Bifunctional Intercalator [N-MeCys^3,N-MeCys^7]TANDEM Binds to the Dinucleotide TpA†

Kerstin Waterloh,‡ Richard K. Olsen,§ and Keith R. Fox*,‡

Department of Physiology & Pharmacology, University of Southampton, Bassett Crescent East, Southampton SO9 3TU, U.K., and Department of Chemistry and Biochemistry, Utah State University, Logan, Utah 84322

Received February 12, 1992; Revised Manuscript Received April 14, 1992

ABSTRACT: The binding of [N-MeCys^3,N-MeCys^7]TANDEM has been examined by DNase I footprinting and diethyl pyrocarbonate modification of several synthetic DNA fragments containing AT-rich regions. DNase I footprinting reveals that at low concentrations the ligand binds preferentially to the center of $(AT)_n$ regions. A fragment containing the tetranucleotide AATT was unaffected by the ligand. Diethyl pyrocarbonate modification of several fragments containing blocks of $(AT)_n$ revealed a pattern in which alternate adenines were rendered more reactive in the presence of the ligand. These reactive adenines were staggered across the two DNA strands in the 3′-direction, consistent with ligand binding to the dinucleotide TpA. In sequences of the type $(TAA)_n$·$(TTA)_m$, binding of [N-MeCys^3,N-MeCys^7]TANDEM resulted in strong modification of the second adenine in the sequence TAA, i.e., the base on the 3′-side of the ligand binding site. Data for binding to $(AT)_n$ are best explained by suggesting that the adenines sandwiched between the quinoxaline chromophores are rendered most reactive to diethyl pyrocarbonate.

The quinoxaline group of antitumor antibiotics bind to DNA by the mechanism of bifunctional intercalation (Waring & Wakelin, 1974). The naturally occurring members of this series, echinomycin (quinomycin A) and triostin A, have long been known to bind selectively to GC-rich DNAs (Wakelin & Waring, 1976; Lee & Waring, 1978a). More recently, footprinting studies have shown that these compounds are selective for the dinucleotide step CpG (Low et al., 1984a; van Dyke & Dervan, 1984). Crystal structures of both antibiotics bound to short oligonucleotides have confirmed this specificity and suggested that sequence recognition arises from the formation of hydrogen bonds between the alanine carbonyls and the 2-amino group of guanine (Wang et al., 1984; Ughetto et al., 1985) with another weaker bond between the alanine NH and the guanine N3. By contrast, TANDEM (Figure 1), a synthetic derivative of triostin A lacking the four N-methyl groups, binds exclusively to AT-rich DNAs (Lee & Waring, 1978b) showing a highly cooperative interaction with poly(dA-dT). Footprinting studies confirmed this preference for AT-rich sequences (Low et al., 1984b, 1986) but could not resolve whether the recognition site was TpA or ApT, since there was an additional requirement for flanking AT sequences. However, on the basis of restriction enzyme inhibition, Low et al. (1986) suggested TpA as the recognition site. This was consistent with the results of experiments showing that TANDEM increased the melting temperature of poly(dG-dT-dA)·poly(dT-dA-dC) but not poly(dG-dA-dT)·poly(dA-dT-dC) (Evans et al., 1982).

The crystal structure of TANDEM (Viswamitra et al., 1981; Hossain et al., 1982) revealed significant differences from those of triostin A and echinomycin. The most notable change was the formation of an internal hydrogen bond between the NH of valine (which bears a methyl group in the natural antibiotics) and the carbonyl of alanine. As a result of this change, the peptide backbone adopts a conformation in which the alanine NH groups are positioned facing toward the DNA helix. Viswamitra et al. (1981) proposed that the antibiotic would recognize the dinucleotide ApT, forming strong hydrogen bonds between the alanine NH groups and the 2-keto groups of the thymines. The involvement of the valine N-methyls in determining the sequence selectivity was supported by footprinting data on the derivative [N-MeCys^3,N-MeCys^7]TANDEM, possessing methyl groups at the cysteines but lacking the N-methyl valine groups, which showed the same binding preference as TANDEM (Low et al., 1986). However, these data suggested that the recognition site was TpA and not ApT. The model for binding at ApT does not explain why echinomycin cannot bind to this site since it possesses the correct functional groups, though in a less favorable configuration. In addition, NMR studies with the oligonucleotide d(CCCGATCGGGG) did not reveal intercalative binding but indicated that TANDEM bound by a novel mechanism within the minor groove (Powers et al., 1989).

In this study, we have determined the precise sequence selectivity of [N-MeCys^3,N-MeCys^7]TANDEM by footprinting experiments with DNase I and diethyl pyrocarbonate on a series of DNA fragments containing blocks of alternating AT residues. We conclude that the ligand is selective for the dinucleotide step TpA, consistent with the results of a recent NMR investigation (Addess et al., 1992).

Materials and Methods

Drugs and Enzymes. [N-MeCys^3,N-MeCys^7]TANDEM was prepared as previously described (Low et al., 1986). Since this drug is only sparingly soluble in aqueous systems, it was stored as a stock solution of 2 mM in dimethyl sulfoxide and diluted to working concentrations in 10 mM Tris-HCl, pH 8.0, containing 10 mM NaCl immediately before use. The final dimethyl sulfoxide concentration was always less than 5% (v/v); this had no effect on either DNase I digestion or

† This work was supported by grants from the Cancer Research Campaign and the Science and Engineering Research Council. K.R.F. is a Lister Institute Research Fellow.

* To whom correspondence should be addressed.
‡ University of Southampton.
§ Utah State University.

0006-2960/92/0431-6246$03.00/0

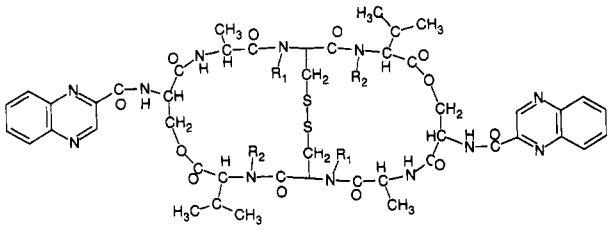

FIGURE 1: Structures of TANDEM ($R_1 = R_2 = H$), [N-MeCys³,N-MeCys⁷]TANDEM ($R_1 = CH_3$, $R_2 = H$), and triostin A ($R_1 = R_2 = CH_3$).

modification by diethyl pyrocarbonate. DNase I was purchased from Sigma and stored as previously described (Low et al., 1984a).

