Effect of abasic linker substitution on triplex formation, Sp1 binding, and specificity in an oligonucleotide targeted to the human Ha-ras promoter

Charles Mayfield^{2,*} and Donald Miller^{1,2}

Bolden Laboratory, Departments of ¹Biochemistry and ²Medicine, University of Alabama at Birmingham, Birmingham, AL and Birmingham VA Medical Center, Birmingham, AL, USA

Received August 30, 1993; Revised December 6, 1993 and Accepted February 26, 1994

ABSTRACT

A region of the human Ha-ras promoter (–8 to –28) which contains two of the three Sp1 binding sites essential for transcriptional activity forms a sequence specific oligonucleotide-directed pur*pur:pyr triple helix. The relative binding of oligonucleotides containing different substitutions, including an abasic propanediol linker, over three potentially destabilizing C:G interruptions in the otherwise poly G:poly C target was examined. DNase I footprint titrations reveal that substitution of the positively charged abasic propanediol linker results in approximately ten fold greater binding than cytosine substitution which in turn provides greater sequence specific binding than substitution of a guanine in the third strand oligonucleotide over the C:G interruptions. Protein binding assays demonstrate that triplex formation by the linker substituted oligomer (HR21Xap) is less effective in inhibiting Sp1 binding than the cytosine substituted oligomer (HR21ap) both to the target sequence as well as an upstream sequence. As an indication of the effect of linker substitution and targeting consensus Sp1 sites on triplex specificity, the relative ability of the Ha-ras promoter targeted oligonucleotides to interact with non-target Sp1 sequences within the Ha-ras promoter as well as in the DHFR promoter and HIV-1 LTR was also investigated. At concentrations which afford complete DNase I protection of the target sequence, HR21ap does not bind to the non-target sequences while HR21Xap interacts weakly only at a distal site in the DHFR promoter. Also, HR21ap as well as HR21Xap are specific in their inhibition of Sp1 binding. These results suggest that the propanediol linker is able to skip over interruptions in a target sequence thereby stabilizing triplex but, slightly compromises sequence specificity and the ability to inhibit Sp1 binding to the Ha-ras promoter.

INTRODUCTION

Purine and pyrimidine rich oligonucleotides have been shown to form pur*pur:pyr and pyr*pur:pyr intermolecular triple helices with predominantly polypurine:polypyrimidine (pur:pyr) sequences (1–6). In these triplex structures the oligonucleotide third strand lies in the major groove of the duplex, forming Hoogsteen hydrogen bonds with specific purine bases of the duplex (1,7). In pyr*pur:pyr triplex the pyrimidine rich third strand is oriented parallel to the purine rich strand of the duplex through the formation of T*A:T and pH dependent C**G:C triplets (1). At physiological pH, stable pur*pur:pyr and mixed pur/pyr*pur:pyr triplexes have been formed with predominantly G*G:C triplets along with A*A:T and T*A:T triplets interspersed (7–9). In pur*pur:pyr and pur/pyr*pur:pyr triple helices the third strand has been shown to bind in an antiparallel orientation with respect to the purine rich duplex strand through reverse-Hoogsteen hydrogen bonding (7–10).

Because of their ability to recognize specific sequences in duplex DNA, triplex forming oligonucleotides offer the potential for gene specific transcriptional repression. Oligonucleotide-directed triplex formation has been shown to inhibit binding of the transcription factor Sp1 both to sites inserted in the murine metallothionein I promoter (6) and to the DHFR promoter (10) and more recently, the binding of NFkB to a interleukin-2 receptor (IL-2R) α regulatory sequence (11). Triplex forming oligonucleotides targeted to a PuF binding site in c-myc (8), to the NFkB site in IL-2R (11,12) and to Sp1 sites in the HIV1-LTR (13) have been shown to repress gene transcription in intact cells. However, a significant limitation to the further application of triplex in biologically occurring sequences is the requirement for predominantly polypurine:polypyrimidine tracts.

While pur:pyr rich sequences occur disproportionately in the promoters of eukaryotic genes (14), there are often pyr:pur base pairs interrupting the pur:pyr tract. Because of the loss of Hoogsteen bonding and/or steric hinderance introduced by the third strand base, interruptions in pur:pyr tracts potentially destabilize triplex formation, decreasing oligonucleotide affinity or preventing triplex formation (15).

*To whom correspondence should be addressed at: 1824 6th Avenue South, Wallace Tumor Institute, Room 520, Birmingham, AL 35294-3300, USA

We have recently demonstrated oligonucleotide-directed triplex formation in a pur:pyr rich region of the *Ha-ras* promoter (−8 to −28) which encompasses the two Sp1 binding sites most proximal to the initiation sites that are required for full transcriptional activity (16). A *Ha-ras* promoter targeted oligonucleotide binds sequence specifically in an antiparallel orientation relative to the purine-rich duplex strand through G*G:C and C*C:G alignments and inhibits Sp1 binding to its target site. However, this oligonucleotide displays relatively poor binding to the target sequence. This low affinity is likely due to the three potentially destabilizing C*C:G interruptions in an otherwise stable G*G:C triplex. In the present study we have examined the effects of different oligonucleotide base substitutions, including a positively charged, abasic linker, over these C:G interruptions on triplex formation and the relative ability of these oligonucleotides to inhibit Sp1 binding to the *Ha-ras* promoter. In addition, to assess the effect of linker substitution and the targeting of consensus Sp1 sites in general on triplex specificity, the relative ability of the *Ha-ras* promoter targeted oligonucleotides to interact with non-target Sp1 sequences within the *Ha-ras* promoter as well as in the DHFR promoter and HIV-1 LTR was also investigated.

