

DNA Recognition by Alternate Strand Triple Helix Formation: Affinities of Oligonucleotides for a Site in the Human p53 Gene†

Wendy M. Olivas and L. James Maher, III*

Eppley Institute for Research in Cancer and Allied Diseases and the Department of Biochemistry and Molecular Biology, University of Nebraska Medical Center, 600 South 42nd Street, Omaha, Nebraska 68198-6805

*Received September 27, 1993; Revised Manuscript Received November 19, 1993**

ABSTRACT: Duplex DNA recognition by oligonucleotide-directed triple helix formation is generally limited to homopurine target domains. Various approaches have been suggested for the relief of this constraint. Artificial DNA sequences have previously been used to show that adjacent homopurine domains on opposite DNA strands can be simultaneously recognized by oligonucleotide probes that switch triple helix recognition motifs between domains. Using assays of electrophoretic mobility and chemical protection, we have explored in detail whether such strategies are of benefit in designing high-affinity probes for a natural DNA sequence in the human p53 gene. This target site contains three adjacent, purine-rich domains on opposite DNA strands. Our results show that (i) a modest but statistically significant enhancement in affinity can be achieved for this sequence by designing an oligonucleotide that simultaneously recognizes all three purine domains, (ii) correction of a pyrimidine interruption in one purine domain does not dramatically alter this result, (iii) the relative energetic and structural contributions attributable to recognition of each purine domain can be assessed using probes with combinations of specific and nonspecific nucleotide sequences, and (iv) probe affinity is not correlated with the apparent number of base triplets for certain complexes. These data suggest that unfavorable free energy changes may be associated with alternation between triple helix motifs using existing strategies. In contrast to artificial DNA sequences optimized for this purpose, a substantial affinity enhancement was not observed using alternate strand DNA recognition at this natural target sequence. We therefore conclude that such enhancement is sequence dependent.

Triple helix formation is being explored as an approach to site-specific duplex DNA recognition. The high specificity and stability of some triple-helical structures suggest an approach to gene-specific therapeutic and diagnostic applications (Moser & Dervan, 1987; Le Doan et al., 1987; Cooney et al., 1988; Maher et al., 1991; Hélène, 1991; Durland et al., 1991; Maher, 1992a). *In vitro* experiments have demonstrated that triple helices can direct the site-specific cleavage of DNA (Moser & Dervan, 1987; Perrouault et al., 1990; Strobel & Dervan, 1991), block DNA binding proteins (Maher et al., 1989, 1990), and inhibit transcription initiation (Cooney et al., 1988; Orson et al., 1991; Young et al., 1991; Maher et al., 1992; Maher, 1992b; Duval-Valentine et al., 1992). Oligonucleotide binding to duplex DNA occurs within the major groove via hydrogen bonds to purine bases in the duplex. Two distinct binding patterns, termed the pyrimidine (Py) motif and the purine (Pu) motif, have previously been characterized (Moser & Dervan, 1987; Beal & Dervan, 1991). The pyrimidine motif derives from the formation of T·A·T and C⁺·G·C base triplets by Hoogsteen hydrogen bonding when the third strand is parallel to the purine strand of the duplex. The purine motif derives from the formation of A·A·T (or T·A·T) and G·G·C base triplets by reverse Hoogsteen hydrogen

bonding when the third strand is antiparallel to the purine strand.

One of the principal limitations of oligonucleotide-directed triple helix formation is the necessity for homopurine domains in the target duplex. Various approaches have been suggested for the relief of this constraint. For example, it has been demonstrated that an oligonucleotide can recognize purine domains on alternate duplex DNA strands for sequences of the form 5'-(Pu)_m(Py)_n-3' or 5'-(Py)_n(Pu)_m-3'. In the first experiments of this type, synthetic oligonucleotides bound in the pyrimidine motif and contained unnatural backbone linkages (either 3'-3' or 5'-5') to satisfy polarity requirements when crossing to the alternate strand (Horne & Dervan, 1990; Froehler et al., 1992). It was subsequently shown that an oligonucleotide utilizing consecutive pyrimidine motif and purine motif sequences could simultaneously recognize alternate strand purine domains without unusual phosphodiester linkages (Jayasena & Johnston, 1992a,b, 1993; Beal & Dervan, 1992). This alternate strand recognition is accomplished by switching motifs, and thus relative polarities, at strand crossovers. The third strand may be of the form 5'-Pu motif:Py motif-3' or 5'-Py motif:Pu motif-3'.

Elegant studies by Beal and Dervan (1992) indicated that alternate strand binding could be improved by adding two nucleotides to the third strand probe at the 5'-Pu motif:Py motif-3' junction. Furthermore, these authors found that the two nucleotides of the third strand on either side of a 5'-Py motif:Pu motif-3' junction do not form triplets. The removal of a nucleotide from the third strand at such a 5'-Py motif:Pu motif-3' junction appeared to somewhat relieve this steric constraint.

It has recently been reported that alternate strand binding to a 5'-A₈C₈A₈-3' duplex target sequence can be accomplished

† Supported in part by Grant 5 P30 CA36727-08 from the National Cancer Institute and Grant GM 47814 from the National Institutes of Health. L.J.M. is a recipient of a Junior Faculty Research Award from the American Cancer Society and a Young Investigator's Award from Abbott Laboratories. W.M.O. is supported by a University of Nebraska Medical Center Emley Fellowship.

* To whom correspondence should be addressed: 402-559-8288; 402-559-4651 (FAX); jmaher@unmc.edu (bitnet).

• Abstract published in *Advance ACS Abstracts*, January 1, 1994.

using 5'-T₈G₈-3' oligonucleotides, but not using 5'-G₈T₈-3' oligonucleotides (Jayasena & Johnston, 1993). These authors concluded that binding to 5'-(Pu)_m(Py)_n-3' duplex target sequences is favored by the alternate strand method, whereas binding to 5'-(Py)_n(Pu)_m-3' duplex target sequences may be more sequence dependent.

We wished to determine whether alternate strand triple helix formation enhances the affinity of oligonucleotide probes for certain naturally occurring DNA sequences of diagnostic significance. We therefore identified a site in the human p53 tumor suppressor gene for further study. This sequence, 5'-AGAG₂A₂GAGA₂TCTC₂GCA₂GA₃G₄AG-3', is of the general form 5'-(Pu)₁₂(Py)₇(Pu)₁₂-3', where the underlined guanosine residue represents a single defect in the homopyrimidine central domain. This sequence is interrupted in a clinically relevant microdeletion associated with a case of hepatocellular carcinoma (Bressac et al., 1991). Oligonucleotide probes that bind individual homopurine domains within this sequence have been shown to allow discrimination of mutant and wild-type alleles at this site (Olivas & Maher, 1994).

