Biochemical Evidence That a D-Loop Is Part of a Four-Stranded PNA−DNA Bundle. Nickel-Mediated Cleavage of Duplex DNA by a Gly-Gly-His Bis-PNA[†]

Matthew Footer,[‡] Michael Egholm,[§] Steve Kron,[‡] James M. Coull,[§] and Paul Matsudaira*,[‡,||]

Whitehead Institute for Biomedical Research, Nine Cambridge Center, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, and Biosearch Products, PerSeptive Biosystems, 500 Old Connecticut Path, Framingham, Massachusetts 01701

Received February 27, 1996; Revised Manuscript Received June 13, 1996[⊗]

ABSTRACT: A peptide nucleic acid (PNA) with improved strand-displacement capability and a site-specific DNA cleavage function is a novel reagent for probing the structure of PNA−DNA complexes in solution. By linking two PNAs in tandem with an aliphatic linker, the bis-PNA forms a bis-PNA−DNA triple-stranded complex having a higher stability to thermal denaturation than conventional monomeric PNAs. When a Gly-Gly-His tripeptide is placed on either the Watson−Crick or Hoogsteen bis-PNA strand, nickel-mediated cleavage is detected at specific sites on the displaced and hybridized DNA strands. Because the displaced strand is cleaved when GGH is placed on either PNA strand, the D-loop must be close to the backbone of the bis-PNA−DNA triplex. Furthermore, the pattern of cleavage on the displaced strand suggests the nickel−tripeptide complex lies in a groove formed by the displaced DNA strand and both PNA strands. These observations suggest that the D-loop is a part of a four-stranded bis-PNA−DNA_2 bundle.

A D-loop is formed when one strand of a double-stranded DNA molecule is displaced by an RNA primer or an oligonucleotide DNA probe. The displaced strand is referred to as a D-loop because in electron micrographs the displaced strand appears as a thin strand which has looped away from a thicker double-stranded DNA hybrid (Brack, 1981). The open structure of a D-loop is confirmed by its sensitivity to single-strand-specific nucleases and its ability to mimic a transcription bubble by supporting transcription (Demidov et al., 1993; Møllegaard et al., 1994; Nielsen et al., 1991). While these observations suggest that a D-loop is physically separate from the hybrid duplex, the displaced strand must nevertheless remain in close proximity to its complementary strand because their lengths are identical. This constraint in DNA path length raises a question whether the D-loop is unstructured as suggested by electron microscopy or is part of a recognizable bundle with the hybrid DNA duplex.

Probing the structure of a D-loop requires a reagent which can form a stable hybrid with double-stranded DNA. Long (>500 base) DNA probes like those used in fluorescence *in situ* hybridization of chromosomal DNA have sufficient base pair stabilizing interactions to remain bound to the double-stranded DNA substrate. Alternatively, shorter DNA probes (15-mer) can form stable complexes with double-stranded DNA molecules if the complexes are decorated with a DNA-binding protein such as RecA (Ferrin & Camerini-Otero, 1991). DNA-binding proteins are not needed to stabilize long probes. Short PNA probes (6−10-mer), DNA analogs in which the DNA bases are linked to a polyamide backbone consisting of oligomers of *N*-(2-aminoethyl)glycine (Dueholm et al., 1994; Egholm et al., 1993, 1992a; Nielsen et al., 1991), are capable of forming exceptionally stable complexes with double-stranded DNA. The stability results from triplex formation between a pair of PNAs and a single DNA strand. The structure of a triplex reveals the PNAs[1] form normal Watson−Crick and Hoogsteen base pair interactions. However, this triple helix is unwound into a new structure, the P-helix, which has features of both A- and B-helices (Betts et al., 1995). This P-helix contains three of the four strands we expect to find when a bis-PNA is bound to double-stranded DNA.

Because of several unique properties, PNAs can form stable complexes with double-stranded DNA under conditions that would dissociate DNA duplexes. First, the neutral charge of the PNA backbone does not repel the phosphate backbone of the complementary DNA strand. Thus a PNA−DNA duplex remains stable in low ionic strength conditions. Second, the triplex is very resistant to denaturation, as much as 40 °C higher melting temperatures compared with a DNA duplex of identical sequence (Egholm et al., 1992a,b; Kim et al., 1993). This stability is explained not only by the presence of a PNA Hoogsteen strand but by additional interactions between amide groups on the PNA Hoogsteen strand and the DNA phosphate backbone (Betts et al., 1995). Stability is further improved by covalently linking the Watson−Crick PNA strand to the Hoogsteen PNA strand to form a bis-PNA (Egholm et al., 1995; Griffith et al., 1995). Linkage in tandem reduces a trimolecular reaction to a bimolecular reaction. Consequently, bis-PNAs form a triplex which has a higher thermal stability than conventional PNAs. Lastly, to eliminate the pH-dependent binding of the Hoogsteen strand, one group (Egholm et al., 1995) has substituted a protonated cytosine analog, pseudoisocytosine. These modifications generate a bis-PNA that binds sequences as

[†] Supported by NIH Grant P50-HG00098.
* To whom correspondence should be addressed.
[‡] Whitehead Institute for Biomedical Research.
[§] PerSeptive Biosystems.
[||] Massachusetts Institute of Technology.
[⊗] Abstract published in *Advance ACS Abstracts,* August 1, 1996.

