DNA Sequence Recognition by the Antitumor Drug Ditercalinium†

Stephen D. G. Crow,‡ Christian Bailly,*,§ Christiane Garbay-Jaureguiberry,|| Bernard Roques,|| Barbara Ramsay Shaw,⊥ and Michael J. Waring*,‡

Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1PD, U.K., INSERM U-524 et Laboratoire de Pharmacologie Antitumorale du Centre Oscar Lambret, IRCL, Place de Verdun, 59045 Lille, France, Département de Pharmacochimie Moléculaire et Structurale, INSERM U266-CNRS UMR 8600, Faculté de Pharmacie, 4, Avenue de l'Observatoire, 75270 Paris, France, and Department of Chemistry, Box 90346, Duke University, Durham, North Carolina 27708-0346

Received December 27, 2001; Revised Manuscript Received April 22, 2002

ABSTRACT: The antitumor drug ditercalinium is a rare example of a noncovalent DNA-binding ligand that forms bisintercalation complexes via the major groove of the double helix. Previous structural studies have revealed that the two connected pyridocarbazolium chromophores intercalate into DNA with the positively charged bis(ethylpiperidinium) linking chain oriented to the wide groove side of the helix. Although the interaction of ditercalinium with short oligonucleotides containing 4−6 contiguous GC base pairs has been examined in detail by biophysical and theoretical approaches, the sequence preference for ditercalinium binding to long DNA fragments that offer a wide variety of binding sites has been investigated only superficially. Here we have investigated both sequence preferences and possible molecular determinants of selectivity in the binding of ditercalinium to DNA, primarily using methods based upon DNase I footprinting. A range of multisite DNA substrates, including several natural restriction fragments and different PCR-generated fragments containing nonconventional bases (2,6-diaminopurine, inosine, uridine, 5-fluoro- and 5-methylcytosine, 7-deazaguanine, 7-deazaadenine, and N^7-cyanoboranoguanine), have been employed to show that ditercalinium selectively recognizes certain GC-rich sequences in DNA and to identify some of the factors which affect its DNA-binding sequence selectivity. Specifically, the footprinting data have revealed that the 2-amino group on the purines or the 5-methyl group on the pyrimidines is not essential for the formation of ditercalinium−DNA complexes whereas the major groove-oriented N^7 of guanine does appear as a key element in the molecular recognition process. The loss of N^7 at guanines but not adenines is sufficient to practically abolish sequence-selective binding of ditercalinium to DNA. Thus, as expected for a major groove binding drug, the N^7 of guanine is normally required for effective complex formation with GC base pairs, but interestingly the substitution of the N^7 with a relatively bulky cyanoborane group does not markedly affect the sequence recognition process. Therefore, the hydrogen bond accepting capability at N^7 of guanines is not sufficient to explain the GC-selective drug−DNA association, and the implications of these findings are considered.

Ditercalinium (Figure 1) is an antitumor 7*H*-pyridocarbazole dimer which bisintercalates into DNA (*1*, *2*). This synthetic drug is most unusual among nonconvalent DNA-binding agents because it binds within the major groove of the double helix. Both X-ray and NMR studies of ditercalinium complexed to the tetranucleotide $d(CGCG)_2$ showed that each pyridocarbazolium ring was able to intercalate into the contiguous CpG steps with the positively charged bis(ethylpiperidinium) linking chain oriented to the wide groove side of the helix (*3*−*6*). This atypical arrangement distinguishes ditercalinium from other dimeric intercalating agents such as the quinoxaline antibiotics (echinomycin, triostin A, TANDEM) which all gain access to the DNA bases via the minor groove. There is one other ligand, elinafide, which is believed to bisintercalate into DNA from the major groove (*7*). In addition, the bisintercalation of ditercalinium into DNA provokes specific structural changes including a significant bend in the helix axis by about 15° toward the minor groove (*6*). Ditercalinium-bound DNA is a privileged target for enzymes involved in locating and repairing damaged sites within large domains of DNA, in particular for the nucleotide excision repair system (*8*, *9*). For this reason, it has been hypothesized that the curious distortion of the DNA conformation induced by ditercalinium may be specifically recognized by the (A)BC exonuclease excision repair system (*6*).

While it is clear from the NMR and crystallographic structural studies that ditercalinium binds avidly to $d(CGCG)_2$ to form bisintercalation complexes, the binding mode is not always homogeneous. The interaction of the drug with DNA

† This work was supported by grants (to M.J.W.) from the Cancer Research Campaign, the Wellcome Trust, the European Union, and the Sir Halley Stewart Trust; (to C.B.) from the Association pour la Recherche sur le Cancer, Grant 4221; and (to B.R.S.) from the National Cancer Institute, Grant CA60139.

* Address correspondence to either of these authors. M.J.W.: e-mail, mjw11@cus.cam.ac.uk. C.B.: e-mail, bailly@lille.inserm.fr.

‡ University of Cambridge.

§ INSERM U-524.

|| Université René Descartes.

⊥ Duke University.

10.1021/bi012207q CCC: $22.00
Published on Web 06/08/2002

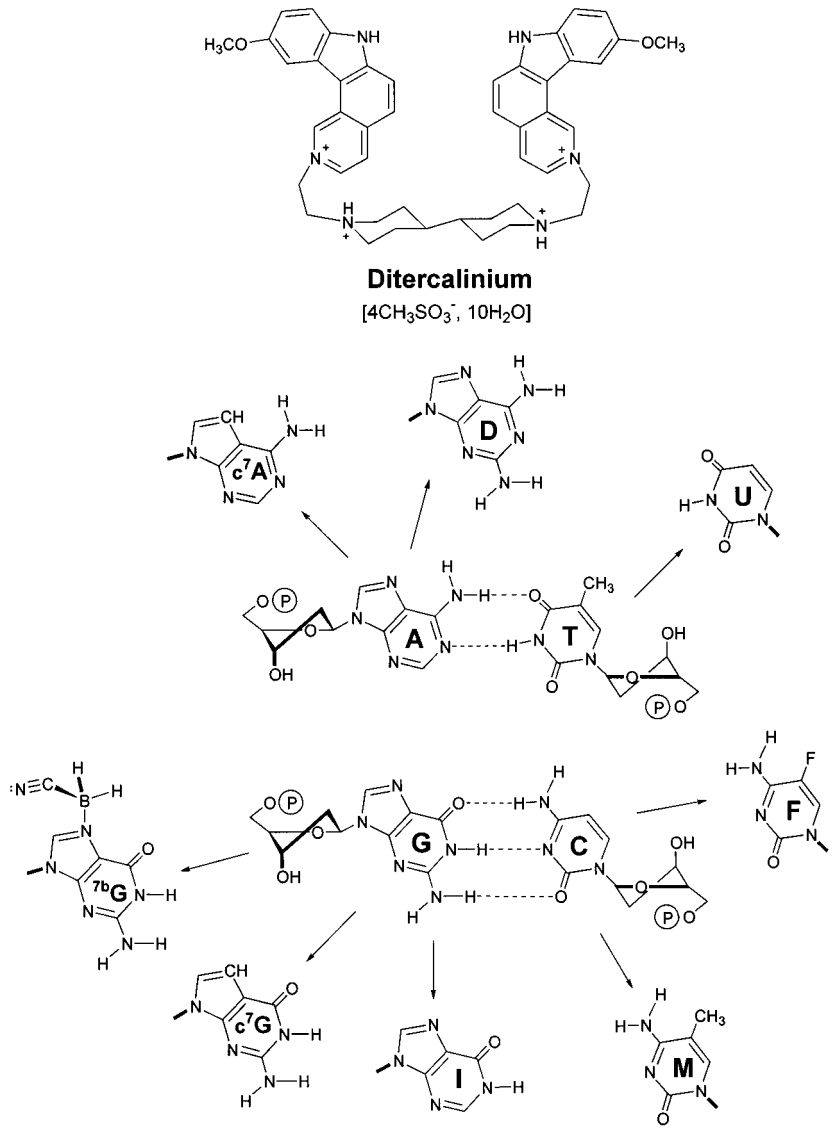

FIGURE 1: Structure of ditercalinium and base pairs. Abbreviations: A, adenine; T, thymine; G, guanine; C, cytosine; D, diaminopurine; U, uracil; I, inosine; M, 5-methylcytosine; F, 5-fluorocytosine; c^7A, 7-deazaadenine; c^7G, 7-deazaguanine; ^{7b}G, *N*^{7}-cyanoboranoguanine.

is tight, with a binding affinity greater than 10^8 M^{-1} (*2*), but ditercalinium can bind to different types of sequences without necessarily forming bisintercalation complexes. For example, upon binding to $d(GCGC)_2$ one of the pyridocarbazole rings intercalates at the CpG step while the other lies stacked upon an external base pair (*10*). The same study also revealed that ditercalinium can interact with the octanucleotide sequence $d(CCTATAGG)_2$ but does not form any specific complexes (*10*). Using DNase I footprinting methodology, Mendoza et al. (*11*) reported that ditercalinium does not protect any specific sequences in random B-DNA, but it does protect GC-rich and mixed sequence regions in a bent DNA fragment from *Crithidia fasciculata*. Thus it is still unclear whether ditercalinium binds to specific sequences containing CpG steps or preferentially recognizes a particular conformation that tends to be associated with those sequences. A recent molecular modeling analysis of ditercalinium bound to $d(GCGCGC)_2$ suggested that the drug might actually recognize and stabilize the A-type DNA conformation that this hexanucleotide is capable of adopting (*12*).