Electrophoretic gel mobility shift titrations were performed to analyze the binding of 14 oligonucleotides designed to recognize this target site with zero, one, or two motif junctions using the strategy suggested by Beal and Dervan (1992). The gel mobility analysis offers significant advantages over affinity cleaving and footprinting methods previously applied to this problem. Mobility shift experiments are sensitive, offering a wide dynamic range of binding measurements while also providing information about the structures of the resulting complexes. We have utilized both types of information, as well as dimethyl sulfate (DMS)¹ footprinting data, to characterize oligonucleotide recognition. Several conclusions are presented. First, for the p53 target sequence, probes designed to simultaneously recognize all three purine domains on alternate strands afforded a modest, but statistically significant, increase in binding affinity relative to probes that bound individual purine domains. Second, there is evidence for unfavorable interactions associated with junctions between third strand triple helix motifs at this site. Third, we found that binding affinity was not always correlated with the apparent number of base triplets in a complex. Together, these results suggest that the design of high-affinity oligonucleotide probes for alternate strand recognition of natural DNA sequences may be more complex and difficult to achieve than for artificial sequences that were optimized for triple helix formation.

MATERIALS AND METHODS

Oligonucleotides. Synthetic oligonucleotides were purified and quantitated as previously described (Maher et al., 1990). Oligonucleotides for target duplexes were annealed as follows: top strand (500 pmol) was mixed with complementary bottom strand (500 pmol), 5 M NaCl (2 μL), and H₂O in a total volume of 42 μL; annealing reaction mixtures were incubated at 75 °C for 12 min and then gradually cooled to 25 °C. For radiolabeling, oligonucleotide duplex (60 pmol) was added to 2.5 μL of 10× polymerase buffer (100 mM Tris-HCl, pH 7.9, 100 mM MgCl₂, 500 mM NaCl, and 10 mM dithiothreitol), 4 μL of [α-³²P]dATP (13 pmol; 3000

Ci/mmol), 1 μL of dGTP, 1 μL of dTTP, 1 μL of dCTP (each at 2 mM), 1 μL of DNA polymerase I Klenow fragment (5 units), and H₂O in a total volume of 25 μL. The reaction was allowed to proceed for 45 min at 22 °C. The resulting labeled duplex oligonucleotides were precipitated with ethanol in the presence of ammonium acetate and resuspended in H₂O.

Electrophoretic Mobility Shift Assays. Binding reaction mixtures contained labeled duplex (50 000 cpm; ca. 0.1 pmol), 1 μL of 10× binding buffer (1 M NaOAc, pH 5.0, 50 mM NaCl, and 100 mM MgCl₂), 1 μL of tRNA (1 mg/mL), 1 μL of probe oligonucleotide (to yield the indicated final concentration), and H₂O in a final volume of 10 μL. Reaction mixtures were incubated at 22 °C for at least 5 h. Reaction mixtures were supplemented with 1 μL of an 80% glycerol solution containing bromophenol blue and were loaded on a 20% native polyacrylamide gel (19:1 acrylamide:bis(acrylamide)) prepared in 0.1 M NaOAc, pH 5.0, and 1 mM MgCl₂. Electrophoresis was performed in this buffer (with recirculation) at 4 °C overnight (9 V/cm). The resulting gel was imaged and quantified by storage phosphor technology using a Molecular Dynamics PhosphorImager.

Analysis of Gel Mobility Shift Titrations. The apparent fraction of target duplex bound by probe oligonucleotide, Q , was calculated for each concentration of probe oligonucleotide using the definition

$$Q = S_{\text{triplex}} / (S_{\text{triplex}} + S_{\text{duplex}}) \quad (1)$$

where S_{triplex} and S_{duplex} represent the storage phosphor signal for triple-helical and duplex complexes, respectively. When multiple triple-helical complexes were observed, the affinity of the dominant complex is reported. If multiple complexes of similar intensity were observed, these were summed in the calculation of S_{triplex} . To correct for background effects, the fraction of duplex in triple-helical form was then converted to a scaled score, F :

$$F = (Q - Q_{\text{low}}) / (Q_{\text{high}} - Q_{\text{low}}) \quad (2)$$

where Q_{low} and Q_{high} correspond to Q values obtained in the presence of 0 or 2.5 μM oligonucleotide probe, respectively. This treatment assumes that the fraction of triple helix in the absence of oligonucleotide probe must be 0 and that the fraction of triple helix approaches 1.0 in binding reaction mixtures containing 2.5 μM specific oligonucleotide probe. These assumptions are generally valid for the experiments described here. Exceptions are noted in the text. Values of the apparent association constant, K_a , were obtained by least squares fitting of the data to the binding isotherm:

$$F = [O]^n K_a^n / (1 + [O]^n K_a^n) \quad (3)$$

where $[O]$ is the oligonucleotide probe concentration, and n is the Hill coefficient (Cantor & Schimmel, 1980).

Measures of statistical significance were performed using the nonparametric Wilcoxon rank-sum test (Bhattacharyya & Johnson, 1977). Statistical significance was determined at the $\alpha = 0.05$ level of probability.

Mobility values (μ) of triple-helical complexes were calculated using the equation

$$\mu = D_c / D_d \quad (4)$$

where D_c and D_d are distances migrated by triple-helical complexes and free duplex, respectively.

Dimethyl Sulfate Footprinting. Top and bottom strands of the wild-type target duplex (500 pmol) were phosphorylated by addition of 10× kinase buffer (2 μL), 0.6 μL ATP (100 mM), 1 μL T4 polynucleotide kinase (10 units), and H₂O in

¹ Abbreviations: bp, base pair(s); cpm, radioactive counts per minute; $\Delta\Delta G$, difference in Gibbs free energy change; DMS, dimethyl sulfate; nt, nucleotide; SEM, standard error of the mean; Tris-HCl, 2-amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride.

a total volume of 20 μ L. Reaction mixtures were incubated at 37 $^{\circ}$ C for 45 min. Phosphorylated complementary strands were combined with 5 M NaCl (2 μ L) and annealed as previously described. The resulting duplex was ligated into plasmid pG₅E4T that had been cleaved by *Hind*III and *Pst*I (Maher et al., 1992b). A clone bearing the desired insertion was confirmed by sequencing. A 2704-bp *Bam*HI-*Kpn*I restriction fragment was then prepared and uniquely end-labeled by either polynucleotide kinase (bottom strand) or the Klenow fragment of DNA polymerase I (top strand). Labeled fragment (50 000 cpm; ca. 0.4 pmol) was incubated with 1 μ L of 10 \times binding buffer (1 M NaOAc, pH 5.0, 50 mM NaCl, and 100 mM MgCl₂), 1 μ L of sheared calf thymus DNA (0.65 mg/mL), 1 μ L of oligonucleotide probe (10 μ M), and H₂O in a total volume of 10 μ L. Binding reactions were incubated overnight at 22 $^{\circ}$ C. Dimethyl sulfate [1 μ L of a 4% (v/v) aqueous solution] was added to each reaction mixture and allowed to incubate for 8 min at 4 $^{\circ}$ C. Reactions were stopped with 5 μ L of stop mix [1.5 M NaOAc, 7% (v/v) β -mercaptoethanol, and 100 μ g/mL tRNA]. Following ethanol precipitation, 100 μ L of 1% (v/v) piperidine was added, and the samples were incubated for 30 min at 90 $^{\circ}$ C. Piperidine was removed by repeated lyophilization. The DNA was then resuspended in 5 μ L of formamide dye mix, electrophoresed on a 20% polyacrylamide sequencing gel (19:1 acrylamide:bis(acrylamide)) containing 7.5 M urea in 0.5 \times TBE buffer, and autoradiographed.