DNA Fragments. The preparation of pUC19 plasmids containing various inserts within the *Sma*I site has previously been described (Waterloh & Fox, 1991a,b; Fox et al., 1991). The plasmid containing the "Dickerson dodecamer" (CGCGAATTCGCG) (Fratini et al., 1982) was obtained by cloning the dodecamer into the *Sma*I site of pUC19 and transformed the plasmid into *Escherichia coli* TG2, as previously described (Waterloh & Fox 1991a). The sequence was confirmed by dideoxy sequencing with a T7 polymerase kit (Pharmacia) and found to contain a trimeric insert, lacking one cytosine, i.e., 5′-CGCGAATTCGCGGCGAATTCG-CGCGCGAATTCGCG. Plasmids were purified using Qiagen columns according to the manufacturer's instructions.

Radiolabeled DNA fragments containing the synthetic inserts were prepared by digesting with *Hin*dIII, labeling at the 3′-end with [α-³²P]dATP using reverse transcriptase, and cutting again with *Eco*RI or *Kpn*I. The labeled fragments of interest (about 75 base pairs) were separated on 7% polyacrylamide gels. For some experiments, the DNA was labeled at the opposite end by cutting first with *Eco*RI followed by *Hin*dIII or *Pst*I after labeling with [α-³²P]dATP using reverse transcriptase.

Footprinting and Gel Electrophoresis. Footprinting experiments with DNase I and diethyl pyrocarbonate were performed as previously described (Low et al., 1984a,b; Portugal et al., 1986). The products of digestion were resolved on denaturing polyacrylamide gels (8% for fragments labeled at the *Hin*dIII end, 12% for fragments labeled at the *Eco*RI end) containing 8 M urea. For experiments with the dodecamer fragment, the denaturing gels also contained 20% formamide. This was required because of the high GC content of this fragment which led to band compressions in the absence of formamide. Gels were then fixed in 10% (v/v) acetic acid, transferred to Whatman 3MM paper, dried under vacuum at 80 °C, and subjected to autoradiography at −70 °C with an intensifying screen. Autoradiographs were scanned with a Joyce-Loebl Chromoscan 3 microdensitometer. Bands were assigned by comparison with dimethyl sulfate–piperidine markers specific for guanine.

RESULTS

DNase I Footprinting. In order to examine the sequence selectivity of [N-MeCys³,N-MeCys⁷]TANDEM, we have performed DNase I footprinting experiments on several DNA fragments containing regions of alternating A and T residues. In each case, we find that cleavage of the $(AT)_n$ blocks is totally abolished by drug concentrations of 20 μM and above. Representative examples are shown in Figure 2 for fragments containing the inserts $(AT)_5GC(AT)_5$ and $(AT)_{10}CGGC(AT)_{10}$. As previously shown, cleavage in the controls is an alternating pattern of bands in which ApT is cleaved with much greater efficiency than TpA (bands from the latter are barely visible). [N-MeCys³,N-MeCys⁷]TANDEM concentrations of 20 μM and above completely abolish cleavage within the inserts. With lower drug concentrations (5 μM), bands at the center of each $(AT)_n$ block are more strongly protected than those toward the ends. This is especially noticeable with $(AT)_{10}CGGC(AT)_{10}$ for which cleavage products are still evident for CpG, GpG, and the first ApT on the 3′-(lower)side. At higher concentrations (25 μM and above), these are protected from cleavage as well.

Figure 2 also shows the effect of [N-MeCys³,N-MeCys⁷]-TANDEM on DNase I cleavage of a clone of the Dickerson dodecamer (CGCGAATTCGCG). It is evident that even at a concentration of 100 μM the ligand has had no effect on DNase I cleavage; bands around the central AATT steps are unaffected by the presence of the ligand. This may suggest that ApT does not constitute the drug binding site.

Although these DNase I footprints confirm the selectivity of [N-MeCys³,N-MeCys⁷]TANDEM for regions rich in A and T and suggest that the ligand binds best in the center of such regions, they reveal nothing new about its absolute base sequence preferences. We have therefore studied the binding in more detail by examining the modification by diethyl pyrocarbonate (DEPC).

Modification by Diethyl Pyrocarbonate (DEPC). Diethyl pyrocarbonate has been widely used as a probe for detecting unusual DNA structures. It has been particularly useful for examining the binding of quinoxaline antibiotics to DNA (Portugal et al., 1988; Mendel & Dervan, 1987; Jeppesen & Nielsen, 1988; Fox & Kentebe, 1990). Adenines on the 3′-side of echinomycin binding sites in the sequences CGA are rendered hyperreactive to DEPC. By examining the reactivity of regions of alternating AT to DEPC in the presence of echinomycin and comparing the data across the two DNA strands, this probe was used for determining the selective secondary binding of echinomycin to ApT (not TpA) (Waterloh & Fox, 1991b; Fox et al., 1991). We therefore hoped that the patterns of reactivity to DEPC would be useful for determining the precise selectivity of [N-MeCys³,N-MeCys⁷]TANDEM. Figure 3 shows the effects of [N-MeCys³,N-MeCys⁷]TANDEM on the modification of several DNA fragments, containing regions of alternating AT, by DEPC. In each case, cleavage in the control is weak and is not affected by the presence of 5% dimethyl sulfoxide (the concentration included with 100 μM ligand). It is immediately apparent that in the presence of [N-MeCys³,N-MeCys⁷]-TANDEM adenines throughout the inserts become much more sensitive to DEPC modification. On closer inspection, it can be seen that not all adenines are modified to the same extent and that in many instances an alternating pattern of cleavage products is evident. For example, with $(TA)_5GC(TA)_5$ products can be seen corresponding to the first, third and fifth adenines on either side of the central GC while products from the intermediate adenines are virtually absent. The simplest explanation for this alternating cleavage pattern is that not every TpA (or ApT) step is equally occupied by the ligand and that the drug molecules are binding within the $(AT)_n$ region in a preferred overall configuration. The reasons for this will be considered further in the Discussion. In several instances, this alternation is more pronounced at the highest ligand concentration. These patterns of DEPC modification are shown in greater detail in Figures 4–6 as histograms representing the relative cleavage at each bond, derived from densitometer traces of the data in Figure 3.