[1] Abbreviations: PNA, peptide nucleic acid; GGH, glycine−glycine−histidine; dsDNA, double- stranded DNA; PCR, polymerase chain reaction; P_i, phosphate.

S0006-2960(96)00486-2 CCC: $12.00

short as only seven purines in a double-stranded target with efficient strand displacement. Despite the high stability of the complexes, single base pair mismatches greatly reduce the melting temperature. Thus, PNAs are able to recognize double-stranded DNA in a very sequence-specific manner (Demidov et al., 1995).

Since their relatively recent appearance PNAs have found diverse use in molecular biology. They have been used for DNA mapping (Demidov et al., 1993, 1994; Nielsen et al., 1991), as transcription promoters (Møllegaard et al., 1994), as restriction enzyme inhibitors (Nielsen, 1993), to isolate active genes (Boffa et al., 1995), as antigene agents (Hanvey et al., 1992), and for mapping rare restriction sites in genomic DNA (Veselkov et al., 1996). Our interest in PNAs is their use as a tool for high-resolution mapping of DNA. The ease with which a PNA can be made and modified, its high stability when bound to DNA, and their sequence specificity make them an ideal candidate as a mapping reagent. A simple approach to map a set of sites is to cleave the DNA at the sites and resolve their sizes by electrophoresis.

We have modified the bis-PNAs to create a synthetic DNA cleavage reagent by coupling a Gly-Gly-His peptide to the end of the bis-PNA. By addition of Gly-Gly-His, the consensus sequence for the copper-binding domain of serum albumin (Bryce et al., 1966; Kruck et al., 1976) to the DNA-binding domain of Hin recombinase one group (Mack, 1991; Mack & Dervan, 1992) was able to cleave DNA at Hin sites. They also showed experimental evidence that the cleaving group of GGH (Hin 139−190) lies in the minor groove as predicted by cocrystal structures of DNA-binding proteins with dsDNA. We show that the Gly-Gly-His peptide on either the Watson−Crick or Hoogsteen strand of a bis-PNA supports sequence-specific nickel-mediated cleavage of not only the DNA Watson strand but also the displaced DNA Crick strand. Furthermore, the pattern of the cleavage sites demonstrates that the D-loop is in close proximity to the bis-PNA−DNA triplex. Our results demonstrate the use of a novel nonenzymatic DNA cleavage reagent which can reveal important insights into the structure of elaborate DNA complexes.

MATERIALS AND METHODS

PNA Probes. Three specially modified PNAs, bis-PNA 6229, bis-PNA 5939, and bis-PNA 6230 (Figure 1), were synthesized as described (Egholm et al., 1995). The PNAs consisted of two 10-mer PNA sequences (a Crick PNA strand complementary to the Watson DNA strand and Hoogsteen PNA strand) linked in tandem by three 8-amino-3,6-dioxaoctanoic acid units to form a bis-PNA. The flexibility attained with the linker allows the bis-PNA to fold back on itself and "clamp" a single DNA strand forming a bimolecular triple helix. A Gly-Gly-His peptide, designed to chelate nickel and induce cleavage of the DNA backbone, is incorporated by standard peptide synthesis chemistry to one end of the PNA strand. To study the specificity of the cleavage reaction, we placed the Gly-Gly-His peptide on the Crick PNA strand (bis-PNA 6230) or on the Hoogsteen PNA strand (bis-PNA 5939 and bis-PNA 6229) and varied the distance between the peptide and the PNA with a linker (bis-PNA 6229 and bis-PNA 6230) or without a linker (bis-PNA 5939). The pH sensitivity in the formation of the triple helix was eliminated by replacing cytosine in the Hoogsteen PNA strand with pseudoisocytosine. As controls, we synthesized two 10-mer PNAs, PNA-C and PNA-H, for determining the relative thermal stability of the bis-PNAs. We also synthesized two 12-mer PNAs (Figure 2) without Gly-Gly-His for comparison. Since Gly-Gly-His does not participate in base pairing and is not expected to present a steric problem, we did not incorporate them into these control PNAs.

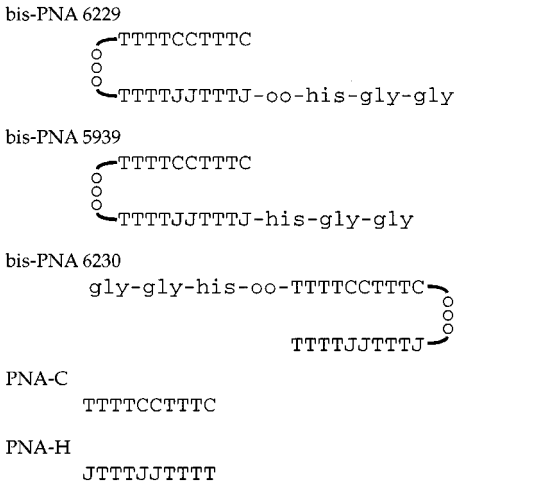

FIGURE 1: Schematic diagram of PNA structures. The PNAs are oriented with their Crick strand at the top and Hoogsteen strand at the bottom. The Gly-Gly-His peptide is located at the N-terminus of the bis-PNA. J represents pseudoisocytosine while O is a hydrophilic linker. PNA-C and PNA-H represent conventional PNAs.

