonucleotides exhibit definitive binding sites within the distal pur·pyr region, in which the only landmarks within the otherwise oligo dG·dC duplex are the occasional C·G interruptions. The 14 base distal targeted oligonucleotides do not 'slide' within the larger 19 bp target.

The proximal pur·pyr sequence of the human DHFR promoter varies from its distal counterpart by the insertion of two individual A·T interruptions. The oligonucleotides targeted to this sequence have utilized A residues aligned with A interruptions and, as with the distal sequence, C residues aligned with C interruptions. Stable shifted triplex bands and appropriate DNAse protection patterns indicate that these alignments are compatible with stable 2pur:1pyr triple helical binding.

The recent report by Bernues, et. al. (24) and follow-up results of Lyamichev, et. al. (26) and Collier and Wells (10) have indicated that specific divalent cations other than Mg^{++}, such as Zn^{++} and perhaps Co^{++} or Mn^{++}, may permit formation of pur·pur·pyr triple helical structures when the purine-rich strand of the target contains multiple frequent or alternating A residues. This important observation may expand the capabilities of the 2pur:1pyr model, and a requirement for such cations may explain the failure of our prox-H with three consecutive A residues to bind (in the presence of magnesium) to its corresponding target within the proximal pur·pyr region of the DHFR promoter.

Consequences of triple helical binding

DNAse 1 is sensitive to the topology of the minor groove, binding across both strands to produce single stranded nicks that generally are staggered by three nucleotides on opposite strands (39,40). Triple helical binding of a purine-rich oligonucleotide within the major groove essentially abolishes DNAse 1 digestion of the two native strands over a sequence approximating the intended double helical target. Since the third strand binds within the major groove and associates primarily with the purine-rich strand, direct steric interference with DNAse 1 binding to the minor groove or cleavage of the pyrimidine-rich strand seems unlikely. Rather, a conformational alteration of the native double helix induced by binding of the oligonucleotide probably is responsible for the observed endonuclease protection. The resulting secondary structure is dependent upon the inherent reactivity of the particular sequences to the binding of the third strand. Francois, et. al. found a similar pattern using another minor groove specific probe, copper-phenanthroline, to measure intermolecular 2pyrimidine:1purine triple helical binding (15). A relatively uniform effect of triple helical binding on double helical conformation is demonstrated by the appearance of DNAse enhancements near the triplex-duplex boundaries on the pyrimidine-rich strand. We conclude that 2purine:1pyrimidine triple helical binding produces conformational alterations which extend several base pairs into the flanking sequences of the native double helix.

Binding of a purine-rich oligonucleotide third strand can block access of the restriction enzyme Hpa II to its recognition sequence in the native double helix. This effect is highly localized: formation of a triple helical structure which overlaps by three base pairs the endonuclease recognition site (alt-dist-14) does prevent digestion, but triplex formation by another oligonucleotide of the same size, targeted two base pairs away from the Hpa II site (dist-14) exhibits no interference. Hpa II digestion of the DHFR promoter in the presence of dist-14 binding, which is associated with conformational alterations extending into the Hpa

II recognition sequence (evidenced by DNAse enhancements) leads us to propose that restriction endonuclease interference is more sensitive to direct steric hindrance than to secondary structural alterations. Hanvey, et. al. found inhibition of restriction endonuclease activity by 2pyr:1pur triplex formation to be highly localized (18). We conclude that binding of a purine-rich oligonucleotide third strand sterically blocks access to the native double helix, with this effect restricted approximately to the sequences actually involved in the triple helical structure.

Utilization of triplex forming oligonucleotides to alter specific molecular interactions

Triplex formation exhibits a lower tolerance for unfavorable sequence alignments than does duplex formation (hybridization), and mismatched triple helical structures display a drastically decreased stability. Although this association of nucleic acids based on sequence recognition is not inherently suited for maximal binding affinity, the apparently very high degree of specificity of triple helical binding may be ideal for molecular intervention. Triplex formation blocks protein binding (16,18, 27), and the pur·pyr targets of the DHFR promoter represent biologically important protein binding sites.

An understanding of the functional architecture of the human DHFR promoter is still developing. The two Sp1 binding G/C box sequence elements contained within the two pur·pyr regions may have a differential influence on expression of DHFR. We have determined (2) previously that the distal G/C box has a higher affinity for Sp1 than its proximal counterpart, and its position (−49 to −40) relative to the transcription initiation site is consistent with potential function as a positive regulator of DHFR expression. The proximal G/C box is located uncharacteristically near to the transcription initiation site (−14 to −5), and separated by 2.5 turns of the double helix, resides on approximately the opposite surface of the DNA from the distal Sp1 binding site. A 'CAA box/ element 3', homologous among the three mammalian DHFR gene promoters which have been studied, is positioned between the two G/C box sequences and overlaps the proximal pur·pyr region. Undoubtedly within the site of assembly of the transcription initiation complex, the distal and proximal purine·pyrimidine sequences of the human DHFR promoter can be targeted specifically for triplex formation by exogenous purine-rich oligonucleotides.

ACKNOWLEDGMENTS

Support was provided by grants from NIH (CA42664 and RO1 CA42337), VA Medical Research Service, American Cancer Society (S.B.) and a Physician Scientist Award 1 K11CA01581-01 (S.B.). We wish to thank Barbara Wilson for her excellent assistance in preparation of the manuscript and Katherine Coleman for her excellent work in preparation of the figures.