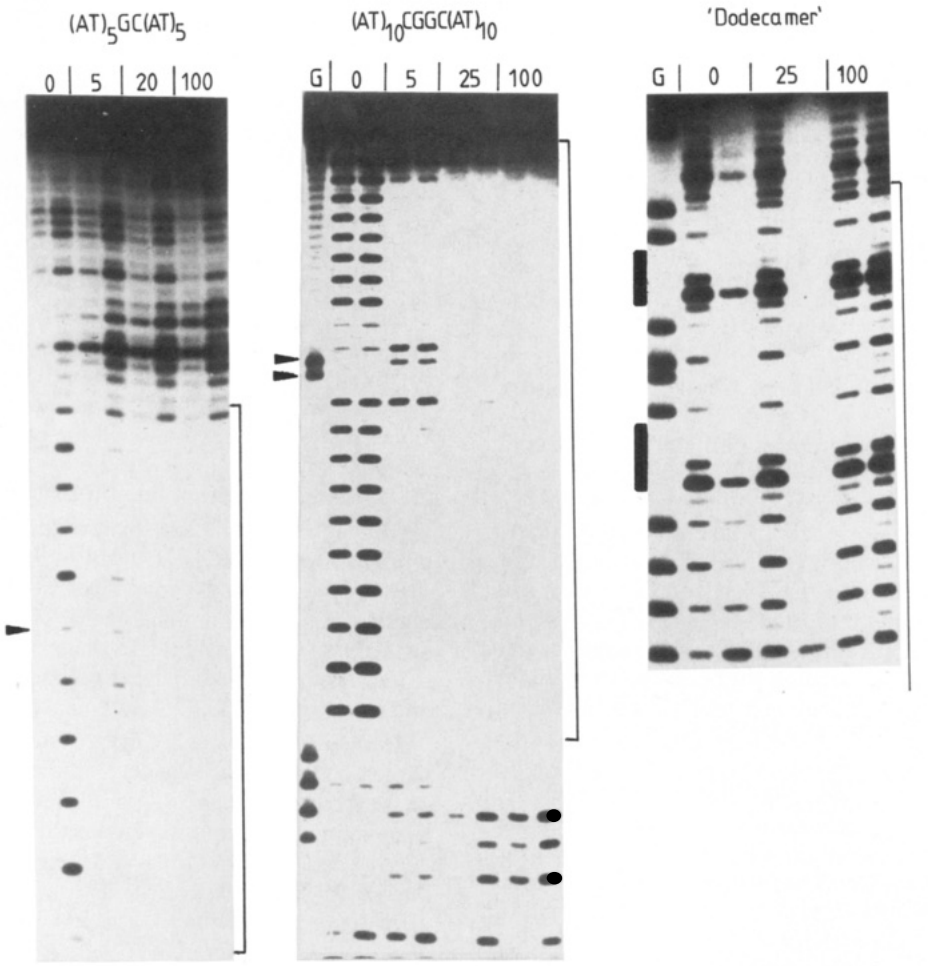

FIGURE 2: DNase I digestion patterns for DNA fragments containing the inserts $(AT)_5GC(AT)_5$, $(AT)_{10}CGGC(AT)_{10}$, and a trimer of the Dickerson dodecamer (CGCGAATTCGCG) in the presence and absence of [*N*-MeCys^3,*N*-MeCys^7]TANDEM. $(AT)_5GC(AT)_5$ corresponds to a *Hin*dIII–*Eco*RI fragment labeled at the 3′-end of the *Eco*RI site. $(AT)_{10}CGGC(AT)_{10}$ and the dodecamer are *Hin*dIII–*Sac*I fragments labeled at the 3′-end of the *Hin*dIII site. Each pair of lanes corresponds to digestion by the enzyme for 1 and 5 min. Micromolar drug concentrations are shown at the top of each pair of lanes. The tracks labeled G are dimethyl sulfate–piperidine markers specific for guanine. The position and lengths of the inserts are indicated by the square brackets. The central guanines of $(AT)_5GC(AT)_5$ and $(AT)_{10}CGGC(AT)_{10}$ are indicated by arrows; the positions of the AATT steps of the dodecamer are shown by the filled boxes. Although the dodecamer insert is a trimer, only two copies are shown; the one at the 3′-end has run off the bottom of the gel.

Figure 4 presents the results for fragments containing a central GC site, flanked by regions of alternating AT. From first looking at the data for $(AT)_5GC(AT)_5$ (Figure 4A and the third and fourth panels of Figure 3), it can be seen that there is an alternating pattern of modification of adenines at the left-hand end. The most reactive adenines are staggered in the 3′-direction across the two DNA strands. This alternating pattern of modification is again evident throughout the sequence $(TA)_5GC(TA)_5$ (Figure 4B and the first panel of Figure 3). However, in the dimeric clone of this insert [i.e., $(TA)_5GC(TA)_{10}GC(TA)_5$, Figure 4C and the second panel of Figure 3] there is an alternating pattern of modification in the $(TA)_5$ regions at the ends, but a more even pattern is in the central $(TA)_{10}$.

It should be noted that both monomeric and dimeric inserts of $(TA)_5GC(TA)_5$ display a pattern in which the odd adenines on the 3′-side are most reactive, while in $(AT)_5GC(AT)_5$ the even adenines show the strongest cleavage. The reasons for this difference will be considered further in the Discussion but may be related to the identity of the first base after the GC step which is T for the former case and A for the latter.

Figure 5 presents patterns of DEPC modification for fragments containing central CG or GG steps flanked by blocks of alternating AT. An alternating pattern is again evident on both sides of the central CG step in $(TA)_5CG(TA)_5$ (Figure 5A). With the sequence $T(AT)_8CG(AT)_{15}$ (Figure 5B and the fifth panel of Figure 3), an alternating pattern of products is evident on the 3′-side in the $(AT)_{15}$ region which fades out toward the end of the insert. In contrast, DEPC modification is much more even on the 5′-side in the sequence $(AT)_8$. A similar effect is seen with $(AT)_{15}GG(AT)_6$ (Figure 5C) with uneven cleavage at the 3′-end within the sequence $(AT)_6$. Modification on the 3′-side of the GG step is fairly even, except at the 5′-end, which shows an alternating pattern of bands which fades out over the first five AT steps. Once again, those fragments with a T following the GC site show strong modification of even adenines while in those followed by A the even ones are most reactive.

Figure 6 shows histograms of DEPC modification of three fragments containing four central GC base pairs flanked by $(AT)_{10}$ on each side. In each case, a similar alternating cleavage pattern is produced; the data are taken from the last three panels of Figure 3. The first adenine on the 5′-side of the GC pairs is more reactive than the second, while on the 3′-side the second adenine is most reactive. Figure 6C shows results for both strands of $(AT)_{10}CCGC(AT)_{10}$; it can be seen that the reactive adenines are staggered in the 3′-direction across the two strands.