Oligonucleotides were obtained from Research Genetics (Huntsville, AL). Plasmid pDP1007 was the gift of David Page (Page et al., 1987). A 128 bp target DNA encoding part of the human ZFY gene was generated by PCR (Woodford et al., 1995) using pDP1007 as a template. Individual strands of PCR product were labeled with [^{33}P]-dATP and polynucleotide kinase prior to PCR. The labeled products were purified by electrophoresis through 3% Nusive 3:1 (FMC BioProducts, Rockland, ME) TAE agarose gels.

Thermal Stability of PNA−DNA Complexes. To measure the melting temperatures of the bis-PNA−DNA complexes, we melted and reannealed the triplexes while monitoring the change in absorbance at 260 nm. All of these studies were performed on an Aviv Model 118DS spectrophotometer using the following conditions: 140 mM NaCl, 10 mM P_i, pH 7.5, oligonucleotides at 2 μM, PNAs at 4 μM, and bis-PNAs at 2 μM. Prior to melting, all the mixtures were heated to 65 °C for 1 h and slowly cooled to 4 °C. Absorbance measurements at 260 nm were recorded from 4 to 94 °C in 2 °C steps and a 2 min equilibration between each measurement. We monitored the progress of reannealing directly after the melts by cooling the solution in 2 °C steps every 2 min The T_m values were determined from the maximum of the first derivative of the plot of the absorbance at 260 nm versus temperature.

Cleavage of DNA with Bis-PNAs. Previous reports (Mack, 1991; Mack & Dervan, 1992) demonstrate that when the Gly-Gly-His sequence is linked to a fragment of the Hin recombinase, the Gly-Gly-His−Hin conjugate protein can cleave DNA in the presence of oxidized nickel. To demonstrate that Gly-Gly-His bis-PNAs can cleave DNA, we studied the effects of nickel-mediated cleavage with bis-PNA 6229. Both strands of the 128 bp target DNA were labeled with ^{33}P using polynucleotide kinase. The labeled

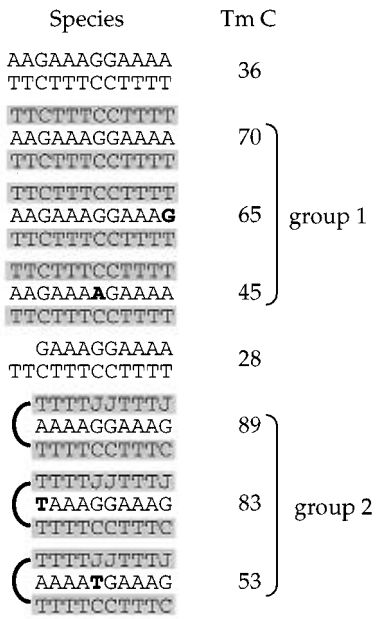

FIGURE 2: Denaturation of DNA−PNA complexes. DNA−DNA complexes have the Watson strands on top (5′ base on left). Conventional (group 1) and bis (group 2) PNAs (highlighted, Hoogsteen top strand) complexed with Watson DNA (5′ base on left) target. Bases in boldface type indicate position of a base pair mismatch (J = pseudoisocytosine). All the bis-PNAs in group 2 are PNA 6229.

target DNA (1.5 pmol) was mixed with bis-PNA 6229 (150 pmol) in 60 μL of buffer A (20 mM P_i, pH 8.0, 20 mM NaCl). After heating to 94 °C and slowly cooling to 25 °C, the bundle was purified by centrifugation through a spin column (Clonetech Chroma Spin 30 column equilibrated with buffer A). In a typical 10 μL cleavage reaction we used 2 μL of purified complex and added nickel(II) acetate to a final concentration of 10 μM and then monoperoxyphthalic acid to a final concentration of 10 μM. After 15 min at room temperature, the reaction was terminated by precipitating the bundle with 7.5 μL of 10 M ammonium acetate and 74 μL of ethanol using 100 ng of λ DNA as a carrier. To increase the cleavage efficiency, some samples were incubated at 90 °C for 30 min in 40 μL of 0.1 M butylamine. All of the samples were evaporated to dryness in a centrifuge concentrator and resuspended in 10 μL of gel sample buffer prior to electrophoresis through a 7.5% polyacrylamide gel. We characterized the point of cleavage with the three different bis-PNAs using 128 bp PCR product that was labeled on either the forward or reverse strand with the nickel, peracid, PNA, and base workup described above. After electrophoresis through a 6% urea−polyacrylamide gel, the cleavage products were visualized with a Fuji BAS 2000 phosphoimager.

RESULTS

Stability of Bis-PNA−DNA Triplexes. To establish the conditions for cleaving double-stranded DNA, we measured the melting temperatures of various PNA_2−DNA (group 1) and bis-PNA−DNA (group 2) complexes. These data are summarized in Figure 2. Among the first group of PNAs, the T_m of a 12-mer PNA_2−DNA triplex was 34 °C higher than the corresponding double-stranded DNA duplex. The effects of single base mismatches were studied by introducing a single change in the target DNA strand. As expected from previous studies (Ajmera et al., 1986; Dedon & Goldberg, 1992; Dedon et al., 1992), mismatches in the target sequence reduce the T_m. For example, a single base mismatch at an end of the target reduces the T_m of the PNA_2−DNA complex by 5 °C while a mismatch in the center of the complex reduces the T_m by 25 °C.