MATERIALS AND METHODS

Oligonucleotide synthesis

Oligonucleotides were synthesized on a Milligen Cyclone DNA synthesizer using phosphoramidite chemistry. The abasic linker substituted oligonucleotide (HR21Xap) was synthesized using N-Fmoc-O¹-DMT-O³-cyanoethoxydiisopropylaminophosphinyl-2-aminobutyl-1,3-propanediol (Clontech). All oligonucleotides were purified by preparative 16–20 % polyacrylamide gel electrophoresis and their structural integrity and purity established by 5' ³²P-labelling and electrophoresis on polyacrylamide gels. Concentrations were determined from absorbance measurements at 260 nm using molar extinction coefficients. Calculated molar extinction coefficients for oligonucleotides were as follows: HR21ap = 232,500 cm^{−1}M^{−1}, HR21Gap = 245,700 cm^{−1}M^{−1}, HR21Xap = 210,600 cm^{−1}M^{−1}. Oligonucleotides were 5' ³²P-labelled using [γ-³²P]ATP and T₄ polynucleotide kinase.

Promoter fragment isolation

The +36 to −157 region of the human c-*Ha-ras* 1 gene was PCR amplified from *Bam*HI linearized plasmid pT24 DNA containing the entire c-*Ha-ras* 1 gene from human bladder carcinoma using Vent DNA polymerase. The 5' PCR primer was designed with a *Eco*RI recognition sequence and the 3' primer with a *Bam*HI recognition sequence. The PCR amplified fragment was purified from low melting agarose gel, digested with *Eco*RI and *Bam*HI, and cloned into pTZ 18R vector. The phagemid was digested with *Eco*RI, ³²P-labeled with [α-³²P]dATP using the Klenow fragment of DNA polymerase, then digested with *Bam*HI. A 186 bp promoter fragment from the DHFR gene containing two Sp1 binding sites was isolated from a subclone of the 1.8 kb *Eco*RI human DHFR clone (17). After digestion of the plasmid with *Eco*RI the fragment was ³²P-labeled with [α-³²P]dATP using the Klenow fragment of DNA polymerase. It was then digested with *Not*I to yield the 186 bp fragment.

A 141 bp HIV-1 LTR fragment containing three Sp1 binding sites was isolated from plasmid pL5'-A1 by digestion with *Bgl*II followed by ³²P-labelling with Klenow and [α-³²P]dATP then

digestion with *Taq*I. All plasmid digests were purified by electrophoresis on 5 % native polyacrylamide gels. The desired labeled fragment was excised from the gel then crushed and eluted overnight in 0.5 M NH₄OAc. The eluate was filtered through a 0.45 μm filter and the DNA ethanol precipitated. The sequence of all fragments was confirmed by Maxam–Gilbert sequencing (18).

Gel mobility shifts

For triplex shifts the synthetic coding strand of the 21 bp c-*Ha-ras* triplex target sequence was 5' ³²P-labeled with [γ-³²P]ATP and annealed to its oligonucleotide complement. Potential triplex forming oligonucleotides were heated at 65°C for 10–15 minutes to reduce self-aggregation of the G-rich oligos then cooled on ice. They were then added to the labeled 21 bp target in 90 mM Tris–borate (pH 8.0)/10 mM MgCl₂ and incubated for 45 minutes at 25°C. Samples were analyzed by electrophoresis on 16 % native polyacrylamide gels at 100 V for 10–15 hrs then exposed at −70°C for autoradiography. Both gel and running buffer contained 90 mM Tris–borate (pH 8.0)/10 mM MgCl₂.

Protein binding assay

For Sp1 shifts oligonucleotides were heated at 65°C then quick cooled on ice and added to ³²P-labeled promoter fragment. Samples were incubated in protein binding buffer consisting of 20 mM Tris (pH 8.0)/100 mM KCl/1.5 mM MgCl₂/1 mM dithiothreitol/8 % (vol/vol) glycerol for 45–60 min at 25°C. The samples were cooled on ice then recombinant Sp1 516C was added and allowed to bind for 30 min on ice. Samples were analyzed by electrophoresis on 5 % native polyacrylamide gels at 150 V in 90 mM Tris–borate (pH 8.5)/2 mM EDTA followed by autoradiography. The gel contained 90 mM Tris–borate (pH 8.5)/2 mM EDTA without Mg²⁺.