The nature and origin of these unusual DEPC modification patterns for regions of alternating AT will be considered

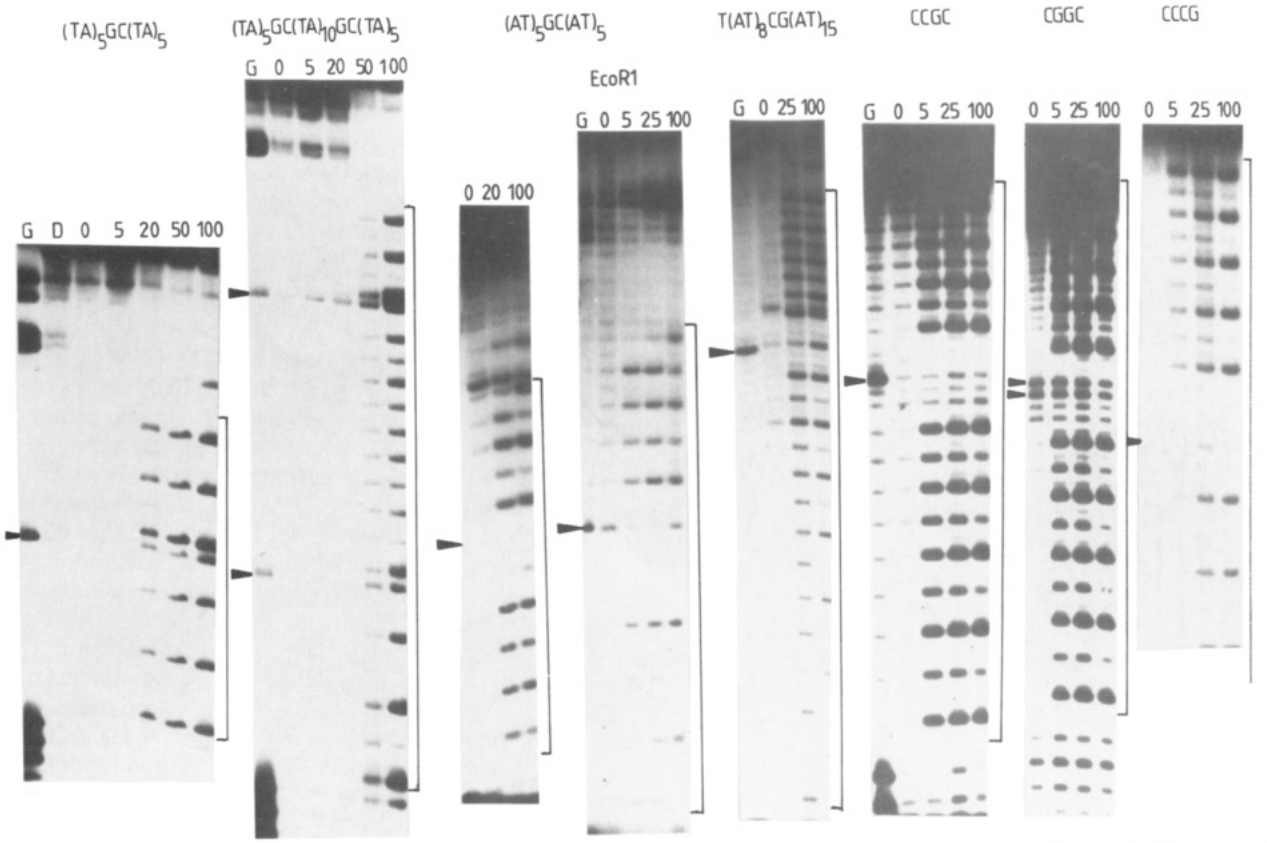

FIGURE 3: Patterns of diethyl pyrocarbonate (DEPC) modification of DNA fragments in the presence of [*N*-MeCys³,*N*-MeCys⁷]TANDEM. Fragments containing the inserts $(TA)_5GC(TA)_5$, $(TA)_5GC(TA)_{10}GC(TA)_5$, $(AT)_5GC(AT)_5$, and $T(AT)_8CG(AT)_{15}$ were obtained by digesting with *Eco*RI and *Hin*dIII and are labeled at the 3′-end of the *Hin*dIII site, except for the fourth panel which is labeled at the 3′-end of the *Eco*RI site. Panels labeled CCGC, CGGC, and CCCG are *Hin*dIII–*Kpn*I fragments containing the inserts $(AT)_{10}CCGC(AT)_{10}$, $(AT)_{10}CGGC(AT)_{10}$, and $(AT)_{10}CCCG(AT)_{10}$, respectively, and are labeled at the 3′-end of the *Hin*dIII site. Micromolar drug concentrations are shown at the top of each lane. Tracks labeled G are dimethyl sulfate–piperidine markers specific for guanine. Square brackets indicate the position and length of the inserts; arrows indicate the position of the central guanines. The track labeled D shows DEPC modification in the presence of 10% dimethyl sulfoxide.

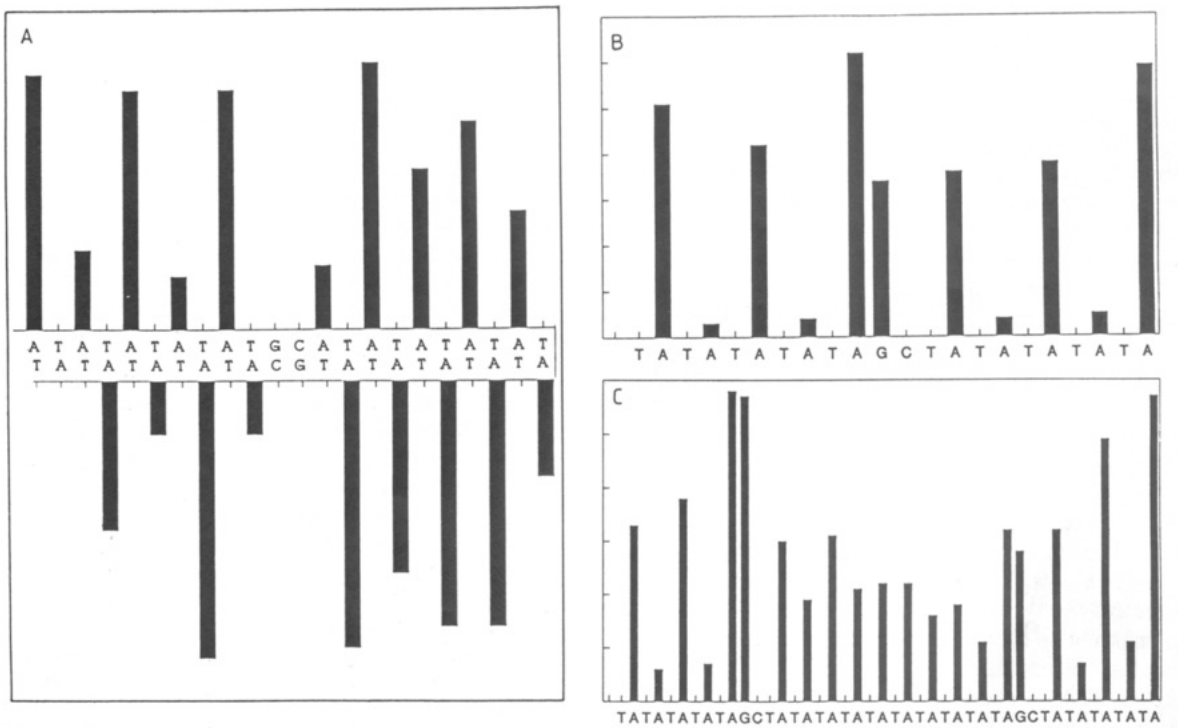

FIGURE 4: Histograms of DEPC modification of the inserts $(AT)_5GC(AT)_5$ (A), $(TA)_5GC(TA)_5$ (B), and $(TA)_5GC(TA)_{10}GC(TA)_5$ (C) in the presence of 100 μM [*N*-MeCys³,*N*-MeCys⁷]TANDEM. Data are taken from densitometer traces of the autoradiographs presented in Figure 3. In (A), the upper strand corresponds to the *Hin*dIII-labeled fragment and is written in usual convention (5′–3′). The lower strand is for the fragment labeled at the 3′-end of the *Eco*RI site and is written from 3′ to 5′.