RESULTS AND DISCUSSION

Experimental Design. Previous studies of alternate strand triple helix formation have employed artificial duplex target sequences optimized for triple helix formation within each binding domain. Typically, this involves selecting A-rich sequences for pyrimidine motif recognition and G-rich sequences for purine motif recognition. We wished to extend these model studies by investigating a naturally occurring target sequence, such as that identified in the p53 tumor suppressor gene. Exon 8 of the wild-type p53 gene contains three adjacent purine-rich domains on alternate duplex strands (denoted as areas I, II, and III in Figure 1A). Area I begins within codon 284 of the p53 cDNA and is composed of 12 purine nucleotides. Area II, on the alternate DNA strand, is composed of six purine nucleotides and a single interrupting pyrimidine near the 5' end. Area III is composed of 12 purine nucleotides.

To facilitate electrophoretic mobility analyses of probe binding, we designed synthetic duplexes containing 38-bp sequences corresponding to wild-type and modified forms of the p53 gene. To assess the importance of the interrupting pyrimidine in area II, the modified form changes this C-G base pair to an A-T base pair (see Figure 1A). To study the association of oligonucleotide probes to one or more purine-rich domain(s) in these target duplexes, oligonucleotides were synthesized to specifically (or nonspecifically) interact with one, two, or all three homopurine domains (Figure 1C). The disposition of probes that individually recognize area I, II, or III is depicted using a cylindrical representation of the DNA double helix in Figure 1B (top). Adjacent termini of probes specific to area I (pyrimidine motif) and area II (purine motif) overlap in the major groove. Probes intended to simultaneously recognize both areas were designed by removing a thymidine residue from the 5'-Pu motif:Pu motif-3' junction, using the strategy devised by Beal and Dervan (1992), thus allowing the probe to more smoothly cross the major groove from area I to area II (Figure 1B, bottom). In contrast, adjacent termini

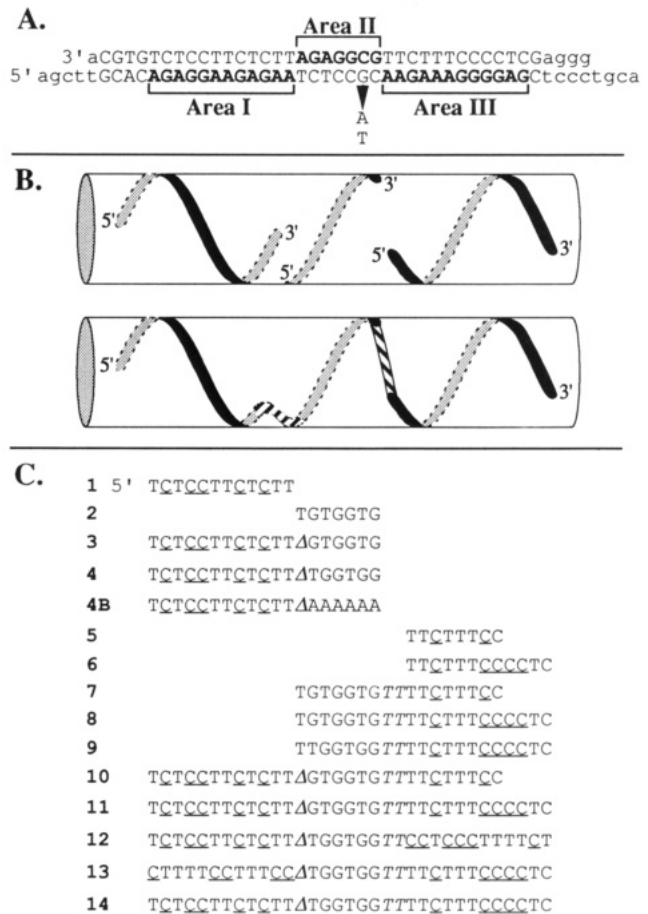


FIGURE 1: (A) Target sequence. The wild-type duplex target sequence is shown, with homopurine areas indicated in bold. Uppercase residues denote sequences derived from the p53 gene, while lowercase residues were added for cloning purposes. The arrowhead indicates a single base pair change made in producing a modified target duplex that lacks the pyrimidine interruption in area II. (B) Probe design. The top illustration depicts the relative position of third strand probes bound to alternate homopurine areas I, II, and III on a cylindrical model of DNA. The bottom illustration emphasizes modifications in a single probe bound to areas I, II, and III. Striped regions represent motif junctions. (C) Nucleotide sequences of probes used for oligonucleotide-directed DNA triple helix formation. Underlined C residues represent 5-methylcytosine. 3'-Terminal cytosines were unmethylated due to synthetic constraints. The deletion of a residue at the area I/area II junction is indicated by a triangle (Δ), while linker residues added at the area II/area III junction are indicated by italics (TT).

of probes to area II (purine motif) and area III (pyrimidine motif) do not overlap, but are somewhat separated. A single probe for the simultaneous recognition of both areas was designed by adding two thymidine residues at the 5'-Pu motif:Pu motif-3' junction, again by using the published strategy (Beal & Dervan, 1992), thus creating a bridge to cross from area II to area III (Figure 1B, bottom). Electrophoretic mobility shift titrations at pH 5.0 in the presence of magnesium were analyzed to measure probe affinities and triple-helical complex mobilities, while DMS footprinting assays were used to confirm details of probe binding.

Triple Helix Recognition of Areas I and II. Probes designed to individually recognize area I or area II of the target duplexes were initially studied. The 12-nt probe 1, designed to recognize area I, bound the wild-type and modified duplexes with association constants ($K_a \pm$ SEM) of $(1.89 \pm 0.084) \times 10^7$ and $(1.95 \pm 0.147) \times 10^7$ M⁻¹, respectively (Figure 2). As expected, the difference between these K_a values is not statistically significant, since the sequence of area I is identical