DNase I footprinting

After heating at 65°C for 10–15 min then cooling on ice oligonucleotides were incubated with ³²P-labeled promoter fragment in 20 mM Tris (pH 7.4) and 10 mM MgCl₂ for 45–60 min at 25°C. Samples were precooled on ice then digested with DNase I for 60 sec on ice. Digestion was terminated by the addition of 20 mM EDTA in 90 % formamide then heating at 95°C for 5 min to inactivate DNase I. Samples were quickly cooled on ice then analyzed on 8M urea/8 % polyacrylamide sequencing gel at 42 W. Bands were visualized by autoradiography and protected sequences identified from gels in which both Maxam–Gilbert sequencing and DNase I footprints were run.

Protein extract preparation

The Sp1 extract was prepared as described by Kadonaga *et al.* (19) using pSp1-516C and *Escherichia coli* strain BL21. The recombinant Sp1-516C was used as a crude extract without further purification.

RESULTS

Oligonucleotide design

Two of the three human c-*Ha-ras* Sp1 binding sites essential for transcription (20) were targeted for oligonucleotide directed triplex formation (Figure 1). These two Sp1 binding sites are most proximal to the transcription start sites and separated by a single base pair. The sequence in this region is not strictly

Figure 1. Map of the human Ha-*ras* promoter showing the three Sp1 binding sites essential for transcriptional activity. Positions indicated are relative to the 5' most transcription initiation start site (+1). The triple target sequence (–8 to –28) is shown in brackets with Sp1 binding sequences underlined.

A

B

Figure 2. (A) Sequences of base substituted oligodeoxynucleotides and their alignment with the Ha-*ras* target site. The C:G interruptions in the target sequence and oligodeoxynucleotide base substitutions are indicated by bold type. Oligodeoxynucleotides containing cytosine, guanine, or a propanediol linker (X) opposite C:G base pairs in the target sequence are designated HR21ap, HR21Gap, or HR21Xap, respectively. (B) Structure of the propanediol abasic linker shown between two guanine bases (GXG).

homopurine/homopyrimidine, but contains three C:G interruptions in an otherwise 21bp G:C tract that encompasses both Sp1 sites. Based on previous studies in our laboratory with a very similar sequence (10,21), the Ha-*ras* targeted triplex forming oligonucleotide HR21ap was originally designed with guanosine to recognize G:C base pairs and with cytosine aligned with the C:G base pairs in the target (Figure 2). The relatively high dissociation constant inferred from gel mobility shift analysis and the lack of DNase I protection below 100 μ M demonstrated that HR21ap has a low affinity for its target sequence. This low affinity is likely due to the C:C:G interruptions within the otherwise stable G*G:C triplex (15). These observations prompted us to investigate other base substitutions in HR21ap,

Figure 3. Gel mobility shift analysis of the relative binding of base substituted oligonucleotides to the Ha-*ras* promoter target. Oligonucleotides containing C substitutions (HR21ap), G substitutions (HR21Gap), or propanediol linker substitutions (HR21Xap) were added to 1.0 nM 32 P-labeled 21bp duplex at the concentrations indicated above each lane. Lane 1 is control DNA containing single strand 21mer and 21bp duplex. S = single strand DNA, D = duplex DNA, T = triplex DNA

Figure 4. DNase I footprint titrations of the base substituted Ha-*ras* triplex forming oligonucleotides demonstrating the relative effect of different base substitutions over C:G interruptions in the target sequence on binding. Oligonucleotides containing either C substitutions, HR21ap (A), G substitutions, HR21Gap (B), or propanediol linker substitutions, HR21Xap (C) were incubated at the concentration indicated above each lane with 100 nM 32 P Klenow labeled promoter fragment followed by DNase I digestion. Control digests (lane 1) were of promoter fragment incubated in the absence of any oligonucleotide. The Ha-*ras* target sequence determined from alignment of Maxam–Gilbert sequencing data ran alongside DNase I digests is indicated by brackets.

including a 2-(4-aminobutyl)-1,3-propanediol linker (Figure 2B), opposite the three C:G interruptions.

Triplex formation by substituted oligonucleotides

The relative binding of oligonucleotides in which cytosines (HR21ap), guanines (HR21Gap), or propanediol linkers (HR21Xap) were aligned with the C:G interruptions was compared by gel mobility shift analysis (Figure 3) and DNase I footprint titrations (Figure 4). In gel mobility shifts HR21ap shifts only a small percentage of duplex at 10.0 μ M. However, HR21Gap, a poly G 21mer which contains guanine substitutions

Figure 5. Gel mobility shift analysis of the effect of triplex forming oligonucleotides on the binding of Sp1 to the 199 bp human *Ha-ras* promoter fragment. A 199bp ³²P Klenow labeled *Ha-ras* promoter fragment (2.5 nM) containing the two Sp1 binding sites/target sequence and an additional upstream Sp1 site was incubated in the absence and presence of the indicated concentration of the respective oligonucleotide followed by incubation with recombinant Sp1. The Sp1-DNA complexes formed are indicated as B¹ and B². Unbound or free 199bp probe is indicated by F.

instead of cytosine opposite the C:G interruptions induces an almost complete retardation of the duplex target at 0.1 μ M. Increasing concentrations of HR21Gap result in very little increase in the amount of duplex shifted with a small amount still being present even at 10 μ M. Unlike HR21Gap, oligonucleotide HR21Xap which incorporates the propanediol linker moieties over the C:G interruptions displays significant concentration dependence in its ability to shift the duplex target. HR21Xap produces very little, if any, shift at 0.1 μ M but, a majority of the target is retarded at 1.0 μ M and a complete transition occurs by 10.0 μ M.