These considerations prompted us to investigate further the binding of ditercalinium to DNA using recently developed complementary technical approaches providing information on both the sequence selectivity of drug binding to DNA and its molecular determinants. First, we applied the standard footprinting methodology that has been extensively used over many years to investigate DNA recognition by small molecules. Several DNA restriction fragments representing a total of about 700 bp[1] were employed to identify a panel of ditercalinium binding sequences, most of which were found to contain contiguous G·C base pairs but not necessarily arranged with alternating purines and pyrimidines. Next, we looked in greater detail at the possible molecular contacts between ditercalinium and G·C base pairs. We synthesized by PCR a series of 160 bp DNA molecules containing various types of nonconventional bases (Figure 1). Specifically, we varied the position of the 2-amino group on the purines and modified the region about the N^7 of purines and the 5-methyl group on the pyrimidines, as represented in Figure 1. In addition, because the binding of ditercalinium to DNA presumably involves interaction with guanine residues in the major groove of the double helix

[1] Abbreviations: bp, base pair; TBE, Tris−borate−EDTA.

(*6*), we investigated the binding of the drug to DNA containing 7-deazapurine nucleotides or N^7-cyanoboranoguanine along the lines of experiments performed several years ago with echinomycin (*13*). The use of DNA substrates containing designed structural modifications has enabled us to evaluate the influence of specific atomic groups of the DNA bases upon the interaction with ditercalinium and to identify the significant role of the N^7-position of the deoxyguanosine nucleotides. These findings provide important mechanistic insight into the action of ditercalinium.

MATERIALS AND METHODS

Drug. The synthesis and analytical characterization of ditercalinium [1,1′-bis[2-(10-methoxy-7*H*-pyrido[4,3-*c*]carbazolium)ethyl-4,4′-bipiperidine]dimethanesulfonate dihydromethanesulfonate] have been reported previously (*2*). The drug was first dissolved in water at 5 mM and then further diluted with water. The stock solution was kept at −20 °C and freshly diluted to the desired concentration immediately prior to use.

Deoxyribonucleoside Triphosphates. Ultrapure dATP, dCTP, dGTP, dITP (2′-deoxyinosine 5′-triphosphate), dTTP, dUTP (2′-deoxyuridine 5′-triphosphate), 5-Me-dCTP (5-methyl-2′-deoxycytosine 5′-triphosphate), and c^7-dGTP (c^7-deaza-2′-deoxyguanosine 5′-triphosphate) nucleotides were from Pharmacia. Similarly, ultrapure c^7-dATP nucleotide (c^7-deaza-2′-deoxyadenosine 5′-triphosphate) was from Boehringer Mannheim, while 5-F-dCTP was from Sierra Bioresearch (Tucson, AZ). ^{7b}dGTP nucleotide (N^7-cyanoborano-2′-deoxyguanosine 5′-triphosphate) was synthesized in the laboratory of B. Ramsay Shaw by the method of Porter et al. (*14*), starting from ^{7b}dG (*15*). dDTP (2,6-diaminopurinosyl 2′-deoxy-5′-triphosphate) was synthesized by Dr. M. Guo (ArQule Inc., Waltham, MA). Nucleoside triphosphates labeled with ^{32}P (γ-ATP, 6000 Ci/mmol; α-ATP, 6000 Ci/mmol) were obtained from NEN.

Biochemicals. Restriction endonucleases (Boehringer), *Taq* polymerase (Promega), DNase I (Sigma), and T4 polynucleotide kinase (Promega) were used according to the suppliers' recommended protocols in the activity buffers provided. The PCR primers 5′ GTTACCTTTAATCCGTTACG (Watson primer) or 5′ GGGCTCGGGAACCCCCACCA (Crick primer) having a 5′-OH were synthesized by the Protein and Nucleic Acid Facility in the Department of Biochemistry, University of Cambridge, U.K.

Preparation and Purification of DNA Fragments. Plasmid pKMp27 (*16*) containing the *Eco*RI−*Eco*RI 169 bp *tyrT* fragment (courtesy of Dr. A. A. Travers, MRC Laboratory of Molecular Biology, Cambridge, U.K.) was isolated from *Escherichia coli* by a standard sodium dodecyl sulfate−sodium hydroxide lysis procedure and purified by banding in CsCl−ethidium bromide gradients. Ethidium was removed by several 2-propanol extractions followed by exhaustive dialysis against Tris−EDTA buffer. The purified plasmid was then precipitated and resuspended in appropriate buffer prior to restriction enzyme digestion. The 169 base pair *tyrT*(A93) fragment used as a template was cut out of the plasmid with *Eco*RI. It was purified by electrophoresis on a nondenaturing 6% polyacrylamide gel. Similar procedures were used to prepare the 3′-end-labeled DNA fragments of 117 bp and 265 bp (*Eco*RI−*Pvu*II digest of plasmid pBS) and the 178 bp fragment (*Eco*RI−*Pvu*II digest of plasmid pLAZ3) (*17*).

Primer Labeling. The 5′-OH Watson or Crick primer was 5′-end labeled with [γ-^{32}P]ATP in the presence of T4 polynucleotide kinase according to a standard procedure for labeling oligonucleotides (*18*).

Polymerase Chain Reaction (PCR). Each reaction mixture contained 1 ng of *tyr*T(A93) DNA template, 0.1 μM each of the appropriate pair of primers (one with a 5′-[^{32}P]-phosphate and one with a 5′-NH_2-terminal group), 250 μM each appropriate dNTP (dATP or dDTP or c^7-dATP; dCTP or 5-Me-dCTP or 5-F-dCTP; dGTP or dITP or c^7-dGTP or ^{7b}dGTP; and TTP or dUTP according to the desired DNA), approximately 2 mM $MgCl_2$, and 1 unit of *Taq* polymerase in a volume of 50 μL containing 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 0.1% Triton X-100, and 1.5 mM $MgCl_2$. To prevent unwanted primer-template annealing before the cycles began, the reactions were heated to 60 °C before addition of the *Taq* polymerase (*19*). Finally, paraffin oil was added to each reaction to prevent evaporation. Samples containing the requisite nucleotides to produce any of the desired DNA species, except those containing either dITP or c^7-dATP, were heated to 94 °C for 3 min. Following this initial denaturation step, 30 amplification cycles were performed, each consisting of 94 °C denaturation for 1 min, 45 °C primer annealing for 2 min, and polymerization at 72 °C for 10 min. After the last amplification cycle was completed, the PCR samples were heated to 72 °C for another 10 min to complete unfinished polynucleotide chains. Following this step the reaction mixtures were slowly cooled to facilitate complete annealing of complementary DNA strands, including two additional 5 min steps, one at 55 °C and one at 37 °C. PCR tubes containing either dITP or c^7-dATP had to be put through a protocol at lower temperatures, as the standard conditions failed to yield product. This alternative PCR program consisted of a 84 °C denaturing step, 45 °C primer annealing step, and 62 °C polymerization step comparable to that used previously to incorporate 7-deazapurine or inosine residues (*13*, *20*). The number of cycles and cycling time as well as the last two annealing steps were identical to the PCR program used for natural nucleotides. The reaction mixtures were then extracted with chloroform to remove the mineral oil, and parallel reactions were pooled. Several extractions with water-saturated 1-butanol were performed to reduce the volume prior to loading the samples onto a 6% nondenaturing polyacrylamide gel. Following electrophoresis for about 90 min, the gel was exposed to X-ray film for about 5 min. After development of the film the bands were located, excised, crushed, and soaked overnight in TN buffer (pH 7.4) at 37 °C. This suspension was filtered through a Millipore 0.22 μm filter, and the DNA was precipitated with ethanol. Following washing with 70% ice-cold ethanol and air-drying of the precipitate, the purified DNA was resuspended in TN buffer (pH 7.4) to a concentration of 4800 cpm/μL (series 900 minimonitor).