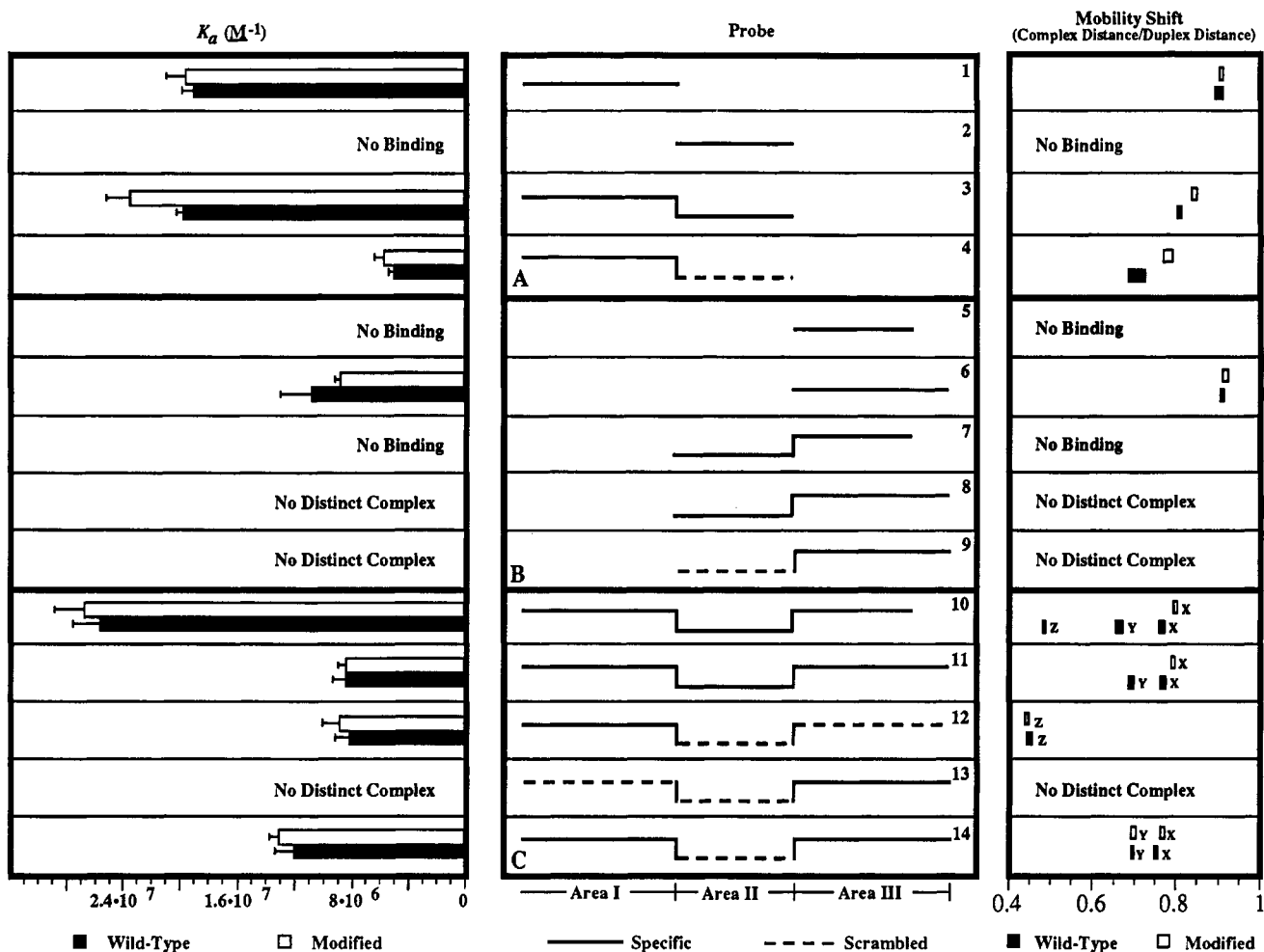


FIGURE 2: Analysis of electrophoretic mobility shift assays: summary data set. The center panel depicts oligonucleotide probes used in mobility shift assays. Solid lines represent sequences specific to target areas, while dashed lines represent sequences intentionally scrambled to provide controls. Probes in box A target area I and/or II. Probes in box B target area II and/or III. Probes in box C simultaneously target areas I, II, and III. The left panel depicts measured association constants for each probe to the wild-type (black bars) or modified (white bars) target duplexes. Each bar is an average (\pm SEM) from at least three experiments. The right panel portrays the mobilities of triple-helical complexes formed by each oligonucleotide probe. Mobility values were calculated by using eq 4 as described in Materials and Methods. A value of 1.0 represents the mobility of the free duplex. The position and size of each box represent an average triple-helical complex mobility (\pm SEM) derived from at least three experiments. Filled boxes denote complexes with the wild-type duplex, while open boxes denote complexes with the modified duplex. Probes that did not bind even at $2.5 \mu M$ concentration are labeled "No Binding" in the left and right panels. Probes that titrated at least a portion of free duplex but produced multiple, indistinct triple-helical complexes are labeled "No Distinct Complex" in the left and right panels.

in both duplexes. Essentially equivalent mobility values of 0.90 ± 0.008 and 0.91 ± 0.004 were obtained for complexes involving probe 1 bound to wild-type and modified duplexes, respectively (Figure 2). These values reflect compact complexes whose electrophoretic migration is only slightly retarded relative to that of the free duplexes. The results of DMS footprinting experiments confirmed specific binding of probe 1 to area I (Figure 3, lane 3 of panels A and B).

In contrast to probe 1, the 7-nt probe 2 (designed to recognize area II) did not bind to either the wild-type or the modified duplex ($K_a \ll 4 \times 10^5 M^{-1}$). Previous studies have shown that strand orientations can become ambiguous for cases of oligonucleotides containing guanosine and thymidine residues when the latter predominate (Sun et al., 1991). However, probe 2 contains an excess of guanosine residues and is therefore unlikely to suffer from this limitation. Apparently, the target sequence of area II is too short and/or it lacks a sufficient number of contiguous guanosine base pairs to promote oligonucleotide binding in the purine motif under these conditions. This result is unaffected by correction of the interrupting pyrimidine of the wild-type duplex to a purine in the modified duplex. It should be noted that binding

reactions were supplemented with 7% (v/v) glycerol prior to electrophoresis. Although there is no evidence that this addition adversely affects triple helix formation, the role of this cosolvent in modifying the stabilities of particular triple-helical complexes has not been investigated.

In spite of the observation that probe 2 did not associate with area II, we wished to test whether alternate strand triple helix formation using probe 3 might permit simultaneous binding to both areas I and II, yielding a greater affinity than with probe 1 alone. The 18-nt probe 3 bound the wild-type and modified duplexes with K_a values of $(1.97 \pm 0.044) \times 10^7$ and $(2.34 \pm 0.165) \times 10^7 M^{-1}$, respectively (Figure 2). The modest increase in the K_a value for probe 3 binding to the wild-type duplex is not statistically significant relative to that for probe 1. However, the observed K_a value for probe 3 binding to the modified duplex represents a statistically significant improvement in affinity relative to that for probe 1. This result suggests that simultaneous binding to areas I and II can enhance affinity if area II lacks triplet mismatches. Mobility values of 0.81 ± 0.004 and 0.84 ± 0.006 were obtained with probe 3 binding to wild-type and modified duplexes, respectively (Figure 2). The gel mobilities of these complexes