The relative sequence specific binding of the HR21 oligonucleotides was compared by DNase I footprint titrations. As shown in Figure 4, all oligonucleotides yield footprints corresponding to protection of the 21 base pair target sequence, defined by the brackets, as well as approximately 7 or 8 base pairs upstream of the target sequence. This 5' flanking sequence, as well as 3 or 4 base pairs downstream of the target sequence, is fully protected by HR21Xap. Preliminary results from DMS footprinting (data not shown) do not show corresponding protection of the flanking regions suggesting that the extended DNase I protection may be due to induced conformational changes and/or occlusion of DNase I and not to any interaction of the oligonucleotide to either side of the target sequence. The footprint titrations reveal that HR21ap provides significant protection of the target sequence from DNase digestion at 100 μ M (Figure 4A, lane 5). At the same concentration HR21Gap provides only partial protection from digestion as shown by the presence of faint bands within the footprint in lane 5 of Figure 4B. Therefore, the gel retardation produced by HR21Gap to the extent observed is apparently not due to sequence specific interactions and may be attributable to nonspecific binding of the poly-G 21mer to the duplex DNA. This effect has also been observed in previous studies in which a poly-G 17mer targeted to the DHFR promoter gave shifts at much lower concentrations than that required for DNase I protection of the target sequence

(22). HR21Xap fully protects against digestion of its target sequence at 10 μ M (Figure 4C, lane 4) while HR21ap or HR21Gap afford no apparent protection at this concentration (lanes 4 of Figures 4A and 4B). Greater protection is afforded by HR21Xap at 10 μ M than either HR21ap or HR21Gap at 10 fold higher concentration.

Relative effect of triplex forming oligonucleotides on Sp1 binding

The relative effect of the triplex forming oligonucleotides on Sp1-516C binding to the 199 bp promoter fragment was determined by gel mobility shift analysis (Figure 5). This fragment contains the two Sp1 sites within the target sequence as well as an upstream Sp1 site. Figure 5, lane 2 shows that all the labeled promoter fragment is bound by Sp1 as evidenced by retardation of all the labeled fragment upon incubation with recombinant Sp1. Incubation of the promoter fragment with a greater amount of Sp1 does not give rise to an additional protein-DNA complex indicating that there is not preferential binding to only some of the three Sp1 sites due to a limiting amount of Sp1. Also, Sp1 has been shown to bind to all three sites simultaneously (20). Therefore, the upper major band (B₂) likely corresponds to all three Sp1 sites being occupied while the lower band (B₁) is likely a result of Sp1 binding to only one or two of the three sites. Incubation of the *Ha-ras* promoter with triplex forming HR21ap at 2.5 μ M (1000 fold molar excess) followed by incubation with Sp1 results in complete abrogation of the top retarded band and a significant loss of the lower band (lane 4). The lack of appearance of an additional complex upon virtually complete inhibition of B₁ and B₂ suggests that Sp1 binding to all three sites are inhibited by HR21ap. HR21Xap at the same concentration also inhibits formation of the upper Sp1 complex but, less effectively than HR21ap while it has little effect on the lower complex (lane 8). Even though HR21Xap displays ten fold greater binding than HR21ap to the target sequence it is less effective in inhibiting Sp1 binding. This effect is also observed in protein binding assays in triplex forming conditions and thus is not the result of a reversal in triplex forming ability of the two oligonucleotides in protein binding conditions (data not shown). HR21Gap, which affords only partial DNase I protection of the target sequence at a 1000 fold molar excess, does not appear to have any effect on Sp1 binding to the same fragment at the same molar excess (lane 6). The control nonsense oligonucleotide which does not form triplex also has no effect on Sp1 binding up to 2.5 μ M (1000 fold excess) (lanes 9 and 10). Also, at the same concentrations labeled oligonucleotide incubated with Sp1 results in less than 10% shift indicating that inhibition of Sp1 binding is not due to interaction with oligonucleotide. This is further demonstrated by the lack of any effect on Sp1 binding by the poly G 21mer HR21Gap. Thus, inhibition of Sp1 binding to the *Ha-ras* promoter appears to be a direct result of triplex formation.