DNase I Footprinting and Electrophoresis. DNase I experiments were performed essentially according to the original protocol of Low et al. (*21*) whereby 2 μL of the labeled DNA fragment was added to 4 μL of the appropriate concentration of drug in TN buffer (pH 7.4) or TN buffer alone and incubated for 30 min at room temperature. The

digestion of the samples was then initiated by addition of 2 μL of a DNase I solution whose concentration had been adjusted to yield a final enzyme concentration of about 0.01 unit/mL in the reaction mixture. Thus, the extent of digestion was limited to less than 30% of the starting material so as to minimize the incidence of multiple cuts in any strand ("single-hit" kinetic conditions). Optimal enzyme dilutions were established in preliminary calibration experiments. After 6 min the digestion was stopped by placing samples on dry ice. The samples were then lyophilized, washed once with 50 μL of water, lyophilized again, and finally resuspended in 4 μL of an 80% formamide solution containing tracking dyes. Samples were heated at 90 °C for 4 min and chilled on ice for 4 min prior to electrophoresis. DNA cleavage products were resolved by polyacrylamide gel electrophoresis under denaturing conditions (0.3 mm thick, 8% acrylamide containing 8 M urea). After electrophoresis (about 2 h at 70 W, 1600 V, in TBE buffer), gels were soaked in 10% acetic acid for 15 min, transferred to Whatman 3 MM paper, dried under vacuum at 80 °C, and placed next to a PhosphorImager storage screen overnight.

Data Reduction. A Molecular Dynamics 425E PhosphorImager was used to collect data from the storage screens exposed to the dried gels overnight at room temperature. Baseline-corrected scans were analyzed by integrating all of the densities between two selected boundaries using ImageQuant for the Macintosh version 1.2 software. The area under each peak was integrated by simple addition of the pixels under the curve. Peak areas were stored on a disk as an ASCII file and subsequently transferred into standard spreadsheet programs (Cricket Graph) using a Macintosh personal computer. In general, the data set consisted of about 90 band positions for each DNA containing normal or modified bases. Peak areas were determined in duplicate and were averaged. In the first stage, peak areas were corrected for slight differences in the total amount of DNA loaded per lane by multiplying the peak area by a specific correction factor determined from the sum of all band intensities in a given lane (see ref *22* for details). These individual corrections yield the "corrected" or "normalized" band intensities which are used to construct the differential cleavage plots in which the data are presented in the form $\ln(f_a) - \ln(f_c)$ representing the differential cleavage at each bond relative to that in the control (f_a is the fractional cleavage at any bond in the presence of the drug, and f_c is the fractional cleavage of the same bond in the control). The results are displayed on a logarithmic scale for the sake of convenience; positive values indicate enhanced cleavage whereas negative values indicate cleavage inhibition (*23*). Each resolved band was assigned to a particular bond within the respective fragment by comparison of its position relative to standard Maxam and Gilbert sequencing standards (G and G + A tracks).

RESULTS

Footprinting Experiments with DNA Restriction Fragments. Three DNA restriction fragments of 117, 265, and 178 base pairs, all 3′-end labeled, were used as substrates for the endonuclease DNase I, which is a sensitive enzyme for mapping DNA-binding sites of small molecules (*24*). The procedure for the preparation, labeling, and purification of the DNA is given in Materials and Methods. A set of typical autoradiographs of the sequencing gels used to fractionate the products of partial digestion of each DNA fragment complexed with increasing concentrations of ditercalinium is presented in Figure 2. The drug markedly affects the cutting of DNA by the nuclease to produce clear footprints. There is thus no doubt that ditercalinium is a sequence-selective binder. Numerous bands in the drug-containing lanes are weaker than the same bands in the drug-free lane, corresponding to attenuated cleavage, while others display relative enhancement of cutting. A densitometric analysis of the autoradiographs is presented in Figure 3. Regions of attenuated DNA cleavage can be discerned around positions 36, 55, 74, 98 (117-mer), 34, 42, 71, 87, 101 (265-mer), and 28, 45, 60, 99 (178-mer). These footprints all coincide with the positions of nucleotide sequences having a high GC content, such as 5′-CGGCC, 5′-CCGC, and 5′-CACCCC. However, several binding sequences encompass the dinucleotide steps CA-TG such as the sequence ACACAGG around position 29 on the 178-mer. The footprints are detectable at low drug concentrations, well below 1 μM, as expected from the high affinity of the drug for double-stranded DNA. Many can be detected at 0.2 μM and become stronger as the drug concentration is raised to 0.5 μM. At higher concentrations, all cleavage of the DNA by the enzyme was abolished (data not shown).

Upon binding of ditercalinium, certain sequences are rendered more susceptible to attack by DNase I than in the control. These regions of drug-induced enhanced cleavage all correspond to AT-rich sequences. For example, DNase I cutting at the sequences 5′-AAAA and 5′-AAATTAA around positions 55 and 76 on the 265-mer is markedly increased in the presence of ditercalinium (Figure 2). Such cleavage enhancement at certain AT sites (but not all) is commonly seen with intercalating drugs and has been attributed to intercalation-induced perturbations of the double helical structure of DNA (*22*−*26*). However, there are two other possible explanations for enhanced nuclease cutting: (a) attractive interaction between the nuclease and the drug, leading to locally increased DNase I concentration around the drug binding sites, and (b) a mass-action effect, due to drug-induced displacement of the nuclease away from the drug-binding sites and resultant increase of the effective nuclease concentration at other sites not recognized by the drug (*27*, *28*). This latter explanation is considered unlikely because it would lead to extensive regions of enhanced cutting all along the DNA, essentially unrelated to nucleotide sequence, whereas what is observed is quite marked enhancement of cleavage at particular sequences which often lie closely flanking the evident binding sites of the ligand. The former explanation would require all enhancements to occur at flanking sequences, probably more or less equally noticeable around all binding sites, which manifestly does not occur either.

In the following set of experiments we took advantage of the availability of substituted DNA species to investigate the molecular determinants of the sequence selectivity of ditercalinium.

Footprinting Experiments with DNA Containing Inosine and/or Diaminopurine. The 2-amino group of deoxyguanosine nucleotides exposed in the minor groove of the DNA double helix represents the main molecular determinant of preferred binding to specific sequences for numerous DNA intercalating agents and groove binders. For example, the