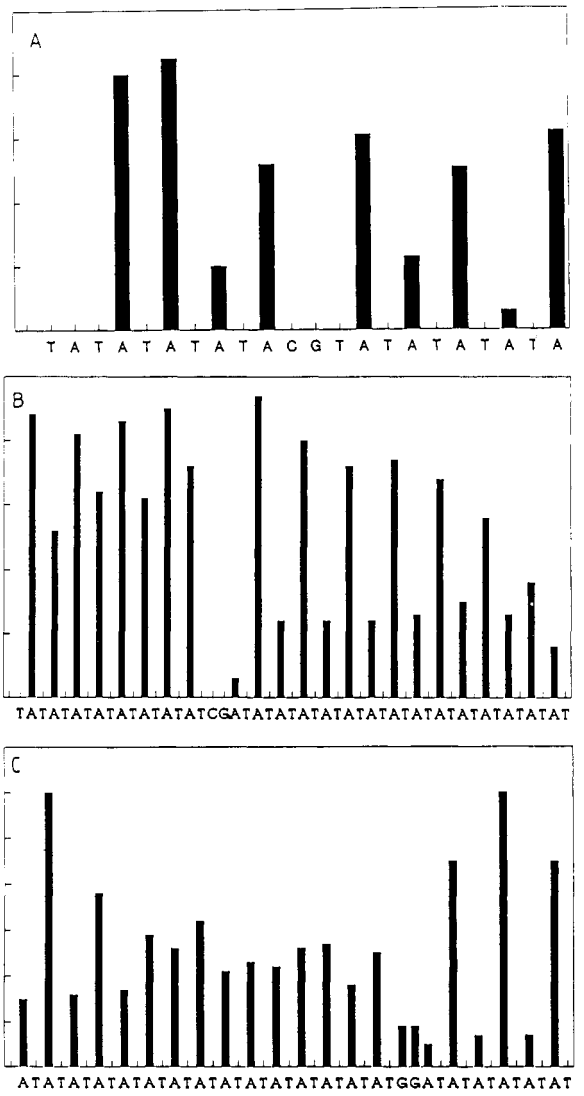

FIGURE 5: Histograms of DEPC modification of fragments containing the inserts $(TA)_5CG(TA)_5$ (A), $T(AT)_8CG(AT)_{15}$ (B), and $(AT)_{15}GG(AT)_6$ (C) in the presence of 100 μM [*N*-MeCys^3,*N*-MeCys^7]TANDEM.

further in the Discussion. In order further to explore the recognition of AT sequences by [*N*-MeCys^3,*N*-MeCys^7]-TANDEM, we have studied its interaction with regions of (ATT)·(AAT) and (TAA)·(TTA).

$(TAA)_4CG(TTA)_4$ and $(ATT)_4CG(AAT)_4$. Figure 7 shows patterns of DNase I cleavage and DEPC modification of fragments containing the inserts $(TAA)_4CG(TTA)_4$ and $(ATT)_4CG(AAT)_4$ in the presence of [*N*-MeCys^3,*N*-MeCys^7]TANDEM. DNase I cleavage throughout the AT sequence is protected by the ligand for both fragments; the only bands remaining within the inserts correspond to cleavage of one or two bases around the central CG steps. This confirms the selectivity for AT-rich regions but reveals little more about the precise selectivity of the ligand. It does, however, reveal that the drug does not have an absolute requirement for alternating AT but can bind well to other sites such as TATT or TTAT.

More information can be gained from the patterns of DEPC modification. From first looking at the dimeric insert of $(TAA)_4CG(TTA)_4$, modification in the control is poor with weak bands evident only for the first adenine of the TAA sequences. In the presence of the drug, there is increased cleavage at every adenine. This is strongest at the second adenine of each of the TAA steps. This is the opposite of the results obtained with echinomycin binding which caused enhanced DEPC modification of the upper adenine (AAT) (Fox et al., 1991). For $(ATT)_4CG(AAT)_4$, cleavage in the control is weak with faint bands evident for the first adenine of the AAT steps. In the presence of the ligand, cleavage of the second, third, and fourth adenines of the AAT steps is greatly enhanced. No enhanced cleavage is evident for the AAT adjacent to the central CG site.

Discussion

Sequence Selectivity. The DNase I footprinting patterns presented in this paper confirm that [*N*-MeCys^3,*N*-MeCys^7]TANDEM binds well to AT-rich DNAs. Each of the DNA fragments show large footprints in the presence of the ligand, extending throughout the AT-rich regions, representing the sum of several overlapping binding sites. These footprints provide little additional information about the sequence selectivity of the compound, but they suggest that the ligand binds best in the center of such regions; bands closer to the central GC steps are only inhibited at higher concentrations. It should be noted that the inhibition of cleavage at central GC sequences does not arise from drug binding to GC sequences but is due to overlap of footprints from the adjacent AT sites, on account of the large size of the probe.

The inability to produce a footprint with the DNA fragment containing the Dickerson dodecamer seems to suggest that the recognition site is not ApT. However, we cannot rule out the possibility that this is due to the nature of the surrounding sequences and that the ligand might bind to an ApT step within a region of strictly alternating AT such as the tetranucleotide TATA (rather than AATT).

Further information about the sequence selectivity can be obtained from the patterns of hyperreactivity to diethyl pyrocarbonate. In most of the DNA fragments investigated, an alternating pattern of hyperreactive adenines is observed; possible reasons for this will be considered below. The pattern of reactivity across the two strands is staggered by one base in the 3′-direction, in contrast to studies with echinomycin binding to regions of alternating AT which gave a 5′-stagger (Fox & Kentebe, 1990; Fox et al., 1991). This 3′-stagger is indicative of ligand binding to TpA, rather than to ApT. Figure 8A illustrates this difference. If we assume that reactive bases are located at adenines adjacent to bound drug molecules, then these will be staggered across the two strands in the 3′-direction for TpA, but in the 5′-direction for ApT. The same conclusion will be obtained if reactive adenines are sandwiched between the intercalated quinoxaline chromophores. The observed 3′-stagger, together with the lack of any footprint with the dodecamer, provides compelling evidence that [*N*-MeCys^3,*N*-MeCys^7]TANDEM binds to TpA rather than to ApT. This is in contrast to the prediction made from the crystal structure of the parent compound TANDEM.