Oligonucleotide specificity for target Sp1 sequence

The relative ability of HR21ap and HR21Xap to interact with the distal Sp1 site in the *Ha-ras* promoter as well as sequences containing Sp1 sites in the DHFR promoter and HIV-1 LTR was investigated by DNase footprinting. The target site in *Ha-ras* contains a consensus Sp1 sequence (GGGGCGGGGC) which is separated by one bp from a sequence which differs from the consensus by only one bp (GGGGCGGGGG) (Figure 2A). The

A

Distal Sp1 Site in Ha-ras Promoter

-148-119

3'-GCCTGCCCGCGCCCCGCCCAGCAGCGTCC-5'

5'-CGGACGGGGCGCGGGGCGGGGCGTGCGCAGG-3'

B

Alignment of Oligonucleotides With Non-Target Sp1 Sequences in DHFR Promoter

[illegible]

C

HIV-1 LTR With Spl Binding Sites

-82 -40

3'-GGTCCCTCCGCACCGGACCGCCCTGACCCCTCACCGCAGGGA-5'

5'-CCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGTCCCT-3'

Figure 6. Non-target sequences containing Sp1 binding sites in the (A) Ha-*ras* promoter, (B) DHFR promoter and, (C) HIV-1 LTR. Illustrated in B are alignments of Ha-*ras* targeted triplex forming oligonucleotides HR21ap (N=C) and HR21Xap (N=X) with the proximal and distal non-target Sp1 sequences in the DHFR promoter. Oligonucleotides are shown aligned with third strand G*C:G mismatches (underlined) and with stretches of the third strand to form identical triplets as in the target sequence (indicated by bold type) leaving 3' and/or 5' ends unbound.

distal region of the Ha-*ras* promoter is also a G:C rich sequence containing a consensus Sp1 site (Figure 6A). Similarly, the distal Sp1 site in the DHFR promoter contains a consensus Sp1 sequence which overlaps by two bp a sequence that differs from the consensus at the same position as in Ha-*ras* (Figure 6B). The HIV-1 LTR contains three non-consensus Sp1 sites within a 33bp sequence (Figure 6C).

Figure 7A shows the full footprints resulting from complete DNase I protection of the target sequence (upper bracket) by HR21ap and HR21Xap at a 1000 and 100 fold molar excess, respectively. However, there is no protection of the similar sequence containing the third distal Sp1 site (lower bracket), indicating that both oligonucleotides are specific for their target sequence and do not bind to the distal site. At the same molar excess which gives full protection of the Ha-*ras* target sequence, neither HR21ap (lane 2) or HR21Xap (lane 3) provide any protection of the proximal DHFR promoter (Figure 7B, lower brackets) or the HIV-1 LTR (Figure 7C). However, in the region containing the distal Sp1 site in the DHFR promoter (Figure 7B, upper brackets) some slight DNase protection is detectable in the presence of HR21Xap (lane 3) but, not in the presence of HR21ap (lane 2). At ten fold greater concentration HR21Xap gives very little additional protection of the DHFR site and still significantly less than full DNase I protection (data not shown). This data demonstrates that while the use of the abasic linker allows for increased binding to its target sequence its ability to also skip over interruptions in a highly homologous non-target sequence results in a slight decrease in specificity. Still, HR21Xap appears to have more than ten fold greater specificity for its target

sequence over the distal DHFR site even though 15 of the 18 bases in HR21Xap can be aligned so as to form G*G:C triplets with the DHFR site (Figure 6B).

Effect of Ha-*ras* targeted oligonucleotides on Sp1 binding to nontarget promoters

The relative effect of the Ha-*ras* targeted triplex forming oligonucleotides on Sp1 binding to targeted Ha-*ras* and nontargeted DHFR and HIV-1 promoters was examined by protein gel mobility shift analysis. Figure 8 shows the inhibition of Sp1 binding to the Ha-*ras* promoter by HR21Xap (lane 2) and HR21ap (lane 3) at a 1000 fold molar excess with respect to the promoter fragment. Consistent with the results from the protein binding assay shown in Figure 5, HR21Xap significantly inhibits formation of the upper Sp1-promoter complex but has little effect on the lower complex while HR21ap completely ablates the upper complex and inhibits much of the lower complex. At identical absolute concentrations of oligonucleotide and molar excesses neither Ha-*ras* targeted oligonucleotide affects Sp1 binding to the DHFR promoter to any significant extent (lanes 4–6). Both oligonucleotides appear to have a small effect on Sp1 binding to the HIV-1 LTR (lanes 7–9). However, comparison of lanes 1–3 with 7–9 reveals that both HR21Xap and particularly HR21ap produce greater inhibition of Sp1 binding to the Ha-*ras* promoter than to the HIV-1 LTR. The results from these studies suggest that the Ha-*ras* targeted oligonucleotides may be sufficiently specific in their inhibition of Sp1 binding. This data also further demonstrates that the inhibition of Sp1 binding to the Ha-*ras* promoter by HR21Xap and HR21ap to the extent

Figure 7. DNase I footprint analysis of promoter binding specificity of *Ha-ras* targeted triplex forming oligonucleotides. A 32 P Klenow labeled 199bp *Ha-ras* promoter fragment (A), 186bp DHFR promoter fragment (B), or 141bp HIV-1 LTR fragment (C) was incubated in the absence (lane 1) or presence of the *Ha-ras* targeted oligonucleotides HR21ap (lane 2) or HR21Xap (lane 3) at 10^3 or 10^2 molar excess, respectively, then subjected to DNase I digestion. The sequence of the target site (upper bracket in A), as well as nontarget Sp1 binding sequences in *Ha-ras* (lower bracket in A), in DHFR and the HIV-1 LTR are indicated by brackets. The Sp1 binding sites are indicated by lines beside the sequence.

observed is due to triplex formation and not to oligonucleotide interference with Sp1 itself.