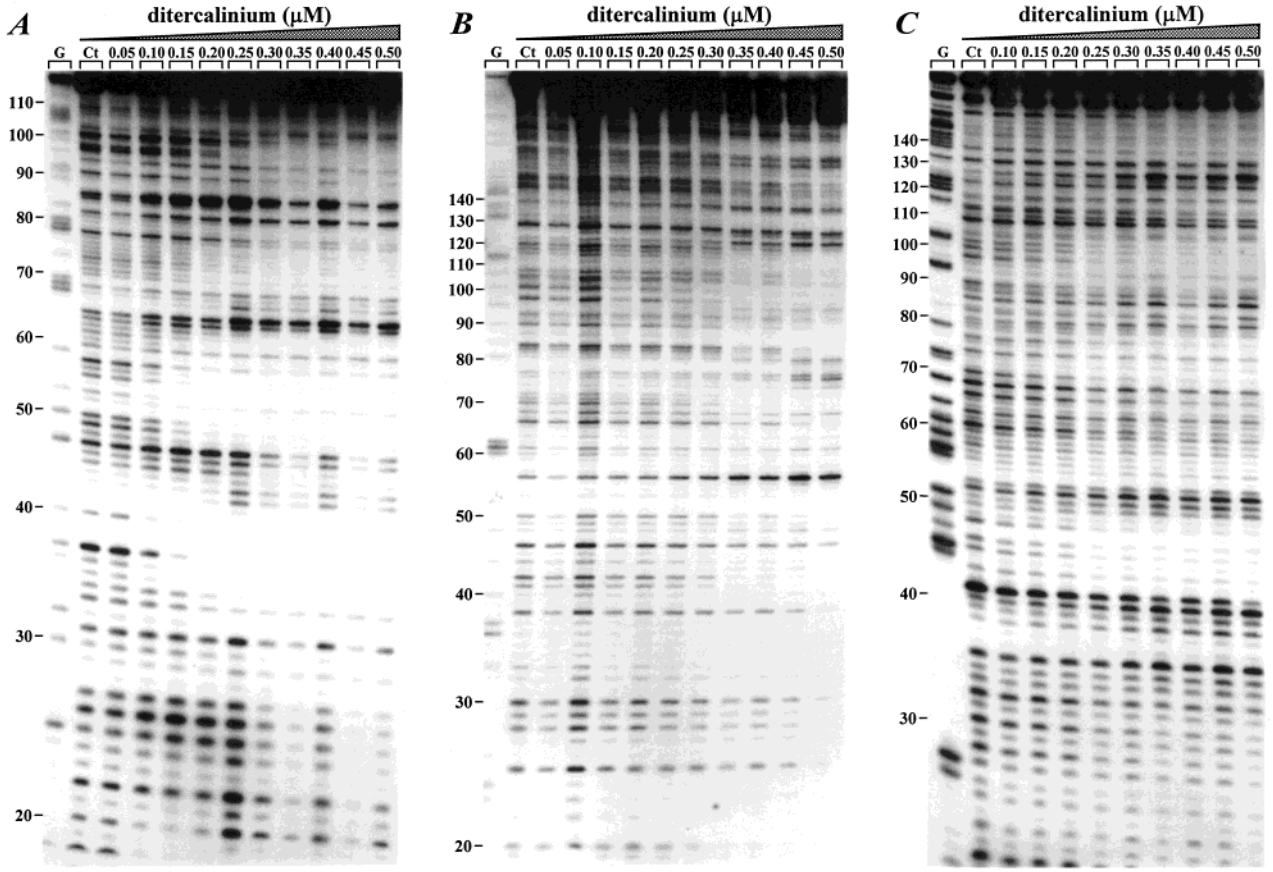

FIGURE 2: Sequence-selective binding of ditercalinium. The gels show DNase I footprinting with three DNA restriction fragments of (A) 117, (B) 265, and (C) 178 base pairs. The 117-mer and 265-mer *Pvu*II−*Eco*RI fragments were cut from the plasmid pBS. The 178-mer *Eco*RI−*Pvu*II fragment was obtained from plasmid pUC12. In each case, the DNA was 3′-end labeled at the *Eco*RI site with [α-^{32}P]dATP in the presence of AMV reverse transcriptase. The products of nuclease digestion were resolved on an 8% polyacrylamide gel containing 7 M urea. The concentration (μM) of the drug is shown at the top of the appropriate gel lanes. Control tracks (Ct) contained no drug. Guanine-specific sequence markers obtained by treatment of the DNA with dimethyl sulfate followed by piperidine were run in the lanes marked G. Numbers on the left side of the gels refer to the standard numbering scheme for the nucleotide sequence of the DNA fragment.

location of the 2-amino group of purine nucleotides determines to a large extent the position of binding sites at AT-rich sequences for the conventional minor groove binder distamycin and at GC-rich sequences for mithramycin (*29*). Similarly, the guanine 2-amino group exerts a dominant influence over the GC-preferred binding of bisintercalators belonging to the quinoxaline family such as echinomycin (*20*). We sought to investigate the influence of the guanine 2-amino group on the binding of ditercalinium to DNA using the same PCR-based strategy developed previously (*30*). Four homologous 160 base pair *tyr*T DNA fragments were synthesized by PCR amplification, each containing either natural bases or inosine residues in place of guanosines (G → I substitution) or 2,6-diaminopurine residues (henceforth abbreviated DAP or D within a sequence for clarity) in place of adenines (A → DAP substitution) or both I and DAP residues. In each case, primers in which the 5′-terminal nucleotide residue bore a 5′-[^{32}P]phosphate or a 5′-NH_2-terminal group were used so as to enable selective labeling of one or the other strand in the PCR product, i.e., the Watson (antisense) strand or the Crick (sense) strand chosen at will. The cleavage patterns of the various modified DNAs were analyzed and compared to normal DNA using DNase I. Similar methodology was used to incorporate the other modified bases (U, M, F, c^{7}A, c^{7}G, ^{7b}G) as discussed below.

Both the A → DAP and G → I base substitutions were found to have relatively little effect on the binding of ditercalinium to DNA. As shown in Figure 4, addition of ditercalinium to *tyr*T DNA containing the canonical base pairs (normal DNA) prevents DNase I from cleaving particular sequences which appear as footprints at the GC-rich sequences around nucleotide positions 75 and 98. The same footprints were detected whether the DNA substrate was labeled on one or the other of the complementary strands (data not shown). Neither the replacement of all guanosines by inosine residues nor the replacement of all adenines by diaminopurine residues abolishes the sequence-specific binding process. The extent of binding is slightly reduced with the DAP DNA, but still a footprint can be detected at the GC-rich sites. These data suggest that the 2-amino group is not a primary element of DNA sequence recognition by ditercalinium.

Footprinting Experiments with DNA Containing Uridine Nucleotides or 5-Fluoro- or 5-Methylcytosines. Next, using the same PCR-based approach, we prepared a series of DNA fragments in which the 5-methyl group pointing into the major groove of the double helix had been either removed from thymines (T → U substitution) or added to cytosine residues (C → M substitution). An example of a footprinting gel obtained with the uridine-containing DNA is shown in

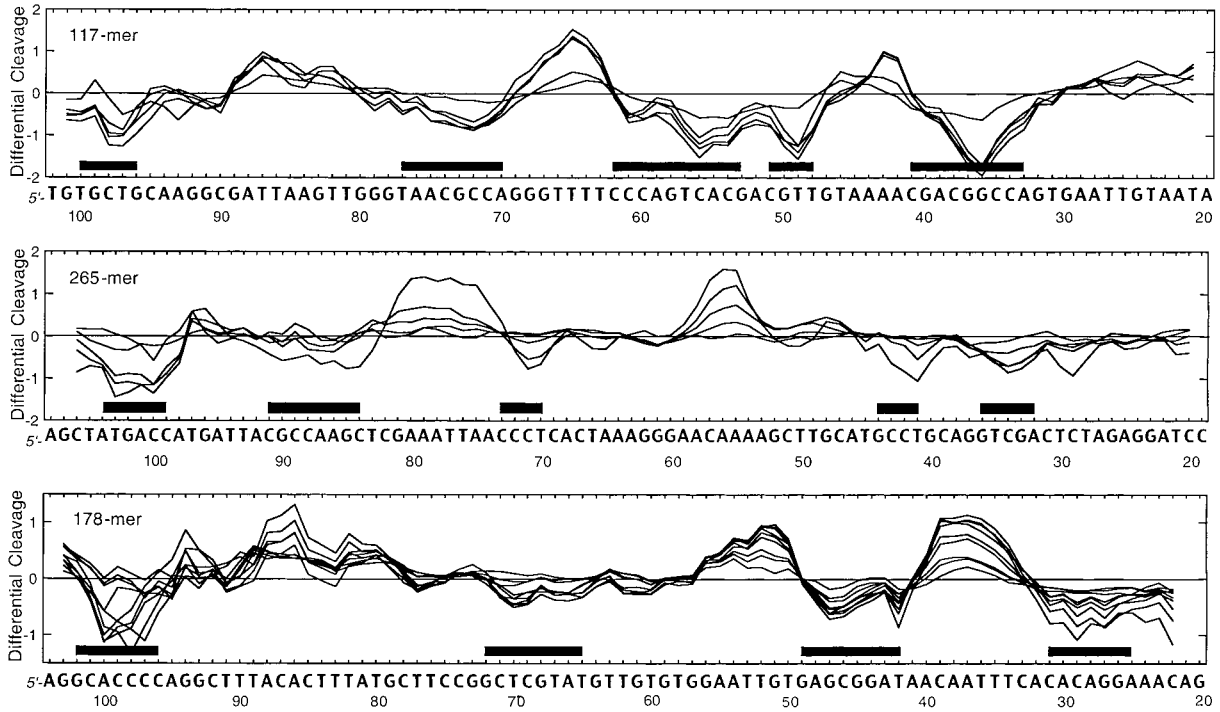

FIGURE 3: Differential cleavage plots comparing the susceptibility of the 117-, 265-, and 178-mer DNA fragments to DNase I cutting in the presence of increasing concentrations of ditercalinium. Negative values correspond to a ligand-protected site, and positive values represent enhanced cleavage. Vertical scales are in units of $\ln(f_a) - \ln(f_c)$, where f_a is the fractional cleavage at any bond in the presence of the drug and f_c is the fractional cleavage of the same bond in the control, given closely similar extents of overall digestion. Each line drawn represents a three-bond running average of individual data points, calculated by averaging the value of $\ln(f_a) - \ln(f_c)$ at any bond with those of its two nearest neighbors. Only the region of the restriction fragments analyzed by densitometry is shown. Black boxes indicate the positions of inhibition of DNase I cutting in the presence of ditercalinium.