REFERENCES

1. VanDyke, M.W. and Dervan, P.B. (1983) Biochemistry, **22**, 2373−2377.
2. Blume, S.W., Snyder, R.C., Ray, R., Thomas, S., Koller, C.A., and Miller, D.M. (1991) J. Clin. Invest., **88**, 1613−1621.
3. Snyder, R.C., Ray, R., Blume, S. and Miller, D.M. (1991) Biochemistry, **30**, 4290−4297.
4. Morgan, A.R. and Wells, R.D. (1968) J. Mol. Biol., **37**, 63−80.
5. Arnott, S. and Bond, P.J. (1973) Nature New Biol., **244**, 99−101.
6. Arnott, S. and Selsing, E. (1974) J. Mol. Biol., **88**, 509−521.
7. Marck, C. and Thiele, D. (1978) Nucleic Acids Res., **5**, 1017−1028.
8. Hanvey, J.C., Klysik, J., and Wells, R.D. (1988) J. Biol. Chem., **263**, 7386−7396.
9. Hanvey, J.C., Shimizu, M., and Wells, R.D. (1988) Proc. Natl. Acad. Sci. USA, **85**, 6292−6296.
10. Collier, D.A. and Wells, R.D. (1990) J. Biol. Chem., **265**, 10652−10658.
11. Belotserkovskii, B.P., Veselkov, A.G., Filippov, S.A., Dobrynin, V.N., Mirkin, S.M., and Frank-Kamenetskii, M.D. (1990) Nucleic Acids Res., **18**, 6621−6624.
12. Moser, H.E. and Dervan, P.B. (1987) Science, **238**, 645−650.
13. Praseuth, D., Perrouault, L., LeDoan, T., Chassignol, M., Thuong, N., and Helene, C. (1988) Proc. Natl. Acad. Sci. USA, **85**, 1349−1353.
14. Lyamichev, V.I., Mirkin, S.M., Frank-Kamenetskii, M.D., and Cantor, C.R. (1988) Nucleic Acids Res., **16**, 2165−2178.
15. Francois, J-C., Saison-Behmoaras, T., and Helene, C. (1988) Nucleic Acids Res., **16**, 11431−11440.
16. Maher, L.J.III, Wold, B., and Dervan, P.B. (1989) Science, **245**, 725−730.
17. Griffin, L.C. and Dervan, P.B. (1989) Science, **245**, 967−971.
18. Hanvey, J.C., Shimizu, M., and Wells, R.D. (1990) Nucleic Acids Res., **18**, 157−161.
19. Strobel, S.A. and Dervan, P.B. (1990) Science, **249**, 73−75.
20. Broitman, S.L., Im, D.D., and Fresco, J.R. (1987) Proc. Natl. Acad. Sci. USA, **84**, 5120−5124.
21. Kohwi, Y. and Kohwi-Shigematsu, T. (1988) Proc. Natl. Acad. Sci. USA, **85**, 3781−3785.
22. Cooney, M., Czernuszewicz, G., Postel, E.H., Flint, S.J., and Hogan, M.E. (1988) Science, **241**, 456−459.
23. Letai, A.G., Palladino, M.A., Fromm, E., Rizzo, V., and Fresco, J.R. (1988) Biochemistry, **27**, 9108−9112.
24. Bernues, J., Beltran, R., Casasnovas, J.M., and Azorin, F. (1989) EMBO J., **8**, 2087−2094.
25. Beal, P.A. and Dervan, P.B. (1991) Science, **251**, 1360−1363.
26. Lyamichev, V.I., Voloshin, O.N., Frank-Kamenetskii, M.D., and Soyfer, V.N. (1991) Nucleic Acids Res., **19**, 1633−1638.
27. Gee, J.E., Blume, S., Snyder, R.C., Ray, R., and Miller, D.M., J. Biol. Chem., in press.
28. Chen, M-J., Shimada, T., Moulton, A.D., Cline, A., Humphries, R.K., Maizel, J., and Nienhuis, A.W. (1984) J. Biol. Chem., **259**, 3933−3943.
29. Mitchell, P.J., Carothers, A.M., Han, J.H., Harding, J.D., Kas, E., Venolia, L., and Chasin, L.A. (1986) Mol. Cell. Biol., **6**, 425−440.
30. Farnham, P.J. and Schimke, R.T. (1986) Mol. Cell. Biol., **6**, 2392−2401.
31. Azizkhan, J.C., Vaughn, J.P., Christy, R.J., and Hamlin, J.L. (1986) Biochemistry, **25**, 6228−6236.
32. Swick, A.G., Blake, M.C., Kahn, J.W., and Azizkhan, J.C. (1989) Nucleic Acids Res., **17**, 9291−9304.
33. Fujii, H. and Shimada, T. (1989) J. Biol. Chem., **264**, 10057−10064.
34. Means, A.L. and Farnham, P.J. (1990) Mol. Cell. Biol., **10**, 653−661.
35. Shea, R.G., Ng, P., and Bischofberger, N. (1990) Nucleic Acids Res., **18**, 4859−4866.
36. Francois, J.C., Saison-Behmoaras, T., Thuong, N.T., and Helene, C. (1989) Biochemistry, **28**, 9617−9619
37. Pilch, D.S., Levenson, C., and Shafer, R.H. (1991) Biochemistry, **30**, 6081−6087.
38. Durland, R.H., Kessler, D.J., Gunnell, S., Duvic, M., Pettitt, B.M., and Hogan, M.E. (1991) Biochemistry, **30**, 9246−9255.
39. Drew, H.R. (1984) J. Mol. Biol., **176**, 535−557.
40. Drew, H.R. and Travers, A.A. (1984) Cell, **37**, 491−502.