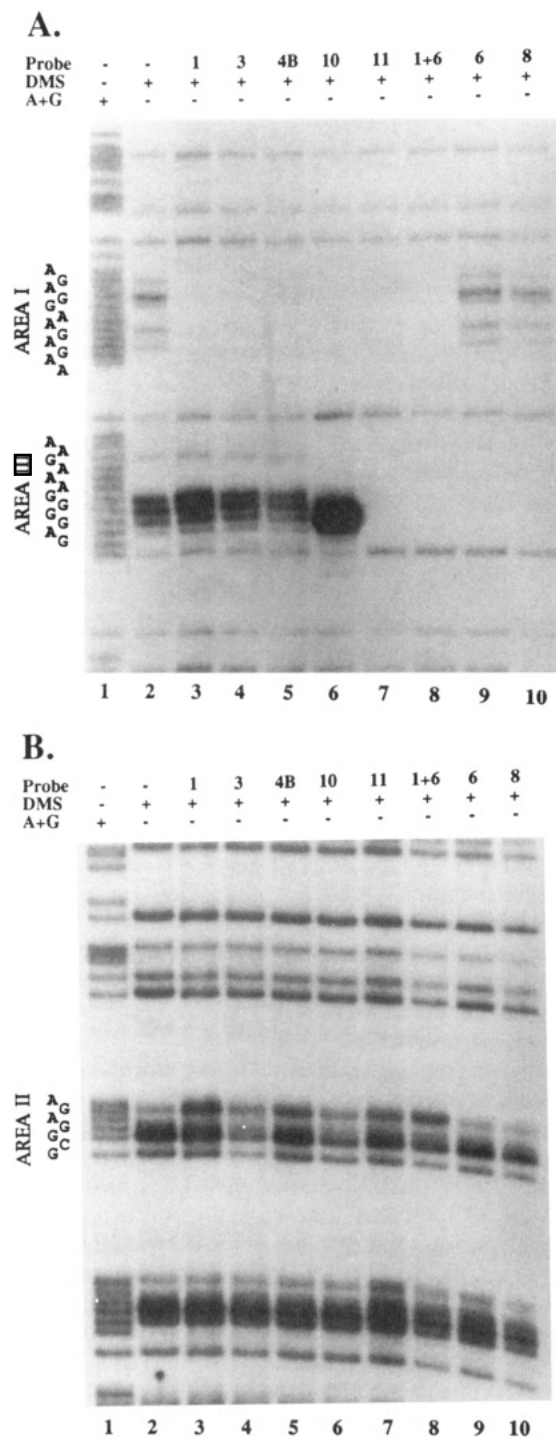


FIGURE 3: Dimethyl sulfate footprinting assays. The wild-type target duplex was subcloned into a recombinant plasmid that was then linearized, labeled, and subjected to DMS modification in the absence or presence of oligonucleotide probes as described in Materials and Methods. Lane 1 of each panel is an A+G ladder (Maxam & Gilbert, 1980). (A) Labeling of top strand. The sequences of areas I and III are indicated. (B) Labeling of bottom strand. The sequence of area II is indicated.

were distinctly lower than those involving probe 1, presumably due to the increased mass of probe 3. Furthermore, the complex of probe 3 with the modified duplex migrated faster than the complex involving the wild-type duplex. This observation emphasizes the sensitive shape-selectivity of the electrophoretic assay. Moreover, the results suggest that the complex involving the modified duplex is more compact, consistent with the observation that the correction of area II enhances binding. The results of DMS footprinting exper-

iments for the wild-type duplex complexed with probe 3 showed strong protection of area I and partial protection of area II (Figure 3, lane 4 of panels A and B).

To further explore whether probe 3 is simultaneously binding both areas I and II, we determined the K_a for the 18-nt probe 4, which has the same base composition as probe 3, but the sequence for recognition of area II has been scrambled. The wild-type and modified duplexes bound probe 4 with K_a values of $(4.92 \pm 0.432) \times 10^6$ and $(5.65 \pm 0.623) \times 10^6$ M⁻¹, respectively (Figure 2). It is notable that these values are ~3–4-fold smaller than the K_a values obtained using probe 3, indicating that significantly weaker complexes were formed. This decrease in affinity confirms that probe 3 is capable of forming productive contacts in both area I and area II. Mobility values of 0.71 ± 0.020 and 0.78 ± 0.008 for probe 4 binding with wild-type and modified duplexes, respectively, were observed. Thus, although they have the same mass as complexes with probe 3, complexes with probe 4 are distinctly less compact and therefore demonstrated more retarded migration in the gel (Figure 2). In addition, DMS footprinting experiments performed with the similar probe 4B showed no evidence of binding to area II (Figure 3, lane 5 of panels A and B). The origin of the slight difference in mobilities between probe 4 bound to wild-type vs modified duplexes is unknown, since the correction of the pyrimidine interruption in area II is not predicted to affect binding of the scrambled sequence. Taken together, the affinity and mobility data for probes 1, 3, and 4 demonstrate that probe 3 is simultaneously binding to areas I and II by alternate strand recognition; however, the enhancement in binding affinity relative to that of probe 1 is modest and only significant for the modified duplex. In particular, although the enhancement is statistically significant, the estimated difference in binding free energies ($\Delta\Delta G = -0.11$ kcal/mol) is small considering the potential addition of six base triplets in the complex with probe 3 relative to probe 1. We interpret this result as indicative of an unfavorable free energy contribution associated with crossing the DNA major groove to bind adjacent purine domains in this target duplex.

Triple Helix Recognition of Areas II and III. The 8-nt probe 5, designed to recognize eight purines in the 5' region of area III, showed no evidence of binding to either wild-type or modified duplex ($K_a \ll 4 \times 10^5$ M⁻¹). However, the 12-nt probe 6, designed to bind the entire area III, yielded K_a values of $(1.07 \pm 0.224) \times 10^7$ and $(8.68 \pm 0.377) \times 10^6$ M⁻¹ with wild-type and modified duplexes, respectively (Figure 2). The difference in these K_a values is not statistically significant, as expected, since the sequence of area III is the same in both duplexes. The increased affinity of probe 6 relative to that of probe 5 presumably derives from additional C⁺·G·C protonated base triplets, stabilized under the acidic (pH 5) assay conditions. The K_a values for probe 6 are ~2-fold smaller than the K_a values obtained with probe 1 binding to area I, even though both probes are 12 nucleotides in length. This statistically significant difference emphasizes the sequence dependence of binding affinities in triple helix formation. Mobility values for probe 6 complexes were 0.91 ± 0.006 and 0.92 ± 0.008 with wild-type and modified duplexes, respectively (Figure 2). These values are similar to those found using probe 1, consistent with the similarity in masses for the two probes. The results of DMS footprinting experiments confirmed that probe 6 specifically protects all of area III except for the 3'-terminal guanosine residue (Figure 3, lane 9 in panels A and B). The lack of protection at this

residue may reflect poor triplet formation by the unmethylated 3'-terminal cytosine in probe 6.

We sought to determine whether alternate strand triple helix formation using probe 7 might permit simultaneous binding to area II and the eight purines at the 5' terminus of area III. As described above, neither probe 2 nor probe 5 bound when tested individually. We found that the 17-nt probe 7 was also unable to bind either duplex ($K_a \ll 4 \times 10^5 \text{ M}^{-1}$). In contrast, probe 8, designed to recognize area II and all of area III, did bind to both wild-type and modified duplexes. However, K_a and mobility values cannot be estimated with certainty for probe 8, because probe binding titrated the amount of free duplex without producing distinct complexes. This result may suggest the formation of multiple complexes that are unstable on the electrophoretic time scale. Binding of probe 8 did not titrate all free duplex, even at a probe concentration of 2.5 μM . By this criterion, the affinity of probe 8 is less than that of probe 6 binding area III alone. The results of DMS footprinting experiments confirmed specific binding of probe 8 to area III (except 3'-terminal guanosine), but not to area II of the wild-type duplex (Figure 3, lane 10 of panels A and B). Thus, for these particular purine domains, attempting alternate strand triple helix formation at a 5'-Pu motif:Py motif-3' junction actually hinders binding.