DISCUSSION

Effect of basic and abasic substitutions over C:G interruptions

The human *Ha-ras* oncogene promoter contains three Sp1 binding sites essential for transcriptional activity (20). Two of these sites are within a 21 bp region immediately upstream (−8 to −28) of the major transcription start sites. In an effort to develop sequence specific DNA binding agents which may inhibit transcription of this gene we have designed triplex forming oligonucleotides targeted to this G:C rich region of the human *Ha-ras* promoter. An oligonucleotide (HR21ap) oriented antiparallel to the purine rich duplex strand was shown previously to bind in a sequence specific manner to the target sequence yet, displayed low affinity. The target sequence in *Ha-ras* contains three C:G interruptions in an otherwise G:C motif which likely contribute to the low affinity of HR21ap by the loss of Hoogsteen bonding and/or steric hinderence introduced by a third strand base at these positions. Initially, cytosine was used in the third strand over these C:G interruptions since cytosine has been used previously in our lab over C:G interruptions in a DHFR promoter target (10,21). Guanines have also been used frequently in different third strand oligonucleotides opposite a number of C:G interruptions and have been shown to form mixed PU/PY-PU:PY triplex with their targets (9).

Recently, Beal and Dervan (23) have determined the relative cleaving abilities of EDTA linked 15mer oligonucleotides with

Figure 8. Gel mobility shift analysis of the specificity of Sp1 inhibition by *Ha-ras* promoter targeted triplex forming oligonucleotides. A 32 P Klenow labeled 199bp *Ha-ras* promoter fragment (lanes 1–3), 186bp DHFR promoter fragment (lanes 4–6), or 141bp HIV-1 LTR fragment (lanes 7–9) at 20.0 nM was incubated in the absence (lanes 1, 4, and 7) or presence of oligonucleotides HR21Xap (lanes 2, 5, and 8) or HR21ap (lanes 3, 6, and 9) at the indicated concentrations and molar ratios followed by incubation with recombinant Sp1-516C.

all possible triplet combinations at a single position in a pur*pur:pyr triplex consisting of otherwise G*G:C and T*A:T triplets. In this triplex oligonucleotides with adenosine, guanine, or cytosine opposite a single C:G base pair yielded little or no cleavage while thymidine substitution gave better but, still weak cleavage. However, substitution of thymidine (or adenine) in the third strand over C:G interruptions in an otherwise entirely G*G:C triplex in the DHFR promoter has been shown to significantly decrease binding in DNase I protection titrations (22). These differences may be attributable to the differences between triplex forming conditions used to optimize differences in cleavage efficiencies and those used with the *Ha-ras* and DHFR targeted oligonucleotides as well as sequence dependent effects. The lack of a specific, stabilizing third strand base to recognize C:G base pairs represents a current limitation to this type of triplex. The three C:G base pairs in the *Ha-ras* target may represent the maximum number of such interruptions that can be tolerated in this type sequence and still maintain triplex formation. We show that substitution of a novel positively charged, abasic linker in a third strand oligonucleotide (HR21Xap) over the three C:G bp in the *Ha-ras* target results in approximately a ten fold, or greater, increase in binding as determined by DNase I footprinting while substitution of a guanosine decreases sequence specific binding. This abasic linker places a propanediol moiety over the C:G interruptions maintaining internucleotide spacing as a partial isostere of the ribose moiety in nucleosides. The propanediol chain contains a 1-aminobutyl side chain in which the amino group exists almost exclusively in the protonated form under triplex conditions. Even though the use of the linker results in the loss of Hoogsteen bonding and base stacking interactions, this loss may be compensated for by an attraction between the positively charged amino group and the negatively charged phosphodiester backbone of the target site.

Recently, Horne and Dervan reported the relative ability of an uncharged abasic nucleoside to bridge over all four possible

base pairs at a single position within an entirely T–A:T triplex (24). Interestingly, although with pyr*pur:pyr triplex, the abasic nucleoside appeared to show the highest affinity compared to other bases when placed over a C:G base pair. The abasic nucleoside was slightly worse opposite the C:G than C but, significantly better than G. The ability of different substitutions in triplex forming oligonucleotides to be placed over interruptions is almost certainly dependent on the type of triplex (pur*pur:pyr or pyr*pur:pyr) as well as flanking sequences and therefore may be different for each individual sequence and triplex motif.

A major limitation to the application of oligonucleotide directed triplex formation to naturally occurring sequences is the requirement for predominantly purine:pyrimidine tracts. Approaches to circumvent this problem have included the identification of bases which may either recognize (G*T:A) (25) or simply 'fill in' with minimal steric hinderance (C*C:G, G*C:G, T*T:A) (9,10) at base pairs for which there is as yet no defined triplets, the synthesis of bidirectional oligonucleotides which could switch strands where polypurine:polypyrimidine tracts alternate (26–29), and the use of abasic residues to 'skip over' base pairs (24). In addition, attempts to expand the number of potential targets amenable to triplex formation has led to the targeting of non-regulatory regions (30). However, a noncovalent triplex within transcribed regions has been shown to be unable to inhibit transcriptional elongation (30). The novel abasic linker described here may increase the number of amenable targets by allowing for potentially destabilizing interruptions in polypurine:polypyrimidine tracts to be bridged over thereby increasing oligonucleotide affinity or even allow for triplex formation at target sites which normally would not form such structure.