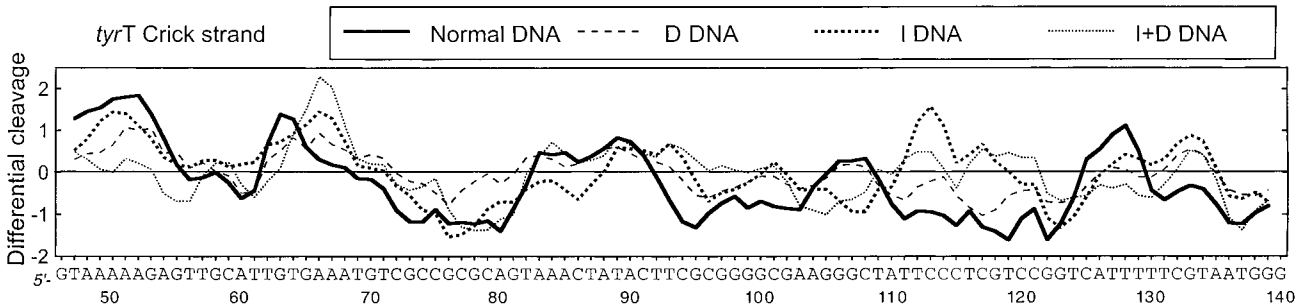

FIGURE 4: Differential cleavage plots comparing the DNase I-mediated cleavage of the Crick (sense) strand of normal, inosine (I), diaminopurine (D), and I + DAP *tyr*T(A93) DNA in the presence of ditercalinium at 0.25 μM. Note that the sequence printed on the *x*-axis corresponds to that of natural DNA, though in the modified DNAs adenine and/or guanosine residues are replaced by diaminopurine and/or inosine residues. Other details as for Figure 3.

Figure 5. Here again, we found that ditercalinium binds almost equally well to the normal and U-containing DNA. The footprints remain located at the same GC-rich sequences, and the intensity of the footprints is only slightly reduced. The plots in Figure 6 reveal that the drug interacts with roughly equal affinity at the sites 5′-AACT·TTGA and 5′-UCAA·UUGA at positions 54−57. The C_{50} values, which approximate to dissociation constants for binding to individual sites, are 0.95 and 1.48 μM for 5′-AACT·TTGA and 5′-UCAA·UUGA, respectively. The DNase I cleavage at the adjacent 5′-TTTTT site at positions 48−52 remains significantly enhanced when the thymine nucleotides are replaced by uridine (Figure 6). Similar footprinting experiments were performed with a DNA substrate in which cytosine residues were replaced by 5-methyl-C (named M in Figure 1), and we also used 5-fluoro-dCTP to prepare a DNA having all cytosines replaced by 5-fluoro-dC residues. The results of DNase I footprinting experiments performed with these DNA substrates were comparable to those obtained with the U-DNA. The autoradiograph (Figure S1) and the corresponding differential cleavage plots in Figure 7 indicate that binding of ditercalinium to GC-rich sequences is not abolished by the incorporation of a major groove methyl or fluoro substituent on C residues. Clear footprints are visible at positions 97, 104, 119, and 124 though their relative intensity differs depending on the type of DNA examined. Thus, the various substitutions do appear to have some effect on the binding of ditercalinium, but probably on its relative affinity for certain sites rather than any marked shift in sequence selectivity. Differing intensities of footprinting between various types of substituted DNA molecules have already been noted at individual sites on the gels shown previously. But what makes this experiment significant is what happens at all four of the footprints analyzed for the

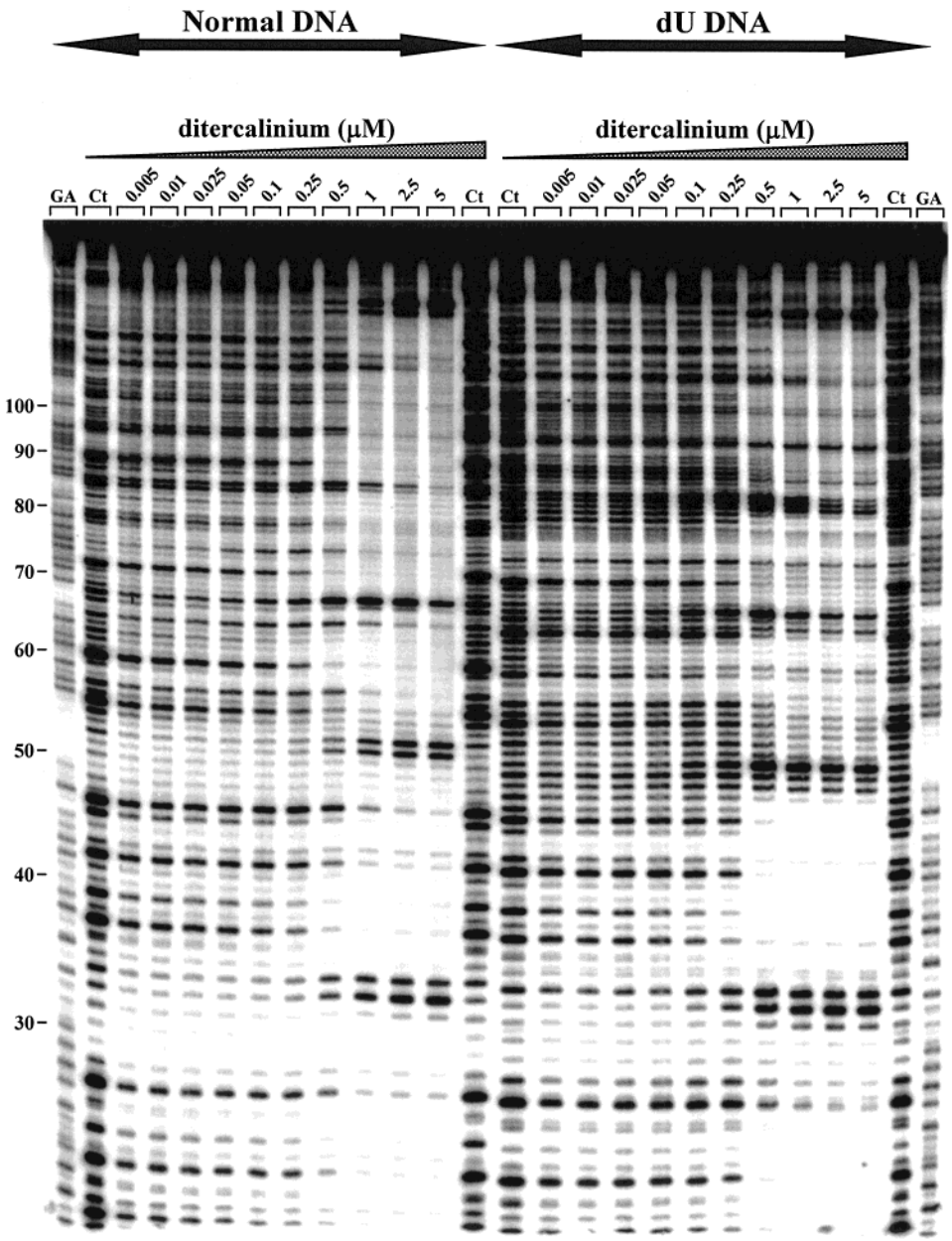

FIGURE 5: Autoradiograph of a high-resolution denaturing gel showing DNase I footprinting of ditercalinium on the Watson (antisense) strand of normal and uridine-containing *tyr*T(A93) DNA. Guanine-specific sequence markers obtained by treatment of the DNA with formic acid followed by piperidine were run in the lanes marked GA. Other details as for Figure 2.

drug and the consistency in the intensity of footprinting depending on the type of DNA used. It appears that the substitution of the proton normally present at the C5 of cytosines with fluorine causes a significant decrease in the affinity of ditercalinium for its preferred binding sites. Conversely, the substitution of the C5 proton with a methyl group, to create what is normally considered methylated DNA, generally has the effect of increasing the affinity of the drug at the same sites.