We next investigated the triple strand formation by two PNA strands linked in tandem. Using T_m as a measure, we found that, unlike conventional PNAs, bis-PNAs (group 2) form a more stable complex with a complementary DNA oligonucleotide. As shown in Figure 2 the melting temperature of a 10-mer bis-PNA−DNA complex is 19 °C higher than a corresponding 12-mer PNA_2−DNA complex. Single base pair mismatches at the end or middle of the sequence decrease the melting temperature by 6 and 35 °C, respectively. By linking the PNA strands in tandem, we confirmed that a bis-PNA−DNA complex was more stable to denaturation than a conventional PNA_2−DNA complex.

We also compared the denaturation temperature with the annealing temperature to determine whether the formation of a triplex with a bis-PNA occurs at a higher temperature than with a conventional PNA. In Figure 3A, the annealing temperature (72 °C) of a bis-PNA−DNA complex is 16 °C lower than the denaturation temperature. Both curves are monophasic. In contrast, a corresponding PNA_2−DNA complex, formed with PNA-C and PNA-H, displays a much greater difference (42 °C) between the melting and annealing temperatures. In addition, the melting transition is multiphasic with at least two intermediate transitions at 63 and 82 °C. This more complex denaturation curve is not seen in the melting and annealing profiles of a short 10-base DNA target (Figure 3B).

Nickel−Gly-Gly-His-Mediated Cleavage of dsDNA. Having established the stability of bis-PNA-DNA complexes in both denaturation and annealing conditions, we investigated the structure of the bundle in a large dsDNA target (Figure 4A). To mark positions in the target that would confirm the binding site, we employed an artificial cleavage approach by coupling a Gly-Gly-His peptide to the end of the PNA (Figure 4B). In the presence of nickel and monoperoxyphthalic acid, the 128 bp target was cleaved into fragments whose sizes are consistent with a cleavage site at the PNA-binding region. Addition of a competing oligonucleotide complementary to the bis-PNA eliminates cleavage by the bis-PNA (data not shown). The yield of cleaved DNA improves with base workup of the DNA after peracid treatment. Under these cleavage conditions, when 100% of the target is complexed with a bis-PNA, approximately 20% of the target is cleaved.

After optimizing the cleavage reaction, we mapped the sites of cleavage to confirm the orientation of the bis-PNA (Figure 5). When Gly-Gly-His is separated from the Hoogsteen strand by two linker molecules (bis-PNA 6229), the cleavage sites lie on both target DNA strands (Figure 5A). Cleavage is favored at one site on each strand with less frequent cleavage at other sites. In addition, there is no evidence of cleavage along the same face of the DNA helix which would lead to a phasing type of cleavage pattern. When the sites are mapped onto a diagram of the DNA target sequence (Figure 5B), cleavage occurs two or three bases 5′ to the position of the nickel−tripeptide complex on the displaced or Watson DNA strands, respectively. Although the bis-PNA lacking a linker (bis-PNA 5939) also directs 5′

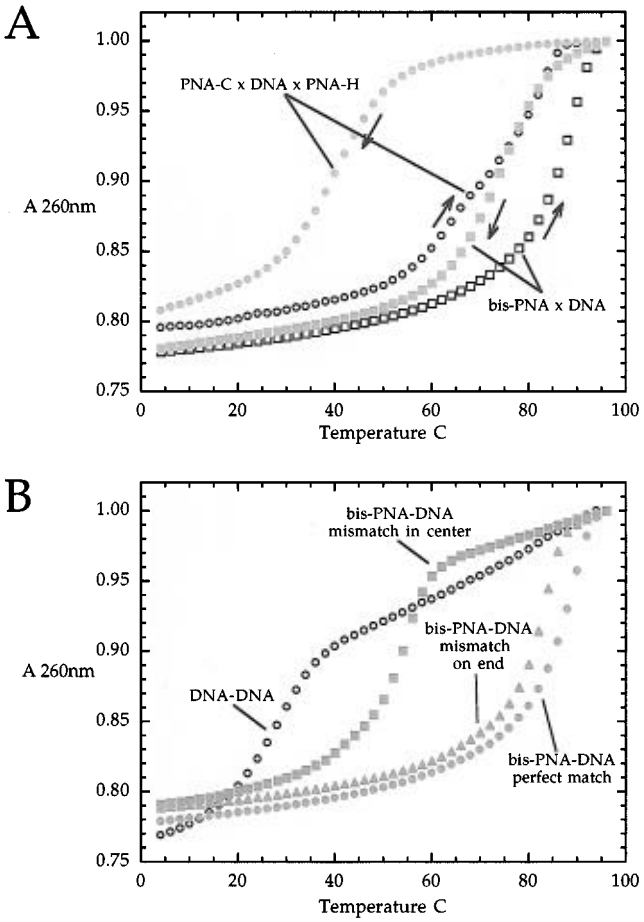

FIGURE 3: Melting T_ms of bis-PNA−DNA complexes. The absorbances are normalized to the measurement taken at 94 °C. (A) The denaturation (up arrows) and annealing (down arrows) curves of PNA_2−DNA (circles) and bis-PNA−DNA (boxes) complexes display hysteresis. A biphasic melting transition is detected for a PNA_2−DNA complex while a bis-PNA−DNA complex displays a monophasic transition whose melting temperature is 20 °C higher. As the temperature is lowered from 94 °C, the bis-PNA−DNA complex reanneals at a higher temperature than a conventional PNA_2−DNA complex, but the difference in T_ms between denaturation and reanealing is much less (16 vs 30 °C). (B) Denaturation curves for a DNA oligo target and either a complementary DNA probe, complementary bis-PNA, or a bis-PNA with an oligo target with either a mismatch at an end or a mismatch in the middle.

cleavage on both DNA strands, the Watson DNA strand is cleaved less frequently while the major cleavage site on the displaced strand is shifted to three bases from the tripeptide moiety.