Effect of targeting Sp1 sites and abasic linker substitution on triplex specificity

Triplex formation is generally thought to be sequence specific. However, with oligonucleotides targeted to consensus sequences for very common transcription factors this specificity could be decreased exponentially by shorter sequences being recognized by only a portion of the triplex forming oligonucleotide or by oligonucleotides of limited base heterogeneity such as HR21ap forming triplex with mismatched sequences. The use of abasic moieties in the oligonucleotide over interruptions in the target motif could pose a further problem by being able to 'skip over' some of the mismatches in non-target sequences thereby significantly decreasing specificity.

Because of their unusual pur:pyr rich nature, Sp1 binding sequences are often attractive targets for triplex formation. However, G:C rich regions containing Sp1 sites and polyG:polyC stretches are characteristic of many eukaryotic promoters in general. Since guanines can be placed over C:G base pairs as well as recognize G:C and T:A base pairs, guanine-rich oligonucleotides may be capable of recognizing non-target sequences even with some mismatches within the triplex. Indeed, it is not clear how far individual motifs can deviate from defined triplets and still maintain triplex formation. In addition, shorter oligonucleotides targeted to consensus Sp1 sites may be capable of binding partially to other non-target consensus Sp1 sites.

The target site in *Ha-ras* contains a consensus Sp1 sequence (GGGGCGGGGC) which is separated by one bp from a second sequence which differs from the consensus by only one bp (GGGGCGGGGG). Similarly, the distal Sp1 site in the DHFR

promoter contains a consensus Sp1 sequence which overlaps by two bp a sequence that differs from the consensus at the same position as in *Ha-ras*. Both the distal Sp1 site in *Ha-ras* and the proximal site in DHFR are consensus Sp1 sequences within GC rich regions. The HIV-1 LTR contains three Sp1 sites within a 33bp sequence. HR21ap can be aligned with the distal nontarget Sp1 sequence such that 15 of the 21 bases allow for G–G:C triplets and 5 of the remaining 6 bases placing either C or G opposite C:G base pairs. In addition, portions of HR21ap and HR21Xap would be able to potentially form the same triplets with DHFR as with the *Ha-ras* target leaving a 3' and/or 5' end of the oligonucleotide unbound (Figure 6B). Durland, *et al.* have shown that triplex forming oligonucleotides targeted to the c-myc promoter can bind in this fashion, leaving a 10 base 3' portion unbound (31). It has been suggested that an oligonucleotide targeted to the mouse insulin receptor promoter may also bind to only a portion of its target sequence with 19 of 38 bases remaining unbound (9). We have found that a 32mer oligonucleotide targeted to a K- promoter sequence forms triplex despite a 10 base unbound end (unpublished results). We have shown that HR21ap, even though it is targeted to consensus Sp1 sites in the *Ha-ras* promoter, does not bind to any of the other non-target Sp1 sites examined at concentrations which give a full footprint at the target sequence. Slight protection by HR21Xap was detected only at the distal Sp1 site in DHFR. Still, HR21Xap affords significantly less protection of this site than its target site at ten fold greater concentration than that which afforded full protection of the target sequence. Therefore HR21Xap appears to display significantly greater binding yet, maintains more than ten fold greater specificity for its target sequence over the distal DHFR site even though 15 of the 18 bases in HR21Xap can potentially form G*G:C triplets with the DHFR site. While the use of the abasic linker allows for increased binding to its target sequence its ability to also 'skip over' interruptions in a highly homologous non-target sequence appears to confer some minimal non-specific binding.

Effect of abasic linker substitution on inhibition of Sp1 binding

We have also shown that linker substitution in triplex forming oligonucleotides influences the ability to block transcription factor binding to the major groove of the target DNA. The abasic linker substituted oligonucleotide HR21Xap inhibits Sp1 binding to the *Ha-ras* promoter but, less effectively than HR21ap at a concentration in which HR21Xap displays ten fold or more greater binding to the target sequence. The reason for this decreased inhibition of Sp1 by HR21Xap is not certain though the three linker moieties in this oligonucleotide could be expected to change its physical characteristics including steric bulk and overall conformational flexibility.

Both HR21ap and HR21Xap are specific in their inhibition of Sp1 as well, having minimal effect on Sp1 binding to the DHFR promoter or HIV-1 LTR. The *Ha-ras* targeted oligonucleotides produce greater inhibition of Sp1 binding to the target *Ha-ras* promoter than to the DHFR promoter or HIV-1 LTR at identical concentrations. By inhibiting binding of Sp1 to sites essential for transcription of *Ha-ras*, the triplex forming oligonucleotides targeted to the *Ha-ras* promoter could serve as transcriptional inhibitors *in vivo* and, based on the critical analysis of their specificity, may have minimal effect on other genes containing sequences highly homologous to the target.