Footprinting Experiments with DNA Containing N^7-Modified Purines. Whereas modifications within the major groove at the C5 position of pyrimidines seemed limited to altering the binding affinity of ditercalinium, we found clear differences in the sequence specificity of the drug using DNA containing alterations at the N^7-position of guanine. Figure 8 depicts a footprinting experiment performed with ditercalinium on the Watson (antisense) strand of *tyr*T DNA containing normal and N^7-modified purines. The results are striking and immediately apparent. Both normal and deaza-dA-containing DNA molecules show signs of concentration-dependent footprinting at many points along the whole length of the molecule, whereas deaza-dG-containing DNA does not show any real signs of drug binding at all. Evidently, the loss of N^7 at guanines but not adenines is sufficient to practically abolish sequence-selective binding of ditercalinium to DNA. This observation is reinforced by the differential cleavage plots in Figure 9 where the footprinting of ditercalinium at 0.5 μM on normal and deaza-dG DNA is compared. The drug produces its most prominent footprints at positions 38 and 54 and between positions 70 and 80 (a particularly GC-rich area) on normal DNA. But the profile is clearly not mimicked with deaza-dG DNA. Ditercalinium shows no evidence of footprinting at all on the unnatural DNA molecule, with few areas registering barely perceptible

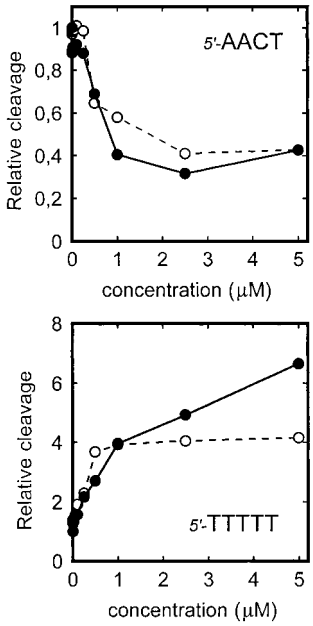

FIGURE 6: Footprinting plots for the binding of ditercalinium to sites 5′-AACT and 5′-TTTTT. The relative cleavage intensity corresponds to the ratio I_c/I_0, where I_c is the intensity of the band at the ligand concentration c and I_0 is the intensity of the same band in the absence of ditercalinium. In the upper panel, symbols ● and ○ refer to cleavage at the sequence 5′-AACT·TTGA and 5′-UCAA·UUGA, respectively. In the lower panel, symbols ● and ○ refer to cleavage at the sequence 5′-TTTTT·AAAAA and 5′-UUUUU·AAAAA, respectively.

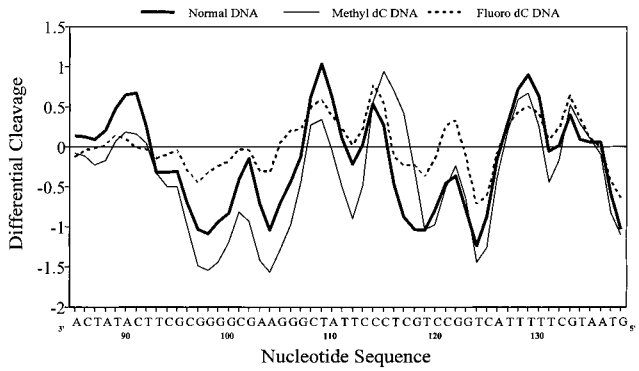

FIGURE 7: Differential cleavage plots comparing the DNase I-mediated cleavage of the Crick (sense) strand of *tyr*T(A93) DNA containing either normal or C-5 modified cytosines in the presence of 2.5 μM ditercalinium. The plot drawn as a thick continuous line refers to normal *tyr*T(A93) DNA. The plots drawn as a thin continuous line and a dashed line refer to the modified *tyr*T(A93) fragments containing either 5-methylcytosine residues (methyl dC DNA) or 5-fluorocytosine residues (fluoro dC DNA), respectively. Other details as for Figure 3.

negative values for differential cleavage (which in any case tend to be exaggerated because of the logarithmic scale of the ordinate).

Similar experiments were carried out with the Crick (sense) strand of *tyr*T DNA labeled instead and including the use of cyanoborano-dG-containing DNA. The results of the footprinting are summarized in Figure 10, which depicts the sequence-dependent effect of 0.5 μM ditercalinium on normal DNA and all three unnatural N^7-modified DNA molecules. As expected, the drug produces strong footprints on the normal DNA at most of the GC-rich sites, most notably around positions 78, 96, 103, 119, and 124. Comparison with

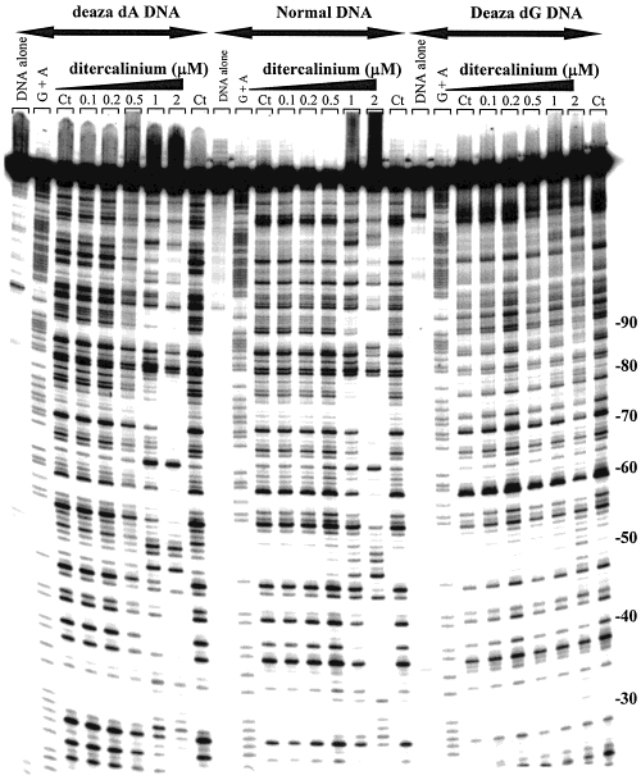

FIGURE 8: DNase I footprinting on the Watson (antisense) strand of *tyr*T(A93) DNA containing either normal or N^7-modified purines in the presence of increasing concentrations of ditercalinium. Deaza-dA and deaza-dG DNA molecules contain N^7-deazaadenine and N^7-deazaguanine bases in place of adenine and guanine, respectively. Guanine-specific sequence markers obtained by treatment of the DNA with formic acid followed by piperidine were run in the lanes marked GA. DNA alone represents samples that have not been in the presence of either ditercalinium or DNase I. Other details as for Figure 2.

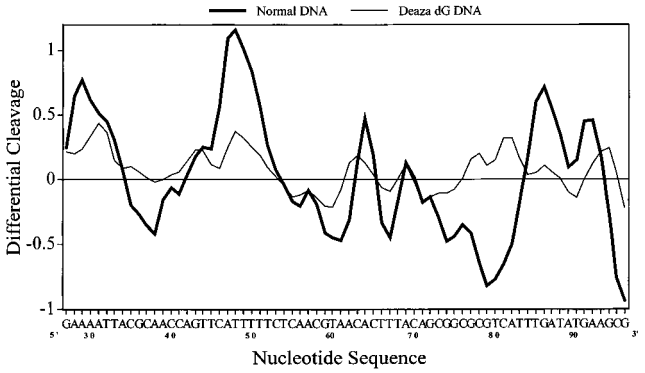

FIGURE 9: Differential cleavage plots comparing the DNase I-mediated cleavage of the Watson (antisense) strand of *tyr*T(A93) DNA containing either normal or N^7-modified purines in the presence of 0.5 μM ditercalinium. The plot drawn as a thick continuous line refers to normal *tyr*T(A93) DNA. The plot drawn as a thin continuous line refers to the modified *tyr*T(A93) fragment containing N^7-deazaguanine residues (deaza dG DNA). Other details as for Figure 3.

the portion of the nucleotide sequence, positions 60−96, that is also present in the cleavage plots for the labeled Watson strand is interesting. Looking at the footprinting profile of ditercalinium on normal DNA, there is good agreement across the two strands with regard to binding sites, particularly positions 78 and 79, and areas of enhanced cleavage, with the occurrence of a slight 3′ shift also apparent. The