Additional evidence for the preference for TpA comes from the DEPC modification patterns seen with fragments containing the sequences $(TAA)_4CG(TTA)_4$ and $(ATT)_4CG(AAT)_4$. In both cases, the pattern is very different from that seen with echinomycin, for which we have previously argued that secondary specific binding occurs at ApT (Waterloh & Fox, 1991b; Fox et al., 1991). This difference is most pronounced for $(TAA)_4CG(TTA)_4$; in the presence of echinomycin, the upper adenine of the TAA repeats (i.e., TAA) is most reactive, whereas with the TANDEM analogue the lower adenine (TAA) is more reactive. It seems then that [*N*-MeCys^3,*N*-MeCys^7]TANDEM and echinomycin have opposite selectivities for TpA and ApT.

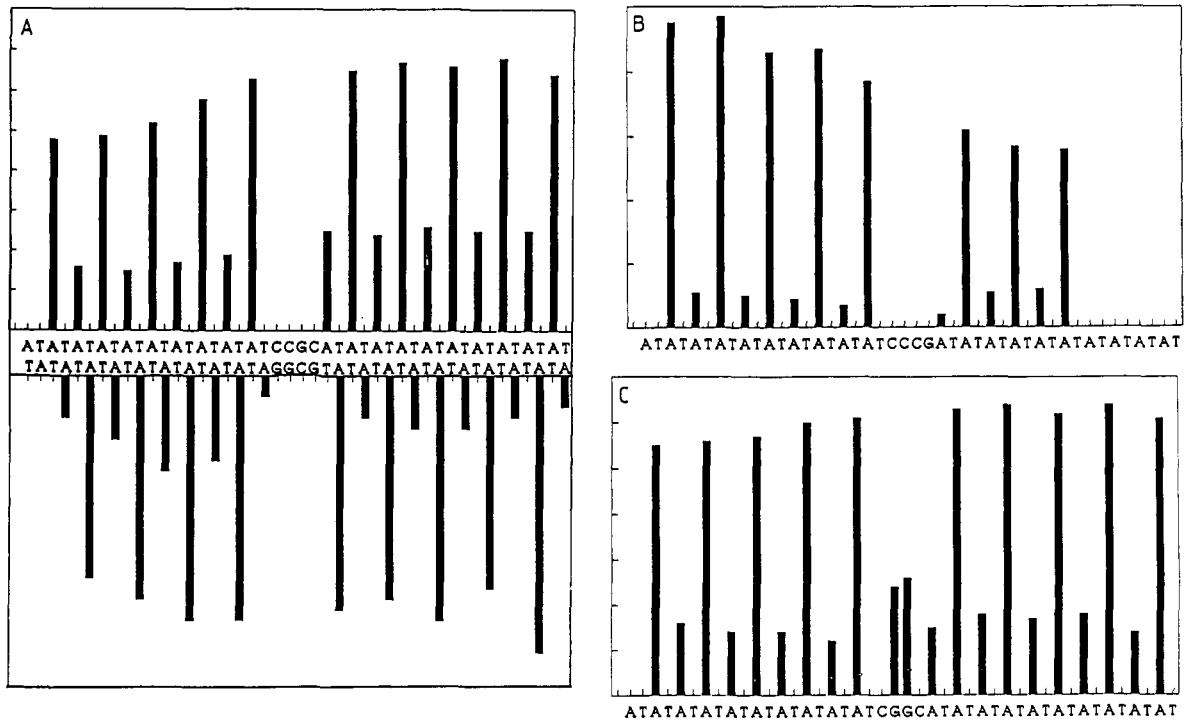

FIGURE 6: Histograms of DEPC modification of fragments containing the inserts $(AT)_{10}CCGC(AT)_{10}$ (A), $(AT)_{10}CCCG(AT)_{10}$ (B), and $(AT)_{10}CGGC(AT)_{10}$ (C) in the presence of 100 μM [N-MeCys³,N-MeCys⁷]TANDEM. In (A), the upper strand (written 5′–3′) is labeled at the 3′-end of the *Hin*dIII site; the lower strand (written 3′–5′) is labeled at the 3′-end of the *Eco*RI site.

Which Adenines are Hyperreactive to DEPC? Previous studies with echinomycin and other natural quinoxaline antibiotics have shown that adenines distal to the CpG binding site (i.e., CGA) are rendered hyperreactive to DEPC (Portugal et al., 1988; McLean & Waring, 1988; Fox & Kentebe, 1990). It therefore seems reasonable to suppose that the same is true for [N-MeCys³,N-MeCys⁷]TANDEM and that the hyperreactive adenines are found adjacent to the sandwiched TpA step. Indeed, this seems to be consistent with the data for $(TAA)_4CG(TTA)_4$ since the second adenine of TAA is rendered hyperreactive (immediately adjacent to the proposed TpA binding site). The same is true for $(ATT)_4CG(AAT)_4$ in which the second adenine of the TAA steps in the second half (i.e., the first adenine of the second, third, and fourth AAT) is rendered hyperreactive to DEPC. However, it is less clear for the fragments containing blocks of alternating AT since each TpA binding site has a thymine on its 3′-side; the adjacent adenine is located on the 5′-side. Previous studies with echinomycin have shown that adenines proximal to echinomycin binding sites (ACG) are not hyperreactive to DEPC (Fox & Kentebe, 1990; Waterloh & Fox, 1991b). In addition there is no a priori reason why the most reactive bases cannot be located at the sandwiched base pairs since the ligand binds from the minor groove, whereas DEPC modifies N7 of adenine in the DNA major groove. We will answer this question by referring to the data for the fragments $(AT)_{10}CCCG(AT)_{10}$, $(AT)_{10}CCGC(AT)_{10}$, and $(AT)_{10}CGGC(AT)_{10}$. In each case, the even-numbered adenines are most strongly modified on the 3′-side of the central block of four GC base pairs. The upper panel of Figure 8B illustrates a possible binding configuration if the modified adenines are located on the 5′-side of the drug bound to TpA. This would involve the binding of four ligand molecules to the $(AT)_{10}$ but cannot explain the enhancement of the final adenine. In contrast, the lower panel of Figure 8B shows the situation if hyperreactive adenines are located at the sandwiched base pairs. This would allow the binding of five ligand molecules and accounts for all the modified bases. In both models, the DNA is fully saturated and all drug molecules are bound at the tetranucleotide step ATAT. There will therefore be almost no difference between the free energy of binding for individual drug molecules in the two models. The binding of an additional ligand, together with the complete explanation of the pattern of hyperreactivity, strongly suggests that the second model is correct. This will be exaggerated if TANDEM binds to such sequences in a highly cooperative fashion. We therefore suggest that the most reactive adenines in the $(AT)_{10}$ sequences correspond to the bases sandwiched between the quinoxaline chromophores.