The pattern of cleavage is completely changed when the nickel−tripeptide complex is placed on the Crick bis-PNA strand (bis-PNA 6230). Because PNAs and peptides are synthesized from the C-terminus, the tripeptide must be placed on the N-terminus of the bis-PNA. This results in the Gly-Gly-His being relocated to the opposite end of the triplex when compared to the bis-PNAs with the Gly-Gly-His on the Hoogsteen strand. At this position, the nickel−tripeptide complex cleaved both strands of the 128 base pair DNA target, and the cleavage sites were clustered near the position of the cleavage reagent. However, a major difference in the cleavage pattern is that cleavage occurs 3′ on the Watson DNA strand and the location of the sites relative to the cleavage reagent had shifted. On the displaced strand cleavage occurred most frequently four bases away while on the Watson strand the site was separated from the cleavage

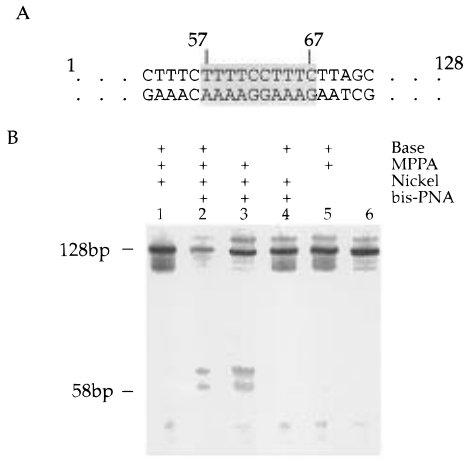

FIGURE 4: Nickel-mediated cleavage site in the DNA target by bis-PNA 6229. (A) The bis-PNAs bind to a 10 base sequence (shaded box) within the 128 bp ds DNA target. (B) The parameters that affect nickel-mediated cleavage of the DNA target duplex. In the presence of PNA, nickel, MPPA, and base (lane 2), the 128 bp target is cleaved into two fragments. No cleavage occurs in the absence of bis-PNA, nickel or monoperoxyphthalic acid (lanes 1, 4, 5, 6).

reagent by one base. Compared with cleavage from the Hoogsteen strand, the nickel−tripeptide complex on the Crick strand was a less efficient cleavage reagent.

DISCUSSION

Our results (Figures 2 and 3) document the behavior of the bis-PNAs under careful melting and annealing conditions. Bis-PNAs exhibit two interesting behaviors when compared to conventional PNAs. First, the hysteresis between melting and annealing is greatly reduced, primarily due to a large increase in the annealing temperature. This change suggests that formation of a bis-PNA−DNA complex occurs more readily than with conventional PNAs. Second, the bis-PNAs do not exhibit a multiphasic melting curve. The multiphasic melting behavior of a conventional PNA_2−DNA complex probably reflects separate T_ms for the dissociation of the Hoogsteen strand and the denaturation of the Watson−Crick strands. The reduced hysteresis and the monophasic melting curve can be explained if the formation of a triplex is simplified from a trimolecular to a bimolecular mechanism. Others have postulated (Demidov et al., 1995; Egholm et al., 1995) that the conversion of a duplex into a triplex is not kinetically limited when the Hoogsteen strand is linked to the Crick strand because once a PNA−DNA Watson−Crick duplex has formed, the Hoogsteen PNA strand is immediately available to complete the triplex. When the target is a long double-stranded DNA, this improved ability to form a triple-stranded complex rapidly reduces competition for the Watson DNA strand by the displaced Crick DNA strand. A practical consequence of this enhanced property is the ability to use bis-PNA at higher temperatures where normal DNA duplex interactions are eliminated.

As a sequence-specific hybridization reagent, a bis-PNA exhibits good sensitivity to a mismatch in the sequence especially when the mismatch occurs at an internal position. However, the melting temperature is only slightly affected when the mismatch is at an end position, as is the case for conventional DNA duplexes. To improve base discrimination by the bis-PNAs, we are currently testing an asymmetric