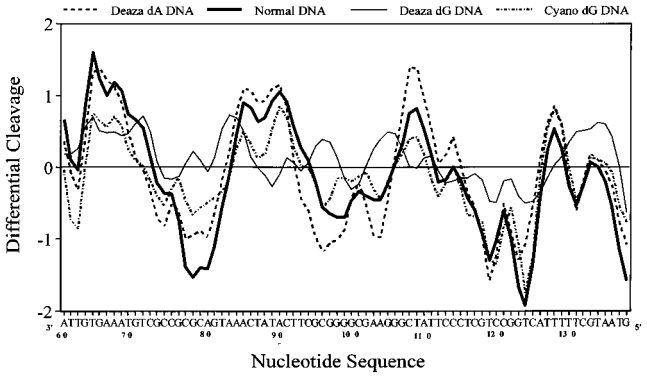

FIGURE 10: Differential cleavage plots comparing the DNase I-mediated cleavage of the Crick strand of *tyr*T(A93) DNA containing either normal or N^7-modified purines in the presence of 0.5 μM ditercalinium. The plot drawn as a thick continuous line refers to normal *tyr*T(A93) DNA. The plots drawn as a thin continuous line, a simple dashed line, and a complicated dashed line refer to the modified *tyr*T(A93) fragments containing either N^7-deazaguanine residues (deaza dG DNA) or N^7-deazaadenine residues (deaza dA DNA) or N^7-cyanoboranoguanine residues (cyano dG DNA), respectively. Other details as for Figure 3.

fact that the drug binds in the major groove should not influence the 3′ shift that was originally noted for minor groove-implicated drugs (*21*), as it is a function of the cleavage agent rather than the drug under investigation (*31*). Once again, the footprinting profile of ditercalinium on the deaza-dA-substituted DNA is remarkably similar to that of the normal DNA. Clear footprints occur at the same positions already noted above, with exact homology evident at the principal footprinting sites around positions 119 and 124. Interestingly though, the footprinting profile of ditercalinium on the cyanoborano-dG-containing molecules is also remarkably similar to that seen with normal DNA. This result is somewhat unexpected, as it was anticipated that this type of unnatural DNA bearing a bulky group on guanines in the major groove would probably behave similarly to deaza-dG DNA. As with deaza dG DNA, an N^7-cyanoborano-dG DNA also lacks the hydrogen bond accepting ability at guanine position N^7 (*14*) that was suggested by earlier workers as a key factor for ditercalinium binding (*6*). Once again though, there is no evidence of sequence-selective binding of the drug to deaza-dG DNA. No footprints are visible around positions 78, 96, or 103, and the possible attenuation of nuclease cleavage at sites 119 and 124 is weak and questionable.

These results totally corroborate those observed with the complementary Watson strand. It appears that the simple loss of the N^7 at guanines, but not adenines, is sufficient to block the sequence-selective intercalation of ditercalinium into the major groove. However, the situation is complicated by the results with cyanoborano-dG DNA, which also lacks the hydrogen bond accepting ability at N^7 of guanines, yet is still able to provide a binding profile for the drug similar to that seen with normal DNA. The implications and possible explanations of these findings are considered in the Discussion.

DISCUSSION

Although the interaction of ditercalinium with DNA has been studied for a number of years, the factors which affect its DNA-binding sequence selectivity are still very poorly understood. To gain insight into the structural basis of this drug−DNA recognition process, we studied the interaction of ditercalinium with a range of specially designed DNA fragments containing several types of modified bases (Figure 1). The methodology of DNase I footprinting, extensively exploited over the past 20 years, provided a tool perfectly suited to this study (*30*, *32*). The first set of footprinting experiments performed with three restriction fragments showed clearly that ditercalinium interacts preferentially with GC-rich sequences and discriminates against homooligomeric runs of A and T, which are often cut more readily by the enzyme in the presence of the drug compared to the control. This conclusion is in perfect agreement with equilibrium dialysis data (not shown) and also agrees satisfactorily with previous observations (*11*). However, this is the first time that the GC preference of ditercalinium has been firmly established using complementary approaches based on real DNA molecules. Most of the previous DNA-binding studies on ditercalinium were performed with short oligonucleotides, only rarely with long DNA fragments that can offer a large range of binding sites such as those used here.

The second set of footprinting experiments performed with DNA samples containing inosine and/or 2,6-diaminopurine provided more detailed information. The data in Figure 4, suggesting that the 2-amino group is not a primary element of DNA sequence recognition by ditercalinium, are notably reminiscent of results previously reported with another bisintercalating drug thought to bind to the major groove of DNA, the antitumor agent elinafide (*7*). Accordingly, we next prepared a series of DNA samples containing uridine or 5-fluoro- or 5-methylcytosine bases to probe the importance of exocyclic substituents on the pyrimidines exposed in the major groove of double helical DNA. The incorporation of a fluorine atom at the C5 position of cytosines appears to reduce significantly the affinity of ditercalinium for its preferred binding sites whereas the incorporation of a methyl group tends to increase the affinity of the drug at the same sites (Figures 5−7). At first sight a simple explanation for these observations might be sought on steric grounds. However, this would be erroneous since the order of increasing van der Waals radius for the three substituents is H < F < Me, whereas the order of increasing affinity appears to be F < H < Me. Clearly, there is no parallel between the two properties. However, a consideration of the electron-donating properties of the three substituents, whereby Me > H > F, does produce a correlation. The various substitutions on cytosine will lead to alterations in the dipole moments, base stacking, and molecular polarizability of G·C base pairs (*33*), which in turn could alter the slide component of base-step conformation (*34*−*36*). In other work reported from this laboratory concerning the winding of substituted DNA molecules around the histone octamer, it was suggested that because the methyl group acts as an electron donor, the effect of adding it to the C5 of cytosines would be to decrease the magnitude of the large dipole present on a G·C base pair, in turn favoring a narrower minor groove (*37*). This would produce a compensatory widening of the major groove, directly attributable to methylation of the cytosines at C5. By extension of this argument, fluorination at C5 of cytosines should cause an increase in the dipole moment, because of the extreme electronegativity of the fluorine atom, and consequently create a narrower major groove. These interpretations will account satisfactorily for

the results on the basis that methylated DNA with its wider major groove would provide easier access or a better fit for improved ditercalinium intercalation, while the opposite holds true for the fluoro-substituted DNA molecules.

The results of the final series of footprinting experiments performed with purine N^7-modified DNA species are of particular interest insofar as they relate to the role of a classically important base pair constituent that is known to be a prime target for the binding of various small molecules to DNA, not least the bifunctional alkylating mustard antitumor drugs, though of course they react to form covalent adducts. It now appears that the N^7 of guanine behaves as some kind of digital regulator for the binding of ditercalinium. Removal of this atom, as in *N*^{7}-deaza-G, is sufficient to abolish the sequence selective binding of the drug, and so far as we are aware, this is the first time such a requirement has been described for a noncovalent interaction with a small molecule. However, the exact origin of this influence is not entirely clear. The almost identical footprinting profiles of ditercalinium on normal and cyanoborano-dG DNA appear inconsistent with the simple idea that sequence selectivity can be attributed to specific hydrogen bonding between the linker and the N^7 of guanines. Molecular dynamics calculations made by Gago and colleagues have delineated the relative importance of hydrogen bonding and stacking interactions in the sequence specificity of ditercalinium (*12*). These authors have espoused the concept that intercalating chromophores are not just simple hydrophobic plates that get sandwiched indiscriminately between the base pairs but, instead, confer additional elements of binding specificity (*35*, *36*). In conformity with this view they attributed the groove preference and sequence selectivity of ditercalinium primarily to its chromophores rather than its linker. The rationale behind their work furnishes useful insight into possible interpretations of the present results.

Specifically, the belief that the chromophores may be endowed with sequence selectivity relies upon the good complementarity of their molecular electrostatic potentials (MEP's) with those of G·C base pairs, particularly at the N^7/O^6 region of guanine. This concept may well explain the footprinting results with cyanoborano-dG DNA. The unnatural ^{7b}G·C base pair in this modified DNA lacks the hydrogen bond accepting ability at the N^7 of guanine, as does deaza-dG DNA. But it should still retain a MEP more similar to that of a normal G·C base pair than does the c^7G·C base pair of deaza-dG DNA. In deaza-dG DNA, the replacement of the N^7 atom of guanine with a much less electronegative carbon atom should alter the MEP quite considerably compared to a normal G·C base pair. This is in contrast to merely using the lone pair of electrons on N^7 in a dative bond to another group, as occurs with cyanoborano-dG DNA, which in theory should not disrupt the MEP so severely. So, if it is indeed the case that the chromophores, and not the linker, are dominant in recognition, then the presumptive dissimilarity in MEP's between the base pairs in cyanoborano-dG and deaza-dG DNA molecules could account for the differing sequence selectivity of ditercalinium. But the question remains whether the MEP's of base pairs containing guanine or deazaguanine (which still possesses a significantly electronegative O^6 atom) are sufficiently dissimilar to effectively abolish all drug binding to deaza-dG DNA. Calculation of MEP's for the altered G·C base pairs would evidently be of interest.