Explaining the Patterns. The arguments presented above suggest that, in regions of alternating AT, TANDEM binds selectively to the dinucleotide step TpA, increasing the reactivity of the sandwiched adenines to modification by DEPC. Is this model sufficient to explain the results for each of the sequences investigated?

$(AT)_{10}$ Surrounding Four GC Base Pairs. We have already explained the DEPC patterns for the 3′-ends of these fragments. The pattern on the 5′-side is the same for all three fragments with the adenine closest to the GC pairs strongly modified, followed by an alternating pattern of enhancements. This is again consistent with saturation of this region by five ligand molecules, each bound to equivalent TpA steps.

$(AT)_{15}GG(AT)_6$. The fragments considered above each contain an odd number (nine) of adjacent TpA steps; as a result, there is only one way of binding five ligand molecules (i.e., to the first, third, fifth, seventh, and ninth TpA steps). In contrast, $(AT)_{15}GG(AT)_6$ possesses five TpA steps on the 3′-side and 14 on the 5′-side. An alternating pattern of DEPC products is seen on the 3′-side, consistent with saturation by binding three drug molecules at the first, third, and fifth TpA steps. On the 5′-side, there is sufficient room for seven drug molecules, but, because there is an even number of TpA steps, these can be bound in several configurations. The extreme examples position the drug molecules at steps 1, 3, 5, 7, 9, 11,

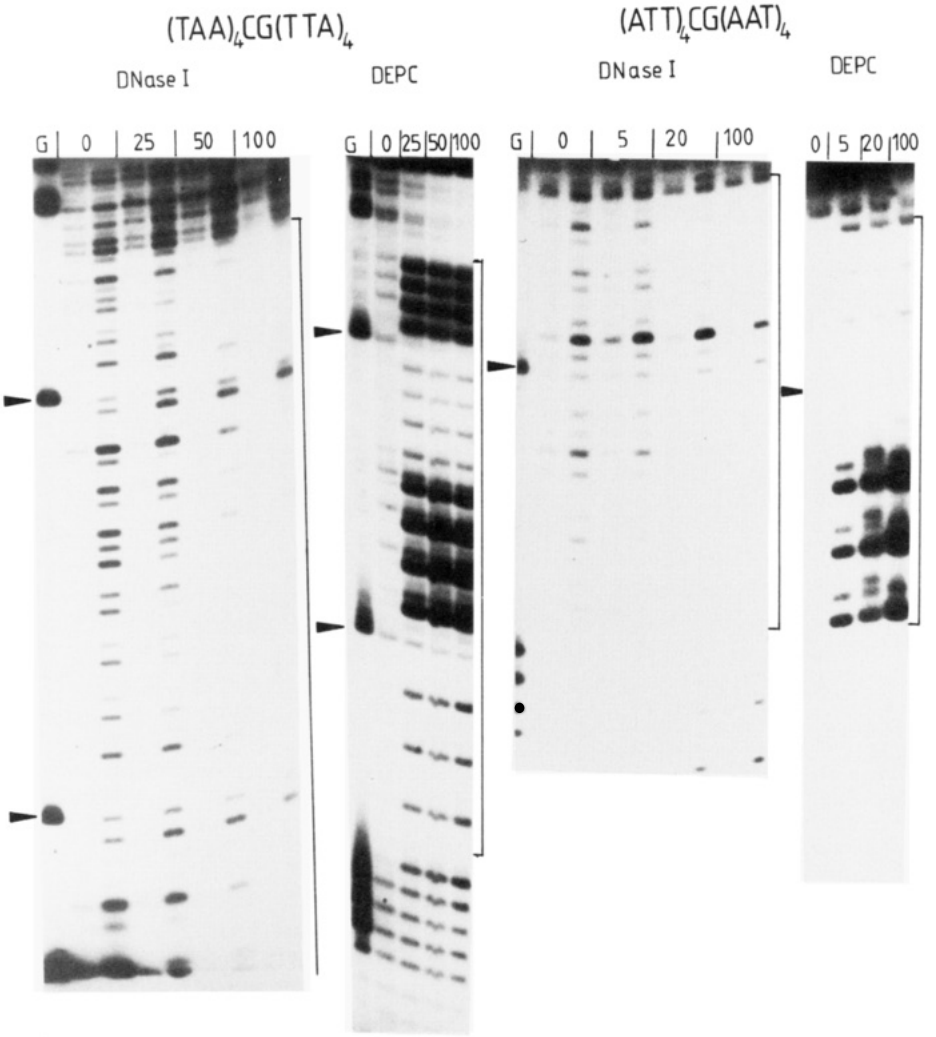

FIGURE 7: Patterns of DNase I digestion and DEPC modification of fragments containing the dimeric insert of $(TAA)_4CG(TTA)_4$ and the monomeric insert of $(ATT)_4CG(AAT)_4$. The former is a *Hin*dIII–*Eco*RI fragment, the latter is a *Hin*dIII–*Sac*I fragment. Both DNAs were labeled at the 3′-end of the *Hin*dIII site. For DNase I, each pair of lanes corresponds to digestion by the enzyme for 1 and 5 mins. Micromolar drug concentrations are shown at the top of the lanes. Square brackets show the position and length of the inserts; arrows indicate the position of the central guanine residues. The tracks labeled G are dimethyl sulfate–piperidine markers specific for guanine.

and 13 or at 2, 4, 6, 8, 10, 12, and 14. Intermediate forms are also possible, with two TpA steps between adjacent molecules, i.e., binding in configurations such as 1, 3, 6, 8, 10, 12, and 14. Since the two extremes are equivalent with each drug molecule bound to the tetranucleotide ATAT, they should be equally occupied, resulting in an even pattern of DEPC enhancements. This is indeed what we observe. However, there is some alternation at the extreme 5′-end, possibly suggesting that the drug molecules line up in phase starting from the end of the sequence.

$(TA)_5GC(TA)_5$. For this fragment an alternating pattern of DEPC products is seen for the monomeric insert. For the dimeric insert, the pattern alternates in the $(TA)_5$ portions, but it is even in the central $(TA)_{10}$. This can be explained in the same way as for $(AT)_{15}GG(AT)_6$. In the $(TA)_5$ sequences, there can only be one way of binding three ligands to the TpA steps (i.e., the first, third, and fifth). However, in the central region of the dimer, with 10 TpA steps, there are several possible configurations, the extremes having drug bound to the odd or even steps. The observed even pattern suggests that these are equally occupied. However, it should be noted that the TpA steps in this fragment are not all equivalent. Those in the center are located in the sequence ATAT, while the two at the ends correspond to CTAT (ATAG). It therefore appears that the drug does not differentiate between CTAT and ATAT, since the first and last adenines are modified to a similar extent.