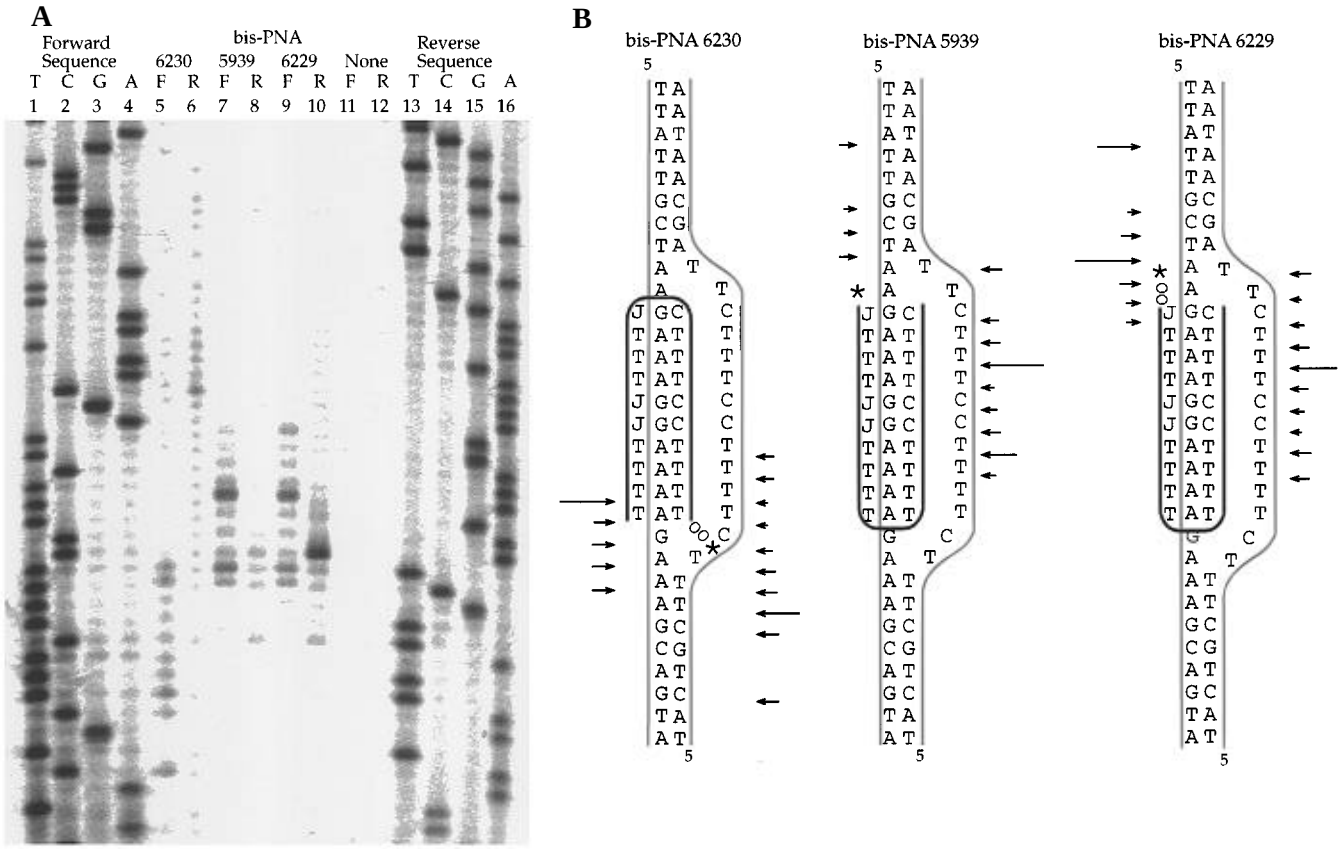

FIGURE 5: Identification of the cleavage sites caused by the nickel−Gly-Gly-His complex located at different positions on a bis-PNA. (A) The cleavage sites are identified on high-resolution gels of target DNA labeled on either the reverse strand (lanes 6, 8, 10, 12) or forward strand (lanes 5, 7, 9, 11). The presence of the Gly-Gly-His peptide on the Hoogsteen PNA strand, bis-PNAs 5939 and 6229 (lanes 7−10), results in cleavage on similar sites on either DNA strand. Without a bis-PNA, no cleavage occurs (lanes 11, 12). The target DNA is sequenced with the "forward" primer (lanes 1−4) or "reverse" primer (lanes 13−16). (B) The location of the cleavage sites is depicted in a schematic diagram. The arrows indicate the site of cleavage; the length of the arrow is approximately proportional to the cleavage frequency at the site.

bis-PNA. The asymmetry lies in the lengths of the two PNA strands; the Crick PNA strand is 12−15 bases long while the pyrimidine-rich Hoogsteen strand is 7 bases in length. This design is intended to improve the sequence specificity, determined by the Crick PNA strand, without affecting the high thermal stability provided by the Hoogsteen PNA strand.

Directed DNA Cleavage by a Bis-PNA−GGH Chimera. The mechanism for cleavage of the DNA backbone by nickel−Gly-Gly-His has not been determined, but it is clear that the mechanism involves an oxidized nickel intermediate and the increased amount of cleavage product upon base workup suggests that more than one path to the final cleavage product exists. Our observations are consistent with a model utilizing a ligand radical produced by an oxidized metal ion intermediate (Hecht, 1986). Cleavage does not go to completion but seems limited to 20% of the DNA molecules. This limit may indicate a strict stereospecific orientation of the square-planar tripeptide−nickel complex with respect to the DNA target. In addition, a limited half-life of the radical due to self-destruction may also compete with DNA cleavage.

Bis-PNA−GGH chimeras display two properties that are superior to other nonenzymatic DNA cleavage reagents. First, cleavage can be targeted to any site because of the sequence-specific binding of the bis-PNA. Second, cleavage is more efficient than with a diffusible free radical, such as that generated by Fe−EDTA, because cleavage is restricted to the three dimensional space around the nondiffusible GGH. In examples where the cleavage reagent can diffuse, the cleavage sites are clustered with a Gaussian distribution, and the clusters are "phased" or spaced with a periodicity (Mack et al., 1990). Phasing reflects that cleavage occurs along the same face of the DNA helix. We do not detect phasing of cleavage sites in our gels. Instead, our data are consistent with those reported by Mack and co-workers (Mack, 1991; Mack & Dervan, 1992) of a small number of cleavage sites, which suggests that cleavage occurs at sites closest to the GGH group.