ACKNOWLEDGEMENTS

This work was supported by NIH Grants: RO1 CA42664, RO1 CA42337, NCI-CN-25428-33, a Veterans Administration Merit Review Grant and Share Foundation Grant (DM), and a NIH Public Health Service Training Grant (HL07457-11) (CM).

REFERENCES

1. Moser, H.E. and Dervan, P.B. (1987) *Science*, **238**, 645–650.
2. Ferdorova, O.S., Knorre, D.G., Podust, L.M. and Zarytova, V.F. (1988) *FEBS Lett.*, **228**, 273–276.
3. Lyamichev, V.I., Mirkin, S.M., Frank-Kamenetskii, M.D. and Burkholder, G.D. (1988) *Nucleic Acids Res.*, **16**, 2165–2178.
4. Peuth, D., Perrouault, L., Le Doan, T., Chassignol, M., Thuong, N. and Helene, C. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1349–1353.
5. Hanvey, J.C., Shimizu, M. and Wells, R.D. (1989) *Nucleic Acids Res.*, **18**, 157–161.
6. Maher, L.J. III, Wold, B. and Dervan, P.B. (1989) *Science*, **245**, 725–730.
7. Beal, P.A. and Dervan, P.B. (1991) *Science*, **251**, 1360–1363.
8. Postel, E.H., Flint, S.J., Kessler, D.J. and Hogan, M.E. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 8227–8231.
9. Durland, R.H., Kessler, D.J., Gunnell, S., Duvic, M., Pettitt, B.M. and Hogan, M.E. (1991) *Biochemistry*, **30**, 9246–9255.
10. Gee, J.E., Blume, S., Snyder, R.C., Ray, R. and Miller, D.M. (1992) *J. Biol. Chem.*, **267**, 11163–11167.
11. Grigoriev, M., Peuth, D., Robin, P., Hemar, A., Saison-Behmoa, T., Dautry-Varsat, A., Thuong, N.T., Helene, C. and Harel-Bellan, A. (1992) *J. Biol. Chem.*, **267**, 3389–3395.
12. Orsen, F.M., Thomas, D.W., McShan, W.M., Kessler, D.J. and Hogan, M.E. (1991) *Nucleic Acids Res.*, **19**, 3435–3441.
13. McShan, W.M., Rossen, R.D., Laughter, A.H., Trial, J., Kessler, D.J., Zendegui, J.G., Hogan, M.E. and Orson, F.M. (1992) *J. Biol. Chem.*, **267**, 5712–5721.
14. Wells, R.D., Collier, D.A., Hanvey, J.C., Shimizu, M. and Wohlrab, F. (1988) *FASEB J.*, **14**, 2939–2949.
15. Mergny, J.L., Sun, J.S., Rougee, M., Montenay-Garestier, T., Barcelo, F., Chomilier, J. and Helene, C. (1991) *Biochemistry*, **30**, 9791–9798.
16. Mayfield, C., Ebbinghaus, S., Gee, J., Jones, D., Rodu, B., Miller, D. (1993) submitted for publication.
17. Goldsmith, M.E., Beckman, C.A. and Cowan, K.H. (1986) *Mol. Cell. Biol.*, **6**, 878–886.
18. Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499–500.
19. Kadonaga, J.T., Carner, K.R., Masiarz, F.R. and Tjian, R. (1987) *Cell*, **51**, 1079–1090.
20. Ishii, S., Kadonaga, J.T., Tjian, R., Brady, J.N., Merlino, G.T. and Pastan, I. (1986) *Science*, **232**, 1410–1413.
21. Blume, S.W., Gee, J.E., Shrestha, K. and Miller, D.M. (1992) *Nucleic Acids Res.*, **20**, 1777–1784.
22. Gee, J.E., Mayfield, C., Ebbinghaus, Blume, S. and Miller, D.M. (1993) submitted for publication.
23. Beal, P.A. and Dervan, P.B. (1992) *Nucleic Acids Res.*, **20**, 2773–2776.
24. Horne, D.A. and Dervan, P.B. (1991) *Nucleic Acids Res.*, **19**, 4963–4965.
25. Griffen, L.C. and Dervan, P.B. (1989) *Science*, **245**, 967–971.
26. Horne, D.A. and Dervan, P.B. (1990) *J. Am. Chem. Soc.*, **112**, 2435–2437.
27. Ono, A., Chen, C.N. and Kan, L.S. (1991) *Biochemistry*, **30**, 9914–9922.
28. Froehler, B.C., Terhorst, T., Shaw, J-P. and McCurdy, S.N. (1992) *Biochemistry*, **31**, 1603–1609.
29. Jayasena, S.D. and Johnston, B.H. (1992) *Nucleic Acids Res.*, **20**, 5279–5288.
30. Young, S.L., Krawczyk, S.H., Matteucci, M.D. and Toole, J.J. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 10023–10026.
31. Durland, R.H., Kessler, D.J., Duvic, M., Hogan, M. (1990) In Pullman, B. and Jortner, J. (eds.), *Molecular Basis of Specificity in Nucleic Acid-Drug Interactions*. Kluwer Academic Press, Boston, pp 565–578.