$(AT)_5GC(AT)_5$. On the basis of the above arguments, we would expect an even pattern of DEPC enhancements on both sides of the central CG since there is an even number (four) of TpA steps. This is indeed what we observe at the right-hand end (Figure 4A). However, an alternating pattern of DEPC modification can be seen at the right-hand end. We have no explanation for this pattern since one would expect that both halves of a symmetrical sequence (cloned into a symmetrical restriction site) would behave identically.

$(TA)_5CG(TA)_5$. This fragment possesses five TA steps on either side of the central CG. As a result, maximal binding is achieved by binding three drug molecules at the first, third, and fifth TpA steps in a unique configuration leading to modification of the sandwiched adenines. This is indeed what we observe. Once again, this must mean that there is little difference between binding at CTAT and ATAT. It is worth remembering that each of these explanations concerns ligand binding at a high concentration (100 μM). The DNase I footprinting data suggest that at low concentrations the ligand binds preferentially to the center of alternating AT sequences, consistent with the more even DEPC modification patterns.

A Model for Binding of TANDEM to TpA. Extrapolation from the crystal structure of TANDEM (Viswamitra et al.,

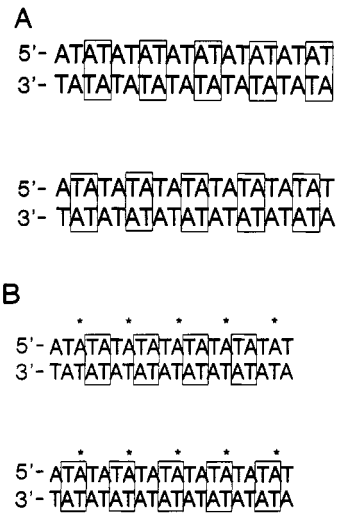

FIGURE 8: (A) Schematic representation of the interaction of the ligand with ApT (upper panel) or TpA (lower panel). The boxes correspond to the drug molecules. (B) Schematic representation for the modification (*) of adenines by DEPC with the ligand bound to the dinucleotide step TpA. The upper panel shows the situation if DEPC modifies adenines situated between drug molecules. The lower panel corresponds to modification of adenines sandwiched between the chromophores of the drug molecule.

1981) led to the suggestion that it would recognize the dinucleotide ApT. We now believe this to be incorrect. What then might be the structural features responsible for interaction with TpA? If we assume that the ligand binds within the DNA minor groove, then a likely hydrogen acceptor would be the 2-keto group of thymine. Comparison with the crystal structures of other quinoxaline antibiotics bound to short oligonucleotides might suggest that the valine NH group is a likely H-bond donor. This is methylated in the naturally occurring quinoxaline antibiotics echinomycin and triostin A thereby preventing them from binding to TpA. However, in the crystal structure of TANDEM, the valine NHs form internal hydrogen bonds to the alanine carbonyls and are presumed not to be available for specific DNA recognition. This is confirmed by a recent NMR study (Addess et al., 1992) which demonstrated that these strong internal hydrogen bonds were retained on binding to DNA. We must therefore look for some other feature of the structure responsible for the selectivity. Addess et al. (1992) suggest that the alanine NHs are involved in the selectivity, consistent with the observation that the TANDEM analogue containing a lactone linkage in this position failed to bind selectively to DNA. They propose that selective binding is achieved by interaction between the alanine NHs and adenine N3. Although this model seems highly plausible, it does not easily explain why TANDEM cannot also recognize CpG (as does echinomycin) by forming similar hydrogen bonds to guanine N3. The most likely reason for this is that there would be interference from the 2-amino group of guanine. A full explanation of this and other factors will have to await the determination of a crystal structure of TANDEM bound to an oligonucleotide.

References

Addess, K. J., Gilbert, D. A., Olsen, R. K., & Feigon, J. (1992) *Biochemistry 31*, 339–350.

Evans, D. H., Lee, J. S., Morgan, A. D., & Olsen, R. K. (1982) *Can. J. Biochem. 60*, 131–136.

Fox, K. R., & Kentebe, E. (1990) *Biochem. J. 269*, 217–221.

Fox, K. R., Marks, J. N., & Waterloh, K. (1991) *Nucleic Acids Res. 19*, 6725–6730.

Fratini, A. V., Kopka, M. L., Drew, H. R., & Dickerson, R. E. (1982) *J. Biol. Chem. 257*, 14686–14707.

Hossain, M. B., van der Helm, D., Olsen, R. K., Jones, P. G., Sheldrick, G. M., Egert, E., Kennard, O., Waring, M. J., & Viswamitra, M. A. (1982) *J. Am. Chem. Soc. 104*, 3401–3408.

Jeppesen, C., & Nielsen, P. E. (1988) *FEBS Lett. 231*, 172–176.

Lee, J. S., & Waring, M. J. (1978a) *Biochem. J. 173*, 115–128.

Lee, J. S., & Waring, M. J. (1978b) *Biochem. J. 173*, 128–144.

Low, C. M. L., Drew, H. R., & Waring, M. J. (1984a) *Nucleic Acids Res. 12*, 4865–4879.

Low, C. M. L., Olsen, R. K., & Waring, M. J. (1984b) *FEBS Lett. 176*, 414–419.

Low, C. M. L., Fox, K. R., Olsen, R. K., & Waring, M. J. (1986) *Nucleic Acids Res. 14*, 2015–2033.

McLean, M. J., & Waring, M. J. (1988) *J. Mol. Recognit. 1*, 138–151.

Mendel, D., & Dervan, P. B. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 910–914.

Portugal, J., Fox, K. R., McLean, M. J., Richenberg, J. L., & Waring, M. J. (1988) *Nucleic Acids Res. 16*, 3655–3670.

Powers, R., Olsen, R. K., & Gorstein, D. G. (1989) *J. Biomol. Struct. Dyn. 7*, 515–556.

Ughetto, G., Wang, A. H. J., Quigley, G. J., van der Marel, G., van Boom, J. H., & Rich, A. (1985) *Nucleic Acids. Res. 13*, 2305–2323.

Van Dyke, M. W., & Dervan, P. B. (1984) *Science 225*, 1122–1127.

Viswamitra, M. A., Kennard, O., Cruse, W. B. T., Egert, E., Sheldrick, G. M., Jones, P. G., Waring, M. J., Wakelin, L. P. G., & Olsen, R. K. (1981) *Nature 289*, 817–819.

Wakelin, L. P. G., & Waring, M. J. (1976) *Biochem. J. 157*, 721–740.

Wang, A. H.-J., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H., & Rich, A. (1984) *Science 225*, 1115–1121.

Waring, M. J., & Wakelin, L. P. G. (1974) *Nature 252*, 653–657.

Waterloh, K., & Fox, K. R. (1991a) *J. Biol. Chem. 266*, 6381–6388.

Waterloh, K., & Fox, K. R. (1991b) *Nucl. Acids Res. 19*, 6719–6724.