We wished to test whether recognition of area II played any role in triple helix formation with probe 8. We therefore designed probe 9, which is identical to probe 8 except that the sequence for recognition of area II has been scrambled. As observed for probe 8, the binding of probe 9 titrated the target duplex without producing distinct complexes, thus preventing determination of the K_a . However, by examination of the extent of duplex titration as a function of probe 9 concentration, it was evident that binding by probe 9 was weaker than that by probe 8. The probable ranking of K_a values for probes 5–9 is therefore $6 > 8 > 9 > (5,7)$. One interpretation of this ordering is that although both probe 8 and probe 9 bind more weakly than probe 6, the binding of probe 8 is stronger than the binding of probe 9 and must therefore involve recognition of both area II and area III by alternate strand triple helix formation. If correct, this curious result implies that increased specific interactions do not necessarily correlate with increased binding affinity (e.g., probe 8 vs probe 6). Alternatively, the binding of probe 9 may be inhibited because of unknown competing intramolecular structures. In any case, because probe 7 did not bind more strongly than probes 2 or 5, and because probe 8 bound less strongly than probe 6, we conclude that crossing the DNA major groove once to simultaneously bind alternate 5'-(Py)_m(Pu)_n-3' domains in the p53 sequence does not confer increased affinity.

Triple Helix Recognition of Areas I, II, and III. After characterizing alternate strand triple helix formation for recognition of either 5'-(Pu)_m(Py)_n-3' or 5'-(Py)_n(Pu)_m-3' domains of the p53 sequence, we wished to determine whether a probe that simultaneously recognizes all three purine domains would bind with an affinity greater than that of a probe specific for only one or two domains. Probe 10, designed to simultaneously recognize area I, area II, and the eight 5'-proximal purine bases of area III, yielded K_a values of $(2.55 \pm 0.191) \times 10^7$ and $(2.66 \pm 0.207) \times 10^7 \text{ M}^{-1}$ with wild-type and modified duplexes, respectively (Figure 2). A binding curve that summarizes the data for probe 10 is shown in Figure 4. Although modest, the increase in K_a values for probe 10 versus probe 1 is statistically significant for both wild-type (1.3-fold increase) and modified (1.4-fold increase) duplexes. When compared to that of probe 3 (designed to simultaneously bind

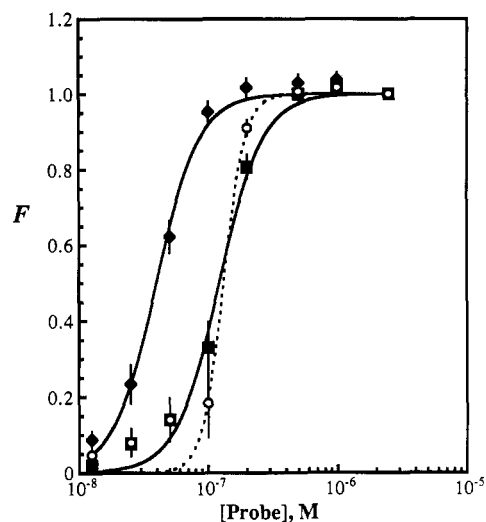


FIGURE 4: Examples of binding isotherms. Binding curves for oligonucleotide probes 10 (\blacklozenge), 11 (\blacksquare), and 12 (\circ) with wild-type target duplex were fitted using data from electrophoretic mobility shift titrations. The fraction of duplex in triple-helical form (F) was calculated from eqs 1 and 2 and fitted to eq 3 as described in Materials and Methods.

areas I and II), the increased K_a value for probe 10 is statistically significant for the wild-type duplex, but not for the modified duplex. Although the affinity of probe 10 is only slightly higher than the affinity of probe 1 or 3, there is a ~ 2 –3-fold increase in the affinity of probe 10 relative to that of probe 6. In fact, probes 5 and 7, which recognize the same portion of area III, did not bind. Thus, the recognition of area I is important for maximum binding affinity. It should be noted that the enhancement of affinity demonstrated by probe 10 could not be duplicated by mixing individual probes specific for combinations of target areas (e.g., 1 + 2, 1 + 6, 2 + 6; data not shown).

We characterized complexes involving probe 10 by examining their relative electrophoretic mobilities. As summarized in Figure 2, binding of probe 10 to the wild-type duplex produced three complexes, designated X, Y, and Z (Figure 5A, lanes 10–18; mobility values of 0.77 ± 0.007 , 0.66 ± 0.011 , and 0.49 ± 0.004 , respectively). In contrast, only a single complex was apparent when probe 10 bound the modified duplex (Figure 5A, lanes 28–36; mobility value of 0.80 ± 0.003). Interestingly, each complex appeared as a distinct band, suggesting that the complexes do not decay or interconvert during electrophoresis. This apparent failure to interconvert during electrophoresis, and the approximately equivalent affinities represented by the multiple complexes, were unexpected observations that remain difficult to interpret. We initially hypothesized that complex X in Figure 5A (the most populated complex for the wild-type duplex) corresponds to probe 10 simultaneously binding areas I, II, and III. Complex Y may correspond to simultaneous binding of only two areas (e.g., I and III), and the slow complex, Z, may correspond to binding of only area I. These assignments are further explored below and would be consistent with the observed absence of the Y and Z complexes in the case of the modified duplex, as the correction of area II may increase the stability of complex X.

For the wild-type duplex, DMS footprinting experiments should report the sum of complexes X, Y, and Z. Lane 6 in panels A and B of Figure 3 confirms that probe 10 fully protects area I, partially protects area II, and fully protects the seven 5'-proximal purines of area III. As discussed above, lack of protection of the eighth purine of area III may be related to