The original idea that it is simple hydrogen bonding between linker and guanines that determines sequence specificity cannot be entirely discounted, especially since some hydrogen bond accepting capability could be retained within a cyanoborane group. For example, it is possible for the cyano nitrogen to form a hydrogen bond with one hydrogen in the amino group of cytosine in an N^{7b}G·C pair (*37*). Although hydrogen bonds cannot be formed in the usual way with the N^7 of cyanoborane-substituted guanines, we have considered whether the terminal cyano nitrogen retains an accessible lone pair of electrons (though the cyanoborane group blocks the N^7 lone pairs).

Modifications were introduced at the C^5 of cytosine in the expectation that the sequence selectivity of ditercalinium might be affected through steric hindrance, but they led to an appreciation of how the electrostatic properties of base pairs could affect the affinity of drug binding via attenuation of groove width. To our knowledge this is the first time an effect of this nature has been considered with a small molecule. However, it is right to question whether these differences in drug binding could be explained in relation to changes in the MEP's of the base pairs, as considered above for the N^7 modifications. It has been shown that the presence of a methyl group at the C^5 of cytosines is sufficient to increase the electron density at the N^2 atom of its complementary guanine base (*38*). However, the N^2 atom is involved in donating a hydrogen bond to the cytosine base, whereas the N^7-position is not. Thus the electron density at the N^7 atom will not be affected to the same degree as N^2. The N^7 (and O^6) atom(s) is (are) supposed to provide the best points of MEP overlap between the drug and G·C base pairs. Hence alterations in MEPs are unlikely to account for the differences in affinity of ditercalinium for binding sites in DNA containing C^5-modified cytosines.

Finally, we are led to question whether the differences in binding of ditercalinium elicited by the C^5 substitutions might be explicable on a similar basis to the effects observed with the N^7 modifications. There have been no published experiments examining whether the loss of the N^7 of guanine can affect groove width, although it has been shown to cause destabilization of the duplex (*39*). Applying similar thinking to that of Buttinelli et al. (*40*), one would expect loss of the N^7 of guanine to decrease the dipole moment of a G·C base pair and thus lead to a wider major groove. This invites comparison with the C^5-substituted DNA where a wider major groove, caused by 5-methylcytosine, seemingly enhanced the affinity of ditercalinium for its binding sites. Yet we found that ditercalinium displays little or no affinity for deaza-dG DNA. Thus the altered groove width argument cannot really be applied to the results seen with N^7 modifications. In conclusion, neither of the two explanations advanced to account for the differences in the binding characteristics of ditercalinium seen with either guanine-N^7- or cytosine-C^5-modified DNA molecules are sufficient to explain all of the data. Perhaps further experiments with a wider range of substituted nucleotides would shed light on the mechanisms that determine binding site selection by this remarkable drug molecule.

ACKNOWLEDGMENT

We thank Dean Gentle and Julie Morgan for outstandingly expert technical assistance.

SUPPORTING INFORMATION AVAILABLE

One figure showing DNase I footprinting of ditercalinium on the *tyr*T(A93) DNA containing either normal or C5-modified cytosines. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Roques, B. P., Pelaprat, D., Le Guen, I., Porcher, G., Gosse, C., and Le Pecq, J. B. (1979) *Biochem. Pharmacol. 28*, 1811−1815.
2. Pelaprat, D., Delbarre, A., Le Guen, I., Roques, B. P., and Le Pecq, J. B. (1980) *J. Med. Chem. 23*, 1336−1343.
3. Delbarre, A., Delepierre, M., D'Estaintot, B. L., Igolen, J., and Roques, B. P. (1987) *Biopolymers 26*, 1001−1033.
4. Delbarre, A., Delepierre, M., Garbay, C., Igolen, J., Le Pecq, J. B., and Roques, B. P. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2155−2159.
5. Gao, Q., Williams, L. D., Egli, M., Rabinovich, D., Chen, S. L., Quigley, G. J., and Rich, A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2422−2426.
6. Williams, L. D., and Gao, Q. (1992) *Biochemistry 31*, 4315−4324.
7. Bailly, C., Braña, M., and Waring, M. J. (1996) *Eur. J. Biochem. 240*, 195−208.
8. Lambert, B., Jones, B. K., Roques, B. P., Le Pecq, J. B., and Yeung, A. T. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6557−6561.
9. Lambert, B., Roques, B. P., and Le Pecq, J. B. (1988) *Nucleic Acids Res. 16*, 1063−1078.
10. Delepierre, M., Milhe, C., Namane, A., Dinh, T. H., and Roques, B. P. (1991) *Biopolymers 31*, 331−353.
11. Mendoza, R., Markovits, J., Jaffrezou, J. P., Muzard, G., and Le Pecq, J. B. (1990) *Biochemistry 29*, 5035−5043.
12. de Pascual-Teresa, B., Gallego, J., Ortiz, A. R., and Gago, F. (1996) *J. Med. Chem. 39*, 4810−4824.
13. Sayers, E. W., and Waring, M. J. (1993) *Biochemistry 32*, 9094−9107.
14. Porter, K., Tomasz, J., Huang, F., Sood, A., and Ramsay Shaw, B. (1995) *Biochemistry 34*, 11963−11969.
15. Sood, A., Spielvogel, B. F., and Ramsay Shaw, B. (1989) *J. Am. Chem. Soc. 111*, 9234−9235.
16. Drew, H. R., Weeks, J. R., and Travers, A. A. (1985) *EMBO J. 4*, 1025−1032.
17. Goossens, J. F., Bouey-Bencteux, E., Houssin, R., Hénichart, J. P., Colson, P., Houssier, C., Laine, W., Baldeyrou, B., and Bailly, C. (2001) *Biochemistry 40*, 4663−4671.
18. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular cloning, A laboratory manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY.
19. Bloch, W. (1991) *Biochemistry 30*, 2735−2747.
20. Marchand, C., Bailly, C., McLean, M. J., Moroney, S. E., and Waring, M. J. (1992) *Nucleic Acids Res. 20*, 5601−5606.
21. Low, C. M., Drew, H. R., and Waring, M. J. (1984) *Nucleic Acids Res. 12*, 4865−4879.
22. Bailly, C., Hamy, F., and Waring, M. J. (1996) *Biochemistry 35*, 1150−1161.
23. Fox, K. R., and Waring, M. J. (2001) *Methods Enzymol. 340*, 412−430.
24. Bailly, C., and Waring, M. J. (1995) *J. Biomol. Struct. Dyn. 12*, 869−898.
25. Fox, K. R., and Waring, M. J. (1984) *Nucleic Acids Res. 12*, 9271−9285.
26. Fox, K. R., and Kentebe, E. (1990) *Nucleic Acids Res. 18*, 1957−1963.
27. Ward, B., Rehfuss, R., Goodisman, J., and Dabrowiak, J. C. (1988) *Nucleic Acids Res. 16*, 1359−1369.
28. Goodisman, J., and Dabrowiak, J. C. (1992) *Biochemistry 31*, 1058−1064.
29. Bailly, C., Payet, D., Travers, A. A., and Waring, M. J. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 13623−13628.
30. Bailly, C., and Waring, M. J. (2001) *Methods Enzymol. 340*, 485−502.
31. Fox, K. R. (1997) in *Drug−DNA Interaction Protocols* (Fox, K. R., Ed.) Vol. 90, pp 1−22, Humana Press, Totowa, NJ.
32. Fox, K. R., and Waring, M. J. (2001) *Methods Enzymol. 340*, 412−430.
33. Sowers, L. C., Ramsay Shaw, B., and Sedwick, W. D. (1987) *Biochem. Biophys. Res. Commun. 148*, 790−794.
34. Hunter, C. A. (1993) *J. Mol. Biol. 230*, 1025−1054.
35. Gallego, J., Luque, F. J., Orozco, M., and Gago, F. (1994) *J. Biomol. Struct. Dyn. 12*, 111−129.
36. Gallego, J., Pascual-Teresa, A. R., Pisaborro, M. T., and Gago, F. (1995) in *QSAR and Molecular Modelling: Concepts, Computational Tools and Applications* (Sanz, F., Giraldo, J., and Manaut, F., Eds.) pp 274−281, Prous Science Publishers, Barcelona.
37. Banks, B. (1992) Ph.D. Dissertation, Duke University, Durham, NC.
38. Das, A., Tang, K. S., Gopalakrishnan, S., Waring, M. J., and Tomasz, M. (1999) *Chem. Biol. 6*, 461−471.
39. Seela, F., and Driller, H. (1986) *Nucleic Acids Res. 14*, 2319−2332.
40. Buttinelli, M., Minnock, A., Panetta, G., Waring, M. J., and Travers, A. A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 8544−8549.

BI012207Q