Triplex and D-Loop Structure Revealed by Nickel-Mediated Cleavage. Because the Gly-Gly-His tripeptide is tethered to the bis-PNA, we are able to probe the structure of the DNA while in a complex with the bis-PNA. The first finding is that the D-loop is in close proximity to the Hoogsteen and Crick PNA strands. Displaced from its former position as the DNA Crick strand, the D-loop must remain near the DNA Watson strand for no better reason than that their path lengths are identical. This constraint of length limits the distance the displaced strand can "loop" away from the Watson strand. To visualize this more clearly, we used the structure of the PNA−DNA triplex (Betts et al., 1995) in a model of the strand-invaded bundle (Figure 6). We assumed that the D-loop must follow the path of its Watson strand complement and therefore must lie on the same side of the helix as the Crick PNA strand. Both Watson and D-loop DNA strands were then fitted by hand to form B-form duplex DNA at opposite ends of the bis-PNA−DNA

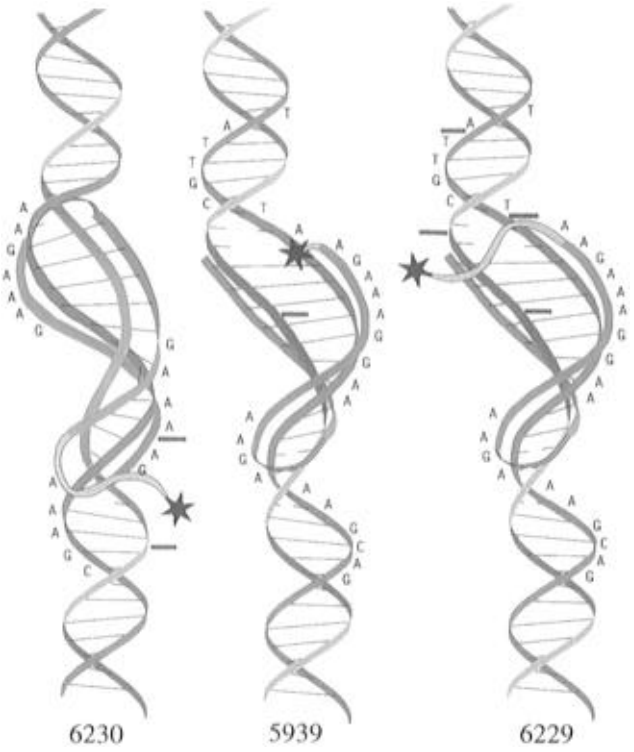

FIGURE 6: Models of the bis-PNA−dsDNA bundle. The target is B-form DNA with Watson (blue with black letters)−Crick (gray) base pairs (black lines). The DNA helix is unwound into a P-helix by the bis-PNA (Crick, green; Hoogsteen, brown) to form a triplex and a displaced DNA strand (gray). Linker molecules connect the PNA Crick and Hoogsteen strands (thin brown line) and the GGH cleavage group (red star) to the bis-PNA (thin pink line). The length of the linker is to scale. Major sites of cleavage (red bars) are located near the position of the GGH group.

triplex. These constraints generate a crude model of a four-stranded bundle in which the D-loop and PNA Hoogsteen strand are adjacent. The GGH tripeptide is clearly accessible to both Watson and D-loop DNA strands, and the majority of cleavage sites on the D-loop and the Watson DNA strand lie on the same side of the bundle as the cleavage group. Our model for the structure is still consistent with the view that the D-loop is an open structure. Since a short D-loop is sensitive to cleavage by various nucleases (Demidov et al., 1993), we do not postulate that this fourth D-loop strand is fixed rigidly in position, but we do argue that, for small D-loops such as ours, the D-loop occupies a space strongly reminiscent of its original place in the helix. Electron micrographs showing D-loops made of large displaced DNA strands may be more indicative of sample preparation artifacts than of the D-loop as it exists in solution. On the basis of our results, we suggest the published electron micrographs show molecules which are abnormally untwisted. Our experiments, using short probes and conditions which allow for native double helix formation in solution, provide insight into D-loop structure as it may exist *in situ*.

The model and P-helix structure helped to solve a problem in interpreting the cleavage data shown in Figure 5. Before the PNA−DNA structure was published we could not reconcile our cleavage data using the standard B-form DNA structure because some of the cleavage sites were located on the back of the helix, a region inaccessible to a tethered cleavage reagent. The cleavage pattern was more sensible if the DNA helix was unwound. Because our cleavage data are better explained using the P-helix which has a pitch of 16 bases per turn, our results provides biochemical evidence in support of the P-helix structure.

The 5′ or 3′ location of the cleavage sites relative to the cleavage reagent provides additional information about the environment around the cleavage reagent. Mack et al. (1991) show that GGH cleaves at sites 3′ to the cleaving group if it lies in the minor groove of B-form DNA. In contrast, we show that cleavage occurs 5′ to the Gly-Gly-His tripeptide (Figure 5B) when it is coupled to the end of either the Crick or Hoogsteen PNA strands. Superficially, this pattern would be consistent with a model that the cleavage reagent lies in the major groove of B-form DNA and certainly with the model of P-form DNA. However, because one of the cleaved strands is the D-loop, the 5′ cleavage pattern provides supporting evidence for our model that the D-loop follows a path like the Crick strand in the major groove of a DNA duplex. As a consequence, the D-loop DNA strand must form some type of "major groove" with the Watson DNA strand. With the Hoogsteen strand present, our data argue that the strand invading bis-PNA is at least a four-stranded bundle with the duplex DNA target.