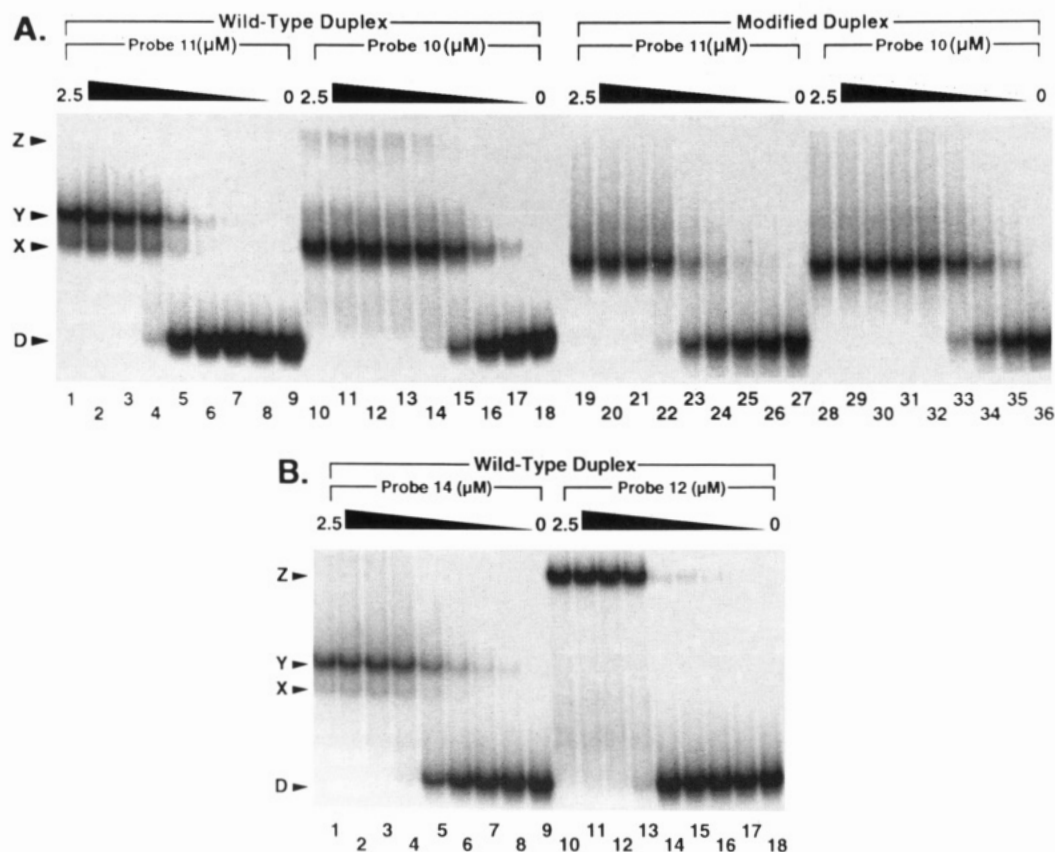


FIGURE 5: Examples of electrophoretic mobility shift assay data. (A) Decreasing micromolar concentrations (2.5, 1.0, 0.5, 0.2, 0.1, 0.05, 0.025, 0.0125, 0) of oligonucleotide probes **11** (lanes 1–9 and 19–27) and **10** (lanes 10–18 and 28–36) were added to labeled wild-type (lanes 1–18) and modified (lanes 19–36) duplex targets. Arrowheads represent the locations of free duplex (D) and triple-helical complexes (X, Y, and Z). (B) Decreasing micromolar concentrations (2.5, 1.0, 0.5, 0.2, 0.1, 0.05, 0.025, 0.0125, 0) of oligonucleotide probes **14** (lanes 1–9) and **12** (lanes 10–18) were added to wild-type target duplex. The equivalence of total radioisotopic signal in each lane was confirmed by storage phosphor imaging.

the lack of methylation of the 3'-terminal cytosine of the probe. The dramatic hypermethylation of unbound guanosine residues in area III distal to the triple helix is particularly notable (Figure 3A, lane 6). This DMS hyperreactivity may reflect structural changes in the duplex at the duplex-triplex junction.

Since recognition of the four 3'-proximal purines in area III was critical for binding (compare probes **6** and **8** with probes **5** and **7**), it was important to determine if recognition of these purines would increase the affinity of probes designed to simultaneously bind all three purine domains. Probe **11**, designed to bind areas I, II, and all of III, yielded K_a values of $(8.36 \pm 0.916) \times 10^6$ and $(8.32 \pm 0.577) \times 10^6 \text{ M}^{-1}$ with the wild-type and the modified duplexes, respectively (Figure 2). A binding curve summarizing the data for probe **11** is shown in Figure 4. Surprisingly, the affinity of probe **11** was reduced relative to that of probe **10**. The measured K_a values for probe **11** are, in fact, statistically indistinguishable from those of probe **6**. The destabilizing effect obtained by adding four 3'-proximal pyrimidines in this context is not easily explained by current models. Perhaps these added nucleotides alter competitive intramolecular structures adopted by probe **11**. As summarized in Figure 2, two complexes involving probe **11** were observed with the wild-type duplex (Figure 5A, lanes 1–9; mobility values of 0.77 ± 0.005 and 0.69 ± 0.005), while only one complex was seen with the modified duplex (Figure 5A, lanes 19–27; mobility value of 0.79 ± 0.002). Again, we hypothesized for the wild-type duplex that the faster migrating complex, X in Figure 5A, represents probe **11** simultaneously binding all three areas, while the more populated complex with intermediate mobility, Y, represents binding to only areas I and III. As we observed for probe **10**,

complex X appears to be the most stable when probe **11** binds to the fully homopurine area II in the modified duplex. A complex at position Z was not observed for probe **11** with the wild-type duplex. The results of DMS footprinting experiments with the wild-type duplex demonstrate specific binding of probe **11** to area I and area III (except the 3' guanosine), but not to area II (Figure 3, lane 7 of panels A and B).

Assignments of Complexes X, Y, and Z. To experimentally verify our hypothetical assignments of complexes X, Y, and Z in Figure 5A, we designed a series of probes containing scrambled nucleotide sequences targeted to one or more domains. To confirm that complex Z of probe **10** reflects binding to area I but not to area II or III, we designed probe **12** to specifically recognize area I but to contain scrambled nucleotides in sequences recognizing areas II and III. Probe **12** yielded K_a values of $(8.03 \pm 1.04) \times 10^6$ and $(8.74 \pm 1.20) \times 10^6 \text{ M}^{-1}$ with wild-type and modified duplexes, respectively (Figure 2). A binding curve summarizing the data for probe **12** is also shown in Figure 4. These K_a values are ~2-fold smaller than those of probe **1**, perhaps reflecting an unfavorable free energy contribution associated with the unbound "tail". The mobility values of the single complexes formed by probe **12** (0.45 ± 0.005 and 0.44 ± 0.004 for wild-type and modified duplexes, respectively) are very similar to those of complex Z formed by probe **10** (Figures 2 and 5B). Because probes **10** and **12** have the same molecular weight, this mobility similarity supports our hypothesis that complex Z represents these probes bound only to area I. It remains unclear why complex Z involving probe **10** does not equilibrate with the fully bound form (presumably complex X) during electrophoresis. We would expect such equilibration to produce a

smear rather than the distinct complexes that we observed.

A surprising characteristic of probe **12** was its extremely sharp (cooperative) binding transition (Figure 4). The least squares fit to the model binding isotherm reproducibly yielded a Hill coefficient of ~ 6 , whereas binding curves for other probes in this study yielded Hill coefficient estimates of ~ 2 . We interpret the latter values to be indicative of the relatively slow association kinetics for probes at low concentrations (Maher et al., 1990). However, the pronounced cooperativity exhibited by probe **12** suggests the involvement of additional structural equilibria. For example, the presence of competing self-complexes has been postulated to increase binding specificity by other authors studying similar oligonucleotides at acidic pH (Roberts & Crothers, 1991).

To eliminate the possibility that any of the complexes X, Y, or Z derive from binding only in area III, we designed probe **13** to be specific to area III but to contain scrambled nucleotides in sequences recognizing areas I and II. We were unable to quantitate the K_a for this probe because it titrated only a small fraction of free duplex at 2.5 μM probe concentration and produced indistinct complexes. This result suggests that binding to area III alone is weak for probes **10** and **13** and that none of the complexes X, Y, or Z represent this interaction.

Having obtained evidence to support our assignment of complex Z, we wished to explore the possibility that complex Y of probes **10** and **11** actually derives from simultaneous binding to areas I and III, but not to area II. Probe **14** was designed to recognize areas I and III, but the nucleotide sequence for recognition of area II was scrambled. The K_a values for probe **14** were $(1.19 \pm 0.132) \times 10^7$ and $(1.30 \pm 0.058) \times 10^7 \text{ M}^{-1}$ with wild-type and modified duplexes, respectively (Figure 2). The larger K_a value for probe **14** relative to that for probe **11** is curious, but is only statistically significant for the modified duplex. As expected for a scrambled area II binding domain, the correction of area II in the modified duplex did not eliminate complex Y for probe **14** as it does for probe **11**. With respect to assigning a structure for complex Y, it was surprising that both complexes X and Y were again observed when probe **14** bound to wild-type and modified duplexes, in contrast to our prediction that only the Y complex would appear (Figure 5B, lanes 1–9). Because probe **14** retains the potential for only two of six triplets with area II, the difference between mobilities of complexes X and Y must not be a simple reflection of binding in area II. As an alternative, we consider it possible that complex X may represent a species with probe binding to area III, tracking along the major groove of area II (with or without forming specific triplets) and then binding to area I. Perhaps complex Y reflects binding in area III, with the probe crossing over the minor groove (avoiding area II) and then binding to area I.

Summary. In an attempt to design oligonucleotide probes with increased affinity for sequences in the human p53 tumor suppressor gene, we have tested alternate strand triple helix formation using strategies proposed by other authors (Beal & Dervan, 1992). The experimental target system reported here differs from the artificial sequences previously employed, in that homopurine domains were not optimized for recognition in the two triple helix motifs and, in fact, include a single pyrimidine interruption in one of the three purine domains. Perhaps because of these characteristics and the rather short length of area II, alternate strand triple helix formation offered only a modest (less than 2-fold) improvement in affinity relative to a single, shorter probe. Improvements in affinity due to one or two strand crossovers were statistically significant in

some cases, particularly for duplexes where the pyrimidine interruption in area II was corrected to a purine. For the wild-type duplex, the ranking of probe affinities was $10 > (3,1) > (14,6,11,12) > 4 > 8 > 9 > (5,7)$. For the modified duplex, the affinity ranking was $(10,3) > 1 > 14 > (12,6,11) > 4 > 8 > 9 > (5,7)$. The sensitivity of the electrophoretic mobility shift assay allowed us to confirm simultaneous occupation of multiple target domains, and to detect the formation of multiple, nonequilibrating complexes for some of the probes tested. Detection of such complexes cannot be easily achieved using affinity cleaving and footprinting approaches alone. By designing probes containing scrambled nucleotide sequences, we were able to assign a reasonable structure for the slow-migrating complex Z. Although the most compact complex (X) appears to represent simultaneous recognition of all three target areas, intermediate complex Y is more difficult to assign and could represent a bound complex in which the third strand does not remain in the DNA major groove in area II.

These results suggest several conclusions. First, electrophoretic gel mobility assays offer unique advantages when applied to characterizing triple-helical complexes. This technique can reveal the presence of multiple complexes that are difficult to detect using other approaches. Second, probe affinity was not uniformly predicted by the apparent number of available base triplets (e.g., probes **8**, **11**, and **12**). Finally, as has previously been suggested by other authors, the utility of alternate strand triple helix formation may be sequence dependent (Jayasena & Johnston, 1993).

ACKNOWLEDGMENT

We acknowledge the excellent technical assistance of D. Eicher and C. Mountjoy. We thank B. Gold and T. Hollingsworth for comments on the manuscript and R. Roberts for helpful discussion.

REFERENCES

- Beal, P. A., & Dervan, P. B. (1991) *Science* 251, 1360–1363.
- Beal, P. A., & Dervan, P. B. (1992) *J. Am. Chem. Soc.* 114, 4976–4982.
- Bhattacharyya, G. K., & Johnson, R. A. (1977) in *Statistical Concepts and Methods*, pp 507–515, John Wiley & Sons, New York.
- Bressac, B., Kew, M., Wands, J., & Ozturk, M. (1991) *Nature* 350, 429–430.
- Cantor, C. R., & Schimmel, P. R. (1980) in *Biophysical Chemistry*, p 864, W. H. Freeman and Company, New York.
- Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J., & Hogan, M. E. (1988) *Science* 241, 456–459.
- Durland, R. H., Kessler, D. J., & Hogan, M. (1991) in *Prospects for antisense nucleic acid therapy of cancer and AIDS* (Wickstrom, E., Ed.) pp 219–226, Wiley-Liss, New York.
- Duval-Valentine, G., Thuong, N. T., & Hélène, C. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 504–508.
- Froehler, B. C., Terhorst, T., Shaw, J.-P., & McCurdy, S. N. (1992) *Biochemistry* 31, 1603–1609.
- Hélène, C. (1991) *Eur. J. Cancer* 27, 1466–1471.
- Horne, D. A., & Dervan, P. B. (1990) *J. Am. Chem. Soc.* 112, 2435–2437.
- Jayasena, S. D., & Johnston, B. H. (1992a) *Biochemistry* 31, 320–327.
- Jayasena, S. D., & Johnston, B. H. (1992b) *Nucleic Acids Res.* 20, 5279–5288.
- Jayasena, S. D., & Johnston, B. H. (1993) *Biochemistry* 32, 2800–2807.

- Le Doan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J.-L., Thuong, N. T., Lhomme, J., & Hélène, C. (1987) *Nucleic Acids Res.* 15, 7749–7760.
- Maher, L. J. (1992a) *BioEssays* 14, 807–815.
- Maher, L. J. (1992b) *Biochemistry* 31, 7587–7594.
- Maher, L. J., Wold, B., & Dervan, P. B. (1989) *Science* 245, 725–730.
- Maher, L. J., Dervan, P. B., & Wold, B. J. (1990) *Biochemistry* 29, 8820–8826.
- Maher, L. J., III, Dervan, P. B., & Wold, B. (1991) in *Prospects for antisense nucleic acid therapy of cancer and AIDS* (Wickstrom, E., Ed.) pp 227–242, Wiley-Liss, New York.
- Maher, L. J., Dervan, P. B., & Wold, B. (1992) *Biochemistry* 31, 70–81.
- Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol.* 65, 499.
- Moser, H. E., & Dervan, P. B. (1987) *Science* 238, 645–650.
- Olivas, W. M., & Maher, L. J., III (1994) *BioTechniques* 16, 128–132.
- Orson, F. M., Thomas, D. W., McShan, W. M., Kessler, D. J., & Hogan, M. E. (1991) *Nucleic Acids Res.* 19, 3435–3441.
- Perrouault, L., Asseline, U., Rivalle, C., Thuong, N. T., Bisagni, E., Giovannangeli, C., & Le Doan, T. (1990) *Nature* 344, 358–360.
- Roberts, R. W., & Crothers, D. M. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 9397–9401.
- Strobel, S. A., & Dervan, P. B. (1991) *Nature* 350, 172–174.
- Sun, J., Bizemont, T. D., Duval-Valentin, G., Montenay-Garestier, T., & Hélène, C. (1991) *C. R. Acad. Sci., Ser. 3* 313, 585–590.
- Young, S. L., Krawczyk, S. H., Matteucci, M. D., & Toole, J. J. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 10023–10026.