ACKNOWLEDGMENT

The authors thank Dr. Noboru Uryu for providing preliminary characterization of PNAs, Dr. James Berger for help with PNA graphics, Dr. Steve Jordan for the P-helix coordinates, Dr. Alex Rich for discussions about DNA triplexes, and Dr. Martha Oakley for advice about chemical cleavage of DNA.

REFERENCES

Ajmera, S., Wu, J. C., Worth, L., Jr., Rabow, L. E., Stubbe, J., & Kozarich, J. W. (1986) *Biochemistry 25*, 6586−6592.

Betts, L., Josey, J. A., Veal, J. M., & Jordan, S. P. (1995) *Science 270*, 1838−1841.

Boffa, L. C., Carpaneto, E. M., & Allfrey, V. G. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 1901−1905.

Brack, C. (1981) *CRC Crit. Rev. Biochem. 10*, 113−169.

Bryce, G. F., Roeske, R. W., & Gurd, F. R. N. (1966) *J. Biol. Chem. 241*, 1072−1080.

Dedon, P. C., & Goldberg, I. H. (1992) *Chem. Res. Toxicol. 5*, 311−332.

Dedon, P. C., Jiang, Z. W., & Goldberg, I. H. (1992) *Biochemistry 31*, 1917−1927.

Demidov, V., Frank-Kamenetskii, M. D., Egholm, M., Buchardt, O., & Nielsen, P. E. (1993) *Nucleic Acids Res. 21*, 2103−2107.

Demidov, V. V., Cherny, D. I., Kurakin, A. V., Yavnilovich, M. V., Malkov, V. A., Frank-Kamenetskii, M. D., Sonnichsen, S. H., & Nielsen, P. E. (1994) *Nucleic Acids Res. 22*, 5218−5222.

Demidov, V. V., Yavnilovich, M. V., Belotserkovskii, B. P., Frank-Kamenetskii, M. D., & Nielsen, P. E. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 2637−2641.

Dueholm, K. L., Petersen, K. H., Jensen, D. K., Egholm, M., Nielsen, P. E., & Buchardt, O. (1994) *Bioorg. Med. Chem. Lett. 4*, 1077−1080.

Egholm, M., Buchardt, O., Neilsen, P. E., & Berg, R. H. (1992a) *J. Am. Chem. Soc. 114*, 1895−1897.

Egholm, M., Buchardt, O., Nielsen, P. E., & Berg, R. H. (1992b) *J. Am. Chem. Soc. 114*, 9677−9678.

Egholm, M., Behrens, C., Christensen, L., Berg, R. H., Nielsen, P. E., & Buchardt, O. (1993) *J. Chem. Soc., Chem. Commun.*, 800−801.

Egholm, M., Christensen, L., Dueholm, K. L., Buchardt, O., Coull, J., & Nielsen, P. E. (1995) *Nucleic Acids Res. 23*, 217−222.

Ferrin, L. J., & Camerini-Otero, R. D. (1991) *Science 254*, 1494−1497.

Griffith, M. C., Risen, L. M., Greig, M. J., Lesnik, E. A., Sprankle, K. G., Griffey, R. H., Kiely, J. S., & Freier, S. M. (1995) *J. Am. Chem. Soc. 117*, 831−832.

Hanvey, J. C., Peffer, N. J., Bisi, J. E., Thomson, S. A., Cadilla, R., Josey, J. A., Ricca, D. J., Hassman, C. F., Bonham, M. A., Au, K. G., Carter, S. G., Bruckenstein, D. A., Boyd, A. L., Noble, S. A., & Babiss, L. E. (1992) *Science 258*, 1481−1485.

Hecht, S. M. (1986) *Fed. Proc. 45*, 2784−2791.

Kim, S. K., Nielsen, P. E., Egholm, M., Buchardt, O., Berg, R. H., & Norden, B. (1993) *J. Am. Chem. Soc.* 6477−6481.

Kruck, T. P., Lau, S., & Sakar, B. (1976) *Can. J. Chem. 43*, 1300−1308.

Mack, D. P. (1991) Ph.D. Dissertation, California Institute of Technology, Pasadena, CA.

Mack, D. P., & Dervan, P. B. (1992) *Biochemistry 31*, 9399−9405.

Mack, D. P., Sluka, J. P., Shin, J. A., Griffin, J. H., Simon, M. I., & Dervan, P. B. (1990) *Biochemistry 29*, 6561−6567.

Møllegaard, N. E., Buchardt, O., Egholm, M., & Nielsen, P. E. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 3892−3895.

Nielsen, P. E., Egholm, M., Berg, R. H., & Buchardt, O. (1991) *Science 254*, 1497−1500.

Nielsen, P. E., Egholm. M., Berg, R. H. & Buchardt, O. (1993) *Nucleic Acids Res. 21*, 197−200.

Page, D. C., Mosher, R., Simpson, E. M., Fisher, E. M., Mardon, G., Pollack, J., McGillivray, B., de la Chapelle, A., & Brown, L. G. (1987) *Cell 51*, 1091−104.

Veselkov, A. G., Demidov, V. V., & Frank-Kamenetskii, M. D. (1996) *Nature 379*, 214.

Woodford, K., Weitzmann, M. N., & Usdin, K. (1995) *Nucleic Acids Res. 23*, 529.

BI960486P